ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 6, No. 11, pp. 2124-2133, November 2016
DOI: http://dx.doi.org/10.17507/tpls.0611.09

Investigating the Relatedness of Five Different Operationalizations of Explicit Knowledge of English Morphosyntax: A Levels of Processing Approach

Amir H. Sarkeshikian
Department of English Language, Science and Research Branch, Islamic Azad University, Tehran, Iran

Mohammad Alavi
Department of English Language, Faculty of Foreign Languages and Literature, University of Tehran, Tehran Iran

***Abstract*—This study aimed to investigate the relatedness of five different operationalizations of explicit knowledge of English morphosyntax within the framework of levels of processing. Two groups of university students, majoring in English translation at BA level, participated in the study. For the purpose of the study, five different operationalizations of the explicit knowledge of English morphosyntax were drawn from the literature and two metalinguistic knowledge tests were constructed. The two measures were piloted; the one with 52 items showed strong overall internal consistency (α=0.98) and the other one with 30 items showed an acceptable overall internal consistency (α=.70). The data analyses using Pearson correlation revealed that there is no high correlation between the five variables, and therefore no serious covariation problem was detected. Moreover, the results of repeated-measures ANOVA showed that there was a significant within-group difference between the performances of each group on the metalinguistic test. The results of the study are discussed within the level of processing (LoP) framework and implications are offered.**

***Index Terms*—levels of processing, metalinguistic knowledge, explicit knowledge, English morphosyntax**

I. Introduction

In the process of second language acquisition, besides acquiring the implicit knowledge of language, learners can develop a kind of explicit knowledge which enables them to reflect on the language and its use intentionally through an element of reflexiveness, called metalanguage, which helps talk about the language itself (Berry, 2005). As such, language acquisition remains far from complete after its initial acquisition as long as its metalanguage is not acquired because it plays a fundamental role in the development of a language (Robinson, 2005).

There are two main approaches to the construct; metalanguage as knowledge and metalanguage as awareness (Berry, 2005). On the one hand, the former, broadly speaking, refers to the explicit knowledge that language users have about a language (Roehr & Gánem-Gutiérez, 2009). The explicitness of the knowledge denotes transferability into awareness, consciousness, potential verbalizability, declarative knowledge, imprecision and inaccuracy, controlled processing and unlimited learning (Ellis, 2009). This kind of knowledge as Kaufman, DeYoung, Gray, Jiménez, Brown and Mackintosh (2010) have asserted, involves “the conscious, deliberate, and reflective learning processes” (p.322).

On the other hand, metalinguistic awareness is the other approach to the study of metalanguage in its own right. Masny (1987, p.59, as cited in Elder, 2009) obviously defined metalinguistic awareness as “an individual’s ability to match, intuitively, spoken or written utterances with his or her knowledge of language.” As the term intuition implies in the definition, metalinguistic awareness involves implicit knowledge, “the knowledge of a language that is typically manifest in some form of naturally occurring language behavior such as conversation” (Ellis, 2001, p. 252), and therefore, implicit learning, which is “typically characterized by a set of automatic, associative, nonconscious, and unintentional learning processes” (Kaufman et al., 2010, p.322).

Despite a bulk of research into metalanguage, ranging from the investigation of the relationship between metalanguage and language proficiency to the levels of metalanguage in good language teachers, to students’ reactions to the metalanguage in one area of English grammar and the sources of those reactions, to the construct validity and unidimentionality of the metalanguage test and to the component analysis of metalinguistic knowledge (Alderson, Clapham & Steel, 1997; Andrews & McNeill, 2005; Berry, 2004; Elder, 2009, Roehr, 2007; Tokunaga, 2010; Wistner, 2010), the relatedness of different operational definitions of explicit knowledge of English morphosyntax has not been investigated within the levels of processing (LoP) framework for memory research (Craik & Lockhart, 1972).

II. Review of the Related Literature

A. Implicit and Explicit Knowledge of Language

While dealing with explicit knowledge, we cannot avoid making a reference to implicit knowledge given the controversy over the interface between them; as a result, there are three different positions in this regard. The "no interface hypothesis" (Krashen, 1981, 1982) holds that implicit knowledge and explicit knowledge are totally separate as they result from different processes. The "strong interface hypothesis" (DeKeyser, 2007) claims that "explicit and implicit knowledge are not fundamentally distinct but, rather, extremes on one continuum" (De Graaff & Housen, 2009, p.734). The "weak interface hypothesis" (Ellis, 2004, 2005) purports that "implicit and explicit knowledge are two separately coexisting knowledge systems and L2 knowledge ideally starts out as implicit knowledge…through the use of instructional tasks that facilitate noticing…. Explicit knowledge can be promoted by means of awareness-raising tasks …." (De Graaff & Housen, 2009, p.734).

