# Nature of L2 Grammatical Proficiency:

# A Multi-Dimensional Approach

Ahmad Reza Lotfi P.O.BOX: 81595-158 Islamic Azad University-Khorasgan Branch, Iran Tel: 98-913-317-8603 E-mail: arlotfi@yahoo.com

Behrouz Nouri Samani (Corresponding author) P.O.BOX: 81595-158 Islamic Azad University - Khorasgan Branch, Iran Tel: 98-311-260-1115 E-mail: Behrouznourisamani@yahoo.com

#### Abstract

This is a study with descriptive hypothesis-testing design investigating the true nature of L2 grammatical proficiency by focusing on its component parts. It examines the performance of 60 female EFL learners between 19 and 30 years of age at 2 levels of proficiency using 2 types of computerized Grammaticality Judgement tasks in different orders (i.e. Written Stimulus GJ task and Oral Stimulus GJ task) for each participant at a 1 week interval for 2 weeks per each task to determine the extent to which L2 learners' performance is affected by a few components of grammatical proficiency chosen and manipulated out of a great many that exist (i.e. written-oral dichotomy [viz. modality]) as well as language proficiency while focusing on Subjacency. Participants rendered GJ tasks on a 2-point interval. Results from analyzing the elicited measurement data indicated that the performance scores of the two proficiency groups in the two Grammaticality Judgement tasks did not differ significantly.

**Keywords:** Grammaticality Judgement task, L2 grammatical proficiency, Oral Stimulus GJ task, Subjacency, Written Stimulus GJ task

#### 1. Introduction

A growing body of research has indicated the importance of syntax ever since the influence of Chomsky's ideas on linguistics. Within syntax, an important syntactic operation exists (i.e. *merger*) which allows words, phrases and clauses to form sentence structures. In fact, it is this operation which allows the formation of grammatical sentences in all languages; that is, words combine to form phrases, phrases combine to form clauses, and clauses combine to form complex sentences. The rationale behind the operation merger is that all languages, English included, have a potentially infinite number of sentences that can be in the form of interrogative sentences, declarative sentences, and the like.

Such complex syntactic structures are the results of the successive merging of pairs of categories to form larger phrases. Such structures serve as projections of overt constituents such as words, phrases, and clauses. Moreover, such syntactic structures consist of covert, null, and empty constituents. For example, one of the important aspects of syntactic structures is the empty category that is of immense significance in the formation of WH-questions. One sample of such empty categories is the trace left behind as the result of different movement types. Traces are of significance in the processing of different sentences. Tree diagrams, which are binary-branching diagrams, represent different syntactic structures and consist of three positions (i.e. *specifier*, *head* and *complement*) (Radford, Atkinson, Britain, Clahsen & Spencer, 2009).

The syntactic structures of sentences are independent of their related meanings, hence the ability to study such syntactic structures independently of their associated semantic contents. However, the sentential ordering of syntactic structures, according to Chomsky, only permits specific orderings, hence the ungrammaticality of some sentences in GJ tasks. A variety of grammar models, which are derivations of UG, try to explain the rationale behind such specific orderings and explore the computational properties of humans responsible for the syntactic

processing of sentences in general. Hence, Universal Grammar is an appropriate choice in this regard because it is of great complexity and power in syntax.

Syntactic computation, which is a major property of human language faculty, is structure dependent. That is, humans have to consider structural hierarchies if they intend to perform syntactic operations. In addition, humans' language faculty can compute and build such binding relations as c-command and indexation that are parts of the resources that constitute their language faculty (Isac & Reiss, 2008). Syntax is a universal combinatory system of humans' linguistic competence, which is an integral part of language faculty and constitutes the basis for all human language instantiations (Ramchand, 2008).

#### 1.1 WH-Questions

Interrogatives are one type of complex syntactic structures and are of two types (i.e. *closed* and *open*). Whereas the former shows subject-auxiliary inversion, the latter shows fronting of the nonsubject interrogative phrase that involves a WH-word with the ability to trigger subject-auxiliary inversion and, as a result, the occurrence of WH-movement. The criterion for classifying all complex syntactic structures which are in the form of questions is the answer types that follow them; that is, questions are either yes/no questions, alternative questions because of the presence of *or*, or WH-questions. WH-questions express a variable in their proposition, and it is the substitution of this variable with a value that gives rise to answers to such WH-questions. In addition, such WH-questions can be in the form of multiple WH-questions (Collins, 2006). In fact, Interrogatives are multiclausal sentences that are syntactically related to each other via subordination because their clauses show unequal syntactic status. *Noun* (i.e. complement), *adjective* (i.e. relative), and *adverb clauses* are three types of subordinate clauses in English (Huddleston & Pullum, 2006).

#### 1.2 Movement

An important syntactic point within WH-questions is movement. Movement is a process in generative grammar that was formulated from the early 1970s. In fact, movement is a syntactic operation by which constituents move and leave behind silent trace copies. Furthermore, movement operations "move as few constituents as possible the shortest distance possible" (i.e. economy principle) (Radford et al. 2009, p.301). In 2000s, traces, which came into existence as the result of movement, were reformulated in terms of *copying* according to which units in one position were duplicated in another (Matthews, 2007).

Movements are of two types, that is A-movements and A-bar movements. In cases where an expression, such as an NP, is moved to a position that can only be occupied by argument expressions, such as the case of passivization, A-movement occurs, whereas in cases where an expression, such as a WH-phrase, is moved to a position that can be occupied by both arguments and adjuncts, A-bar movement occurs (Radford, 2004). Another type of syntactic movement is head movement, such as the cases where a V moves to I (Hawkins, 2001).

Movement can occur both overtly (i.e. before spell-out) and covertly (i.e. after spell-out at LF). Different causes trigger different movement types such as *feature checking* that triggers WH-movements. One of the major characteristics of movements is the resultant *chains* that are formed between the moved WH-phrases (i.e. head of the chain) and the traces left behind by such movements (i.e. foot of the chain). Moreover, movements are constrained by a UG principle termed the subjacency principle (Radford, 1997). The syntactic form of movement acts as a distinguishing factor among languages and, as a result, differs from language to language. Whereas in some languages like Japanese and Chinese, questions are formed without movement, in others, such as English, they require movement. Variable movements exist in other languages such as the case of Turkish (Cook, 1993).

