Implicit Knowledge, Explicit Knowledge, and Achievement in Second Language (L2) Spanish

Xavier Gutiérrez University of Windsor

Abstract

Implicit and explicit knowledge of the second language (L2) are two central constructs in the field of second language acquisition (SLA). In recent years, there has been a renewed interest in obtaining valid and reliable measures of L2 learners' implicit and explicit knowledge (e.g., Bowles, 2011; R. Ellis, 2005). The purpose of the present study was to examine the nature of the knowledge representations developed by two groups of learners of Spanish as a L2 at different levels of proficiency. The results show that the two groups differed with respect to their implicit and explicit knowledge of Spanish and also regarding the relationship between measures of those representations and the measures of L2 achievement used with each group.

Résumé

Les connaissances implicites et explicites de la langue seconde (L2) sont deux concepts centraux dans le domaine de l'acquisition de la langue seconde (ALS). Les dernières années ont vu un regain d'intérêt pour la recherche de moyens valides et fiables de mesurer les connaissances implicites et explicites des apprenants de L2 (par exemple, Bowles, 2011; R. Ellis, 2005). Le but de la présente étude a été d'examiner la nature des représentations des connaissances développées par deux groupes d'apprenants de l'espagnol comme L2 à différents niveaux de compétence. Les résultats démontrent qu'il existe des différences marquées entre les deux groupes au niveau de leurs connaissances implicites et explicites de l'espagnol, et aussi en ce qui concerne la relation entre les mesures de ces représentations et les mesures de performance en L2 utilisés avec chaque groupe.

Introduction

Implicit and explicit knowledge of the second language (L2) are two central constructs in the field of second language acquisition (SLA). Implicit knowledge of the L2 is often defined as the intuitive and procedural knowledge that is normally accessed automatically in fluent performance and that cannot be verbalized. In contrast, explicit knowledge is understood as the conscious and declarative knowledge of the L2 that is accessed during controlled processing and that is potentially verbalizable (Bowles, 2011; R. Ellis, 2005). Research about these two constructs has mainly focused on their role in language learning and language use, as well as the relationship between these two types of representations.

Regarding their role in L2 learning and use, there seems to be consensus in that implicit knowledge is at the core of automated language processing and that the development of these types of representations is the ultimate goal of L2 acquisition (Doughty, 2003). On the other hand, there is certain disagreement as to what the role of explicit knowledge is. Some authors (e.g., Krashen, 1981) attribute a very limited role to these types of representations; namely, that of a monitor or editor for L2 production under very specific circumstances. Others (e.g., Hinkel & Fotos, 2002; Johns, 2003; Mitchell, 2000) point out that successful performance in uses of language such as writing requires access to explicit knowledge. R. Ellis (1994) argues that explicit knowledge may play a facilitative role in L2 acquisition in that it may accelerate the establishment of links between form and meaning. N.C. Ellis (2005) notes that explicit knowledge of the L2 plays an important role in linguistic problem-solving: "when automatic capabilities fail, there follows a call recruiting additional collaborative conscious support" (p. 308).

With respect to the relationship between implicit and explicit knowledge, the debate revolves around whether they constitute two separate knowledge systems or, on the other hand, they interact at the representational level and thus one can be converted into the other. The noninterface position (e.g., Hulstjin, 2002; Krashen, 1981; Paradis, 1994) claims that these two types of knowledge constitute two different systems with different processing mechanisms and, as such, neither implicit knowledge can become explicit nor explicit knowledge can be converted into implicit. Conversely, according to the strong interface position (e.g., DeKeyser, 1998, 2003) explicit knowledge can become implicit and vice versa. Finally, there are different versions of a weak interface position, which claims that explicit knowledge can become implicit under certain conditions¹. Irrespective of their relationship at the representational level, implicit and explicit knowledge do interact at the performance level, since L2 learners often resort to both types of representations in L2 production (N.C. Ellis, 2005).

It is widely accepted that the development of implicit knowledge is central to the development of L2 proficiency. However, as noted above, it is not clear how explicit knowledge contributes to the acquisition of the L2. Indeed, there are plenty of studies that have examined the relationship between explicit knowledge and L2 proficiency operationalized in different ways. For example, Alderson, Clapham, and Steel (1997) only found low correlations between different measures of explicit knowledge and measures of L2 proficiency in French. In a study about learners of Italian as an L2, Sorace (1985) found that learners with higher levels of explicit knowledge performed better in two oral

¹ See R. Ellis (2005) for a detailed discussion of these positions.

production tasks than those with lower levels of explicit representations. Likewise, in a study about third-person singular possessive determiners in English, White and Ranta (2002) found a positive correlation between explicit knowledge and performance in an oral production task for both the control and the treatment group at the pre-test stage. However, the correlation was not significant for the treatment group after they received formal instruction about such determiners. Elder and Manwaring (2004) found that explicit knowledge was associated with better performance in a Chinese-language course for some of the groups in their study, but not for others. In a study about the psychological factors that mediate the availability of explicit knowledge, Hu (2002) found that, when correct explicit knowledge of six structures of English was available, the participants performed better in spontaneous writing tasks. Roehr (2008) examined the relationship between explicit knowledge of L2 German and L2 proficiency measured as knowledge of grammar and vocabulary, and found a strong positive correlation between the two. Renou (2001) also found a strong positive correlation between explicit knowledge of L2 French and proficiency in this language as measured by a test containing sections on listening and reading comprehension, and sections on general knowledge of vocabulary, grammar and structures. Finally, in a study about the validity of a metalinguistic knowledge test (MKT), Elder (2009) found significant correlations between the scores on that test and the different sections of both the Test of English as a Foreign Language (TOEFL) and the International English Language Testing System (IELTS), but not between the MKT and most sections in the Diagnostic English Language Needs Assessment (DELNA). Elder attributes the latter result to the high level of proficiency of the test takers, who may have been more likely to resort to their implicit knowledge of the L2 in that test.

As Elder and Ellis (2009) note, there is a dearth of studies that examine the relationship between the two types of knowledge, implicit and explicit, and L2 proficiency. One such study is Han and Ellis (1998). In this study, the authors examined the implicit and explicit knowledge representations of 48 learners of English as an L2 by means of several measures (an oral production task, three grammaticality judgement tests and a metalinguistic knowledge test) and investigated the relationship of these measures with two general measures of L2 proficiency (the TOEFL and the Secondary Level English Proficiency Test – SLEP). Han and Ellis found that the SLEP test was significantly correlated with measures of both implicit and explicit knowledge, whereas the TOEFL test was mainly correlated with measures of explicit knowledge. They noted that the TOEFL places greater emphasis on sections that favour the use of explicit knowledge (two thirds of the marks are given to those sections) whereas the SLEP, in comparison, "might be considered a more balanced measure of implicit/explicit knowledge" (Han & Ellis, 1998, p. 13). They also found that neither of the measures of L2 proficiency was correlated with the metalinguistic knowledge test, and concluded that ability to verbalize language rules does not seem to be important for L2 proficiency, whereas analyzed explicit knowledge seems to play a significant role.