As a result of the abovementioned differing opinions on the two types of knowledge, Ellis (2004, 2009) decided to make a tradeoff and characterized explicit knowledge as a kind of knowledge which is "conscious", "declarative", "imprecise and inaccurate", broad and deep, "accessible through controlled processing", exploited as a tool while failing to intuitively judge grammaticality and "learnable", refuting the claims that explicit knowledge is an attitude, a practice or an activity or a pedagogic construct.

B. Levels of Processing (LoP): A Framework for Investigating Metalanguage

Craik and Lockhart (1972) developed a framework for the processing levels in the memory, arguing for the hierarchical analysis of stimuli in a series of stages or levels. In their words, the "conception of a series or hierarchy of processing stages is often referred to as "depth of processing" where greater "depth" implies a greater degree of semantic or cognitive analysis" (Craik & Lockhart, 1972, p.675). They suggested that the processing in the memory could be classified under two types: "Type 1 processing (or maintenance rehearsal) maintained processing at the same level of analysis, whereas Type II processing (or elaborative rehearsal) involved deeper or more extensive processing of the stimulus" (Lockhart & Craik, 1990, p.88). For them, the second type of processing is expected to lead to an enhancement in memory if mnemonic performance is a function of the depth of analysis.

Lockhart and Craik (1990) asserted that humans' hierarchically organized cognitive system processes incoming stimuli at different levels of analysis, "with the products of early (or shallow) sensory analyses serving as the input to later (or deeper) semantic analyses" (p.88). They assumed that the deeper the analysis, the more attentional resources are required, and therefore, the more durable memory trace will be recorded. They thought of the memory trace as a record of such normal cognitive processes as "comprehension, categorization, or discrimination", not as a by-product of a memory-encoding process only committed to memory.

Reviewing the literature shows that the levels of processing framework has crossed the borders of memory research such domains as brain hemispheres (e.g., Bitan, Lifshitz, Breznitz & Booth, 2010), cognitive deficits (e.g., Froger, Taconnat, Landré, Beigneux & Isingrini, 2009), educational psychology (e.g., Chew, 2010), reading and prose comprehension (e.g., Nassaji, 2003), attentional system of language (Talmy, 2007, 2008) and selective attention (e.g., Andersen, Müller & Martinovic, 2012).

Hence, it is clear that assessing the metalinguistic knowledge of morphosyntax within the framework of processing levels may have implications for how we should teach and evaluate grammatical knowledge and translation performance on the following grounds:

Firstly, the distinction between implicit and explicit knowledge of linguistic competence entails that language pedagogy and testing should require performance on such tasks that requires L2 learners to demonstrate their explicit knowledge of grammar as an asset, let alone as a potential source of construct irrelevance variance in discrete-item grammar tests. L2ers may resort to three types of knowledge described as "Other Knowledge, Explicit Linguistic Knowledge, and Implicit Linguistic Knowledge" (Bialystok, 1978, pp.71-72). As Bialystok (1978) postulated elsewhere, "These are, of course, hypothetical constructs …, and … each is considered to contribute in some unique way to the attainment of language proficiency…" (p.72). In the same paper, Bialystok defined explicit knowledge of language as "all the conscious facts the learner has about the language and the criterion for admission to this category is the ability to articulate those facts" (p.72) and assigned three functions to it:

1. A "buffer" for information about a language, presented in an explicit situation, such as new syntactic or morphological forms in a classroom
2. A "store" for explicitly represented information such as the case in which even native speakers consciously attend to distinguish "lie" from "lay" (example adopted from Bialystok, 1978)
3. An "explicit articulatory system" whereby the implicitly represented linguistic knowledge may be made explicit by noticing structures and their constraints

Secondly, the framework may be used to develop criteria for the evaluation of different levels of morphosyntactic metalanguage at which second language learners can "understand and produce both the morphological and syntactic forms of the language," "…interpret and express meanings from inflections…, derivations…, syntax … (e.g., subjunctive mood) or show focus, emphasis or contrast (e.g., voice and word order)", "…identify the direct language function associated with language use" and "…encode a wide range of pragmatic meanings" (Purpura, 2004, p.94).

Thirdly, processing metalanguage may involve data-driven processing (DDP) or conceptually driven processing (CDP) (Jacoby, 1983); the former refers to “simple maintenance rehearsal of instances of input in memory, and the later to the “elaborative rehearsal and the activation of schemata or higher-order relations from long-term memory” (Robinson, 1995, p.299). DDP, as Robinson (1995) put, requires “accumulation and rehearsal of instances encountered with input in the memory and may lead to the development of simple patterns of association between co-occurring items” (pp.301-302) and CDP demands “the elaboration of input following activation of the schemata” (p.302). Accordingly, in the process of instructed second language acquisition and translation as well, besides the data-driven implicit knowledge of language, learners can, to a greater extent, develop a kind of conceptually-driven elaborative knowledge which enables them to intentionally reflect on the language and its use through metalanguage.