#### 1.3 Subjacency

Subjacency, which was formulated by Chomsky in the late 1970s, refers to the syntactic constraints that block the movement of a unit if that unit crosses more than one bounding node. WH-movement is constrained by subjacency and refers to the movement of WH-forms to the beginning of a sentence or even a clause resulting in the formation of WH-traces. The extracted constituents make short-step moves because long moves give rise to ungrammaticality in constructions. In addition, the Q morpheme of complementizer requires WH-phrases to move into the clause-initial specifier CP-position.

#### 1.3.1 Subjacency in English and Persian

Whereas IPs and NPs (DPs) are the bounding nodes in English, NPs and CPs are the bounding nodes in Persian. Hence, WH-movement must be, in English, to the nearest empty CP to avoid subjacency violations, whereas it must be, in Persian, to the nearest empty IP to avoid subjacency violation. Subjacency is an example of "subtle parameter variation" that extends far beyond L2 input and cannot be developed by L1 grammar. As for bounding nodes, parametric variation is seen across languages (White, 1989, p.111). In some languages where WH-forms are not preposed, such as Persian, WH-forms are called *in-situ;* however, in English such WH-in-situ questions are used as echo questions. Feature strength triggers operator movements in English. In English, the WH-feature of *probe* is strong; hence, the noninterpretable feature of the probe attracts its interpretable counterpart on *goal* overtly, that is before spell-out, and the moved WH-form c-commands its trace (Radford, 1997).

Checking requirements trigger the movement of operator expressions. Checking checks and deletes the noninterpretable formal features, which have no semantic content, upon reaching LF because such features are the viruses of the system, unlike the interpretable formal features that are reserved and interpreted. WH-phrases either move to check (i.e. the operation *attract*, or rather *enlightened self interest*) or to be checked (i.e. the operation *greed*) (Lotfi, 2003).

#### 1.4 WH-Islands

Many of the UG principles are parameterized because they function with varying degrees of effects in different languages. As for the principle of subjacency, it should be noted that, when it comes to the issue of movement, certain construction types are islands; that is, extraction is impossible out of them. Such syntactic islands fall into four categories, namely *object noun complements, subject noun complements, complex NPs*, and *WH-islands*. Accordingly, certain general constraints limit the freedom of the WH-movement rules such as the A-over-A condition, subject condition, the complex noun phrase constraint (CNPC), and WH-island constraint. All of these constraints can be reduced to the cases of subjacency (White, 1988).

#### 1.5 The Nature of Grammar

Grammar is a property of individual minds the nature of which individuals are unaware of (Isac & Reiss, 2008). In fact, the grammar of a language refers to all the rules and principles of that language and is observationally adequate, that is it provides a systematic description of sentence formation in different languages (Haegeman, 1991).

Every language is made up of linguistic ingredients, and such ingredients are arranged based on a set of rules that are called grammar. One part of grammar that deals with the sentence structures is called syntax. As such, the major concern of syntax is *word order*, and those linguists who work within generativism have taken into account syntactic universals as the best piece of evidence for UG. Such universals have even been of constructive contribution in SLA research because they apply to all language types (Magni, Scalise & Bisseto, 2009).

Grammar is a single and complex unit that generates syntactic structures and assigns representations to such structures at two levels, namely phonological form (PF) and logical form (LF). Different opinions have existed about the grammar structure since the start of generative linguistics in the 1950s; however, all such different opinions have agreed on one fact, that is grammar is abstract and is not located in space or time. A speaker's grammar is his knowledge object because such a grammar contains information about his language. In addition, such a grammar is represented in the speakers' minds. Grammars serve in the speaker/hearers' mental processes and mediate the production and perception of speech. Moreover, the study of grammar is highly effective in understanding the nature of linguistic ability of humans because when people speak a language, they follow the formal rules of it, such as rules of syntax (George, 1989).

Understanding the nature of grammar enables researchers to understand the true nature of language. Whereas descriptive grammar describes a speaker's basic linguistic knowledge, the rules that he already knows, and represents the unconscious linguistic knowledge of speakers because of being a model of the mental grammar that every speaker has, prescriptive grammar legislates what a grammar should be (Fromkin, Rodman & Hyams, 2003).

#### 1.6 Research Question

Does the written-oral dichotomy of GJ tasks affect EFL participants' performance on WH-questions?

#### 1.7 Research Hypothesis

There will be no relationship between the written-oral dichotomy of GJ tasks and EFL participants' performance on WH-questions.

#### 1.8 Statement of the Problem

The specific problem under study in this research is to determine the extent to which EFL learners performance in L2 could be attributed to their competence; in other words, the problem is to determine whether L2 learners' access to UG is blurred due to interference emanating from such test-performance variables as written mode and oral mode that are manipulated and selected from among a great many that exist. Lastly, the problem is to

determine if GJ tasks are apt testing instruments for research in SLA and UG-accessibility. Different psychological processes are triggered by different modes of input. Hence, the purpose of the study is to determine whether written and oral modes of stimuli (i.e. WH-questions) presentations yield similar or dissimilar performance types in female EFL participants when such participants are tested on subjacency violations in GJ tasks, and, in case of the existence of dissimilar performance types, which mode is of more positive effects on female EFL participants.

#### 1.9 Significance of the Study

This study has enormous research-oriented significance first and foremost. In addition, this study is of tremendous cognitive significance because it deals with L2 learners' sentential processing via written and oral modes of input presentations. This study is also of significance to test developers because it can fill the gap as to the reliability and validity of GJ tasks in addition to assisting test developers with devising more reliable testing instruments for effectively measuring grammaticality judgements of L2 learners. Each mode will examine L2 learners' grammatical proficiency from a different angle to see if different approaches trigger distinct levels of comprehension of English WH-questions on L2 learners; a case that, if proved, can lead to a major shift in SLA researchers' view towards L2 learners' grammatical proficiency because grammatical proficiency will no longer be viewed as a unified solid whole, but a unit made of numerous component parts with each part being triggered by a distinct task type. The approach to solving the problem is to statistically analyze the within- and between-group performance of the two proficiency levels across two WH-movement GJ tasks over two consecutive weeks.