In the paper mentioned above, Elder and Ellis (2009) reported two studies that examined the relationship between different measures of implicit and explicit knowledge and the TOEFL and IELTS, respectively. In the first study, these authors used one measure of implicit knowledge (a timed grammaticality judgement test - GJT) and two measures of explicit knowledge (an untimed GJT and an MKT). They found that the scores obtained on the explicit knowledge measures by 111 learners of English as a L2, as well as their scores on all the sections of two versions of the TOEFL test loaded on the same factor in a

principal components analysis, whereas the scores of the implicit knowledge measure loaded on a different factor. In the second study, Elder and Ellis examined a different cohort of learners of English as a L2 with respect to their performance on two measures of implicit knowledge (an elicited imitation task and a timed GJT) and two measures of explicit knowledge (an untimed GJT and an MKT), and the relationship of these scores with those obtained on the different sections of the IELTS test. Unlike the first study, the results showed that both implicit and explicit knowledge of the grammatical structures covered in the tests correlated with the proficiency measures. Elder and Ellis note that, although both implicit and explicit knowledge correlated with the four language skills tested in the IELTS, implicit knowledge correlated more strongly with the oral skills (speaking and listening) and explicit knowledge correlated more strongly with the written skills (writing and reading). The authors offer three possible interpretations to explain the different results obtained in the first and second studies: (a) the differences between the two proficiency measures used in each study with respect to the type of knowledge they encourage; (b) the different measures of implicit knowledge used in each study; and (c) the fact that they looked at total scores on the measures of implicit and explicit knowledge in the first study, whereas they examined the specific scores on the 17 grammatical structures covered in those measures in the second study.

The Study

Given the dearth of studies noted above, the purpose of the present study was to examine the nature of the knowledge representations developed by two groups of learners of L2 Spanish at two different levels of proficiency, as well as the relationship between these knowledge representations and the learners' achievement² in the L2. More specifically, the study sought to answer the following research questions.

- 1. Are there any differences in learners' performance on tests of implicit and explicit knowledge?
- 2. Are there any differences between the two proficiency groups with respect to their implicit and explicit knowledge representations?
- 3. Are there any correlations between the different tests of implicit and explicit knowledge and the measures of L2 achievement?

Participants

The study took place at a Canadian university and the participants were 53 students (43 female and 10 male) enrolled in two university-level courses of L2 Spanish who returned a signed consent form. The lower proficiency group consisted of 30 participants (22 female and eight male) who were near the end of their third term of Spanish language instruction at the university. In general, their proficiency level was close to the A2 level of the CEFR (Consejo de Europa, 2002). Seventeen of these participants had only taken Spanish language courses at the university, whereas thirteen of them had also taken Spanish at other institutions (mainly in high school). Nineteen participants described the instruction they had received at the university as a mixture of formal and informal, whereas eleven of them described it as mainly formal. Finally, twenty-two participants in this group stated that they use Spanish outside the classroom, mainly in activities such as listening to music,

 $^{^{2}}$ As noted by an anonymous reviewer, the term achievement might be preferable to the term proficiency because the measures used in the present study are the regular end of term assessment instruments, rather than more general measures of proficiency.

watching TV or movies, speaking with friends, and reading books, magazines and newspapers.

The higher proficiency group consisted of 23 participants (21 female and two male). These students were near completion of their fifth term of instruction in Spanish, and their proficiency level was close to the B1 level of the CEFR. In addition to taking Spanish language courses at the university, thirteen of the 23 participants in this group reported having taken courses at other institutions. About half of the group described the instruction they had received at the university as mainly formal, whereas the other half described it as a mixture of formal and informal. Finally, eighteen participants in this group reported using Spanish outside the classroom in similar activities as the lower proficiency group.

Instruments and Procedures

The participants completed a language background questionnaire, a timed grammaticality judgment test (TGJT), an untimed grammaticality judgment test (UGJT), a metalinguistic knowledge test (MKT), an oral proficiency test, and a written proficiency test. The language background questionnaire provided personal information about the participants, as well as information about their exposure to Spanish (see previous section). Based on previous research (Bowles, 2011; R. Ellis, 2005), the timed GJT was intended to be a measure of implicit knowledge, whereas the untimed GJT and the MKT were used as measures of explicit knowledge³. Finally, the oral test and the written test were used as measures of L2 achievement. These two tests were the assessment instruments used at the end of the Spanish language training courses that the participants were enrolled in.

The timed GJT consisted of 64 sentences covering 16 grammatical structures (see Appendix for a list of structures and sample sentences) about which the participants had received explicit instruction during the university Spanish-language training courses that they had taken or were taking at the time of the study. For each of the grammatical structures, there were two grammatical sentences and two ungrammatical ones. The sentences in the timed GJT were presented on a self-paced PowerPoint slide show. The participants were reminded of the speeded nature of the test and were instructed to indicate whether the sentences were grammatical or ungrammatical on an answer sheet. The amount of time that the sentences remained on the screen varied between 6 and 9 seconds depending on sentence length. These times were arrived at after examining previous research in which sentences were shown for 3 or 3.5 seconds (Bialystok, 1979; Han, 2000) or up to 6.24 seconds (Loewen, 2009). For the present study, between 3 and 6 seconds were allowed for sentence processing, and an additional 3 seconds for writing the response on the answer sheet. Every sixteen sentences, the participants were given a 15-second break.

In a second session, the participants completed the untimed GJT, which consisted of the same 64 sentences as the timed GJT. The participants received a sheet with the instructions and the sentences and a separate answer sheet on which they were asked to indicate whether the sentences were grammatical or ungrammatical. There was no time limit to complete the untimed GJT.

The MKT contained 16 ungrammatical sentences covering the same grammatical structures as the GJTs. In each sentence, the ungrammaticality had been underlined and the participants were asked to explain the rule that had been violated. The participants could

³ There is no guarantee, however, that learners will resort exclusively to implicit or explicit knowledge in such tests.

provide their explanations in English and they were informed that they did not have to correct the error. This test was administered during a third session and there was no time limit to complete it.

As part of the learners' regular evaluation in their language courses, at the end of the term the participants' proficiency in L2 Spanish was assessed by means of an oral and a written test. Obviously, these two tests, as well as the criteria for evaluating them were different for the two proficiency levels. The oral test for the low proficiency group consisted of a short in-class presentation on a free topic for which the learners could prepare beforehand. The written test for this group contained two sections, one with several exercises on grammar and structures, and one in which the learners' had to write a short text on a pre-specified topic. With respect to the higher proficiency group, the oral test was a five-minute oral conversation conducted with the course instructor. Unlike the lower proficiency group, the learners in this group were not informed about the topic of the conversation. The written language proficiency test for this group included listening, reading, writing, and grammar components. In the listening section, the learners heard two audio files and had to respond to a series of comprehension questions about each. Likewise, the reading section contained two short texts about which the learners had to answer a number of questions regarding comprehension. The writing section asked the learners to write a short text in response to one of two topics. Finally, the grammar section contained 'fill-in-the-blank' type exercises. Although this test is not a standardized measure of L2 Spanish proficiency, it is modeled after the Diploma de Español como Lengua Extranjera (DELE), the official test for Spanish as an $L2^4$.