In view of the gap in the literature regarding the processing levels of English morphosyntactic metalanguage, this study, as part of an ongoing project, is an attempt to answer the following research questions:

RQ: Is there any statistically significant within-group difference between English translation students’ performance on five different types of English morphosyntactic metalinguistic question, viz, error explanation, error correction, grammatical function identification, rule use justification and rule explanation?

III. Method

A. Experiment 1

1. Participants

The participants were 30 university students, majoring in English translation. The students’ language proficiency was not taken into account since all participating students had passed compulsory English courses at high schools 2 hours a week for four consecutive years as a requirement of Iranian high school curriculum, and passed the entrance university exams, held nationwide for admission purposes. Moreover, they all had passed two compulsory four-credit grammar courses at the university where they received explicit instruction on English grammar. At the same time, language proficiency measures have been found to be uncorrelated with metalinguistic knowledge tests (Alderson et al., 1997)

2. Variables

The variables of the study were five qualitatively distinct types of question which expected test-takers to explicitly show their knowledge of English morphosyntactic metalanguage by explaining why an ill-formed sentence is ungrammatical, correcting grammatically erroneous sentences, identifying the grammatical function, justifying the use of grammatical structures and explaining the function that rules play in sentences. In line with Craik (2002), the measure of depth of processing involved in performing the five orienting questions would be the numerical output from functions that test-takers apply to the processes involved in performing them; thus, level or depth of processing is “appropriately thought of as the end product of research, not the starting point” (Craik, 2002, p.92).

3. Instrument

A metalinguistic knowledge test was developed by the researchers. To that end, 10 TOEFL tests released by the ETS were selected from which the researcher randomly selected 50 structures on which participants had received instruction. The test of English morphosyntactic metalanguage knowledge (Henceforth called MMKT) was to assess the test-takers’ explicit knowledge of English morphosyntactic metalanguage. The test-takers were allowed to complete the test in 100 minutes.

In a non-balanced design, 15 items were allocated to error explanation (subtest 1), 15 items to error correction (subtest 2), 11 items to identification of grammatical function (subtest 3), 8 items to rule explanation (subtest 4) and 3 items to justification of rule use (subtest 5). The types of question were based on the differential operationalizations of the construct of English Morphosyntactic Metalanguage in the literature. Each section contained ten multiple-choice items. Each item will be worth one point. All items will be in multiple-choice format so that they can be objectively scored. As explicit linguistic knowledge can be viewed as an analyzed knowledge (Bialystok, 1994), discrete-point tests with separate-and-explicit or selected-response tasks may be useful for assessment purposes (Rea-Dickins, 2001). The measure was piloted, showing strong overall internal consistency (α=0.98).

4. Results

The results of the descriptive statistics, calculated for the data set, showed that the distribution of scores on four subtests is approximately symmetric as the skewness values fell well within the range of −/+1.00 (Table I). However, the subtest 4 did not meet the assumption of normality (skewness = +1.807), which may be due to its difficulty level, and therefore depth of processing. However, the transformation was not done since it usually increases the difficulty of the interpretation of the data and may alter the results of the study (Leech, Barret, & Morgan, 2005).

TABLE I
DESCRIPTIVE STATISTICS FOR FIVE SUBTESTS OF MORPHOSYNTACTIC METALINGUISTIC KNOWLEDGE TEST

	N	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Subtest 1	30	.3867	.15403	-.350	.427	-.727	.833
Subtest 2	30	.4267	.19777	-.190	.427	-.734	.833
Subtest 3	30	.2259	.16110	.811	.427	.384	.833
Subtest 4	30	.1542	.23370	1.807	.427	2.869	.833
Subtest 5	30	.3111	.33828	.565	.427	-1.039	.833
Valid N (listwise)	30						

Then, to check for the collinearity problems by highly correlated subtests, the relationships among the scores on the subtests were explored as correlations. Table II shows the results of correlations between the subtests of the MMKT. The subtest 1, the explanation of ungrammatical structures subtest, was moderately correlated with the subtest 2, the error correction subtest (p <.01, r= .64), and partially with the subtest 3, the identification of grammatical function subtest (p <.05, r= .39). Moreover, subtest 4 was moderately correlated with subtests 3 (p <.01, r= .59) and 5 (p <.01, r= .55).

