The Iranian EFL Journal

April 2011

Volume 7

Issue 2

Chief Editors:

Dr. Paul Robertson

Dr. Rajabali Askarzadeh Torghabeh

The Iranian EFL Journal   April 2011   Volume 7 Issue 2

Iranian EFL Journal

Iranian EFL Journal

Iranian EFL Journal

Table of Contents

Foreword

Welcome to the second edition of the year 2011. The Iranian EFL Journal is a bi-monthly journal from 2011. The journal has had strong growth over the last few years with a monthly readership now exceeding 2000 readers. For a journal examining the topic of EFL/ESL, Literature and Translation studies, the growth and readership has been pleasing. Statistically, readers are coming from almost 80 countries. In the second issue of volume 7 we present 15 articles for your reading. In the first article, the author Reza Pishghadam in a groundbreaking article has introduced applied ELT as a new approach to second/foreign language studies. He believes that ELT is not a part of applied linguistics any more, requiring a fresh look at its principles. In the second article, Keivan Zahedi and Mohammad Ali Shams have explored the IELTS candidates' use of formulaic sequences in the writing tasks.

In the next article, Masoud Sharififar and Mahboubeh Akbarzadeh have investigated demotivation of English learners to identify common demotivating factors among university students. In the fourth article, Mehdi Mardani and Ahmad Moinzadeh have investigated the effect of vocabulary learning strategy on retention and recall of idioms in Iranian EFL context. In the next article, Aram R. Sadeghi and Ali Asadi have examined the applicability of critical discourse analytical tools in an EFL classroom. In the next article, Sima Khezrlou, Ali Akbar Khomeijani Farahani and Fateme Layeghi have explored the way adult Persian-speaking foreign language learners of English attempted to resolve ambiguities of relative clause type. In the next article, Seyyed Mohammad Reza Amirian and Reza Bagheri Nevisi investigated the status of Overt Pronoun Constraint (OPC) in Persian. In the next article, Abbas Pourhosein Gilakjani and Seyedeh Masoumeh Ahmadi have examined the effect of intrinsic and extrinsic motivation on Iranian EFL learners' language learning.

In the next article, Sanaz Ghobadi Mohebi and Ebrahim Khodadady have investigated the beliefs, students usually held about language learning, based on the beliefs about

language learning inventory questionnaire and comparing them with teachers. Naser Atasheneh and Ahmad Izadi have explored the similarities and differences between Iranian EFL learners' use of English and Persian refusals, using role play scenarios. Moreover, Ali Rahimi and Nabi. A Ebrahimi have introduced the field of learning environment research to ELT practitioners. In the next article, Reza Keshavarz and Rasool Roozegar have introduced a method for assessing English language ability of the learners. Mahboobeh Abhaji Ezabadi determined the effect of the construction of mental images on FL learners' ability to recall narrative passages.

In the next article, Roya Khoii and Nazli Shamsi have investigated the effect of the test method on trait by comparing the construct validities of two different formats of error-identification grammar. In the last article of the issue, Mansoor Fahim and Shahla Azarnioushi have attempted to examine whether there is any relationship between the critical thinking ability of language learners and their performances using rule driven/ discovery learning approaches to teaching grammar.

We hope you enjoy this edition and look forward to your readership.

Title

Introducing *Applied ELT* as a New Approach in
Second / Foreign Language Studies

Author

Reza Pishghadam (Ph.D.)
Ferdowsi University of Mashhad, Iran

Bio Data

Reza Pishghadam holds a Ph.D. in TEFL in Ferdowsi University of Mashhad, Iran. His research interests are: Psychology and Sociology of language education.

Abstract

Following the huge amounts of studies done on ELT, and due to its rich and expanding literature which is gaining more momentum, this study intends to claim that ELT has found a firm theoretical foundation, ready to be applied to other disciplines. In fact, this study attempts to introduce applied ELT as a new counterpart to theoretical ELT, holding the idea that ELT theorizers and practitioners should draw their attention to this new area in the field. Applied ELT opens a new and promising horizon for researchers to channel their studies into this new aspect of ELT. Moreover, this paper considers English language learning classes to be sites of dealing with issues related to life qualities. In the end, a new map of doing research in the field has been drawn to guide the researchers to apply ELT to different domains of knowledge.

Keywords: Theoretical ELT, Applied ELT, Linguistics, English language, Life syllabus.

Introduction

Granted the fact that different disciplines generally have two aspects of theoretical and applied, it seems that ELT theorizers and practitioners underscoring the theoretical aspect of ELT, have disregarded its applied part. Digging into the history of ELT, one can testify the fact that it has emerged out of the findings of theoretical linguistics (Berns & Matsuda, 2006). When the term applied linguistics first came into existence in the 1950s, it was virtually synonymous with language teaching (Strevens, 1992), and it was later that the term found more comprehensive meaning. That is why, nowadays ELT is studied as a branch of applied linguistics.

In my view, ELT has grown in maturity over years, establishing an independent identity for itself. It does not play second fiddle to applied linguistics any more. A cursory look at the literature of ELT reveals the fact that hundreds of journals and thousands of books are being published all over the world dealing with different aspects of language teaching and learning. All these materials aim at deepening our understanding and opening our eyes to the realities of language teaching and learning.

In the course of time, ELT in an interdisciplinary endeavor, has freed itself from the dominance of theoretical linguistics, embracing the findings of other fields of study including, psychology, sociology, neurology, computer, etc. to enrich itself. All of the above-mentioned fields have made ELT develop exponentially, smoothing the way for teachers and learners to teach and learn a second language more effectively. Thus, due to its rich literature and interdisciplinary nature, it is fair to say that ELT has now gained scientific acceptability, forming its theoretical background, ready to be applied to other fields.

Therefore, this study is intended to introduce and emphasize on the applied aspect of ELT, claiming that ELT, as an independent field of study, has the potentiality to be applied to other domains of knowledge. To espouse the claim, I first provide the readers with an overview of theoretical ELT, and then I shed more light on the concept of applied ELT, introducing life syllabus.

Theoretical ELT

The spread of English and the related expansion of its utility generated keen interests in how to improve English language learning and teaching. Since language learning deals with issues regarding language and learning, it was quite common that ELT practitioners had to refer to the findings of linguistics and its branches like

psycholinguistics and sociolinguistics to advance and broaden their understanding of English language learning and teaching. That is why, the first generation of ELT practitioners, such as the proponents of Audiolingualism and Situational language teaching have drawn heavily on the findings of linguistics to form new methods and concepts in ELT. In fact, ELT was considered to be a part of applied linguistics, which had to wait for the prescriptions and proscriptions of it to improve its status. English language teachers were considered to be merely consumers of the findings of other disciplines, especially linguistics (Schmitt, 2002).

Three main explanations can be provided to show why ELT at the outset of its genesis leaned towards linguistics. First, as Mackey (1965) has pointed out language teachers to dissociate themselves from the subjective language teachers tried to associate themselves with linguistics which was objective and scientific. Since linguistics deals with the most detailed descriptions of language, linguistic approach was considered to be responsible for teaching methodology (Corder, 1973).

Second, the revolutionary work of Chomsky (1957) made linguistics gain more popularity among English language teachers. Chomsky made ELT more linguistic-based. Chomsky, as a linguist, in 1959 in his well-known article criticized Skinner (1957), who was a psychologist. In the article, Chomsky by questioning the major tenets of behaviorism which are repetition, memorization, trial and error, reinforcement, and conditioning, emphasized on the role of mind in language learning. This article made linguistic findings find their own way more in language learning issues. Moreover, Chomsky (1965), by putting forward the idea of modularity warranted the heavy existence of linguistics in ELT. Modularity hypothesis means that language is independent of other sections of the mind. Thus, to understand more about language, there was no need to study other sections of the brain. This micro-linguistic approach made ELT theorizers form their theories based on this outlook for more than two decades.

Third, according to the major principles of reductionism, any discipline to be accepted seriously as a science, it had to conform to the standards set by major disciplines (Pishghadam & Mirzaee, 2008). Thus ELT to be considered as a discipline was to utilize the results obtained in linguistics or psychology, to be considered as a scientific field of study. In fact, we can claim that ELT for some time was imprisoned by other fields of study, especially linguistics to be like them. Of course, this

imprisonment did not take much time until the first signs of resistance were observed and the whisper of freedom was heard.

The situation changed when Widdowson (1979) by making a distinction between *Applied linguistics* and *Linguistics applied* drew English language teachers` attention to the fact that ELT was not just to consume the results of other disciplines, especially linguistics, but it must develop its own identity. According to him, teachers` responsibility is not just to follow blindly the orders made by linguistics; they are expected to reflect on their teaching, forming new theories which are educationally relevant to language. In fact, Widdowson (1979) by questioning the common belief of the time that linguistic sciences can "provide good descriptions" for teaching of languages (Halliday, McIntoch, & Strevens, 1964, p. 167), moved linguistics from the mainstream to the margin of ELT.

Later, the idea of method which was the application of other disciplines in ELT was put into question by Prabhu (1990), Allwright (1992), and Kumarvadivelu (1994). They all tried to encourage teachers to be more empowered, reflecting on their own teaching, and move towards autonomy in language teaching. These authors called for the death of methods, holding that teachers are expected to be more active in theorizing ELT. Gradually, the ideas of action research and reflective teacher came to the fore of language teaching, moving the field to a postmethods era in which teacher is held more accountable for his/her own teaching (Williams & Burden, 1997).

During this postmethods era, interdisciplinary and transdisciplinary studies were given more attention. ELT researchers tried to use other fields of study more directly to expand the ELT domain. For instance, the psycholinguistic books and studies which have linguistic basis were replaced with studies concerning *the psychology of language teachers* (Williams & Burden, 1997), *psychology of language learner* (Dornyei, 2005), and *the psychology of second language acquisition* (Dornyei, 2009). These studies made ELT move in different ways and directions to get fatter in theory and practice. In fact, in this respect and over the years, ELT has employed a cornucopia of ideas from other fields to broaden its perspectives. These ideas have given rise to lots of ideas and ideals in ELT, leading to lots of theories of how to learn and teach English.