Scoring Procedures

The responses to the timed and the untimed GJTs were scored in terms of correct and incorrect answers. Several participants left some items unanswered, and these could arguably be interpreted as either the participants not being sure about the grammaticality of the sentence or the participants not having had enough time to judge the sentence. However, blank responses accounted only for 7.33% and 1.83% of all the responses to the timed and untimed GJTs, respectively. Therefore, it was decided to count such responses as incorrect. Correct responses received 1 point and incorrect responses received 0 points. The scores in both the timed and untimed GJT were calculated for the test as a whole (64 points maximum), but also for grammatical and ungrammatical sentences separately⁵ (32 points maximum).

The MKT was also scored in terms of correct or incorrect answers. Correct answers consisted of adequate rule formulations without considering the use of technical metalanguage (i.e., rules formulated adequately with no technical metalanguage were considered correct). Imprecise or incomplete rule formulations were considered incorrect. Eighteen randomly selected MKTs (around 33%) were scored by a second rater. Inter-rater agreement was 89.93% and all disagreements were resolved through discussion.

The measures of L2 achievement were graded by each course instructor and the scores were part of the learners' final evaluation for the course. The oral test for both levels

⁴ See <u>http://diplomas.cervantes.es/aprender_espanol/informacion_dele.html</u>.

⁵ As it is discussed in the Results section, there are grounds for considering grammatical and ungrammatical sentences in GJTs as measures of different types of knowledge representations. Indeed, the grammatical and ungrammatical sentences in both the timed and untimed GJTs loaded on different factors in exploratory and confirmatory factor analyses.

of proficiency was rated for ability to convey ideas, pronunciation, fluency, and grammatical accuracy. With respect to the written test for the lower proficiency group, the exercises on the grammar and structures section were of the fill-in the blank type and were scored on the basis of correct or incorrect answers. The writing section of this test received a general score that included topic content, text organization, language use (vocabulary and grammar), and mechanics. Regarding the written test for the higher proficiency group, the listening and reading sections were scored on the basis of correct or incorrect responses depending on whether the learners demonstrated comprehension of the text content. The writing and grammar sections of this test were scored in the same way as those in the test for the lower proficiency group. For this study, only the total percentage scores for each test were made available.

Results

The reliability of the timed and untimed GJTs and the MKT was calculated using Cronbach's alpha. Table 1 shows the reliability coefficients for these measures. Cronbach's alpha coefficient was .783 for the timed test, .820 for the untimed one, and .849 for the MKT, which indicates that the three tests were reliable (Field, 2009).

| Test | Items | Participants | Reliability |
|-------------|-------|--------------|-----------------|
| | | | coefficient |
| Timed GJT | 64 | 53 | $\alpha = .783$ |
| Untimed GJT | 64 | 53 | $\alpha = .820$ |
| MKT | 16 | 53 | $\alpha = .849$ |

Table 1. Reliability coefficients for the three tests.

GJT: grammaticality judgment test; MKT: metalinguistic knowledge test

Douglas (2001) emphasizes the importance of demonstrating that the instruments used actually measure what one intended them to measure. In addition, based on the findings in his studies about the construct validity of tests of implicit and explicit knowledge (R. Ellis, 2005; R. Ellis & Loewen, 2007), Ellis expressed his concerns regarding what type of knowledge representation GJTs actually measure. In a personal communication⁶, he noted that grammatical sections in timed and untimed GJTs might be a measure of implicit knowledge, whereas ungrammatical sections in both tests might constitute a measure of explicit knowledge. He suggested carrying out exploratory and confirmatory factor analyses to verify this hypothesis. Following Ellis's recommendations, the scores on the GJTs were split into grammatical and ungrammatical sections, and a principal components factor analysis and two confirmatory factor analyses were carried out'. With respect to the principal components factor analysis, the Kaiser-Meyer-Olkin measure confirmed that the sample was adequate for the analysis (KMO = .64), and the results of Bartlett's test of sphericity, $X^2(10) = 104.46$, p < .001, indicated that the correlations among the five measures were sufficiently large for the analysis (Field, 2009). The initial factor analysis, presented in Table 2, yielded five components, two of which had eigenvalues higher than 1. These two components accounted for 75.38% of the total

⁶ March 28, 2011.

⁷ Gutiérrez (2012) offers a rationale for analyzing grammatical and ungrammatical sentences in GJTs separately and a detailed discussion of the results of factor analyses and other statistical tests in this regard.

variance. Table 3 shows the loadings for the two factors retained. The ungrammatical sections of the timed and untimed GJTs and the MKT loaded strongly on the first component, whereas the grammatical sections of the two GJTs had negligible loadings on that first component. Conversely, the grammatical sections of the timed and untimed GJT had strong loadings on the second component, while the loadings on the second factor for the ungrammatical sections of both GJTs and the MKT were insignificant.

| Component | Eigenvalue | Variance ^a | Cumulative ^a | | | | | |
|-----------|------------|-----------------------|-------------------------|--|--|--|--|--|
| 1 | 2.757 | 55.133 | 55.133 | | | | | |
| 2 | 1.012 | 20.246 | 75.380 | | | | | |
| 3 | .749 | 14.988 | 90.367 | | | | | |
| 4 | .247 | 4.936 | 95.304 | | | | | |
| 5 | .235 | 4.696 | 100.000 | | | | | |
| 0 - | | | | | | | | |

Table 2. Principal component factor analysis

^a Percentage

Table 3. Loadings for principal components factor analysis.

| U | | 2 | |
|---------|-------------|-------------|--|
| | Component 1 | Component 2 | |
| UGJT Un | .939 | .026 | |
| TGJT Un | .782 | 054 | |
| MKT | .717 | .019 | |
| TGJT Gr | 030 | 953 | |
| UGJT Gr | .037 | 917 | |

Rotation method: Direct oblimin; UGJT: untimed grammaticality judgment test; TGJT: timed grammaticality judgment test; MKT: metalinguistic knowledge test; Un: ungrammatical; Gr: grammatical.