The primary theoretical implication of the study will be research-bound largely because researchers will get familiar with the nature of GJ tasks and its practical applications in L2 research and UG-accessibility; that is, different researchers will be able to determine whether GJ tasks can be used for effectively measuring different components of participants' grammatical proficiency, such as modality. Moreover, researchers can determine the degree to which GJ tasks are successful in reliably measuring different component parts of L2 learners' grammatical proficiency. The linguistic implication of the study will be to gain a better understanding towards the true nature of L2 grammatical proficiency (i.e. whether grammatical proficiency is a unified whole unit as has always been claimed to be the case by different linguists and SLA researchers such as Lydia White, or a unit consisting of different component parts where each part contributes differently to its functioning). The pedagogical implications of this study will be finding out better ways of improving the grammatical proficiency of L2 learners via adopting new teaching methods that can target different grammatical proficiency component parts more effectively and lead to a better development and a higher proficiency level of L2 learners' grammatical proficiency. In addition, teachers can adopt more effective drills and exercises in the course of teaching L2 learners different principles of subjacency in their L2 because such teachers have now got more familiar with the nature of L2 learners' grammatical proficiency. Language teachers can also change their stance towards the issue of language proficiency because, given the results of this study, they will get a better understanding towards the degree to which L2 learners' language proficiency level can be of significant effects in improving the performance of L2 learners in different tests.

The theoretical implications of the study can change our approach towards the current methods of teaching WH-questions at schools, universities, and like academic settings. Moreover, they can lead to the revision of the established theories of L2 learners' grammatical proficiency and the learning of WH-questions because subjacency is one aspect of how languages work, and facts about the existence of differences in recognizing the written and oral modes of presenting different WH-questions on the L2 learners' part, where subjacency has been violated, can lead to the development of new teaching methods where written and oral modes (i.e. modality) of teaching can be distinguished more efficiently in terms of their instructive features. The resultant formative assessment of such tasks can be used for curriculum evaluation, curriculum development, modification of subsequent learning experiences, and prescription of remedies for group and individual deficiencies.

Lastly, the findings of this study can assist test developers to find out whether they can use GJ tasks for reliably testing the nature of L2 learners' grammatical proficiency. In addition, such findings can encourage test developers to develop more reliable tests and design more meticulous tests for testing the overall nature of L2 grammatical proficiency.

#### 2. Literature Review

Much research in subjacency (e.g., Schachter, 1988) has indicated that different L2 learners' access to the UG principle of subjacency and the proper observance of it in GJ tasks can be either attributed to UG or the effect of L2 learners' L1, that is in cases where L2 learners' L1 resembles L2 (e.g. English) in terms of WH-movements,

L2 learners' observance of subjacency could be either due to UG or their L1, such as the case of Dutch native speakers learning English, whereas in cases where L2 learners' L1 does not resemble L2 in WH-movement, L2 learners' observance of subjacency could only be due to UG, such as the case of Korean native speakers learning English. In addition, such researchers have only tested their L2 learners' knowledge of subjacency under uniform and like conditions and have oftentimes concluded that subjacency is a part of the L2 learners' unconscious knowledge, or rather grammatical proficiency although it may not be consistently obeyed. Other researchers (e.g., Bley-Vroman, Felix & Loup, 1988) have only emphasized UG availability while rendering GJ tasks on subjacency and have concluded that this availability of UG is, at times, only overridden by other factors in performance (White, 1990).

Research on the effect of proficiency in the observance of subjacency (e.g., White, Travis & Maclachlan, 1992) has indicated that all the High-Intermediate adult Malagasy learners of English and only half of the Low-Intermediate group has performed like a native English-speaking control group. In addition, results have proved that L1 is not the only source of the learners' UG-like knowledge, and UG principles remain available in adult L2 acquisition. This obtained result has, however, contradicted Bley-Vroman's belief that states that access to UG is only via the L1. Moreover, it has opposed the belief of those researchers who claim no access to UG at all such as Clahsen and Muysken (1986) (White, 1996).

There is one major caveat to these research types. Although some researchers (e.g., Schachter, 1988) have investigated the source of L2 learners' access to the subjacency principle in L2 to be either UG, L1 or both, and other researchers (e.g., Bley-Vroman, Felix & Loup, 1988) have mostly concluded that subjacency is part of L2 learners' grammatical proficiency; few, if any, researchers have reported the actual nature of L2 grammatical proficiency. The few researchers who have attempted to relate UG-accessibility to maturity effects (e.g., Johnson & Newport, 1991) have considered the matured and developed L2 grammatical proficiency to be a unified solid whole; no researcher has reported grammatical proficiency to be a unit made of numerous component parts with each part being triggered by a distinct task type. Thus, research to this time has measured only the subjacency performance of L2 learners on WH-movement GJ tasks focusing either on poficiency, source of L2 knowledge, the relationship between implicit-explicit knowledge, and performance with the assumption that grammatical proficiency is a unified solid whole, irrespective of their implications of their research findings on the nature of L2 grammatical proficiency.

In a study, White (1988) examined native speakers of French learning English as a second language. Whereas Ś, S, NP and PP are the bounding nodes in English, French bounding nodes are Ś, NP and PP. If L2 learners are guided by UG, they should not violate island constraints in English. The groups were an experimental group of 18 Low-Intermediate and Intermediate French-speaking subjects who were undergoing short-term intensive ESL instruction and a control group of 14 adult English-speaking subjects. The other experimental group consisted of 25 High-Intermediate and Advanced adults who were taking a night course in ESL at a college in Montreal. Both groups reported little exposure to English outside the classrooms. She used two different GJ tasks to establish whether L2 learners were aware that certain forms were impossible. Moreover, a comprehension task was used. She acted according to the logic of Otsu (1981) which stated that in order to test for the presence of some UG principle, one has, first, to show that the learner has reached an appropriate level of sophistication in the interlanguage.

Hence, the tests consisted of both ungrammatical and grammatical sentences with equivalent levels of embedding. To establish an independent measure of the subjects' ESL proficiency level, she administered a cloze test to them. One of the tasks was an unpaced task with a multiple-choice format where 80 sentences were grouped into groups of 3 or 4 with each group being preceded by a context sentence that was not to be judged.