TABLE II
CORRELATION BETWEEN THE FIVE SUBTESTS OF MORPHOSYNTACTIC METALINGUISTIC KNOWLEDGE TEST

Subtests	1	2	3	4	5
Error explanation	-				
Error correction	.641**	-			
Grammatical function identification	.393*	.237	-		
Rule use justification	.099	-.030	.595**	-	
Grammatical function explanation	.038	-.243	.283	.554**	-

* p <.05. **p < .01.

Although there was no highly correlated pair of variables in the data set, the degree of correlation coefficient decrease by controlling for the subtest 1 was larger than that of correlation decrease by controlling for any other variable, meaning that the subtest 1 affected the relationship between the subtests (Table III). Nevertheless, the partial correlation matrix shows that the subtest 2 (i.e., the error correction subtest) had negative relationship with all other four subtests.

TABLE III
PARTIAL CORRELATION BETWEEN FOUR SUBTESTS OF ENGLISH METALINGUISTIC KNOWLEDGE TEST

Control variable	Subtests	1	2	3	4	5
Error explanation	Error explanation	-				
	Error correction	-.141	-			
	Function identification	.049	-.021	-		
	Rule use justification	.022	-.122	.608	-	
	Grammatical function explanation	.087	-.349	.291	.553	-

* p <.05. **p < .01.

Then, a repeated-measures ANOVA, recommended for "one independent variable with two or more levels that are repeated measures" (Leech et al, 2005), was run to assess whether there were significant difference among the performances of test-takers on the five levels of the morphosyntactic metalinguistic knowledge test. Table IV shows that all four multivariate tests of the within-subjects effect were significant, $F(4,26) = 8.67$, $p< .05$. It indicates that there was a significant difference in how the subtests were answered. The Mauchly test of sphericity was significant, thus the data violated the sphericity assumption of the univariate approach to repeated measures analysis of variance (Table V).

TABLE IV
MULTIVARIATE TESTS

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's Trace	.582	9.034[a]	4.000	26.000	.000	.582	36.137	.997
Wilks' Lambda	.418	9.034[a]	4.000	26.000	.000	.582	36.137	.997
Wilks' Lambda	1.390	9.034[a]	4.000	26.000	.000	.582	36.137	.997
Roy's Largest Root	1.390	9.034[a]	4.000	26.000	.000	.582	36.137	.997

a. Exact statistic
b. Computed using alpha = .05
c. Design: Intercept
Within Subjects Design: MMKT

As such, the Greenhouse-Geisser test, a commonly recommended F-test, was used for drawing conclusions about the within-subject effects. Even with Greenhouse-Geisser adjustment, the within-subjects effects of performance on the subtests were significant, $F(1.99, 57.6) = 9.24$, $p < .001$. This means that the test-takers' performance on the five subtests were significantly different (Table IV). Thus, the null hypothesis is rejected by all multivariate tests.

TABLE V
MAUCHLY'S TEST OF SPHERICITY

Measure:MEASURE_1

					Epsilon[a]		
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
MMKT	.130	55.913	9	.000	.497	.533	.250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept Within Subjects Design: MMKT

TABLE VI
TESTS OF WITHIN-SUBJECTS EFFECTS

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
factor1	Sphericity Assumed	1.513	4	.378	9.240	.000	.242	36.961	.999
	Greenhouse-Geisser	1.513	1.987	.761	9.240	.000	.242	18.361	.970
	Huynh-Feldt	1.513	2.133	.709	9.240	.000	.242	19.707	.977
	Lower-bound	1.513	1.000	1.513	9.240	.005	.242	9.240	.836
Error(factor1)	Sphericity Assumed	4.748	116	.041					
	Greenhouse-Geisser	4.748	57.625	.082					
	Huynh-Feldt	4.748	61.850	.077					
	Lower-bound	4.748	29.000	.164					

a. Computed using alpha = .05

Then, the polynomial contrast was chosen on the assumption that the subtests can be ordered according to the hierarchical level of processing. Despite a significant linear trend, $F(4,116) = 9.24$, $p <.05$, $\eta^2= .24$, this trend was not documented over and above an exploratory venture since more proofs are needed to evaluate the fit of the trend. Given that the population effect size were not known, the observed power computed by the SPSS software was not taken into account since it was a function of significant criterion, sample size and population effect size.

B. *Experiment 2*

1. Participants

The participants were 30 bachelor students, majoring in English translation and English language teaching, from two universities in Iran. The students' language proficiency was not taken into account for the reasons explained in the first experiment. The Test-takers demonstrated their explicit morphosyntactic metalanguage by answering five qualitatively distinct types of question like in Experiment 1.