In sum, ELT has been under the dominance of linguistics and its branches over years, and English language teachers were supposed to conform to the standards

imposed by other disciplines. Later, this transmitter teacher gave way to a transformative teacher, who could form theories and concepts more related to ELT. English language teachers were not merely to import ideas from other fields of study to the way of teaching and learning, but they were to employ the results of other disciplines to localize and domesticate the findings so that they could be more applicable in language learning and teaching. I think this process of scientification has moved on until ELT has formed a very strong basis of ELT theories.

Applied ELT

As mentioned, ELT has already formed its theoretical foundation, ready to be applied to other fields of study. I believe that over the last decades, ELT theorizers and practitioners have moved their orientation from consumer to autonomous. Now it`s time for them to play a producer role. What is meant exactly by applied ELT is the application of ELT in other fields. Generally, we employ other disciplines'` findings including, psychology, sociology, neurology, linguistics, physics, etc. to enrich language teaching and learning theories. However, applied ELT is to reverse the direction, taking a more contributory role.

The aim is to make other disciplines take a new and fresh look at ELT, trying to employ the findings in their studies. Few studies have already been conducted to take the issue into account. For example, in a study, Pishghadam (2008) has shown that literary discussion in a foreign language learning class can enhance the critical thinking abilities of the language learners. Critical thinking is a concept which is discussed generally in psychology; however, I believe that this issue can be noted in a language learning class. In the same vein, Khaza`ifar, Pishghadam, and Ziai (in press) have indicated that English language reading materials can be designed in a way that critical abilities are developed. Hosseini, Pishghadam, and Navari (2010) have also revealed how a language learning class can increase emotional intelligence competencies. They have shown that language learning classes have the ability to help individuals overcome their anxiety, manage their stress, and foster interpersonal competencies. In another study, Pishghadam and Saboori (2011) have shown that English language teachers in Iran have positive attitudes towards the American culture, and try to act like native speakers. The authors have suggested that these teachers can alienate students from their own home culture. This claim has been in line with (Pishghadam & Navari, 2009), who claim that contrary to the Bakhtinian

beliefs when two cultures come together, the two cultures are not necessarily enriched. If English language teachers are not well-trained in dealing with cultural issues, cultural derichment is inevitable. These studies imply that English language teachers can enrich or derich learners` home culture. In fact, due to the comparative nature of language classes in which two cultures come into a close contact, English language teachers play a pivotal role in fostering national identity of the learners. Therefore, English language learning classes have the potential to be the sites for developing the cultural and national identity of the learners.

As all of these studies indicate, English language learning classes can provide us with new ways of exploring ELT. Seemingly, applied ELT owes much to the unique nature of language learning classes, especially in EFL contexts. Since learners do not have direct contact with English after class, English learning in EFL contexts is more difficult than ESL contexts. Thus the learners have to spend more time on learning English, taking it more seriously. Unlike other subjects which are studied at school, an English language learning class has a very different atmosphere in which lots of human abilities can be nurtured and developed in addition to language learning. Due to the unique nature of English language learning class, I believe that it can offer a great deal of opportunities, based on which we can conduct lots of studies.

Unique features of ELT classes

English language learning classes have some unique features, which are rarely found in other subjects of studies at school or university. Some of these features have been presented below.

In these classes a number of topics are discussed; these topics include all things about life, music, culture, politics, society, science etc. We normally do not have any class at school or anywhere to have discussions of these types. In addition to informative role of the topics, they can provide the learners with more food for thought about life. Based on these discussions, learners can reflect more on their life, trying to change the status quo.

Another important feature of English language learning classes which is not much visible in other classes at school is the overuse of pair works and group works. This way of interaction and exchange of information can help the dynamicity of the class, enhancing different communicative abilities in students. The situation makes the class be more dialogic and free from more monologic and narrative classes.

Moreover, in these classes the learners face another culture more directly. By reading another culture, they get more acquainted with their own home culture. This situation can help the schizophrenic nature of the class, in which two big personalities (cultures) struggle for more dominance. Home culture and the foreign culture can have dialogue in class, helping the learners to create or maintain identities.

Another explanation for the unique nature of the classes is that students get more familiar with the structures and words of another language. According to Vygotsky (1978), grammar is a tool for enhancing the higher order abilities of human beings. Grammar can make individuals think more logically. Besides, words of two languages are different which may convey different types of understanding, shedding more light on our vague grasp of things. Thus this linguistic contact of two languages makes the setting more unique.

Another important feature of English language learning classes is that students might take more freedom to express themselves. By speaking in another language, one can project his/her own true identity, especially in non-democratic settings. When one speaks in another language, he/she feels free to say something they cannot express in their mother tongue owing to social or political reasons.

Since English is the lingua franca and the language of science, it is a necessity for all people to learn it. Each individual who is willing to enjoy life, travel other countries, do more business, and to push back the frontiers of knowledge, and to have access to all reliable scientific sources including the Internet, or famous universities must learn English. This literacy nature of English makes all people take it seriously, investing a tremendous amount of time and money to learn it.

Last but not least, is the fun nature of the class. Watching movies, listening to different songs, discussing different topics, using computers, the Internet, mobiles, and different kinds of tasks make the class be a fun. Fun and learning come together to create an embracing atmosphere in class.

In a nutshell, the following are some unique features of English language learning classes:

- Discussing a large number of social, scientific, and political topics,

- Holding pair work and group work in class,

- Comparing two cultures,

- Getting acquainted with the words and grammar of another language,

- Speaking in another language in which one can show their own real self,

- Taking language learning very seriously,

- Having a funny friendly atmosphere for learning.

Life syllabus

Another important aspect of applied ELT is that it goes beyond the typical linguistic syllabus which is normally used in ELT, presenting a kind of syllabus which considers life issues as its first priority. In this new syllabus, language learning is more purposeful and educational, revolving around more important issues of life. To better illustrate this new type of syllabus, I first go over the modern and postmodern eras of ELT.

Pishghadam and Mirzaee (2008) have pointed out that ELT in the western culture lives in the postmodern era. They believe that during the modern era of teaching English, there was a search for finding the best method, native speaker was the ideal model for teaching, and having native-like accent and structure were appreciated. However, during the postmodern era of ELT, the idea of method was considered to be colonial, and it was put into serious question. Native speaker was not considered to be the proper yardstick to follow any more, and violations of British and American pronunciations and structures were allowed. In fact, the idea of *World English* (modernist view) was replaced with the notion of *World Englishes* (postmodernist view).

The coinage and promotion of the term *World Englishes* is mainly associated with Kachru (1982). The underlying philosophy of Kachruvian approach argues for the "importance of inclusivity and pluricentricity in approaches to linguistics of new varieties of English" and deals with some other related topics including creative writing, contact linguistics, critical linguistics, pedagogy, and the sociology of language (Bolton, 2004, p. 367). In addition, in an attempt to empower new Englishes, this theory calls the labels *native speaker* and *native and standard English* into serious question, and denies any special status for them.

If we look closely into different eras of ELT, we witness that we have passed over a very strict adherence to standards in language learning to a more lenient way of learning a language. Once using mother tongue in class was banned, learners had to mimic British and American English as closely as possible; the language was not authentic, and learning English was very laborious and burdensome. However, later use of mother tongue was allowed in class, British and American English was not

sacred cows any more, and the real-life language was employed. Materials were supposed to be more informative in a way that learners learn via language. Learners were supposed to learn something new about life via learning English. In fact, ELT changed its focus from *only-language* classes to *language-and-life* classes.

Moreover, when we look through the methods proposed over a long time, we see that the riddle of language acquisition has never been demystified. Although lots of headway has been made to facilitate second language leaning, we are still facing lots of obstacles while learning another language. It is still a burden for people to learn a second language. Thus, if it is supposed to learn English with difficulty, why should not we focus on learning English along with other issues of life to get more satisfaction of ELT classes?

In applied ELT, I believe that we should transcend discussions over language and linguistics, entering into issues regarding life qualities. Now, it`s time for ELT to enter into not *language-and-life* but to *life-and-language* classes. A language learning class, in the first place, must be a class in which life issues are noted and taken into consideration. Language should be epiphenomenal to life. Due to the unique nature of English language learning classes which were already discussed, the classes are supposed to first, enhance critical abilities, creativity, social intelligence, emotional intelligence, etc., and then teach a language.

What I intend to convey is not the idea that language should not be taught, but I believe that teachers must design their linguistic syllabus around the life syllabus. In the life syllabus, we determine which aspect of life is going to be targeted, e.g. creativity, then we design our linguistic syllabus in a way to achieve this goal. It means that language must be at the service of enhancing life qualities. Therefore, ELT researchers are to move beyond the content or form issues of language, exploring the ways of this new challenge.

Concluding remarks

The purpose of this article was to introduce applied ELT as a new paradigm in the field through which to map the future of studies in ELT. As we have already stated, ELT has gone through different stages to become an independent and scientific field of study. The process of scientification has started from consuming the notions formed by other disciplines to becoming more autonomous by localizing the imported

theories which are more related to ELT. Later, interdisciplinary studies helped ELT to be enriched in theory, establishing a unique identity for itself. In fact, ELT has grown to such an extent that it is now full-fledged separate discipline, which must be treated as such.

I think that ELT has now lots of things to share with other disciplines. It can come to the aid of other disciplines to resolve their problems, and get them to be more expanded. In other way, applied ELT opens new horizons for researchers in the field, showing a novel way of dealing with ELT issues. ELT is now capable of forming new theories which can be exported to other disciplines. In addition, the unique setting of English language learning can provide the researchers with more promising areas of research.