Unlike the usual multiple-choice tests, any sentence had to be read and judged separately. Forty out of 80 sentences were relevant to sbjacency and ECP, and 21 of the 40 sentences were ungrammatical, whereas the other 19 were grammatical and of the same complexity to the ungrammatical ones. The other GJ task was a paced test; that is, subjects received a written version of the test sentences and also heard them on a tape simultaneously. The tape was paced so that only three seconds were allowed for responding to each sentence. In the comprehension task, pairs of sentences were presented to the subjects that were followed by a question to be answered, and subjacency violation was focused on.

Grammatical sentences were included to see whether subjects had reached a level of sophistication in their interlanguage grammar, whereas ungrammatical sentences were included to see whether subjects were aware that extraction out of certain structures was impossible. Testing was done in class time, and the groups took the

tests in different orders for a counterbalance. The scores of the two adult groups on the cloze test were not significantly different from each other.

According to the results, the adult groups showed a much higher level of accuracy when compared with the adolescents on both grammatical and ungrammatical sentences. All in all, the results showed that adolescents performed poorly on all tasks; few differences were seen in their performance on the grammatical and ungrammatical sentences. Both adult groups' performance types on most of the grammatical sentences suggested that they had mastery of complex structures. The hypothesis concerning subjacency was supported in adults because they were bound by subjacency and ECP; in addition, they even did not accept impossible errors. Subjacency was observed because all adults were accurate on the CNPC and extractions from various noun complements. Adults of Group 1 performed low on WH-islands that could be attributed to their L1 parameter. In contrast, adults of Group 2 performed well on WH-islands which could suggest that parameter-resetting is not impossible for L2 learners to achieve. Despite the fact that cloze-test results showed adults of Groups 1 and 2 to be at the same level of ESL proficiency, such adults differed in their performance on WH-islands.

As for ECP, both adult groups observed ECP as it applied to extractions from subjects, but such groups showed the greatest inaccuracy on that-trace violations. ECP operates in a similar fashion in both French and English. Hence, the emergence of certain difficulties cannot be attributed to a failure to reset a parameter. That-trace acts differently from ECP for groups (White, 1988).

#### 3. Method

#### 3.1 Participants

Sixty female EFL university students of the Islamic Azad University of Khorasgan within the age range of 19 to 30 participated in our study. Participants were randomly selected from the accessible EFL student population of the university and divided into two proficiency groups (i.e. High- and Low-Intermediate groups) as determined by the grammar and listening modules of Oxford Placement Test (OPT) (2004). Participants were permanent residents of Iran and lived in Esfahan. Participants were treated according to the ethical principles of psychologists and code of conduct (American Psychological Association, 2001).

#### 3.2 Materials

Two two-choice Grammaticality Judgement tasks with the same number of WH-questions were used to tap female EFL participants' knowledge of forming WH-questions. The tasks examined the effects of two independent variables on the EFL participants' performance in WH-question GJ tasks; that is, the written-oral dichotomy (i.e. modality) of the way stimuli were presented in two GJ tasks in addition to the L2 proficiency level of participants participating in the experiment.

The first GJ task, Written Stimulus GJ task, was a 30-question GJ task based on which female EFL participants were required to recognize the grammatical and ungrammatical WH-questions pertaining to the already-presented written sentences (i.e. written stimuli); that is, we instructed the participants to determine whether the messages conveyed via the on-screen presented sentences (i.e. stimuli) would be grammatically questioned by the WH-questions would pop up four seconds subsequent to the simuli. Having worn headphones over their ears, participants were under instruction in the 30-question Oral Stimulus GJ task to determine whether the orally presented sentences (i.e. stimuli) would be grammatically questioned by the WH-questions that would pop up four seconds subsequent to the stimuli. Having worn headphones that would pop up four seconds subsequent to the stimuli. We used The Authorware Software (2005), an SAT-based software (i.e. Speed and Accuracy Tradeoff), to develop the two GJ tasks.

The notebook computer used for collecting data from participants was of the DELL brand with such processor traits as Intel(R) Core<sup>™</sup> 2Duo CPU, and T9300@ 2.50 GHz; memory (RAM) capacity as 4.00 GB; and system type as 32-bit Operating System.

The tasks tested different cases of subjacency: (a) A-Over-A condition where WH-elements are extracted out of an object noun complements; (b) subject condition where WH-elements are extracted out of a subject noun complements; (c) complex noun phrase constraint (CNPC) where WH-elements are extracted out of complex NPs; and (d) WH-Island constraint where WH-elements are extracted out of a wh-clause. In all such subjacency cases more than one bounding node (NP, PP, S or Ś) is crossed.

The GJ tasks included both grammatical and ungrammatical WH-questions of equal complexity, vocabulary and structure. In addition, filler items were included in all the tasks. Participants were instructed in the two GJ tasks to select only one response for each question, namely either acceptable or not acceptable. The tasks investigated syntactic islands in English; that is, those construction types that are islands with respect to movement.

#### 3.3 Design and Procedure

The research strategy in his study was to manipulate the independent variables (i.e. Written Stimulus and Oral Stimulus GJ tasks) to determine their effects on the dependent variable (i.e. female EFL participants' performance on the two GJ Tasks). We also used a repeated measures experimental research design in this study to test the hypothesis and the problem via administering all the two GJ tasks to all the participants over two weeks.

Testing occurred over two consecutive sessions for each participant. In the first session, one of the two GJ tasks was administered to each participant; in the second session, the other GJ task was administered to the same participant. That is, in the first session, the Written Stimulus GJ task was used for the first group, whereas Oral Stimulus GJ task was used for the second group; in the second session, Oral Stimulus GJ task was used for the first group, whereas Written Stimulus GJ task was used for the second group; in the second group.

The control features used were the randomization in the selection of participants and counterbalancing that was implemented by changing the order of not only tasks administered to each participant, but also items in the same tasks used for different participants. To gather the required experimental data, we dedicated a coding sheet to each participant so that we would be able to record the scores of each participant.

Having received clearly worded instructions on how to take the tasks, each participant would take her seat and use the notebook computer to take part in the experiment. In the Written Stimulus GJ task, participants, first, received a written stimulus in the form of a sentence for each item followed by a four-second-later related WH-question to be judged for grammaticality. Each written stimulus and its respective WH-question would remain on the notebook computer screen until answered. Each item was a two-choice question. We instructed the participants to left-click for acceptability and right-click for unacceptability. In the Oral Stimulus GJ task, the stimuli were in the form of oral sentences; that is, each question item was, first, orally presented and followed by a related four-second-later WH-question to be judged for un/acceptability.