In addition to the principal components analysis, two confirmatory factor analyses (CFA) were carried out. The first CFA tested the model predicted in the principal components analysis discussed above; namely, that the grammatical sections of both GJTs would load on an implicit knowledge factor, and that the ungrammatical sections of both GJTs and the MKT would load on an explicit factor. The second analysis tested an alternative model (Isemonger, 2007). In this case, the model was that predicted by the principal components analysis in R. Ellis (2005) and verified by the CFA in Bowles (2011); namely that the grammatical and ungrammatical sections of the timed GJT would load on an implicit knowledge factor, and that the grammatical and ungrammatical sections of the untimed GJT and the MKT would load on an explicit factor. These two models are illustrated in Figures 1 and 2, respectively

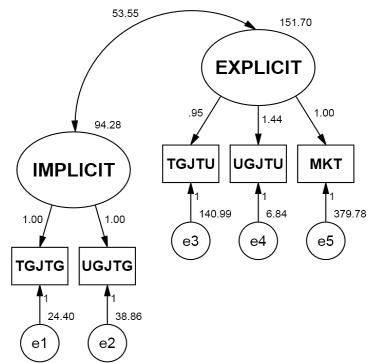


Figure 1. Grammatical/Ungrammatical Model. TGJTG: timed GJT, grammatical section; UGJTG: untimed GJT, grammatical section; TGJTU: timed GJT, ungrammatical section; UGJTU: untimed GJT, ungrammatical section; MKT: metalinguistic knowledge test.

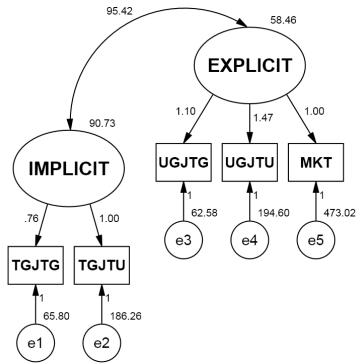


Figure 2. Timed/Untimed Model. TGJTG: timed GJT, grammatical section; TGJTU: timed GJT, ungrammatical section; UGJTG: untimed GJT, grammatical section; UGJTU: untimed GJT, ungrammatical section; MKT: metalinguistic knowledge test.

Several metrics were examined in order to analyze the goodness of fit of both models. First, a non-significant value of the chi-square (X^2) statistic indicates an acceptable fit of the model. In addition, values of .95 or higher in the normed fit index (NFI) and the Tucker-Lewis index (TLI) also indicate a superior fit for the model. Values of .10 or higher in the root mean square error of approximation (RMSEA) indicate a poor fit of the model, whereas values of .05 or lower indicate a good fit. Finally, values of .05 or higher in the significance of close fit (PCLOSE) indicate an acceptable fit of the model. As Table 4 shows, all the values indicate that the Grammatical/Ungrammatical Model (i.e., the model in which the grammatical sections load on the implicit factor and the ungrammatical sections load on the explicit factor) is an acceptable fit. However, none of the criteria discussed above is met for the Timed/Untimed Model (i.e., the model in which the two sections of the timed GJT load on the implicit factor and the two sections of the untimed GJT load on the explicit factor).

| Table 4. Summary of the mode | el of fit for th | ne solution | s in Figu | res 1 and 2. | |
|------------------------------|------------------|-------------|-----------|--------------|-----|
| | | | <u> </u> | | |
| M l . l | v^2 | NICI | TII | DMCEA | DOL |

| Model | X^2 | NFI | TLI | RMSEA | PCLOSE | df | |
|---|----------|------|------|-------|--------|----|--|
| Grammatical/Ungrammatical | 4.406 | .959 | .990 | .044 | .421 | 4 | |
| Timed/Untimed | 31.787** | .706 | .292 | .366 | .000 | 4 | |
| NEL normad fit index: TLL: Tueker Lewis index: DMSEA: root mean square error of | | | | | | | |

NFI: normed fit index; TLI: Tucker-Lewis index; RMSEA: root mean square error of approximation; PCLOSE: significance of close fit; ** p < .00

Given the results of the principal components analysis and the confirmatory factor analysis, the grammatical sentences in the timed and untimed GJT were considered to be measures of implicit knowledge and the ungrammatical sentences in both GJTs and the MKT were considered measures of explicit knowledge. Therefore, the subsequent statistical tests performed on the data for this study were done on the scores on the GJTs split into grammatical and ungrammatical sections.

Tables 5 and 6 show the descriptive statistics for the lower proficiency group and the higher proficiency group, respectively. Specifically, the tables show the mean (in percentage), standard deviation, skewness ratio, and kurtosis ratio for the total scores on the timed and untimed GJTs, for the grammatical and ungrammatical sections on the timed and untimed GJTs, for the MKT, for the oral tests, and for the written proficiency tests. With respect to the lower proficiency group, Table 5 shows that this group performed at an acceptable level on the grammatical sections in both the timed and untimed GJT (69.17% and 74.79%, respectively). However, they performed rather poorly on the ungrammatical sections of both tests and on the MKT (39.69%, 52.45%, and 46.46%, respectively). Furthermore, the standard deviations show greater variation for those three measures than for the grammatical sections of the GJTs. Likewise, as Table 6 shows, the higher proficiency group performed better on the grammatical sections of both GJTs (75.54% and 81.79%) than on the ungrammatical sections and on the MKT (52.72%, 60.36%, and 57.34%). As was the case for the lower proficiency group, the standard deviations also show greater variance in those three measures than in the grammatical sections of the GJTs. With respect to the two achievement measures, the lower proficiency group obtained a high average on the oral test $(86.67\%)^8$ and a medium one on the written test (65.72%). Conversely, the higher proficiency group obtained more balanced averages on both tests

⁸ See discussion below.

(77.66% and 70.33%). The skewness and kurtosis ratios are used as tests of the assumption of normality and are calculated dividing the skewness and kurtosis statistic by their respective standard errors. In small samples, ratios lower than 1.96 indicate that the data are normally distributed (Field, 2009). As Tables 5 and 6 show, most of the data fit the assumption of normality⁹. The total score for the untimed GJT in the lower proficiency group and the oral test in that group, however, are not normally distributed.

| Test & items | Mean (%) | SD | skewness | kurtosis | N |
|--------------|----------|-------|----------|----------|----|
| | | | ratio | ratio | |
| TGJT Total | 54.43 | 9.01 | -0.651 | 0.293 | 30 |
| TGJT Gr | 69.17 | 10.40 | -0.194 | -0.959 | 30 |
| TGJT Un | 39.69 | 13.34 | 0.300 | -1.052 | 30 |
| UGJT Total | 63.39 | 9.83 | 2.272 | 2.104 | 30 |
| UGJT Gr | 74.79 | 11.34 | -0.365 | -0.008 | 30 |
| UGJT Un | 52.45 | 14.74 | -0.628 | 1.055 | 30 |
| MKT | 46.46 | 18.94 | -0.356 | -0.670 | 30 |
| Oral test | 86.67 | 15.32 | -4.520 | 3.826 | 30 |
| Written test | 65.72 | 15.95 | -0.841 | 0.349 | 30 |
| marm 1 | | IIOIT | | | |