2. Instrument

Another metalinguistic knowledge measure was developed by the researchers. To that end, 30 structures were randomly selected from 10 TOEFL tests released by the ETS. Test-takers were allowed to complete the test in 60 minutes. In a non-balanced design, 7 items were allocated to identification of grammatical function (subtest 1), 7 items to the error correction (subtest 2), 6 items to error explanation (subtest 3), 6 items to rule explanation (subtest 4) and 4 items to justification of rule use (subtest 5). The question types were based on the same operationalizations of the construct in the literature. Each item will be worth one point. All items will multiple-choice so that they could be objectively scored. The measure was test-piloted and showed acceptable overall internal consistency (α=0.70). The reduction in reliability can be attributed to the decrease in the number of items or the non-balanced design of the measure.

3. Results

First, the data were subjected to descriptive statistics. The skewness for all subtests was well within the range of −/+1.00 (Table VII), meaning the distribution was approximately symmetric. Then, to check for the collinearity problems created by highly correlated variables, bivariate correlations between the subtests of the MMKT were calculated (Table VIII).

TABLE VII
DESCRIPTIVE STATISTICS FOR FIVE SUBTESTS OF MORPHOSYNTACTIC METALINGUISTIC KNOWLEDGE TEST

	N	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Subtest 1	30	.5238	.25260	.105	.427	-.654	.833
Subtest 2	30	.6571	.17835	.155	.427	-.074	.833
Subtest 3	30	.6278	.20846	.248	.427	-.072	.833
Subtest 4	30	.4111	.26527	.312	.427	-.304	.833
Subtest 5	30	.4583	.23747	.096	.427	-.206	.833
Valid N (listwise)	30						

As the correlation matrix indicates, the subtest 1 (i.e., the explanation of ungrammatical structures subtest) was partially correlated with the subtest 2 (i.e., the error correction subtest) (p <.05, r= .40) and moderately with the subtest 4, (i.e., the justification the use of grammatical structures subtest) (p <.01, r= .53). At the same time, the other subtests were not correlated.

TABLE VIII
CORRELATION BETWEEN THE FIVE SUBTESTS OF MORPHOSYNTACTIC METALINGUISTIC KNOWLEDGE TEST

Subtests							5
Error explanation		-					
Error correction	.406*	-					
Grammatical function identification		.088	.336	-			
Rule use justification	.535**	.357	.212	-			
Grammatical function explanation		.253	.145	.343	.144	-	

* p <.05. ** p < .01.

TABLE IX
PARTIAL CORRELATION BETWEEN FOUR SUBTESTS OF ENGLISH METALINGUISTIC KNOWLEDGE TEST

Control variable	Subtests	1	2	3	4	5
Error explanation	Error explanation	-				
	Error correction	-.295	-			
	Function identification	.187	.329	-		
	Rule use justification	-.053	.182	.196	-	
	Grammatical function explanation	.236	.048	.333	.011	-

* p <.05. ** p < .01.

Despite the lack of any highly correlated pair of variables in the data set, the results of partial correlations demonstrate that the degree of correlation decrease by eliminating the subtest 1 was larger than that of correlation decrease by eliminating any other subtest (Table IX).

Then, as recommended by Leech et al. (2005), a repeated-measures ANOVA was run to assess whether there were significant differences among the performances of test-takers on the five subtests of the morphosyntactic metalinguistic knowledge test.

TABLE X
MULTIVARIATE TESTSC

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's Trace	.571	8.667[a]	4.000	26.000	.000	.571	34.669	.996
Wilks' Lambda	.429	8.667[a]	4.000	26.000	.000	.571	34.669	.996
Wilks' Lambda	1.333	8.667[a]	4.000	26.000	.000	.571	34.669	.996
Roy's Largest Root	1.333	8.667[a]	4.000	26.000	.000	.571	34.669	.996

a. Exact statistic
b. Computed using alpha = .05
c. Design: Intercept Within
Subjects Design: *MMKT*

Table X shows that all four multivariate tests of the within-subjects effect were significant, indicating that there is a difference in how the subtests are answered, $F(4,26) = 8.67$, $p< .05$, $\eta^2=.57$. The Mauchly test of sphericity was not significant, thus the data did not violate the sphericity assumption of the univariate approach to repeated measures analysis of variance (Table XI).