By proposing the idea of applied ELT, I hope that researchers place more premium on studies related to the applied ELT. Two types of studies can be targeted by applied ELT: First, ELT researchers are expected to reflect on how ELT findings can be utilized in other fields of studies. This can be a new kind of challenge for all researchers to think up the ways to apply ELT to other disciplines. Second, the researchers are supposed to figure out how to design life syllabus for English language learning classes. Life syllabus is a new type of syllabus around which linguistic syllabus is designed. As already mentioned, language learning must be mixed with issues related to life qualities. In fact, language learning setting should do service to issues of concern in life. This is my hope that this new proposed approach to second language studies is taken more into account by ELT researchers, opening new horizon for them to conduct more research in the field.

References

Allwright, D. (1992). Exploratory teaching: Bringing research and pedagogy together in the language classroom. *Revue de Phonétique Appliquée, 103,* 101-117.

Berns, M., & Matsuda, P. K. (2006). Applied linguistics: Overview and history. In K. Brown (Ed.), *The Encyclopedia of language and linguistics* (2nd ed.; pp. 394–405). Oxford, UK: Elsevier.

Bolton, K. (2004). World Englishes. In A. Davis, & C. Elder (Eds.). *The handbook applied linguistics* (pp. 367-396). Oxford: Blackwell.

Chomsky, N. (1957). <u>*Syntactic structures*</u>. Mouton de Gruyter.

Chomsky N. (1959). <u>A review of B. F. Skinner's verbal behavior</u>. *Language, 35*, 26-58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Dornyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition.* Lawrence Erlbaum Associates, Inc., Publishers.

Dornyei, Z. (2009). *The psychology of second language acquisition*. Oxford: OUP.

Halliday,M. A. K., McIntosh, A., & Strevens, P. D. (1964) *The linguistic sciences and language teaching.* Harlow: Longman.

Hosseini, A., Pishghadam, R., & Navari, S. (2010). Tasire classhayaye zaban dar afzayesh hooshe hayajani. *Language and Literature Studies, 42*, 1-11.

Kachru, B. B. (1982). *The other tongue: English across cultures*. Urbana, IL: University of Illinois Press.

Khaza`ifar, A., Pishghadam, R., & Ziai, S. (in press). Ahamiayte tafakore enteghadi dar khandan. *Farhangestan Letter.*

Kumaravadivelu, B. (1994). The postmethod condition: Emerging strategies for second/foreign language teaching. *TESOL Quarterly*, *28*, 27-48.

Mackey. W. F. (1965). *Language teaching analysis*. London and Harlow.

Pishghadam, R., & Mirzaee, A. (2008). English language teaching in postmodern era. *TELL*, *2*, 89-109.

Pishghadam, R. (2008). Afzayeshe tafakore enteghadi az tarighe mobahesye adabi. *Journal of Literature, 48,* 153-167.

Pishghadam, R., & Navari, S. (2009). *Cultural literacy in language learning: Enrichment or derichment.* Paper presented in International Conference on Languages, Malaysia.

Pishghadam, R., & Saboori, F. (2011). A Qualitative analysis of ELT in the language institutes of Iran in the light of the theory of 'World Englishes'. *Journal of Language Teaching and Research*, *2*, 569-579.

Prabhu, N.S. (1990). There is no best method-why? *TESOL Quarterly, 24*, 161-176.

Schmitt, N. (2002). *An introduction to applied linguistics*, London: Arnold.

Skinner, B.F. (1957). *Verbal behavior*. Acton, Massachusetts: Copley Publishing Group

Strevens, P. (1992). Applied linguistics: An overview. In W. Grabe & R. B. Kaplan (eds.), *Introduction to applied linguistic*s. (pp. 13-31). Addison-Weley Publishing Company.

Vygotsky, L. (1978). *Mind in society.* US: President and Fellows of Harvard College.

Widdowson, H.G. (1979). *Explorations in applied linguistics*. Oxford: Oxford University Press,

Williams, M., & Burden, R.L.(1997). *Psychology for language teachers: A social constructivist approach*. CUP.

Title

Exploring the Iranian Candidates' Use of
Formulaic Sequences in the Writing Module of IELTS

Authors

Keivan Zahedi (Ph.D.)

Shahid Beheshti University, Tehran, Iran

Mohammad Ali Shams (Ph.D. Candidate)

Ferdowsi University of Mashhad, Iran

Bio Data

Keivan Zahedi is an assistant professor of linguistics at Shahid Beheshti University and Institute for Cognitive Science Studies, Tehran, Iran. His research interests include biolinguistics, clinical linguistics, discourse analysis and phraseology.

Mohammad Ali Shams is a Ph.D. Candidate of TEFL at Ferdowsi University of Mashhad, Mashhad, Iran. His research interests include narrative intelligence, IELTS studies, psychology of language learning, and phraseology.

Abstract

The aim of this study is to explore the IELTS candidates' use of formulaic sequences in the writing tasks. Investigating a corpus created from 136 scripts responding to IELTS Academic and General Training Tasks 1 and 2, and using a framework proposed by Mackenzie (2000) for the identification of formulaic sequences, the researchers intend to describe the effect of the frequency and appropriacy of the use of formulaic sequences on the candidates' obtained scores on the three IELTS writing criteria namely Cohesion, Coherence, and Lexical Resource, in particular, and on the writing scores, in general. The results of computer and manual analysis of the writing samples suggest that appropriate use of formulaic sequences have a positive effect on the writing scores while exotic and frequent use of these sequences compromises the quality of the writings. The findings of this study contribute to development and refinement of rating criteria of the Writing Module of IELTS, and can also be used to

promote learning and teaching of formulaic sequences in IELTS preparation programs.

Introduction

Formulaic language plays an important role in learning a foreign language (Ellis, 1994; Wray, 2005; Schmitt, 2006). Although various aspects of processing (Conklin and Schmitt, 2008), learning (Doughty and Long, 2003), and using spoken (Weinert, 1995) and written (Cumming, 2002) formulaic sequences have already been studied by both linguisticians and TEFL scholars, many features of lexical phrases still need to be investigated. The significance of formulaic sequences in the process of learning has been frequently mentioned by many researchers (cf. Wong-Fillmore, 1976 cited in Ellis, 1994; Ellis, 1997a; Mackenzie, 2000; Overstreet and Yule, 2001; Schmitt and Carter, 2004). The role of formulaic language in assisting learner to achieve native-like proficiency, particularly when being tested in high-stake exams, cannot be denied. Formulaic sequences are very important to language use and are remarkably widespread in discourse (Conklin and Schmitt, 2008), and quite expectedly, this feature of language use receives focused attention from test developers, specially designers of high-stake exams who serve a 'gate-keeping' function. Learners' use of formulaic sequences in IELTS, particularly in the writing module, plays a central role in determining their scores; this, even if not assimilated as a salient fact, calls for in-depth analysis and multi-faceted research.

The findings of this study may be of great use to different groups namely test developers, language teachers, IELTS candidates, and second language acquisition researchers. Bachman (1990) states that, besides performing their conventional functions, language tests can provide SLA researchers with exciting opportunities to explore learners' interlanguage. Although this may be neither claimed nor wanted by IELTS designers, one cannot ignore the possibility of studying the use of formulaic sequences in the writing module; this is an unfulfilled potential that can be finely developed to expand our knowledge about formulaic language and its decisive role in the process of language learning and use.

As it could be implied by looking at the place of formulaic language in writing and testing studies in the last decade, after years on the periphery of linguistics,

formulaic language now seems to be taken more seriously. The attention directed to this issue by ESOL's board of examinations shows that many aspects of the use of formulaic sequences in IELTS, particularly writing module, need to be investigated. Following the accelerating trend of studying formulaic language, especially in testing contexts, the present study aims to investigate the effect of the use of formulaic sequences on the writing scores obtained by IELTS candidates by exploring the functional contribution of these items to the text cohesion and coherence and their place in relation to effective use of lexical resources. Learners' appropriate use of formulaic sequences in the writing module of IELTS can increase their score considerably. This study is meant to describe this effect in detail, to present the relevant applications for testing writing, and to suggest further implications for teaching of formulaic sequences.

Theoretical framework

The existing literature about formulaic sequences covers a wide range of studies from theoretical articles on the origin of language (Wray, 1998) to longitudinal case studies that trace the development of formulaic sequences in learners' interlanguage (Schmitt et al., 2004; Li and Schmitt, 2009). A number of studies, mostly those related to teaching of formulaic sequences, are mainly concerned with the place of formulae in the process of learning (cf. Rava, 1998; Wray, 2000; Biber, Conrad and Cortes, 2004). Some others merely work on practical applications of formulaic language in other fields such as translation (Wray, Cox and Lincoln, 2004) and artificial intelligence (Wray, 2002). In the following sections, theoretical and practical aspects of the use of formulaic sequences will be partially discussed.

Identification of formulaic sequences

Ellis (1997a), reviewing a number of SLA studies, maintains that frequency analysis has proved a valuable tool for investing variability in learner language and for describing the sequences of language acquisition. He states that the studies that have employed frequency analysis hold a central place in SLA research, and that studies of acquisition sequence based on frequency analysis have remained popular and have continued to contribute to theory development. Ellis does not ignore the fact that frequency analysis is not without its problems. The major problem, as Ellis mentions, with frequency analysis is that studying the frequency of certain linguistic elements in

the learners' interlanguage (to trace the development of a particular language behavior) should be carried out through longitudinal projects that are very time consuming. For example, Li and Schmitt (2009) spent 10 months to study the development of formulaic sequences in Amy's writings. However, pseudo-longitudinal studies overcome this problem. In such studies, samples of learners' language are collected from groups of learners of different proficiency levels at a single time.