Table 5 Descriptive statistics for the lower proficiency group

TGJT: timed grammaticality judgment test; UGJT: untimed grammaticality judgment test; MKT: metalinguistic knowledge test; Gr: grammatical; Un: ungrammatical

| Test & items | Mean (%) | SD | skewness | kurtosis | N |
|--------------|----------|-------|----------|----------|----|
| | | | ratio | ratio | |
| TGJT Total | 64.13 | 12.62 | 0.538 | -0.986 | 23 |
| TGJT Gr | 75.54 | 10.93 | 0.395 | -1.155 | 23 |
| TGJT Un | 52.72 | 18.23 | -0.245 | -0.723 | 23 |
| UGJT Total | 70.72 | 14.44 | -0.393 | -0.180 | 23 |
| UGJT Gr | 81.79 | 11.05 | -0.692 | -0.966 | 23 |
| UGJT Un | 60.36 | 21.17 | -1.295 | 0.076 | 23 |
| MKT | 57.34 | 27.16 | -1.056 | -0.847 | 23 |
| Oral test | 77.66 | 16.31 | 0.104 | -1.206 | 23 |
| Written test | 70.33 | 12.68 | -0.326 | -1.359 | 23 |
| TOTAL 1 | | | | | |

| Table (Desering | time statistic | a fan tha hiahan | |
|-------------------|----------------|------------------|--------------------|
| I able o. Descrit | Duve statistic | s for the higher | proficiency group. |
| | | | |

TGJT: timed grammaticality judgment test; UGJT: untimed grammaticality judgment test; MKT: metalinguistic knowledge test; Gr: grammatical; Un: ungrammatical

To address research question 1 (i.e., whether there were differences between the learners' performance on the implicit and explicit knowledge measures), paired-samples ttests were calculated for the two proficiency levels separately. As Tables 7 and 8 indicate, both groups showed significant differences with large effect sizes between the two measures of implicit knowledge (i.e., the grammatical sections in the timed and untimed GJTs) and the three measures of explicit knowledge (i.e., the ungrammatical sections in both GJTs and the MKT).

⁹ The examination of Q-Q plots and histograms confirmed that most of the data were normally distributed.

| | Paired differences | | | | |
|-------------|--------------------|------------|----------|----|-----|
| | | Std. Error | | | |
| Test | Mean (%) | Mean | t | df | r |
| TGJT Gr vs. | 29.48 | 2.87 | 10.261** | 29 | .89 |
| TGJT Un | | | | | |
| TGJT Gr vs. | 16.72 | 2.91 | 5.737** | 29 | .73 |
| UGJT Un | | | | | |
| TGJT Gr vs. | 22.71 | 3.66 | 6.198** | 29 | .75 |
| MKT | | | | | |
| UGJT Gr vs. | 35.10 | 2.90 | 12.088** | 29 | .91 |
| TGJT Un | | | | | |
| UGJT Gr vs. | 22.34 | 3.27 | 6.840** | 29 | .79 |
| UGJT Un | | | | | |
| UGJT Gr vs. | 28.33 | 3.68 | 7.708** | 29 | .82 |
| MKT | | | | | |

Table 7. Paired-samples *t*-tests for the lower proficiency group.

** p < .001; * p < .05; TGJT: timed grammaticality judgment test; UGJT: untimed grammaticality judgment test; Gr: grammatical; Un: ungrammatical; MKT: metalinguistic knowledge test

Table 8. Paired-samples *t*-tests for the higher proficiency group.

| | Paired dif | ferences | | | |
|-------------|------------|------------|---------|----|-----|
| | | Std. Error | | | |
| Test | Mean (%) | Mean | t | df | r |
| TGJT Gr vs. | 22.83 | 3.41 | 6.703** | 22 | .82 |
| TGJT Un | | | | | |
| TGJT Gr vs. | 15.19 | 3.91 | 3.876** | 22 | .64 |
| UGJT Un | | | | | |
| TGJT Gr vs. | 18.21 | 5.57 | 3.267* | 22 | .57 |
| MKT | | | | | |
| UGJT Gr vs. | 29.08 | 3.43 | 8.468** | 22 | .87 |
| TGJT Un | | | | | |
| UGJT Gr vs. | 21.44 | 3.63 | 5.897** | 22 | .78 |
| UGJT Un | | | | | |
| UGJT Gr vs. | 24.46 | 5.56 | 4.397** | 22 | .68 |
| MKT | | | | | |

** p < .001; * p < .05; TGJT: timed grammaticality judgment test; UGJT: untimed grammaticality judgment test; Gr: grammatical; Un: ungrammatical; MKT: metalinguistic knowledge test

Independent-samples *t*-tests were computed in order to address research question 2 (i.e., whether there were differences between the lower and the higher proficiency groups in terms of the scores obtained in the measures of implicit and explicit knowledge). Table 9 shows that the differences between the two proficiency levels were significant with medium and close to medium effect sizes for the grammatical sections in both the timed and

untimed GJT and for the ungrammatical section in the timed GJT, but not so for the ungrammatical section in the untimed GJT and the MKT.

| | Group | | | | | | | |
|------------------|-------|------|-------|------|-------|--------|------|-----|
| | Lov | v P. | Hig | h P | | | | |
| Test | М | SE | М | SE | t | $d\!f$ | sig | r |
| TGJT Gr | 69.17 | 1.90 | 75.54 | 2.28 | 2.163 | 51 | .035 | .29 |
| TGJT Un | 39.69 | 2.44 | 52.72 | 3.80 | 3.007 | 51 | .004 | .39 |
| UGJT Gr | 74.79 | 2.07 | 81.79 | 2.31 | 2.252 | 51 | .029 | .30 |
| UGJT Un | 52.45 | 2.69 | 60.36 | 4.41 | 1.603 | 51 | .115 | .22 |
| MKT ^a | 46.46 | 3.46 | 57.34 | 5.66 | 1.640 | 37.52 | .109 | .26 |

Table 9. Independent-samples *t*-tests.