TABLE XI
MAUCHLY'S TEST OF SPHERICITY

Measure:MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
MMKT	.667	11.113	9	.269	.835	.957	.250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept Within Subjects Design: MMKT

TABLE XII
TESTS OF WITHIN-SUBJECTS EFFECTS

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
factor1	Sphericity Assumed	1.346	4	.337	8.759	.000	.232	35.034	.999
	Greenhouse-Geisser	1.346	3.341	.403	8.759	.000	.232	29.265	.996
	Huynh-Feldt	1.346	3.829	.352	8.759	.000	.232	33.534	.999
	Lower-bound	1.346	1.000	1.346	8.759	.006	.232	8.759	.816
Error(factor1)	Sphericity Assumed	4.458	116	.038					
	Greenhouse-Geisser	4.458	96.898	.046					
	Huynh-Feldt	4.458	111.032	.040					
	Lower-bound	4.458	29.000	.154					

a. Computed using alpha = .05

Nonetheless, both multivariate tests and Greenhouse-Geisser F-test were used for drawing conclusions about the within-subject effects. Even with Greenhouse-Geisser correction, the within-subjects effects of performance on the subtests were significant, $F(3.34, 96.90) = 8.75$, $p < .001$, $\eta^2=.23$, as all four multivariate tests suggested the rejection of the null hypothesis (Table XII). The polynomial contrast was chosen on the assumption that the subtests can be ordered from the deepest to the shallowest level of processing. Despite a significant linear trend, $F(1,29) = 37.1$, $p <.05$, $\eta^2= .56$, this trend is not documented for the same reason mentioned for the first experiment. Moreover, as power figures may be misunderstood, the observed power calculated is not taken into consideration.

IV. DISCUSSION

Based on the results of the statistical analysis, there are statistically significant differences in how the five qualitatively different types of metalinguistic questions in both experiments were answered. The results of correlational analyses in both experiments showed that the variables of the study were not highly correlated. Nevertheless, the first subtest (i.e., error explanation) was moderately and partially correlated with subtests 2 and 3 (i.e., error correction and grammatical function subtests), in the first experiment and partially and moderately correlated with the subtests 2 and 4 (i.e., error correction and rule use justification subtests) in the second experiment. When Subtest 1 was controlled, none of the variables were correlated. It can be assumed in instructed second language acquisition, the ability to explain grammatical errors may help "refine … schemata ..." (Frank & Wilson, 2000, p.9).

Regardless of the strength of correlation coefficients, in the first experiment the ability to correct errors was negatively correlated with all other variables, while in the second experiment it was not. In line with Yonelinas and Levy (2002), it may be argued that that in experiment 1 the subjects induced their implicit metalinguistic information about the morphosyntactic structures for recognition and correction, whereas in experiment 2 the subjects recollect explicit morphosyntactic metalanguage about the underlined part of a sentence and choose from among four distracters one of which is the correct. These results may conform to a dual-process recognition memory model for metalinguistic knowledge (Yonelinas & Levy, 2002).

In view of the results, the first subtest influenced the way other constructs were correlated, and when it was controlled, the correlation coefficients decreased so that none of them were significant anymore. It seems that explanation promotes "learning, beyond simply memorizing or passively encoding" (Williams & Lombrozo, 2010, p.777). In fact, the results are consistent with the research findings that suggest attempts to explain may lead to greater learning as compared with such strategies as receiving feedback without explanations and thinking aloud (Amsterlaw & Wellman, 2006). As such, when asked to explain grammatical rules, learners need to interpret them in terms of their linguistic schemata and the constructed explanatory information result in the "explicit recognition of generalizations that underlie what is being explained" (Williams & Lombrozo, 2010, p.778).

Given the significant correlation coefficients before calculating partial correlations in both experiments, it seems that the ability to explain errors may involve the L2 mental representations that may influence performance on other task. This finding seems to confirm Frank and Wilson (2000), who stated that engaging in explanation may lessen "cognitive dissonance" or increase the compatibility and coherence of mental representations, besides solving internal anomalies.

Finally, all four multivariate tests suggested the test-takers' performance on the five subtests were significantly different. Based on the within-subjects effects and significant linear trend by the polynomial contrast, it can be concluded that the subtests demand the schemata that may be ordered hierarchically (Craik & Lockhart, 1972). However, more proofs are needed to evaluate the fit of the trend because the observed power is a function of significant criterion, sample size and population effect size. Therefore, it might be misleading to discuss the power observed by the ANOVA as O'Keefe (2007) has argued, power analysis is "a useful supplement to *p* values and confidence intervals, but only when based on population effect magnitudes of independent interest" (p.295).

V. Conclusion

The development of metalinguistic knowledge is different than that of the primary linguistic skills since it is acquired later than the primary linguistic skills (Kuafman et al., 2010). The metalinguistic knowledge, and the ways in which it is manifested are dependent on the second language learner, and on the experiences that s/he brings with her/him to the SLA process. Put this way, metalinguistic behavior is broad-range, covering much more than merely language about language.