According to Wray (2000), corpus-based estimates about the formulaic sequences are built upon the assumption that they could be identified by virtue of their being more frequent than other word strings; indeed frequency is a central definitional criterion while dealing with formulaic sequences. Wray believes that although this is not an unreasonable starting place, there are some difficulties with it. One is that there are undoubtedly some formulaic sequences that are widely accepted as such by native speakers but which are actually not widely accepted in normal discourse. These include those associated with a specific famous story or cultural event, e.g. *All for one and one for all*. Another difficulty with using frequency as a means of spotting formulaicity is that it forces us to assume that any sequence of words that is repeated a few times is formulaic, that is, that we will not generate the same sentence from scratch very often without then keeping a copy while for later use. There may be some reasons for favoring this idea, but it certainly is not safe to assume it. The difficulty, in short, is that formulaic and non-formulaic language may sometimes look identical, and frequency counts may not be a reliable means of differentiating them. She adds that most of the other candidate identificational criteria focus on the immature or interlanguage forms of learners, where fossilized errors or the correct use of a construction otherwise not within the scope of the speakers' grammatical competence can indicate that the sequence has not been created from scratch at the time. It is probable that a satisfactory means of identification will entail more than one diagnostic, and this makes it particularly important to understand as fully as possible what formulaic sequences are, so that there is no danger of circularity between the definitional and identificational criteria (Wray, 2000). She finally comes to the conclusion that there is no definite answer to the questions seeking for a comprehensive definition for formulaicity; the openness and unusual flexibility of Wray's (2000) definition clearly reflects her attitude in this regard. Actually taking frequency as the major criterion for identifying instances of formulaic language has

been recently challenged by more scholars; Thorne, Reinhardt and Golombek (2008) do not believe frequency to be a valid criterion for identification of formulaic sequences. They emphasize that while directive constructions may correspond to formulaic sequences and lexical bundle, the functional role of directive constructions rather than their frequency or distribution alone should be taken as the main criterion to identify formulaic sequences in writing. Their emphasis on the functional status of formulaic sequences drags the argument of the identification of these linguistic elements into the realm of pragmatics.

The most comprehensive account of the approaches to defining and identification of formulaic sequences has recently been introduced by Li and Schmitt (2009). They state that there have been four main approaches to defining and identifying the formulaic elements in language. The first approach takes frequency as a sign of formulaicity and uses 'association measures' such as t-score to study the significance of word frequencies. The second approach involves the study of all word combinations of certain grammatical forms, for instance, combinations such as *perfectly natural*, regardless of whether they are 'formulaic' in any defined sense. The third approach focuses specifically on 'collocations' as they are defined in the so-called 'Russian school' of phraseology according to which collocations are typically identified as those combinations that cannot be easily altered without a significant change in meaning particularly the pragmatic response (e.g. the phrase *commit* + [something wrong or illegal] is a collocation because *commit a lie/deceit/delinquency* are arbitrarily blocked). A fourth approach relies on proficient speakers' intuitions whether a piece of language is formulaic or not. Given the priorities of each research projects, one of the above approaches could be adopted for exploring the use of formulaic sequences in learner corpora.

The place of formulaic sequences in the wiring module of IELTS

One can easily see the increasing attention of the IELTS center and its associated research groups to the role of formulaic sequence in the writing module by taking a brief look at *Research Notes* recent issues. Adolph's (2003 cited in Shaw, 2004) study on the acquisition of lexical clusters, King's (2003 cited in Shaw, 2004) analysis of collocational boundaries in L1 and L2, and Frank's (2003 cited in Shaw, 2004) report on mnemonic strategies for memorizing lexical phrases are all instances of IELTS

funded research projects whose findings contribute to the knowledge about learning and use of formulaic sequences.

Barker and Taylor (2005) explore vocabulary expansion, particularly development of formulaic language in teaching and testing contexts and observe the enlargement of learners' formulaic lexicon in their test performance. They claim that since the teaching and testing standards for appropriate use of formulaic sequences in writing do not correctly match, testees' performance in writing is unfairly underestimated by some raters. Although drawing such conclusion from the data found in their study does not satisfy the critical reader, but one cannot deny that the mismatch between instructive framework in teaching and evaluative and sometimes even subjective frameworks for testing formulaic sequence should be considered as an acute problem which needs to be, at least partly, alleviated.

Green (2005) examines the effect of study recommendations given to EAP learners on their scores on the IELTS Academic Writing tasks. Based on the results of his study, Green argues that adult learners should be given direct instructions on the use of formulaic sequences while being provided with the relevant contextual explanations and functional justifications for the use of these items. One important implication which can be grasped from the writing samples taken from the subjects of Green's study is that if the learners are provided with adequate treatment in the course of learning, they are likely to be able to use the formulaic sequences in their writings more appropriately, and thus gain better scores on this module of IELTS. The major demerit of Green's study is that he does not suggest any practical method to achieve the goal which he considers as strategic and desired i.e. making the connection between teaching and testing of formulaic language.

Moore and Morton (2005) point out the dimensions of difference between university writing assignments and IELTS writing tasks. They found that while the IELTS items clearly share features with university scripts, the form of writing they prescribe is closer to certain public forms of discourse. This difference was prominently observed in the types of formulaic sequences used by learners. While presenting novel findings, Moore and Morton do not go beyond a general explanation about the importance of formulaic language in determining the tenor of written discourse, although in suggestions for further research they admit that the nature of formulaic sequences used in writing tasks needs to be almost thoroughly inspected in

another study. As stated before, the present study aims to focus on the representation of formulaic sequences in the IELTS writing tasks.

Kennedy and Thorp (2006) analyzed 130 scripts responding to the same task from IELTS Academic Writing Task 2 in order to investigate the linguistic nature of the answers at three proficiency levels -8 (expert user), 6 (competent user), 4 (limited user). They report eleven main findings in relation to similarities and differences in the three levels. Among the others, one finding is particularly relevant to the present study; it was found that level 8 candidates use more idiomatic language than level 6/4, possibly more frequently than natives would. Of course, Kennedy and Thorp consider the formulaic sequences used by more advanced testees to be misplaced in the formal register the question demands. This study also revealed that the use of discourse markers (formulaic sequences that are used to preserve text's cohesion and coherence) was significantly more frequent in level 8 scripts. Kennedy and Thorp mention their lack of information about the testees' linguistic background as a major weak point of the study.

Shaw (2006), in the fifth (concluding) part of an informative article published in *Research Notes* reporting the results of IELTS center's revision in writing module assessment criteria and scales, maintains that the inclusion of descriptors for legislating the use of formulaic language were considered positive by IELTS examiners because it helps them a great deal with the problem of marking memorized or potentially memorized scripts. However, as Shaw admits, this still remains an area of concern. Obviously, more detailed work on the assessment of formulaic sequences used in IELTS writing tasks will benefit the examiners substantially.

Li and Schmitt (2009), in a quite recent longitudinal study, examined the development of formulaic sequences in an academic writing corpus created by a Chinese MA student of TEFL (Amy), and deliberately recorded the change in patterns of the use of formulaic sequences in the corpus. Having the opportunity of tracing the trend of change in Amy's writings (a distinguishing feature which other studies lack), Li and Schmitt provide the field with interesting findings. This study examines the use of formulaic sequences in more detail, although it is limited to a small corpus created by an individual.

Having reviewed the relevant research projects and mentioned their weak and strong points, now we turn to the present study and what it pursues to accomplish.

Among the others, Howarth's (2003 cited in Shaw, 2004) extensive survey of current methods for assisting learners to remedy their collocational errors is remarkable. Howarth argues that collocational or phraseological competence is perhaps one of the highest levels of linguistic proficiency that learners can attain. While, he contends, for many purposes absolute native naturalness in not always required (or even possible) there are some learners who aim at high levels of achievement in this area for occupational purposes; as he correctly states, this is an area receiving great attention in examinations such as IELTS. Shaw (2004) emphasized that incorporating collocational knowledge into testing, particularly at higher proficiency levels, is clearly a challenging enterprise. As he expects, IELTS research center continues its research into describing written proficiency using corpora in order to explore this area to ensure that IELTS remains relevant to the needs of test takers and users.

Method

Participants

Participants of this study were of two groups: students at IELTS writing preparation classes, and IELTS real candidates. The IELTS writing tasks and scores that we are basing my research on are of two types:

1) IELTS scripts and scores which were taken by real candidates and which were made available to me to work on. These include:

a) 6 scripts responding to IELTS General Training Writing Task 1

b) 6 scripts responding to IELTS General Training Writing Task 2

c) 6 scripts responding to IELTS Academic Writing Task 1

d) 6 scripts responding to IELTS Academic Writing Task 2

2) Sample writing tasks which were produced by candidates who were given a mock IELTS writing exam. The scripts were later on sent to a non-practicing IELTS examiner to be scored. These include:

a) 10 scripts responding to IELTS General Training Writing Task 1

b) 27 scripts responding to IELTS General Training Writing Task 2

c) 22 scripts responding to IELTS Academic Writing Task 1

d) 53 scripts responding to IELTS Academic Writing Task 2

It is worth mentioning that all the scripts were anonymous.

Instrumentation

The main instrument used in this study consists of specimens of IELTS Academic and General Training Writing Tasks (2003 updated January 2005) approved by British Council, Cambridge ESOL and IDP: IELTS Australia.

The specimen materials used in the present study were basically published to be used by IELTS candidates as a practicing device that could help them to estimate their score on the real exam. However, as mentioned in the specimen booklet, a high score on these specimen papers does not guarantee that the same standard will be reached in the real test.

Since the effect of the topical knowledge on the use of formulaic sequences was not meant to be investigated in the present study, a range of different versions of the writing tasks were selected to be administered in the mock exam.

Results and discussion

The results of this study are organized in the following 3 sections. First, the characteristics and distinguishing features of the four groups of scripts regarding the extent to which they have met the introduced criteria for writing tasks (coherence, cohesion, and lexical resources) are discussed. In the second part, candidates' use of different categories of formulaic sequences, regarding both frequency of items and appropriacy of use, is reported. Finally, the information related to the most relevant categories of formulaic sequences to cohesion, coherence, and lexical resources, extracted from the tables reporting frequency and appropriacy of the items, and the mean of the scores given by the IELTS examiner to the candidates' success in meeting the writing tasks' criteria are put together in separate tables to allow one to make more valid comparisons between the four levels.