^a Variances are homogeneous in all tests except the MKT. TGJT: timed grammaticality judgment test; UGJT: untimed grammaticality judgment test; Gr: grammatical; Un: ungrammatical; MKT: metalinguistic knowledge test

Research question 3 addresses the issue of the possible relationships between the different measures of implicit and explicit knowledge and the two measures of L2 achievement. Pearson product moment coefficients were computed in order to examine this question. Tables 10 and 11 show the correlation matrices for the lower and higher proficiency groups, respectively. Table 10 shows that, for the lower proficiency group, none of the measures of implicit and explicit knowledge correlated with the scores on the oral test, whereas only the scores on the ungrammatical section of the untimed GJT and those on the MKT correlated significantly with the written proficiency test. It is worth noting that the correlation between the ungrammatical section in the timed GJT and the written test was very close to significance (p = .52) and had a medium effect size (r = .358). Regarding the higher proficiency group, Table 11 shows that all measures of implicit and explicit knowledge except the MKT correlated significantly with the oral test whereas all measures correlated significantly with the written test.

| Table 10. Correlations between measures of implicit and explicit knowledge and measures | |
|---|--|
| of L2 achievement for the lower proficiency group | |

| Test | TGJT Gr | TGJT Un | UGJT Gr | UGJT Un | MKT |
|--------------|---------|---------|---------|---------|--------|
| Oral test | 249 | .071 | 250 | .270 | .160 |
| Written test | .022 | .358 | .020 | .603** | .660** |

** p < .01 (2-tailed); TGJT: timed grammaticality judgment test; UGJT: untimed grammaticality judgment test; MKT: metalinguistic knowledge test; Gr: grammatical; Un: ungrammatical

Table 11. Correlations between measures of implicit and explicit knowledge and measures of L2 achievement for the higher proficiency group

| or 22 while the industry promotion of brown | | | | | | | |
|---|---------|---------|---------|---------|--------|--|--|
| Test | TGJT Gr | TGJT Un | UGJT Gr | UGJT Un | MKT | | |
| Oral test | .602** | .652** | .675** | .644** | .144 | | |
| Written test | .527** | .510* | .564** | .733** | .652** | | |

* p < .05 (2-tailed); ** p < .01 (2-tailed); TGJT: timed grammaticality judgment test; UGJT: untimed grammaticality judgment test; MKT: metalinguistic knowledge test; Gr: grammatical; Un: ungrammatical

Discussion

This study examined the nature of the knowledge representations developed by two groups of learners of Spanish as a L2 and the relationship between these representations and measures of L2 achievement. The analyses carried out on the data yielded several important findings. First, the factor analyses computed on the different measures of implicit and explicit knowledge indicated that, for this sample, the grammatical sections in a timed and untimed GJT constitute a measure of implicit knowledge whereas the ungrammatical sections in both GJTs and the MKT are measures of explicit knowledge. This finding has important implications for research on the measurement of implicit and explicit knowledge. In previous studies that have used factor analysis (Bowles, 2011; R. Ellis, 2005; R. Ellis & Loewen, 2007; Han & Ellis, 1998), the timed GJT has been identified as a measure of implicit knowledge and the untimed test as a measure of explicit knowledge, although R. Ellis (2005) found that the untimed GJT also loaded on the implicit knowledge factor. Consequently, he reanalyzed the data in that same study and in a subsequent one (R. Ellis & Loewen, 2007) only using the scores for the ungrammatical section of the untimed GJT, which only loaded on the explicit factor. However, no previous study has analyzed grammatical and ungrammatical sections in GJTs separately. As discussed in detail elsewhere (Gutiérrez, 2012), the results of the present study in this regard confirm R. Ellis' concern regarding GJTs discussed above and point to the need to analyze separately the scores of grammatical and ungrammatical sections in GJTs in future research on the construct validity of these instruments.

With respect to the differences between the implicit and explicit knowledge representations of the participants (research question 1), Tables 5 and 6 show that both groups obtained higher scores on the measures of implicit knowledge (grammatical sections in the timed and untimed GJT) than on those of explicit knowledge (ungrammatical sections in both GJTs and MKT). Indeed, both groups performed rather poorly on all explicit knowledge measures. This finding is confirmed by the results of the paired-samples *t*-tests presented in Tables 7 and 8, which show significant differences with large effect sizes between the measures of implicit knowledge and those of explicit knowledge for both groups. The fact that both groups performed better on the former than on the latter is rather unexpected since, given the type of instruction that these learners have received, one would expect them to have developed higher levels of explicit representations than of implicit ones. Loewen (2009) obtained similar results regarding the timed GJT, but not the untimed one. He explained the significant differences in the timed test and the lack of differences in the untimed one on the basis of three processing operations discussed in R. Ellis (2004) and their interaction with the speeded nature of the test. R. Ellis argues that learners may engage in three processing operations when judging the grammaticality of sentences: 1) semantic processing (i.e., "understanding the meaning of a sentence" (R. Ellis, 2004, p. 256); 2) noticing (i.e., deciding whether there is something ungrammatical in a sentence); and 3) reflecting (i.e., identifying what is incorrect and possibly determining why it is incorrect). As Loewen explains, learners engage in semantic processing and noticing when determining the grammaticality of a sentence. However, they may engage in reflecting once they have determined that the sentence is ungrammatical. According to Loewen, the

learners' in his study had sufficient time for semantic processing and noticing in the timed GJT, but not for reflecting, and this may explain the significant difference between grammatical and ungrammatical sentences. Conversely, they had enough time for all three processes in the untimed GJT. This interpretation was supported by the differences in response times between the timed and untimed GJT. Although the time constraints of the timed GJT undoubtedly played a role in the performance of the participants in the present study, such constraints do not explain the significant differences found regarding the untimed test, since the mean differences between the grammatical and ungrammatical section of that test were very similar to those in the timed test. The use of a computerized version of the tests in which response times are recorded would help clarify these findings since it would provide information regarding potential differences between grammatical and ungrammatical sections in terms of processing time.

Research question 2 addressed the differences between the two proficiency groups with respect to their implicit and explicit knowledge. As shown in Tables 5, 6, and 9 above, the higher proficiency group obtained higher scores than the lower proficiency group on all the measures of implicit and explicit knowledge. However, the results of the independent-samples *t*-test indicate that the differences between the two groups were only significant for the grammatical section in both the timed and untimed GJT, and for the ungrammatical section in the timed GJT, with low to medium effect sizes. These results indicate that the higher proficiency group developed significantly more implicit knowledge representations of the grammatical structures tested than the lower proficiency group. However, the differences between the two groups with respect to explicit knowledge representations are less considerable. This is particularly so regarding the two measures with no time constraints (i.e., the ungrammatical section in the untimed GJT and the MKT). Thus, when time pressure is absent, the performance of the lower proficiency group is closer to that of the higher proficiency group.

The two groups in this study behaved quite differently in terms of the relationship between the measures of implicit and explicit knowledge and those of L2 achievement (research question 3), most likely because of the different L2 achievement tests that both groups completed. Since it is not possible to compare the two groups regarding the relationship between implicit and explicit knowledge and the L2 achievement measures, the following discussion focuses on each group separately. With respect to the lower proficiency group, Table 10 shows that none of the measures of implicit and explicit knowledge correlated with the oral test. It is worth noting that these learners had the opportunity to prepare their oral presentation beforehand, and that the instructor of this group reported having been very lenient when rating the learners' performance on the oral test. These two facts combined likely account for the high percentage average (86.67%) that this group obtained and for the severely negatively skewed distribution of the scores (see Table 5 above). Indeed, 23 learners in this group obtained a score higher than the average in the oral test whereas the remaining seven obtained a score lower than the average. This fact might explain why none of the measures of implicit and explicit knowledge correlated with this test.