For that reason, the term metalinguistic should be applied to "a set of problems which share certain features" (Bialystok & Ryan, 1985, pp.230-231). By forming propositional networks of explicit knowledge of language, L2 learners can connect the different properties of structures and enhance their L2 proficiency. Hence, any type of elaboration is better than none for encoding and retrieving information; however, some elaborations are … better than others" (p.156). Accordingly, explaining why a sentence is ungrammatical, why a certain grammatical structure is used and what the function of a structure is may help in different ways build the schemata the explicit knowledge of language. Moreover, specifying what the grammatical function of word(s) is and what syntactic relations exist between linguistic units help metalanguage knowledge be organized, and therefore, be retrieved more effectively.

Accordingly, developing morphosyntactic metalanguage through different explanatory tasks may help learners notice and remedy the gaps or holes in their interlanguages. It is because the level of processing influences the extent to which metalinguistic knowledge is encoded in the memory (Robinson, 1995), and by focusing L2 learners' attention on different aspects of metalanguage, language learning may, therefore, improve in different ways since "familiarity with basic metalinguistic principles for describing structural patterns and structural analogies would probably aid hypothesis testing by directing attention to relevant features of the input to be noticed" (Robinson, 1995, p.320).

Considering the limitations of this study, further research should be done in order to find the standard questions whose processing levels are more obvious, and even more tasks can be identified to recruit other levels. Future researchers need to consider the effect of contextual probabilities of structures, their recency, frequency and salience as remembering the metalinguistic knowledge may be a function of variables such as teaching method, metalinguistic rhetoric used by the teacher, L1 use, types of analysis, and teachers' language awareness. Finally, the future theories or models of pedagogical grammar would be required to consider both the metalinguistic and the mental processes underlying them.

REFERENCES

[1] Alderson, J. Ch., C. Clapham & D. Steel. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research* 1.2, 93–121. doi: 10.1177/136216889700100202.

[2] Andersen, S.K., M.M. Müller & J. Martinovic. (2012). Bottom-up biases in feature-selective attention. *Journal of Neuroscience* 32.47, 16953–16958.

[3] Andrews, S. & A. McNeill. (2005). Knowledge about language and the 'Good Language Teacher'. In N. Bartels (ed.), *Researching applied linguistics in language teacher education*. New York: Springer Science+Business Media LLC, 159–178.

[4] Amsterlaw, J. & H.M.Wellman. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognition and development* 7.2, 139–172.

[5] Berry, R. (2004). Awareness of metalanguage. *Language Awareness* 13.1, 1–16. doi: 10.1080/09658410408667082.

[6] Berry, R. (2005). Making the most of metalanguage. *Language Awareness* 14.1, 3–20. doi: 10.1080/09658410508668817.

[7] Bitan, T., A. Lifshitz, Z. Breznitz & J.R. Booth. (2010). Bidirectional connectivity between hemispheres occurs at multiple levels in language processing, but depends on sex. *J Neurosci* 30.35, 11576–11585. doi:10.1523/JNEUROSCI.1245-10.2010.

[8] Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning* 28.1, 81–103.

[9] Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition* 17, 157–168.

[10] Bialystok, E. & E. B. Ryan. (1985). A metacognitive framework for the development of first and second language skills. In D.L. Forrest-Pressley, G.E. Mackinnon & T.G. Waller (eds.), *Meta-cognition, cognition, and human performance*. New York: Academic Press, 207–252.

[11] Chew, S.L. (2010). Improving classroom performance by challenging student misconceptions about learning. *Observer* 23.4. http://www.psychologicalscience.org/index.php/publications/observer/2010/pril-10/improving-classroom-performance-by-challenging-student-misconceptions-about-learning.html (accessed 9/6/2015).

[12] Craik, H.L.M. (2002). Levels of processing: Past, present … and future? *Memory* 10.5, 305–318. doi: 10.1080/09658210244000135.

[13] Craik, F.I.M. & R.S. Lockhart. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

[14] De Graaff, R. & A. Housen. (2009). Investigating the effects and effectiveness of L2 instruction. In M. H. Long & C. J. Doughty (eds.), *The handbook of language teaching*. New York: Blackwell, 726–755.

[15] DeKeyser, R. M. (2007). Practice in a second language: Perspectives from applied linguistics and cognitive psychology. Cambridge: Cambridge University Press.

[16] Drosopoulos, S., U. Wagner & J. Born. (2005). Sleep enhances explicit recollection in recognition memory *Learn and Memory* 12, 44–51.

[17] Ellis, R. (2009). Implicit and Explicit Learning, Knowledge and Instruction. In R.S. Ellis, S. Loewen, C. Elder, J. Erlam, J. Philip & H. Reinders (eds.), *Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol: Multilingual Matters, 3–26.

[18] Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition* 27.2, 141–172. doi: 10.10170S0272263105050096.

[19] Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning* 54.2, 227–275.

[20] Ellis, R. (2001). Some thoughts on testing grammar: an SLA perspective. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. O'Loughlin (eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies*. Cambridge: Cambridge University Press, 251–263.