Characteristics of the Four Levels

Obviously, more advanced learners can obtain higher writing scores because they can meet the Writing Module's standards to a greater extent. In this section, by focusing on the examiner comments on the scripts, the researcher aims to investigate the differences between the candidates' success in accomplishing the writing tasks in detail. It is worth mentioning that in the examiner discourse, the candidates' use of

lexical resources is discussed in terms of three aspects namely accuracy, collocation, and sophistication.

Table 1 shows the means of the scores given by the IELTS examiner to the scripts' cohesion, coherence, and lexical resources. The tabulated data shows the numerical aspects of the descriptive comments mentioned above.

Table 1
The Means of the Scores Given by the IELTS Examiner to the Scripts' Coherence and Cohesion and the Candidates' Use of Lexical Resources at the Four Proficiency Levels

The extent to which the candidates in the 4 proficiency levels can meet the **IELTS Writing Module' criteria**	Level 4 and less	Level 4.5 to 5.5	Level 6 to 7	Level 7.5 to 9
Cohesion	5	6	7	8.5
Coherence	3.5	4	6	8
Lexical Resources	4	5.5	6.5	8

Frequency of the Formulaic Sequences

Since the numbers of words in the four groups of scripts are not equal, in addition to raw frequency of the formulaic items in each category, the relative frequency of the items is also presented. The probable comparisons between the uses of items in different groups have to be made based on the number of these items per 10,000 words i.e. their relative frequency.

Table 2 shows a list of the categories of formulaic sequences which is sorted from the most frequent category (Connectors) to the least frequent one (Qualifiers). This table also includes the total frequency numbers both in each category and in each level of proficiency.

Table 2

Seven Categories of Formulaic Sequences Sorted Based on the Observed Frequencies in the Writing Samples

Numbers show the frequency of items per 10,000 words	Level 4 & less	Level 4.5 to 5.5	Level 6 to 7	Level 7.5 to 9	Total
Connectors	40.06	**77.54**	64.80	38.75	**221.15**
Summarizers	**38.32**	32.31	23.14	8.02	**101.79**
Lexical Phrases	15.68	19.38	20.83	**45.44**	**101.33**
Evaluators	10.45	**25.85**	11.57	9.35	**57.22**
Fluency devices	**8.71**	7.75	5.79	1.33	**23.58**
Exemplifiers	8.71	14.22	**15.04**	6.68	**44.65**
Qualifiers	**5.23**	2.58	1.16	2.7	**11.67**
Total	**127.16**	**179.63**	**142.33**	**112.27**	

Appropriacy of the Use of Formulaic Sequences

Since mere frequency results are not believed to be enough to judge the effect of candidates' use of formulaic sequences on the writing tasks, the examiner's attitude toward the appropriacy with which these sequences have been used was also investigated. It is worth mentioning that assessing the formulae in terms of appropriacy was accomplished as a completely independent process after the rating process was finished. This is important because the researcher did not want the examiner to be sensitive to the use for formulaic sequences (more than usual) while rating the scripts based on the criteria recommended by IELTS developers.

The instances of formulaic sequences were explicitly introduced to the examiner; he assessed the appropriacy of the use of each item based on its relevance to the cotext and the extent to which it was serving its usual function in the text. The final results of this assessment are presented in Table 3.

Table 3

The Appropriacy of the Use of Formulaic Sequences (7 Categories) at Four Proficiency Levels

The average degree of appropriacy of the use of the 7 categories of **Formulaic Sequences** in the 4 groups of **IELTS** candidates	Level 4 & less				Level 4.5 to 5.5				Level 6 to 7				Level 7.5 to 9			
	Very Appropriate	Appropriate	Less Appropriate	Not Appropriate	Very Appropriate	Appropriate	Less Appropriate	Not Appropriate	Very Appropriate	Appropriate	Less Appropriate	Not Appropriate	Very Appropriate	Appropriate	Less Appropriate	Not Appropriate
Lexical Phrases			●				●			●			●			
Connectors			●			●				●				●		
Fluency devices				●			●				●		●			
Exemplifiers			●			●				●				●		
Qualifiers			●				●			●				●		
Evaluators				●			●				●		●			
Summarizers				●			●			●				●		

Formulaic Sequences and IELTS Writing Criteria

The interpretations presented in this section are meant to shed light on the hypotheses of this research. By focusing on the relationships between formulaic sequences used

by the candidates and three main characteristics of the scripts (cohesion, coherence and lexical resources), the researcher tries to explore the role that formulaic language plays, or can play, in determining one writing score.

Conclusion

This study was meant to cover those areas of the use of formulaic sequences in the learners' interlanguage, particularly EFL writing scripts, which have not been extensively studied before to provide the field with more comprehensive findings regarding the effect of formulaic sequences in the Writing Module of IELTS. To this end, the learners' use of formulaic sequences in the writing tasks were explored based one of the most exhaustive categorizations of these sequences in the literature (Mackenzie, 2000), and the results were juxtaposed with the scores obtained by the candidates in three areas of writing namely Cohesion, Coherence, and Lexical Resources. The main findings of the study and the plausibility of the suggested hypotheses will be discussed in the following section.

Findings

The findings of the present study are as follows:

1) Among the seven categories of formulaic sequences proposed by Mackenzie (2000), 'Connectors' is the most frequent, and 'Qualifiers' is the least frequent in the IELTS Writing performances.

2) Level 4 and less candidates make more use of the formulaic sequences that function as *summarizers*, *fluency devices*, and *qualifiers*; level 4.5 to 5.5 candidates make more use of formulaic sequences that function as *connectors* and *evaluators*; Level 6 to 7 candidates make more use of the formulaic sequences that function as *exemplifiers*; and level 7.5 to 9 candidates make more use of *lexical phrases*.

3) Generally, formulaic sequences are observed more frequently in the scripts at level 4.5 to 5.5 while the candidates at level 7.5 to 9 use these sequences less frequently compared to other groups.

4) Although the frequency of the items belonging to those categories of formulaic sequences that contribute to the cohesion of writings namely connectors, exemplifiers, and summarizers in level 7.5 to 9 is much less than the other levels, the candidates in this level receive the highest cohesion scores.

5) Despite the fact that the candidates at <u>level 4 and less</u> use those formulaic sequences that contribute to the coherence of the writing tasks more than the other levels, they receive the lowest coherence scores.

6) In general, the distances between the scores given to the candidates at the four levels in the three areas of IELTS Writing Module criteria namely Cohesion, Coherence, and Lexical Resources seem to be more affected by the appropriacy of the use of formulaic items rather than their frequency.

Based on the above findings, now we can discuss the research hypotheses and the extent to which they may be corroborated by the available evidence. To the researcher, none of the hypotheses can be easily supported or rejected in the way they are formulated; they can be considered to be supported by the findings of this study only if one condition is met for the use of formulaic sequences. Frequency of these sequences in the writing tasks produced by IELTS candidates does not necessarily guarantee higher writing scores, in general, and higher scores for cohesion, coherence, and lexical resources, in particular. If the formulaic sequences are not used appropriately i.e. in a way that each formulaic sequences serves its relevant function, they would not have a positive effect on the scores. The formulaic sequences have to be contextually relevant to make the desired effect on the organization of the text. If formulaic sequences are not used appropriately and accurately enough, they would have a negative effect on the cohesion and coherence of the script let alone having a positive effect, as it was the case for a considerable number of instances of the use of formulaic sequences in the writings of low-level writers. In such cases, cohesive markers function conversely and decrease the script' cohesion; organizers do not serve their function and weaken the unity (coherence) of the text; lexical phrases are misused because the candidates are not collocationally competent enough to use them at the right time and the right place. By and large, inappropriate use of formulaic sequences would have a negative effect on the IELTS candidates' writing scores.

Applications and Implications

The insights into the texts of IELTS writers which were emerged in this research can be considered by the IELTS partners and IELTS professionals when contemplating changes to the IELTS writing component. These changes may happen in the structure of the test itself or might affect the rating standards in the writing module.

The attention of the non-practicing IELTS examiner that participated in this study focused on the use of formulaic sequences in the scripts and more importantly, their appropriacy and contextual relevance. Such special focus is not, at least officially, conventional in the process of rating the IELTS writing tasks. Due to the importance of the frequency and appropriacy of the use of formulaic sequences, refining the criteria on which scripts are usually scored may help the IELTS developers and raters to measure one's writing ability more accurately.

Rethinking the standards of the Writing Module of IELTS is not a new topic; it has been continuously discussed in *Research Notes* published by Cambridge ESOL (cf. section 2.15). If collocational competence, as an important part of learners' developing interlanguage, is going to be tested more accurately, the significance of developing new standards for identification and evaluation of the use of formulaic language cannot be ignored. The analytic framework used in this study to identify and evaluate the use of formulaic sequences can be considered as a forward step to provide IELTS raters with more stringent criteria for assessing candidates writing skill. The framework used here can be seen a model, although still not detailed enough, to measure candidates' formulaic performance.

The findings of this study carry educational implications for the learning and teaching of formulaic sequences, particularly in IELTS preparation courses. Although some methods implicitly contribute to teaching lexical phrases, prefabricated patterns, or useful chunks of language, a systematic approach to teaching formulaic sequences in IELTS preparation classes has not been developed yet. The candidates are usually exposed to some instances of formulaic sequences without receiving the necessary knowledge to use them appropriately.

The frequent and inappropriate use of some categories of formulaic sequences in the IELTS writing tasks indicates that the candidates are familiar with several formulae through explicit teaching or incidental learning, but their collocational and functional competence have not been broadened in the classrooms. As a result, candidates do make use of formulaic sequences but most of the time, but fail to use them at the right time and at the right place. Functions of formulaic sequences need to be toughed, preferably indirectly, in IELTS preparation classes for writing.

This study was based on a corpus-based investigation; using the facilities generated in the field of corpus linguistics, specially computer programs that are used to identify and explore many aspects of the written corpora, teachers in IELTS

preparation courses can provide the candidates with ample opportunities to see different forms of formulaic sequences in a meaningful context and practice their use to understand the proper mode and function of these sequences. Providing the students with softwares such as *Wordsmith 3.0* and devising exercises for developing their formulaic competence is one the most practical ways to apply the knowledge created in this study in the classroom.