Each participant took part in the computerized GJ tasks separately in a totally empty university class where no distractions existed at all. Further, an attempt was made to observe the same class lighting and time of the day for administration, amount of words used in giving instructions, silence of the class, and other ambient conditions for each participant so that each participant would take the tests in nearly similar testing conditions.

#### 3.4 Statistical Procedures

The actual statistical steps carried out to conduct this study included the use of the parametric test of two-way repeated measures ANOVA. We used the test of two-way ANOVA to test the hypothesis (where the effects of two independent variables [i.e. written and oral modes of stimuli presentations in GJ tasks] on female EFL participants' performance [i.e. dependent variable] were tested).

Since the intention was to compare participants' performance on more than one GJ task, we used the repeated measures ANOVA to be able to compare the overall performance of the High-Intermediate group with the overall performance of the Low-Intermediate group. Each participant in both High- and Low-Intermediate groups did all the two GJ tasks. As a result, we had four sets of results to be compared for each group; that is, two sets of scores from the performance of High-Intermediate group in both Written Stimulus and Oral Stimulus GJ tasks.

Each group was analyzed separately via using repeated measures ANOVA in order to measure how a single group performed on multiple measures (i.e. the two GJ tasks) and to measure if the effects of those measures were independent. In addition, the overall performance of one group was statistically compared with the overall performance of the other group to see if proficiency was of any effects on the overall performance of the groups.

#### 4. Results

Having entered the obtained data into the variable SPSS data view window with their descriptive labels, we obtained the indispensable general tendencies in the data and the overall spread of the scores via analyzing *measures of central tendency and variability* as shown in Table 1.

The numerical values of *mean* indicated that HG-EXP-WRI did better than the other three independent variables because its mean value was weighted in its favor; hence, proficiency level had been of effect in this regard because the High-Intermediate group had performed better than the Low-Intermediate group in both of its performance types. In addition, the mean values tell us that both High- and Low-Intermediate groups have performed better in the written mode of GJ tasks than the oral mode, an issue that supports the prediction of the

posed hypothesis in this study; however, the question remains as to whether such differences in favor of the written modes are statistically significant. In cases where standard deviation values are smaller than mean values, such as the case in Table 1, mean values can better capture the behavior of participants.

The numerical values of standard deviation, as shown in Table 1, give us exact ideas of the typical behavior of the two groups and show that HG-EXP-WRI has the least variability (i.e. more homogeneous in terms of performance on the Written Mode of GJ tasks), whereas LG-ORL-WRI has the most variability and is, as a result, less homogeneous. Hence, HG-EXP-WRI shows the most degree of homogeneity, whereas LG-ORL-WRI shows the least degree of homogeneity.

Moreover, the calculated standard deviation statistics, portrayed in Table 1, show how scores are dispersed around the means. Examining the standard deviation and mean values in relation to one another, we found that the two performance types of the Low-Intermediate group (i.e. LG-ORL-WRI and LG-EXP-WRI) were more or less equally dispersed from the means, whereas the two performance types of the High-Intermediate group (i.e. HG-ORL-WRI and HG-EXP-WRI) were differently dispersed from their means.

In addition, the differences between the calculated mean statistics of two Low-Intermediate *means* (i.e. 17.03 and 16.00) and two High-Intermediate *means* (i.e. 17.60 and 17.37) showed that proficiency level had been of effects on the performance of groups; that is, there was more clustering of scores around the midpoint and, as a result, less variability in the two performance types of the High-Intermediate group than Low-Intermediate group. Hence, the more the proficiency level of a group, the more the homogeneity between the performance types of that group under different conditions (i.e. GJ Tasks).

The level of significance (i.e. alpha level) set as the criterion and standard for the rejection and acceptance of hypotheses at the onset of this study was 5% (.05). Hence, the results of the two-way repeated measures ANOVA tests and the univariate *F*-tests would be analyzed at the .05 level of significance to indicate that a 95% probability existed that the difference between the different performance types of the groups had been due to the experimental treatment rather than to sampling error.

In order to determine the effect of each two-level independent variable (i.e. group [proficiency level], treatment [GJ tasks] and the interaction of these two) on the participant's dependent variable (i.e. female EFL participants' performance on Written Stimulus and Oral Stimulus GJ tasks) and see if these effects are significant, we used the multivariate tests, as shown in Tables 2, 3, and 4, to focus on the independent variables and their interactions. We used the multivariate analysis of variance to assess the effects of treatment, group, and group-treatment interaction in terms of repeated observations of the same participants in two groups of different proficiency levels.

The grouping variable of group, as shown in Table 2, was the multivariate test of the two-level independent variable of proficiency level that distinguished the two groups. This effect was not significant, F(1, 58) = 4.08 > 1.295, indicating that the two levels of proficiency (i.e. High- and Low-Intermediate proficiency levels) did not differ significantly from each other on the weighted aggregate of the two independent variables (i.e. proficiency levels) in terms of having different effects on the High- and Low-Intermediate female EFL participants' performance.

The grouping variable of treatment, as shown in Table 3, was the multivariate test of within-subjects factor (i.e. the two-level independent variable of presenting two GJ tasks to the two groups over two weeks in a counterbalanced way). This effect was not significant either, F(1, 58) = 4.08 > 2.237, indicating that the two levels of the within-subjects factor (i.e. Written Stimulus and Oral Stimulus GJ tasks) did not differ significantly from each other on their weighted aggregate in causing significant differences in the female EFL participants' performance on GJ tasks in either of the groups.

Tests of significance of the multivariate effects of group-treatment interaction, as shown in Table 4, described the multivariate effects of the repeated measures factor and the interaction of this factor with group. In this case, neither of the effects was significant, F(1, 58) = 4.08 > .790, suggesting there did not exist multivariate effects for both treatment and the group-treatment interaction in either of the proficiency groups.