[21] Elder, C. (2009). Validating a test of metalinguistic knowledge. In R.S. Ellis, S. Loewen, C. Elder, J. Erlam, J. Philip & H. Reinders (eds.), *Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol: Multilingual Matters, 113–138.

[22] Frank C. K. & R.A. Wilson. (2000). Explaining explanation. In F. C. Keil & R. A. Wilson (eds.), *Explanation and cognition*. Massachusetts: MIT Press, 1–18.

[23] Froger, C. L., L. Taconnat, K. Landré, K. Beigneux & M. Isingrini. (2009). Effects of level of processing at encoding and types of retrieval task in mild cognitive impairment and normal aging. *J Clin Exp Neuropsychol* 31.3, 312–321. doi: 10.1080/13803390802112554.

[24] Jacoby, L.L. (1983). Remembering the data: Analyzing interactive processes in reading. *Verbal learning and verbal behavior* 22.5, 485–508.

[25] Kaufman, S.B., C.G. DeYoung, J.R. Gray, L. Jiménez, J. Brown & N. Mackintosh. (2010). Implicit learning as an ability. *Cognition* 116, 321–340. doi:10.1016/j.cognition.2010.05.011.

[26] Krashen, S.D. (1981). Second language acquisition and second language learning. Oxford: Pergamon.

[27] Krashen, S.D. (1982). Principles and practice in second language acquisition. Oxford: Pergamon.

[28] Leech, N.L., K.C. Barrett, & G.A. Morgan. (2005). SPSS for intermediate statistics: Use and interpretation (2nd edn.). Mahwah: Lawrence Erlbaum Associates.

[29] Lockhart, R.S. & H.L.M. Craik. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology* 44.1, 87–112.

[30] Nassaji, H. (2003). Higher–level and lower–level text processing skills in advanced ESL reading comprehension. *MLJ* 87.2, 261–276. doi: 10.1111/1540-4781.00189.

[31] O'Keefe, D.J. (2007). Post hoc power, observed power, a priori power, retrospective power, prospective power, achieved power: Sorting out appropriate uses of statistical power analyses. *Communication Methods and Measures* 1.4, 291–299. doi:10.1080/19312450701641375.

[32] Purpura, J. (2004). Assessing grammar. Cambridge: Cambridge University Press.

[33] Robinson, M. (2005). Metalanguage in L1 English-speaking 12-year-olds: Which aspects of writing do they talk about? *Language Awareness* 14.1, 39-55. doi:10.1080/09658410508668819.
[34] Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning* 45.2, 283-331.
[35] Roehr, K. (2007). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics* 29.2, 173–199. doi:10.1093/applin/amm037.
[36] Roehr, K. & G. A. Gánem-Gutiérrez. (2009). Metalinguistic knowledge: A stepping stone towards L2 proficiency? In A.G. Benati (ed.), *Issues in second language proficiency*. New York: Continuum, 79–94.
[37] Talmy, L. (2008). Aspects of attention in language. In N. Ellis & P. Robinson (eds.), *Handbook of cognitive linguistic and second language acquisition*. New York: Routledge, 27–38.
[38] Talmy, L. (2007). Attention phenomena. In D. Geeraerts & H. Cuyckens (eds.), *The Oxford handbook of cognitive linguistic*. Oxford: Oxford University Press. 264–293.
[39] Tokunaga, M. (2010). Metalinguistic knowledge of low-proficiency university EFL learners. In A. M. Stoke (ed.), *JALT 2009 Conference Proceedings.* Tokyo: JALT.
[40] Williams, J.J. & T. Lombrozo. (2010). The role of explanation in discovery and generalization: evidence from category learning. *Cognitive Science* 34, 776–806.
[41] Wistner, B. (2010). Applying the Rasch model in the validation of a test of metalinguistic knowledge. *JALT* 61, 101–107.
[42] Yonelinas, A.P. & B.J. Levy. (2002). Dissociating familiarity from recollection in human recognition memory: Different rates of forgetting over short retention intervals. *Psychonomic Bulletin & Review* 9.3, 575–582.

Amir H. Sarkeshikian received his PhD in TEFL from Science and Research Branch, IAU, Tehran, Iran. He is an assistant professor at English department, faculty of humanities, Qom branch, IAU, Qom, Iran. He has taught graduate and postgraduate courses in TEFL for over eleven years. He has published some papers, besides presenting at national and international conferences.

Mohammad Alavi received his PhD in language testing from Lancaster University. He is an Associate Professor of Applied Linguistics at English Department, Faculty of Foreign Languages, University of Tehran, Iran. He has published papers in such accredited journals as Language Testing, Educational Assessment and RELC.