Based on the findings of the present study, although high-level writers receive the highest cohesion and coherence scores, they do not make use of overt formulaic sequences which are more frequently used by low-level writers to ensure the unity of their scripts. One logical assumption could be that high-level writers use devices other than overt discourse markers to organize their ideas; it would be an appropriate further research objective to discover what such alternative devices might be.

It is also an intriguing question that how writers move from low-levels (with respect to use of formulaic sequences) to the higher levels. In pedagogic terms, a comparison of a typical high-level writing with a low-level writing might be a useful exercise for the IELTS writing classroom; this could set the ground for an experimental research on the teachability of formulaic sequences in writing. The IELTS Writing Module criteria include another item i.e. Grammatical Range and Accuracy, which was not relevant to the scope of the present study. Investigating the grammatical accuracy of the formulaic sequences used by the candidates at different proficiency levels would reveal some other features of the use of these sequences in relation to learners' interlanguage. It would also be an interesting opportunity to test and challenge the opposing theories made about the relationship between exemplar-based knowledge and rule-governed knowledge. The researcher's tentative impression is that there are meaningful differences between the frequency and appropriacy of the formulaic sequences used by Persian candidates' of IELTS and the candidates from other nationalities. Comparing Iranian and non-Iranian candidates' response to IELTS writing tasks would provide us with useful information about the weak and strong pints of the IELTS preparation classes in Iran.

This study was based on a relatively small learner corpus which was not created in an authentic testing context. Further research using a more comprehensive data collected across different linguistic backgrounds, various learner styles, and several

educational settings can provide the field with more interesting findings about learning and use of formulaic sequences.

References

Ari, O. (2006). Review of three software programs designed to identify lexical bundles. *Language Learning & Technology, 10*(1), 30-37.

Bachman, L. (1990). *Fundamental considerations in language testing.* New York: Oxford University Press.

Barker, F., & Taylor, L. (2005). *Vocabulary in teaching and testing contexts: Insights from corpus analysis*. (Research Notes. NO. 22). Cambridge: Cambridge ESOL.

Biber, D., Conrad, S., & Cortes, V. (2004). Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*, 377-405.

Carrell, P. (1982). Cohesion is not coherence. *TESOL Quarterly, 16*, 479-487.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than non-formulaic language by native and nonnative speakers? *Applied Linguistics*, *29*(1), 72-89.

Crewe, W. (1990). The illogic of logical connectors. *ELT Journal, 44*, 316-325.

Cumming, A. (2002). Assessing L2 Writing: Alternative constructs and ethical dilemmas. *Assessing Writing, 8*, 73-83.

Cumming, A., Kantor, R., & Powers, D. E. (2002). Decision making while rating ESL/EFL writing tasks: A descriptive framework. *The Modern Language Journal, 86*(1), 67-96.

Dörnyei, Z., Durow, V., & Zahran, K. (2004). Individual differences and their effects on formulaic sequence acquisition. In N. Schmitt (Ed.), *Formulaic sequences acquisition, processing, and use* (pp. 87-106). Amsterdam: John Benjamins Press.

Doughty, C. J., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology, 7*(3), 50-80.

Ellis, R. (1984). *Classroom second language development.* Oxford: Pregamon.

Ellis, R. (1994). *The study of second language acquisition.* New York: Oxford University Press.

Ellis, R. (1997a). *Analyzing learner language.* New York: Oxford University Press.

Ellis, R. (1997b). *Second language acquisition.* New York: Oxford University Press.

Erman, B., & Warren, B. (2000). The idiom principle and the open-choice principle. *Text, 20*, 29-62.

Field, Y., & Yip, L. (1992). A comparison of internal cohesive conjunction in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal, 23*(1), 15-28.

Flowerdew, L. (1998). Integrating 'expert' and interlanguage computer corpora findings on causality: Discoveries for teachers and students. *ESP Journal, 17*, 329-345.

Green, A. (2005). EAP study recommendations and score gains on the IELTS academic writing test. *Assessing Writing, 10*, 44-60.

Hakuta, K. (1976). A case study of Japanese child learning English as a second language. *Language learning, 26*, 321-351.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* London: Longman.

Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. *International Review of Applied Linguistics in Language Teaching, 3,* 205-219.

Kennedy, C., & Thorp, D. (2006). A corpus-based investigation of linguistic responses to an IELTS Academic Writing Task 2. In M. Milanovic & C. Weir (Series Eds.) & L. Taylor & P. Falvey (Vol. Eds.), *Studies in language testing: Vol. 19. IELTS collected papers: Research in speaking and writing assessments* (pp. 316-377). Cambridge: Cambridge University Press.

Khalil, A. (1990). A study of cohesion and coherence in Arab EFL college students' writing. *System, 17*, 359-371.

Krashen, S. (1982). *Principles and practice in second language acquisition.* Oxford: Pregamon.

Li, J., & Schmitt, N. (2009). *The acquisition of lexical phrases in academic writing: A longitudinal case study. Journal of Second Language Writing, 18*, 85-102.

Lightbown, M., & Spada, N. (1999). *How languages are learned.* New York: Oxford University Press.

Lyons, J. (1968). *Introduction to theoretical linguistics.* Cambridge: Cambridge University Press.

Mackenzie, I. (2000). Improvisation, creativity, and formulaic language. *The Journal of Aesthetics and Art Criticism, 58*, 173-179.

Miyakoshi, T. (2004). Collocations and second language acquisition: The acquisition of English adjectival constructions. *Working Papers in Linguistics, 35*(1), 1-21.

Moore, T., & Morton, J. (2005). Dimensions of difference: A comparison of university writing and IELTS writing. *Journal of English for Academic Purposes, 4*, 43-66.

Myles, F., Hooper, J. & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom language learning. *Language learning, 48*, 323-363.

Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 125-152). Amsterdam: John Benjamins Press.

Neuendorf, K. A. (2002). *The content analysis guidebook.* London: Sage Publication. Overstreet, M., & Yule, G. (2001). Formulaic disclaimers. *Journal of Pragmatics, 33*, 45-60.

Phongphio, T., & Schmitt, N. (2006). Learning English multi-word verbs in Thailand. *Thai TESOL Bulletin, 19*, 122-136.

Rava, S. (1998). A proposal for teaching writing. *The French Review, 72*(1), 58-68.

Schauer, G. A., & Adolphs, S. (2006). Expressions of gratitude in corpus and DCT data: Vocabulary, formulaic sequences, and pedagogy. *System, 34*, 119-134.

Schmitt, N. (1998). Measuring collocational knowledge: Key issues and an experimental assessment procedure. *Review of Applied Linguistics, 119,* 27-47.

Schmitt, N. (2000). Key concepts in ELT: Lexical chunks. *English Language Teaching Journal, 54*, 400-401.

Schmitt, N. (2006). Formulaic language: Fixed and varied. *Estudios de Lingüística Inglesa Aplicada, 6*, 13-39.

Schmitt, N. (2007). Current perspectives in vocabulary learning and teaching. In J. Cummins & C. Davison (Eds.). *The International Handbook of English Language Teaching: Vol. 2* (pp. 827-842). London: Springer.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*, 329-363.

Schmitt, N., & Carter, N. (2000). Lexical phrases in language learning. *The Language Teacher, 24*(8), 6-10.

Schmitt, N., Dörnyei, Z., Adolphs, S., & Durow, V. (2004). Knowledge and acquisition of formulaic sequences: A longitudinal study. In N. Schmitt (Ed.), *Formulaic sequences acquisition, processing, and use* (pp. 55-68). Amsterdam: John Benjamins Press.

Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-derived recurrent clusters psycholinguistically valid? In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, processing, and use* (PP. 127-151). Amsterdam: John Benjamins Press.

Shaw, S. (2004). *Phraseology and metaphor in EAP.* (Research Notes. No. 17). Cambridge: Cambridge ESOL.

Shaw, S. (2006). *IELTS writing: Revising assessment criteria and scales (conclusion).* (Research Notes. NO. 24). Cambridge: Cambridge ESOL.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, *64*, 429-458.

Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics, 18*, 268-286.

Thorne, S. L, Reinhardt, J., & Golombek, P. (2005). Mediation as objectification in the development of professional discourse: A corpus-informed curricular innovation. In J. Lantolf & M. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 113-150). London: Equinox.

Tseng, W. T., & Schmitt, N. (2008). Towards a self-regulating model of vocabulary learning: A structural equation modeling approach. *Language Learning*, *58*, 357-400.

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences acquisition, processing, and use* (pp. 153-172). Amsterdam: John Benjamins Press.

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics, 16*, 180-205.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language & Communication, 18*, 47-67.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics, 21*, 463-489.

Wray, A. (2002). Formulaic language in computer-supported communication: Theory meets reality. *Language Awareness, 11*, 114-131.

Wray, A. (2005). Needs only analysis in linguistic ontogeny and phylogeny. In A. Cangelosi (Ed.), *Proceedings of the Second International Symposium on the Emergence and Evaluation of Linguistic Communication.* Society for the study of artificial intelligence and the simulation of behavior (AISB), Hatfield, April 2005.

Wray, A., & Grace, G. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua, 117*, 543-578.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication, 20*, 1-28.

Wray, A., Cox, S., Lincoln, M., & Tryggvason, J. (2004). A formulaic approach to translation at the post office: Reading the signs. *Language & Communication, 24*, 59-75.

Title

The Effect of Explicit Training of Metacognitive Vocabulary Learning Strategies on

Recall and Retention of Idioms by Advanced EFL Students

Authors

Mehdi Mardani (Ph.D. candidate)
University of Isfahan, Iran
Ahmad Moinzadeh (Ph.D.)
University of Isfahan, Iran

Bio Data

Mehdi Mardani, Ph.D. candidate of TEFL at Isfahan University, Isfahan, Iran. His main research interests are: reading comprehension, vocabulary acquisition, and testing. He has published in some international journals including Journal of Language, Society, and Culture, and Reading in a Foreign Language. He has also submitted an article to the international journal of Language Assessment Quarterly, and another one for Iranian EFL Journal.