As far as the written test is concerned, Table 10 shows that the grammatical sections of the timed and untimed GJT did not correlate with that test. On the contrary, the ungrammatical section of the untimed GJT and the MKT significantly correlated with the written test, and although the ungrammatical section of the timed GJT did not show a significant correlation, the *p* value approached significance (p = .52), as noted above. It is

worth recalling the fact that the written test for this group consisted of a grammar and structures section and a written composition section. Therefore, it is not surprising that the scores on this test showed significant (and almost significant) correlations with those on the three measures of explicit knowledge, but not with the measures of implicit knowledge, since this type of test likely encourages drawing on explicit representations rather than on implicit ones. These results are consistent with those in other studies. Roehr (2008), as noted earlier, found a significant correlation between explicit knowledge and L2 proficiency as measured by a test of grammar and vocabulary knowledge. Elder & Ellis (2009) also found that measures of explicit knowledge, but not measures of implicit knowledge, were related to scores on the TOEFL test, which also seems to encourage use of explicit representations.

Regarding the higher proficiency group, Table 11 shows that the scores on the oral test were significantly correlated with the grammatical and ungrammatical sections in both the timed and untimed GJT, but not with the MKT. These results indicate that oral production in the L2 correlates with implicit knowledge and with analyzed explicit knowledge, but not with the ability to verbalize language rules. A relationship between oral production and implicit knowledge is expected, but one between oral production and explicit knowledge is less so. Given the criteria used for scoring the oral test (ability to convey ideas, pronunciation, fluency, and grammatical accuracy), however, it is not surprising that measures of both types of knowledge correlated with this test, since some of these factors (e.g., the automaticity required to convey ideas fluently) may favour the use of implicit knowledge, particularly at this level of proficiency. Indeed, as noted earlier, Sorace (1985) and White and Ranta (2002) also found that oral production was related to explicit knowledge.

With respect to the written test, Table 11 shows that all measures of implicit and explicit knowledge used in this study were significantly correlated with the scores obtained in that test. This finding is also unsurprising given the fact that this test seems to be more balanced than that used with the lower proficiency group, since it contained different sections (listening, reading, writing and grammar) that likely favour use of one or another type of knowledge representations. Unfortunately, separate scores for each of the sections were not made available and, therefore, it was not possible to establish which sections of the written test correlated with implicit or explicit knowledge, or with both. As discussed above, Han and Ellis (1998) and Elder and Ellis (2009) also found that both types of knowledge correlated with more balanced tests of L2 proficiency such as the SLEP in the former study and the IELTS in the latter. In discussing the findings in Elder and Ellis, R. Ellis (2009) notes that the TOEFL test seems to be heavily biased towards 'cognitive academic language proficiency' (CALP; Cummins, 1983) whereas the IELTS test is more oriented towards 'basic interpersonal communication skills' (BICS). R. Ellis argues that "implicit knowledge can be expected to be more clearly required for BICS while explicit knowledge will be of greater importance for CALP" (2009, p. 339). This distinction might also explain the differences found between the two proficiency groups in this study.

The present study is not exempt from limitations. First, given the questions regarding what type of knowledge GJTs measure, the study could have benefited from having used an additional measure of implicit knowledge that does not raise those concerns. For example, the use of an elicited oral imitation test, which has been identified as a valid measure of implicit knowledge (Erlam, 2006), in conjunction with the grammatical sections

of the GJTs would have solved this limitation. In addition, although two measures per factor is the minimum required in a two-factor model, three measures per factor are usually recommended for confirmatory factor analysis (Klein, 2005). Thus, the use of an elicited oral imitation test would have balanced the number of measures to three per factor. Another limitation is the fact that the study used intact classes and, consequently, relied on achievement measures of L2 Spanish rather than true proficiency measures, although the test used for the higher proficiency group was designed following the DELE as a model. Finally, as noted earlier, specific results for the different sections of the written tests of L2 proficiency were not available. Having those results would have allowed establishing more detailed relationships between the measures of implicit and explicit knowledge and different aspects of L2 proficiency.

Conclusion

The study reported in this paper shows different aspects of the knowledge representations developed by learners of Spanish as a L2 at two levels of proficiency. More specifically, it shows that, despite having received extensive explicit instruction about Spanish grammatical structures, the participants performed rather poorly on all measures of explicit knowledge. This is an unexpected and important finding, particularly because the majority of the tests that assessed the participants' L2 proficiency were strongly correlated with this type of knowledge, especially the written test used with the lower proficiency group. This finding, particularly regarding the results of the MKT, could be explained by the fact that these learners have probably received very little formal grammar instruction in school. In fact, based on my personal observations in language classes, learners of L2 Spanish often struggle to understand grammar instruction. Furthermore, as expected, it also shows that the higher proficiency learners performed better than the lower proficiency ones in all measures of implicit and explicit knowledge. However, the differences were not significant for the explicit knowledge measures without time constraints. Finally, the study also shows that different types of language achievement measures correlate differently with the measures of implicit and explicit knowledge that the learners completed. This is an important finding since, if we agree that the main goal of second language acquisition is the development of implicit knowledge, we need to ensure that the L2 proficiency tests we use actually provide a measure of such knowledge representations (Doughty, 2003). Clearly, the written test used for the lower proficiency group in this study does not. Conversely, more balanced measures like the one used with the higher proficiency group might give us a better indication of the relationship of implicit and explicit knowledge with L2 proficiency. In any case, as R. Ellis (2009) notes, there is a need for future research that provides valid and reliable measures of implicit and explicit knowledge and that examines the relationship between both types of representations and L2 proficiency. The present study is another step in this direction.

Acknowledgements

This research was funded by a Humanities and Social Sciences Research Grant from the University of Windsor. I would like to thank the students that participated in the study and the instructors that allowed me access to their classes. My appreciation also goes to Jelena Primorac for rating some of the tests, to Sandra Gidak for her help in coding the data, and to Rod Ellis for his thoughtful suggestions. I am also grateful for the helpful comments made by an anonymous reviewer.

References

- Alderson, J. C., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research*, *1*-2, 93-121.
- Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality. *Language Learning*, 29, 81-103.
- Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition, 33*, 247-271.
- Consejo de Europa (2002, June). Marco Común Europeo de Referencia para las Lenguas: Aprendizaje, Enseñanza, Evaluación. Retrieved from http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf.

Cummins, J. (1983). Language proficiency and academic achievement. In J. Oller (Ed.) *Issues in language testing research* (pp. 108-130). Rowley, MA: Newbury House.