The two-way repeated measures ANOVA, which compares participants' performance on more than one task, is a robust parametric test. One of the assumptions underlying parametric tests is that data points are independent, but this assumption does not hold for a repeated measures design because, in repeated measures designs, data for different conditions have come from the same people, hence the relatedness of data from different experimental conditions. This issue gives rise to an additional assumption according to which the relationship between pairs of

groups is equal, that is the variances of the differences between levels of the repeated measures factors are equal in a within-subjects design. This assumption is called *sphericity* or *circularity*.

This assumption resembles the homogeneity of variance assumption in between-subjects ANOVA. Sphericity assumption should not be violated because its violation causes loss of power in terms of an increased probability of a Type II error and not being able to compare the F statistic to the normal tables of F. In addition, the software used cannot calculate a significance value. However, the SPSS software package includes a procedure called Mauchly's test that determines if the assumption of sphericity has been violated. First, we ran the Mauchly's test to determine whether the condition of sphericity had been met; that is, it tested the hypothesis that the variances of the differences between conditions were equal. Table 5 shows the findings of running Mauchly's test for the posed hypothesis in this study.

If Mauchly's test statistic is significant (i.e. a probability value less than .05), it will indicate the existence of significant differences between the variance of differences, and, as a result, the violation of sphericity ensues; in this case, we cannot trust the *F*-ratios. If, on the other hand, Mauchly's test statistic is nonsignificant (i.e. a probability value more than .05), it will indicate lack of significant differences between the variance of differences, and, as a result, the observance of sphericity. Since the significance values, as shown in Table 5, were <.05, we concluded that there was a significant difference between the two GJ tasks. In addition, there was a significant difference between the two proficiency levels, but Mauchly's test did not tell us which GJ task or proficiency level differed from the other. To decide which of the three corrections to use (i.e. whether to use Green-House Geisser, Huynh-Feldt or Lower-Bound), we checked the Epsilon value in Table 5.

Epsilon is a descriptive statistic that indicates the degree to which sphericity has been violated. In cases where the epsilon value is more than 0.75 (i.e. >0.75), such as the case in Table 5, we must use the Huynh-Feldt correction, whereas in cases where the epsilon value is less than 0.75 (i.e. <0.75), we must use the Greenhouse-Geisser correction. Since in this study the epsilon values from Mauchly's test values are 1.000 (i.e. Table 5) and all of the three values are E > 0.75, we used the less conservative Huynh-Feldt corrected values to adjust both within-subject and between-subject degrees of Freedom as shown in Tables 6, 7, and 8. Having used the corrections in Table 5 (i.e. under the heading of epsilon) to adjust the degrees of freedom for the averaged tests of significance, we, as a result, used the corrected tests, as shown in Tables 6, 7, and 8 (i.e. tests of within-subjects effects).

Applying the Huynh-Feldt correction epsilon value, as shown in Table 6, we obtained no significant differences between the two proficiency groups in terms of their performance on the two GJ tasks. That is, the corrected *F*-ratio of the Huynh-Feldt correction procedure showed the significance value of (.265) according to which High- and Low-Intermediate proficiency groups showed no significant differences in their performance in the two GJ tasks over two weeks while rendering WH-questions. Hence, we were not capable of obtaining an understanding as to which proficiency group showed significant effects in its performance in the two GJ tasks while rendering WH-questions.

Applying the Huynh-Feldt correction epsilon value, as shown in Table 7, we obtained no significant differences in the effects of different treatments of GJ tasks (i.e. Written Stimulus and Oral Stimulus GJ tasks) on the performance of female EFL participants while rendering WH-questions. That is, the corrected *F*-ratio of the Huynh-Feldt correction procedure showed the significance value of (.146) according to which no significant differences existed between the two GJ tasks in terms of their effects on the female EFL participants' performance while rendering WH-questions in Written Stimulus and Oral Stimulus GJ tasks. Hence, we were not capable of obtaining an understanding as to which of the two GJ tasks was of the significant effects on the performance of female EFL participants while rendering WH-questions. The exact significance levels in Table 7 are the results of comparing the *F*-ratios against a critical value for 1 and 29 degrees of freedom.

Since the statistical analysis used is the two-way repeated measure ANOVA, where one factor is varied between subjects (i.e. proficiency level) and the other within subjects (i.e. Written Stimulus and Oral Stimulus GJ tasks), we analyzed the interaction between a within-subject factor (i.e. the two treatments of Written Stimulus and Oral Stimulus GJ tasks) and a between-subject factor (i.e. group in terms of the two proficiency levels [High-and Low- Intermediate levels]) for the two experimental groups, but since there was violation of sphericity, we, using the Table 8, were not capable of looking at the interaction without a correction, hence the application of the Huynh-Feldt correction epsilon value, as shown in Table 8.

Applying the Huynh-Feldt correction epsilon value, as shown in Table 8, we obtained no significant interaction between the within-subjects factor (i.e. the two treatments of Written Stimulus and Oral Stimulus GJ tasks) and the between-subject factor (i.e. group in terms of the two proficiency levels) for the two experimental groups.

That is, the corrected *F*-ratio of the Huynh-Feldt correction procedure showed the significance value of (.381) according to which no significant interaction existed between the within-subjects factor (i.e. the two treatments of Written Stimulus and Oral Stimulus GJ tasks) and the between-subject factor (i.e. group in terms of the two proficiency levels) for the two experimental groups. The nonsignificant interaction indicated that the groups were not changing in terms of performance over different treatment times. Hence, we were not capable of obtaining an understanding as to whether the interaction of group and treatment was of any significant effects on the ability of female EFL participants while rendering WH-questions in either of the proficiency groups. The exact significance levels in Table 8 are the result of comparing the *F*-ratios against a critical value for 1 and 29 degrees of freedom.

Hence, Mauchly's test indicated that the assumption of sphericity had been violated (chi-square = .000, p <.05), therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (epsilon = 1.000). The results show that the performance scores of the two proficiency groups (i.e. High-Intermediate and Low-Intermediate proficiency groups) in the two GJ tasks (Written Stimulus and Oral Stimulus Grammaticality Judgement tasks) did not differ significantly, and, as a result, the posed hypothesis was not rejected, Group F(1.000, 29.000)= 1.295, P>.05; Treatment F(1.000, 29.000)= 2.237, p>.05; Group\*Treatment F(1000, 29.000)= .790, P>.05. Post hoc tests revealed that the performance scores of the two proficiency groups in Written Stimulus and Oral Stimulus Grammaticality Judgement tasks did not show significant differences. That is, neither group's performance was significantly better than the other group in the two GJ tasks (i.e. Written Stimulus and Oral Stimulus GJ tasks) (p>.05).