Ahmad Moinzadeh, associate professor at University of Isfahan, Isfahan, Iran from 1995 to 2011. At present, he runs Linguistics, Applied Linguistics, Second Language Acquisition, and First Language Acquisition courses. He has published several articles in the field of TEFL and has supervised many dissertations and theses in this field.

Abstract

The purpose of the present study is to investigate the effect of vocabulary learning strategy on retention and recall of idioms in Iranian EFL context. For this purpose, 50 advanced female students were selected and assigned to control and experimental groups on the basis of their TOEFL scores. The procedure of the study consists of two phases. In the first phase it is tried to measure short term recall. For this purpose during the two sessions of each week some part of the class time was allocated to

teaching idioms, and the last session of the same week, the idioms that were taught during the previous sessions were tested through a (short-term test) to measure their short term recall. A t-test was run to compare the short term recall of the two groups. In order to measure long term recall, three weeks after the treatment, a post-test was administered to show the difference between the gained mean score of the pre- and post-tests on long term retention test. A t-test was run to compare the performance of the two groups in terms of their long term retention. The results indicated that treatment did have effect on both short term and long term recall of idioms.

Introduction

Two common themes in foreign language learning are language learning strategies and vocabulary learning, and they have taken up too much room in SLA research. Vocabulary, according to Zimmerman (as cited in Coady and Huckin, 1997), is central to language and language learning. As a subcategory of vocabulary, idioms are believed to be the stumbling block for foreign and second language learning. Strategies, however, are believed to be facilitator of learning, storage and recall of information. There is now convincing evidence that people who use these strategies are more successful than the ones who do not use them or in a lower rate (Oxford, 1990). This means that appropriate and tactful use of language learning strategies helps learners overcome most of their learning problems.

But, a problem that challenges most EFL/ESL students (and sometimes teachers) is that some of the words they learn or memorize will soon be forgotten or hard to retrieve in real context of use. This problem raises the question of "why students cannot remember the meaning of words after the first encounter". Hulstijn (1997) provides the most likely answer to this question by asserting that students do enough for the immediate comprehension but not enough for the retention over time. Retention over time requires learners to make effort to link the words with their meanings. The term retention or recall is vague term which needs to be elaborated

As far as the definitions of recall and retention are concerned it should be pointed out that to date there is no unanimous definition on the time interval between the first exposure and the second encounter. In other words, there is no certain amount of time

that can be considered as short term recall or long term retention. In this study recall and retention have been operationally defined in the following ways. What is meant by *immediate recall* is the ability of learners to remember the material immediately at the end of treatment in each session. This is based on the Laufer's (2007) definition for short-term recall: it is usually measured immediately after performing the task that is supposed to lead to retaining some information, in our case- vocabulary, or after a short intervention. You can test the retention of these words immediately, or make the learners do another task for 10-20 minutes and then test the target vocabulary. For long term retention some people administer a test a month or even 3 months; some people repeat measurement several times to check how much learners retain at different points of time. But there are practical problems with testing vocabulary long after initial exposure. So, long-term retention, in short, is operationally defined as testing students three weeks after instruction (B. Laufer, personal communication, October 29, 2007).

Language Learning Strategies

Instruction of language learning strategies, especially vocabulary learning strategies, will provide learners with a kind of autonomy. As a result, autonomous learners become more highly motivated. This means that motivated students are more likely to seek out opportunities to use the target language and make maximum use of them. As Wenden (1991) reminds us, there is a proverb that says: Give a man a fish and he eats for a day but teach him how to fish and he eats for a lifetime. Applying this proverb to language teaching we can conclude that if students are provided with strategies to work with they will reach a level of autonomy at which they will be able to work out the answer to their problems. According to Knowles (1997), one of the main aims of education, in general, and language teaching, in particular is to help students to develop a sense or attitude that learning is a lifetime process and requires skills of self-directedness; one who is equipped with such appropriate skills and strategies to learn a language in a self-directed way is an autonomous learner.

Research in Language Learning Strategy and vocabulary learning

Oxford (1990) asserts that "learning strategies are steps taken by students to enhance their own learning" (p. 1). It is important to point out that there are tens of different language learning strategies. They are covered under different names, classification

systems for learning strategies or taxonomies by different researchers (Rubin, 1975; Stern, 1992; O'Malley & Chamot, 1990; and Oxford, 1990). The most famous models for the teaching of language learning strategies favor either a direct teaching model, or an indirect model. Generally, in direct or explicit training, learner's attention is directed towards the strategy being taught. According to William and Burden (1997), "this model is preferred to indirect teaching where learners are not told the purpose of the tasks (p.162)". Accordingly, in the present study language learning strategies are presented directly.

A number of qualitative studies in the field of lexical acquisition have explored the strategies that successful L2 learners employ in acquisition of words. Most have used quantitative research and tried to find patterns. Pioneering work in this area was undertaken by Ahmad (1989, as cited in Riazi, Sadeghi and zare, 2005). Other works in this area are (Carter, R., & M. McCarthy, 1988; Gu & Johnson, 1996; Eslami Rasekh and ranjbari, 2003; Gu, 2003; Marefat and Shirazi, 2003), to mention a few. The area of vocabulary language learning strategies has been investigated (ex. Flevel, J., 1976; Oxford, R.,1990; O'Malley, J. M. & A. U. Chamot.,1995; Williams, M. & Burden, R., 1997; Ellis, R. 1999; Oxford, R., 2003;); but, as it was pointed out earlier, most research has examined the effectiveness of one or two strategies and few have adopted a holistic approach. Royer (1973) found that rote repetition is less effective than structured review; Baxter (1980) argued that dependence on bilingual dictionaries.

What oxford (1990) strongly believe in is that indirect strategy training reinforce the direct ones. The result is that learners using vocabulary learning strategies will recall information netter than those who do not use. The reason of this, as Hall (2004) indicates, is that the learners make use of his knowledge of his memory system. Following this line of research Marefat and Shirazi (2003) investigated the impact of teaching direct learning strategies on the retention of vocabulary by Iranian EFL learners. The results of their study showed that learners' use of strategy in short-term retention outweighs that in long term retention. Eslami Rasekh and Ranjbary (2003) also have investigated the effect of metacognitive learning strategy training through the use of explicit strategy instruction on the development of lexical knowledge of EFL students. The result of their study showed that explicit strategy training has a significant positive effect on the vocabulary learning of EFL students. What is easily felt is the fact that there is a common sense between expert in either fields of strategy

training and lexical approach that strategy training will have a positive effect on learning, retention and recall of idioms.

Purpose of the Study

This study is an attempt to make a compromise between proponents of the view of language as rule-based and those of memory-based view. Applying this dichotomy to the present study, we see that there have been controversies among theorists. On the one hand, Chomsky, as the main supporter of rule-based system, emphasizes the capacity of human beings to create and interpret an infinite number of sentences that are unique and have never been heard or produced before, and on the other hand, people like Wilkins, Lewis, and Skehan as supporters of the memory-based view argue that without grammar very little can be conveyed, but without word nothing can be conveyed (Thornbury, 2002). In fact, the latter group believes that although a rule-based view is much more generative, in real use it is the lexical items that are stored in the memory which are active. In other words, most of the things we say are repetitive. In fact, we are not to downgrade the potentials of a rule-based system, but what we want to say in this study is that most of the language, especially spoken language comprises prefabricated expressions.

Lexical competence or vocabulary knowledge is supposed to be the building block of any language. Helping students to increase their repertoires in prefabricated chunks is one solution to the issue of fluency, a feature of language that even some advanced learners cannot master. One way to achieve this is through integrating direct instruction of vocabulary learning strategies. Thus, if the results of this study demonstrate a significant effect on retention of lexical phrases, especially idioms, it can be suggested to language teachers to adopt a strategic approach to arouse students' awareness of the role of vocabulary language learning strategies. The approach which has vocabulary as its central point is known as lexical approach. This approach, according to Richards and Rodgers (1986), is derived from the belief that the building blocks of language learning and language teaching are not grammar, functions, notions, or other units, but lexis. In fact, Lewis (1993) proposed an approach to language teaching which moved vocabulary to the forefront of ELT. Based upon the afore-mentioned discussions, the following research hypotheses stand out:

H 1. Direct teaching of vocabulary learning strategies has no significant effect on long retention of idioms by Iranian EFL students.

H 2. Direct teaching of vocabulary learning strategies has no significant effect on short-term recall of idioms by Iranian EFL students.

Method

Participants

50 female students whose age ranged from 16 to 24 participated in this study. According to the placement test of the institution, the subjects enjoyed the same proficiency level. In fact, they had passed several terms of EFL instruction and at the time of the study they were at the advanced level. This signifies that they have been able to obtain the acceptable scores to fulfill the requirement of the institute. However, in order to double-check the homogeneity of the two groups in terms of proficiency the TOEFL test was administered. The result of the t-test also indicated that they enjoyed the same level of proficiency before the treatment phase. The mean score of the two groups was M= 64.55. One group consisting of 25 subjects was randomly assigned to the control group and the other to the experimental group. More specifically, the subjects who had already been divided into two groups by the institute were assigned to the control and experimental groups, based on their scores on TOEFL.

Instrumentations

Four instruments were used in this study, as follows:

- the TOEFL test,
- vocabulary learning strategies questionnaire (VLS Version 3, by Gu and Johnson, 1997),
- four teacher-made immediate recall tests of idioms, and
- a long term recall test of idiom.

The TOEFL language proficiency test

The TOEFL proficiency test was used to ensure the homogeneity of the subjects in terms of their language proficiency. Another important reason why the TOEFL was administered was that it was used as a criterion for estimating the reliability and

validity of the teacher-made tests for measuring short term and long term retention. The TOEFL was piloted with a group of 20 students similar to the subjects of the main study. It consisted of three parts and all parts were in the form of Multiple-choice questions. TOEFL enjoyed a high reliability: R= .83.