- DeKeyser, R. M. (1998). Beyond focus on form. Cognitive perspective on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), Focus on form in classroom second language acquisition (pp. 42-63). Cambridge, UK: Cambridge University Press.
- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Oxford, UK: Blackwell Publishing.
- Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 256-310). Oxford, UK: Blackwell Publishing.
- Douglas, D. (2001). Performance consistency in second language acquisition and language testing: A conceptual gap. *Second Language Research*, *17*, 442-456.
- Elder, C. (2009). Validating a test of metalinguistic knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders, *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 113-138). Bristol, UK: Multilingual Matters.
- Elder, C., & Ellis, R. (2009). Implicit and explicit knowledge of an L2 and language proficiency. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders. *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 167-193). Bristol, UK: Multilingual Matters.
- Elder, C., & Manwaring, D. (2004). The relationship between metalinguistic knowledge and learning outcomes among undergraduate students of Chinese. *Language Awareness*, 13, 145-162.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit knowledge. *Studies in Second Language Acquisition*, 27, 305-352.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford, UK: Oxford University Press.
- Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, *54*, 227-275.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. A psychometric study. *Studies in Second Language Acquisition*, 27, 141-172.

- Ellis, R. (2009). Retrospect and prospect. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders, *Implicit and explicit knowledge in second language learning*, *testing and teaching* (pp. 335-353). Bristol, UK: Multilingual Matters.
- Ellis, R., & Loewen, S. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005): Responding to Isemonger. *Studies in Second Language Acquisition*, 29, 119-126.
- Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, 27, 464-491.
- Field, A. (2009). Discovering statistics using SPSS (3rd ed.). Thousand Oaks, CA: Sage.
- Gutiérrez, X. (2012). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. Manuscript submitted for publication.
- Han, Y. (2000). Grammaticality judgment tests: How reliable and valid are they? *Applied Language Learning*, 11, 177-204.
- Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research*, *2*, 1-23.
- Hinkel, E., & Fotos, S. (2002). From theory to practice: A teacher's view. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classrooms* (pp. 1-12). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hu, G. (2002). Psychological constraints on the utility of metalinguistic knowledge in second language production. *Studies in Second Language Acquisition*, 24, 347-386.
- Hulstijn, J. (2002). Toward a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, *18*, 193-223.
- Isemonger, I. M. (2007). Operational definitions of explicit and implicit knowledge: Response to R. Ellis (2005) and some recommendations for future research in this area. *Studies in Second Language Acquisition*, 29, 101-118.
- Johns, A. M. (2003). Genre and ESL/EFL composition instruction. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 195-217). Cambridge, UK: Cambridge University Press.
- Klein, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.
- Krashen, S. D. (1981). Second language acquisition and second language learning. Oxford, UK: Pergamon.
- Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders, *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 94-112). Bristol, UK: Multilingual Matters.
- Mitchell, R. (2000). Anniversary article. Applied linguistics and evidence-based classroom practice: The case of foreign language grammar pedagogy. *Applied Linguistics*, 21, 281-303.
- Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393-419). San Diego, CA: Academic Press.
- Renou, J. (2001). An examination of the relationship between metalinguistic awareness and second-language proficiency of adult learners of French. *Language Awareness*, 10, 248-267.

- Roehr, K. (2008). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics*, *29*, 173–199.
- Sorace, A. (1985). Metalinguistic knowledge and language use in acquisition-poor environments. *Applied Linguistics*, *6*, 239-254.
- White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness*, 11, 259-290.

Appendix

List of Structures and Sample Sentences in the GJTs Determiner-noun agreement

- (1) *El programa de televisión es muy interesante.* The TV program is very interesting.
- (2) **La problema más importante es la falta de dinero.* The most important problem is the lack of money.

Subject-verb agreement in present indicative

- (3) *Tú cantas en la ducha por las mañanas.* You sing in the shower in the morning.
- (4) **Yo escribe una carta a mi amiga.* I write a letter to my friend.

Subject-verb agreement in preterit

- (5) Anna cenó con sus padres el sábado.
 Anna had dinner with her parents on Saturday.
- (6) **Gloria y Susana estuvimos en Granada por dos semanas.* Gloria and Susana were in Granada for two weeks.

Subject-verb agreement in simple future

- (7) Nosotros iremos al cine mañana. We will go to the movies tomorrow.
- *Ustedes viajaréis a Costa Rica en verano.
 You will travel to Costa Rica in the summer.

Stem-changing verbs in present indicative

- (9) Mis hermanos duermen 8 horas cada noche. My siblings sleep 8 hours every night.
- (10) **Yo almorzo a las 2 cada día.* I have lunch at 2 every day.

Irregular verbs in preterit

(11) Andrés vino a mi casa anteayer.Andrés came to my house the day before yesterday

(12) **Ustedes tenieron una reunión ayer por la tarde.* You had a meeting yesterday afternoon.

Ser or estar with location

- (13) *La oficina de la profesora está en Lambton Tower.* The professor's office is in Lambton Tower.
- (14) **La sala 202 es en el segundo piso.* Room 202 is on the second floor.

Ser or estar with profession

- (15) *Stephen Harper es el Primer Ministro de Canadá.* Stephen Harper is the Prime Minister of Canada.
- (16) **Kobe Bryant está un deportista muy famoso.* Kobe Bryant is a very famous sportsman.

Ser or estar with place of origin

- (17) *Nosotros somos de Buenos Aires, somos argentinos.* We are from Buenos Aires, we are Argentinean.
- (18) *Mis padres están de Venezuela, nacieron en Caracas.

My parents are from Venezuela, they were born in Caracas.

Ser or estar with personality traits

- (19) *Yo soy un poco tímido.* I am a bit shy.
- (20) **Mis amigos están muy amables.* My friends are very kind.

Ser or estar with mood or conditions

- (21) *Vosotros estáis un poco nerviosos.* You are a little bit nervous.
- (22) **Tú eres muy contenta.* You are very happy.

Imperfect tense with ongoing actions

- (23) *Tus padres iban al teatro cuando se encontraron con Jorge*. Your parents were going to the theatre when they ran into Jorge.
- (24) **Elena caminó por la calle cuando vio un accidente.*

Elena was walking on the street when she saw an accident.

Imperfect tense with habitual actions

- (25) *Todos los días nosotros comíamos cereales para desayunar.* Every day we used to eat cereal for breakfast.
- (26) **Cuando era pequeña yo jugué a fútbol cada día.* When I was a kid I used to play soccer every day.

Preterit with completed actions

- (27) *La semana pasada tú trabajaste 50 horas.* Last week you worked 50 hours.
- (28) **En el año 2007 yo visitaba Santo Domingo.* In 2007 I visited Santo Domingo.

Present subjunctive with noun clauses

- (29) *Te aconsejo que bebas dos litros de agua cada día.* I advise you to drink two liters of water every day.
- (30) **Espero que ellos pueden venir a mi fiesta de cumpleaños.* I hope that they can come to my birthday party.

Present subjunctive with adverbial clauses

- (31) *Ustedes pueden ir al cine con tal de que me dejen en paz.* You can go to the movies as long as you leave me alone.
- (32) **No puedes ver la televisión a menos que tú terminas la tarea.* You can't watch TV unless you finish your homework.