#### 5. Discussion

The results obtained from the statistical analysis of the performance data of Persian female EFL participants support Schachter's (1988) assertion that in cases where L2 learners' L1 does not resemble L2 in WH-movement, such as the case of Persian participants in this study, L2 learners' observance of subjacency could only be due to UG and not L2 learners' L1. Participants of this study did show signs of L1 (i.e. Persian) interference while rendering WH-Questions; that is, in cases where L1 and L2 showed similarities such as the case of some of the filler items used, participants violated subjacency constraints in English by showing signs of L1 interference, even most of the High-Intermediate participants, whereas in cases where no such similarities existed such as the case of the main WH-questions used in the GJ tasks, participants observed the subjacency constraints in English properly, and recognized the grammatical and ungrammatical WH-questions in English based on such bounding nodes as IPs, and NPs (DPs), hence the positive role of UG in this regard and also the confirmation of the stance adopted by Bley-Vroman, Felix & Loup (1988). The results of this study strongly oppose the findings of such researchers as Clahsen and Muysken (1986).

The findings of this study stand in contrast to the stance held by such researchers as White, Travis & Maclachlan, (1992) because this study has proved that proficiency level does not cause any significant differences between the performance of the High- and Low-Intermediate groups. In addition, this study supports the position of such researchers as Johnson & Newport (1991) because L2 grammatical proficiency is indeed a unified solid whole; otherwise, modality would have caused significant differences in the performance of the female EFL participants in the two GJ tasks used.

The findings of this study oppose the results of the study conducted by White (1988) because, unlike the case of White's study where the group of High-Intermediate and Advanced participants outperformed the group of Low-Intermediate and Intermediate participants significantly, the High-Intermediate group by no means significantly outperformed the Low-Intermediate group, hence the insignificant effect of proficiency level on the performance of female EFL participants.

#### 6. Conclusion

The rationale behind the obtained linguistic theories of the hypothesis in this study is to explain an aspect of the process of SLA in a measurable way (i.e. the degree to which written and oral modes of input presentation affect the EFL learners' ability to observe subjacency). We could not establish an interaction between SLA theory-based *ex*planations (e.g. SLA ss a cognitive process) and the data gathered by administering the GJ tasks (Written Stimulus and Oral Stimulus GJ tasks) on such constructs because no significant differences were seen between the written and oral modes of receiving input, which are cognitive-processing issues, in the performance of High- and Low-Intermediate female EFL participants. In fact, the performance of Low-Intermediate group bore a close resemblance to the performance of High-Intermediate group in both Written Stimulus and Oral Stimulus GJ tasks.

Hence, the obtained findings from the hypothesis point to the conclusion that the generative view of language can be adhered to (i.e. language is a symbolic system which is independent of cognition) because the female EFL participants did not show significant differences in their performance in the Written Stimulus and Oral Stimulus GJ tasks in this study. In other words, the written-oral dichotomy (i.e. modality) of the GJ tasks failed to trigger participants' cognitive-processing capabilities despite the fact that participants were in two proficiency levels (i.e. the reason is that participants primarily focused on the symbolic aspects of WH-questions and not the semantic-cognitive-processing relations between written-oral stimuli and WH-questions); a fact that leads us to conclude that female EFL participants are primarily under the influence of the symbolic aspects of their L2. Hence, language is indeed highly complex, and such learning strategies as inductive learning strategies, deductive strategies, or rather general problem-solving strategies can by no means lead to language acquisition.

The overall conclusion drawn from the results and discussion of the hypothesis is that the examination of the nature of female EFL participants' grammatical proficiency from the written-oral standpoint does not result in viewing modality as a distinct component part of female EFL participants' grammatical proficiency. That is, the adoption of a separate modality approach towards the investigation of the nature of L2 grammatical proficiency does not trigger distinct levels of comprehension of English WH-questions on EFL female participants' part. Hence, this issue cannot lead to a major shift in SLA researchers' long-held view towards EFL learners' grammatical proficiency because the hypothesis attests to the long-held view of grammatical proficiency as a unified solid whole and not a unit made of numerous component parts with each part being triggered by a distinct task type; otherwise, EFL female participants would have shown significant differences in their performance in the Written Stimulus and Oral Stimulus GJ tasks.

Accordingly, such independent variables as the written mode of input presentation, such as instruction via using written contents, or the oral mode of input presentation do not have significantly different effects of the EFL female participants' comprehension and learning of different English grammatical concepts and functions such as subjacency; that is, High- and Low-Intermediate EFL female participants perform similarly across written and oral independent variables because they use the same cognitive-linguistic processing and psychological bases for processing such variables. Written and oral modes of receiving linguistic input do not have different effects on female EFL participants' language processing because similar cognitive-processing mechanisms are triggered by such modes of linguistic input. Hence, the written-oral dichotomy does not serve as a separate subcomponent of L2 grammatical proficiency because the written and oral forms of linguistic input have similar effects on female EFL learners' performance in the Written Stimulus and Oral Stimulus GJ tasks. In fact, teachers can adopt a single teaching method for the practical use of written and oral teaching stimuli while teaching subjacency operations to EFL learners of different proficiency levels.

The findings of this study can also improve and modify the theoretical foundations of *testing* towards the use of GJ tasks because GJ tasks have proved not to be of that much practical use in significantly differentiating between the written and oral knowledge modes of L2 learners; hence, researchers should not use this type of test if they intend to measure the differences between the written and oral performances of their L2 learners. Furthermore, different schools, language institutes and universities can revise their use of GJ tasks and make use of other tasks. In fact, such academic organizations should reevaluate the success of GJ tasks in the extent to which such tasks can properly elicit the use of a target structure, such as WH-questions. Moreover, GJ tasks have proved not to be efficient whatsoever.

#### References

Collins, P. (2006). Clause types. In: B. Aarts & A. McMahon (Eds.), *The handbook of English Linguistics* (pp: 180-197). Blackwell Publishing.

Cook, V. (1993). Linguistics and Second Language Acquisition. The Macmillan Press Ltd.

Dekeyser, R. (2003). Implicit and explicit learning. In: C.J. Doughty & M.H. Long (Eds.), *The handbook of Second Language Acquisition* (pp: 313-348). Malden: Blackwell Publishing Ltd.

Ellis, R. (2003). Task-based language learning and teaching. Oxford University Press.

Fromkin, V., Rodman, R., & Hyams, N. (2003). An introduction to language (7th ed.). MA: Thomson Heinle.

George, A. (1989). How not to become confused about linguistics. In: A. George (Ed.), *Reflections on Chomsky* (pp: 90-110). Oxford: Blackwell Publishers.

Haegeman, L. (1991). Introduction to Government and Binding Theory. Cambridge: Cambridge University Press.

Hawkins, R. (2001). Second language syntax: A generative introduction. Oxford: Blackwell Publishing.

Huddleston, R., & Pullum, G. K. (2006). Coordination and Subordination. In: B. Aarts & A. McMahon (Eds.), *The handbook of English Linguistics* (pp: 198-219). Blackwell Publishing.

Isac, D.,& Reiss, C. (2008). *I-Language: An introduction to Linguistics as Cognitive Science*. Oxford University Press.

Lotfi, A.R. (2003). Intonation contours, yes/no questions, and minimalist syntax: A unitarianist perspective. *Cognitive Sciences*, 70-83. Esfahan: Azad University at Esfahan.

Matthews, P.H. (2007). Oxford concise dictionary of linguistics (2<sup>nd</sup> ed.). Oxford: Oxford University Press.

Radford, A. (1997). Syntax. A minimalist introduction. Essex: Cambridge University Press.

Radford, A. (2004). *Minimalist syntax: Explaining the structure of English*. Cambridge: Cambridge University Press.

Radford, A., Atkinson, M., Britain, D., Clahsen, H., & Spencer, A. (2009). *Linguistics: An introduction* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.

Ramchand, G. C. (2008). Verb meaning and the lexicon: A first-phase syntax. Cambridge University Press.

Scalise, S., Magni, E., & Bisetto, A. (Eds.) (2009). Universals of language today: Studies in natural language and linguistic theory. Springer.

Skehan, P. (2008). Interlanguage and language transfer. In: B. Spolsky & F.M. Hult (Eds.), *The handbook of educational linguistics* (pp: 411-423). MA: Blackwell Publishing Ltd.

White, L. (1988). Island effects in Second Language Acquisition. In: S. Flynn & W. O'Neil (Eds.), *Linguistic theory in Second Language Acquisition* (pp: 144-172). Dordrecht: Kluwer Academic Publishers.

White, L. (1989). Universal Grammar and Second Language Acquisition. Philadelphia: John Benjamins Publishing Company.

White, L. (1990). *Second Language Acquisition and Universal Grammar*. SSLA (PP: 121-133). USA: Cambridge University Press.

White, L. (1996). Universal Grammar and Second Language Acquisition: Current trends and new directions. In: n.d. (eds.), *Handbook of Second Language Acquisition* (pp: 85-120). Academic Press, Inc.

Independent			
Variables	N	Mean	Standard Deviation
LG_ORL_WRI	30	16.00	3.991
LG_EXP_WRI	30	17.03	3.978
HG_ORL_WRI	30	17.37	3.378
HG_EXP_WRI	30	17.60	2.811

Table 1. Sample Descriptive Statistics

Table 2. Multivariate Tests

	Effect	Value	F	Hypothesis df	Sig
Group	Pillai's Trace	.043	1.295	1.000	.265
	Wilks' Lambda	.957	1.295	1.000	.265
	Hotelling's Trace	.045	1.295	1.000	.265
	Roy's Largest Root	.045	1.295	1.000	.265

 $\dagger p < .05$ , one-tailed

ISSN 1923-869X E-ISSN 1923-8703

#### Table 3. Multivariate Tests

	Effect	Value	F	Hypothesis df	Sig
Treatment	Pillai's Trace	.072	2.237	1.000	.146
	Wilks' Lambda	.928	2.237	1.000	.146
	Hotelling's Trace	.077	2.237	1.000	.146
	Roy's Largest Root	.077	2.237	1.000	.146

 $\dagger p < .05$ , one-tailed

Table 4. Multivariate Tests

	Effect	Value	F	Hypothesis df	Sig
Group* Treatment	Pillai's Trace	.027	.790	1.000	.381
	Wilks' Lambda	.973	.790	1.000	.381
	Hotelling's Trace	.027	.790	1.000	.381
	Roy's Largest Root	.027	.790	1.000	.381

 $\dagger p < .05$ , one-tailed

#### Table 5. Mauchly's Test of Sphericity

				Epsilon			ilon
Within Subjects		Approx. Chi		Green-House		Huynh-	Lower-
Effect	Mauchly's W	Square df	Sig.	Geiss	er Feldt	Bound	
Group	1.000	.000	0		1.000	1.000	1.000
Treatment	1.000	.000	0		1.000	1.000	1.000
Group*Treat	ment 1.000	.000	0		1.000	1.000	1.000

Table 6. Tests of Within-Subjects Effects

	Type III Sum					
Sour	ce	of Squares	df	Mean Square	F	Sig.
Group	Sphericity Assumed	28.033	1	28.033	1.295	.265
	Huynh-Feldt	28.033	1.000	28.033	1.295	.265
	Lower-Bound	28.033	1.000	28.033	1.295	.265

## Table 7. Tests of Within-Subjects Effects

Ту	Type III Sum									
Source		of Squares	df	Mean Square	F	Sig.				
Treatment	Sphericity Assumed	12.033	1	12.033	2.237	.146				
	Greenhouse-Geisser	12.033	1.000	12.033	2.237	.146				
	Huynh-Feldt	12.033	1.000	12.033	2.237	.146				
	Lower-Bound	12.033	1.000	12.033	2.237	.146				

## Table 8. Tests of Within-Subjects Effects

Type III Su	ım					
Source		of Squares	df	Mean Square	F	Sig.
Group*	Sphericity Assumed	4.800	1	4.800	.790	.381
Treatment	Greenhouse-Geisser	4.800	1.000	4.800	.790	.381
	Huynh-Feldt	4.800	1.000	4.800	.790	.381
	Lower-Bound	4.800	1.000	4.800	.790	.381