Vocabulary learning strategies questionnaire

In this study we adopted the questionnaire of vocabulary learning strategies (VLQ, Version 3, see Appendix 4) devised by Gu and Johnson (1996). The questionnaire consists of 91 items. Serving as a consciousness-raising activity, the questionnaire also provided a general assessment of the types of strategies that learners were familiar with; it also elicited students' beliefs about vocabulary strategies and their self-reported vocabulary learning strategies. The first part of this questionnaire includes beliefs about vocabulary learning strategies. The second part includes types of vocabulary learning strategies, which are classified into eight headings, each with its respective subheadings. Namely, "*beliefs about vocabulary learning", "metacognitive regulation", "guessing strategies", "note-taking strategies", "dictionary use strategies", "memory strategies: rehearsal, "memory strategies: encoding", and "activation strategies".*

Idiom long-term recall test

In order to tap the recall ability of the subjects a 40-item multiple-choice idiom pre/post test which was developed and standardized by the researcher, was used. These items had passed through different filters before being included in the final version of the test. It should be noted that the same test was used as posttest for measuring long term recall of idioms three weeks after the program to measure long-term retention of idioms. Though the same test was used for pre and post test, in order to minimize the effect of test familiarity, the same items were rearranged in two different ways, so that the two tests would not look very much alike. Furthermore, the students were not told that the same test would be used for pre and post test. Though at first it may seem to violate the issue of ethics and human consent, but there was no other choice. Moreover, the interval between the two administrations was long enough to minimize this test effect. Using KR-21, which is a conservative estimate of reliability, reliability index of the test was calculated which was R= 0.70. Because it is an underestimate of the true reliability we can safely interpret it as a significant index.

In other word, this formula never overestimates the true reliability. Also the criterion-related validity index of the test was calculated to be V= 0.75. This index is satisfactory.

Idiom short-term recall tests (quiz)

In order to measure the short-term recall of idioms four short-term-recall tests or quizzes of idioms, each comprising of ten items, were developed by the researcher. In fact, each test measured the recall of the ten idioms which were taught during the previous two sessions in the same week. These tests were immediately administered at the end of each week after the materials of the previous sessions were rehearsed (Appendix 4). The reliability indices of the four tests were: R1= 0.53, R2= 0. 51, R3= 0.45, and R4= 0.47, respectively. Again, as in the case of long-term test, the reliabilities of these tests were underestimated because of the use of KR-21 formula which is a conservative one. The criterion-related validity indices of these quizzes were: V1= 0.60, V2= 0.42, V3= 0.60, and V4= 0.29, respectively.

Research Procedure

During the pilot study the idiom tests were constructed and piloted, and during the main phase of the study these tools were applied to measure the effect of the independent variable on dependent variable. In the pilot study a list of idioms including 50 items were selected based on the following criteria:

- selection of the most frequently heard or seen idioms
- consulting some experienced colleagues
- lack of a word-for-word equivalent in participants' L1
- Teachability

During the main phase of the study both the experimental and control groups used New Interchange as the main book of the course; however, part of the class time was allocated to teaching idioms as a supplementary material to the two groups. The only difference was that the experimental group received language learning strategy instruction as treatment and the control group received a placebo treatment. The students in the experimental group were given an introductory lesson about language learning strategies. Then strategies were incorporated into the teaching material, based on the CALLA model of O'Malley and Chamot (1990) which includes five stages:

- preparation,
- presentation,
- practice,
- evaluation,
- and expansion.

Each session five idioms were taught (three sessions a week) and in the last session of the same week we had a full rehearsal of the idioms. Immediately a quiz or test for measuring short-term recall was administered. This procedure was followed for four weeks, and two weeks after the instruction the long-term retention test was administered to measure retention. The results and findings are presented below.

Results

In order to probe the first null hypothesis which states that direct teaching of vocabulary learning strategies has no significant effect on long term retention of idioms, the two groups participated in the long-term retention test. The same test was used as the pretest and posttest. In order to compare the mean scores of the control and experimental groups they gained mean score for each group was computed by subtracting the posttest mean score from the pretest mean score. The descriptive statistics for the two groups are presented in Table 1.

Table 1. Descriptive Statistics (gained score)

	Groups	N	Mean	Std. Deviation	Std. Error Mean
Gained score	Experimental	25	5.8000	4.38748	.87750
	Control	25	1.0000	1.89297	.37859

Next, an independent t-test was run to compare the gained mean scores of the two groups, the results of which indicate that there is a significant difference between the scores of the two groups on the retention test. As can be observed in Table 2, the amount of t-observed value at 48 degree of freedom is greater than the critical t-value which is 2.70 (i.e. *t= 5.02, df=48, p< .05).*

Thus, the first null hypothesis stating that *"direct teaching of vocabulary learning strategies has no effect on long-term retention of idioms of Iranian advanced EFL*

learners" is rejected and it can be concluded that direct strategy instruction does have a significant effect on long term retention of idioms

Table 2. Independent t-test (long-term tests gained score)

| | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
| | F | Sig. | t | df | Sig. 2(tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
								Lower	Upper
Equal Variance assumed	30.172	.00	5.023	48	.000	4.800	.9557	2.878	6.722
variance not assumed			5.023	32.636	.000	4.800	.9557	2.855	6.745

To probe the second null hypothesis stating that direct teaching of vocabulary learning strategies has no significant effect on short-term recall of idioms by advanced Iranian EFL students, four quizzes measuring short-term retention of idioms were given to the subjects. In order to compare the mean scores of the experimental and control groups on the four tests an independent t-test was run, as displayed in Table 3

Table 3. Independent t-test (mean score of the four tests)

| | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
| | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
								Lower	Upper
Equal variances assumed	3.763	.058	9.011	48	.000	2.17000	.24081	1.6858	2.65419
variance not assumed			9.011	43.217	.000	2.17000	.24081	1.6844	2.65558

As can be seen, the t-observed value is 9.01. This amount of t-value at 48 degree of freedom is greater than the critical t-value, i.e. 2.70; in statistical terms: *t= 5.02, df= 48, p< .05.* Thus, the second null hypothesis stating that *"direct teaching of vocabulary learning strategies has no effect on short-term recall of idioms of advanced Iranian EFL learners"* is rejected and it can be concluded that direct strategy instruction has a significant effect on short-term recall of idioms by the students. Concerning the second research question, in order to test the gradual increase or decrease of the means of the two groups from Test 1 to Test 4 a repeated measure ANOVA was run, the results of which are shown in Table 4.

Table 4. Repeated ANOVA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Overall mean	238.555	7	34.079	34.023	.000
Intercept	8646.125	1	8646.125	8631.739	.000
GROUP	235.445	1	235.445	235.053	.000
WEEK	.495	3	.165	.165	.920
GROUP * WEEK	2.615	3	.872	.870	.458
Error	192.320	192	1.002		
Total	9077.000	200			
Corrected Total	430.875	199			

.

As can be observed, the F-value of 235.05 has a p-value of .000 which indicates that the difference between the overall mean scores of the two groups is significant. The results are displayed in Graph 1 below. As Graph 1 illustrates, it can be concluded that the experimental group (the upper line) has outperformed the control group on the four short term tests. The general pattern of movement of the two groups indicates that only the experimental group has a gradual movement from test 1 to test 2, no increase from test 2 to test 3, and a sharp increase from test 3 to 4. In other words, although there are differences in the performance of our experimental group from week to week, overall this group outperformed the control group.

Graph 1. Performance of experimental and control groups on tests from week 1 to 4.
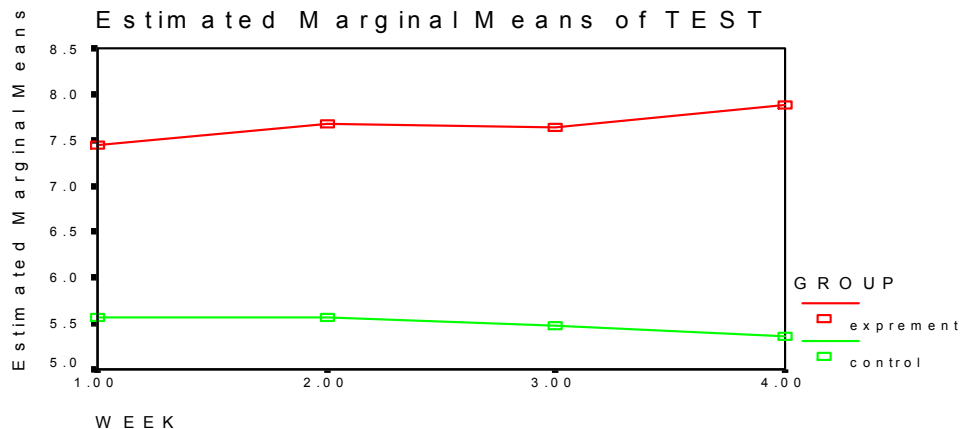
The findings of the present study indicate that direct instruction of strategies has a significant effect on the recall and retention of idioms. The t-values of 5.02, and 9.01 obtained for the retention and recall, respectively, and the F-value of 235.05 all were higher than their related critical values.

Discussion

These findings also provide further confirmation for the results of the previous studies which conclude that language learning strategies, in general, and vocabulary learning strategies, in particular, have significant effects on retention and recall of words and idioms (Cohen & Aphek, 1980; Laufer & Shmueli, 1997; **Marefat & Shirazi, 2003;** Atai, Akbari, and Afzali Shahri, 2004). The findings of the study conducted by Atai, et al. (2004), for instance, showed that sentence writing, as a word-focus vocabulary learning strategy had a significant effect on both immediate recall of the idioms and their delayed recall. **Marefat and Shirazi** (2003) also examined the effect of teaching direct learning strategies (memory, cognitive, and compensation) and their subcategories on vocabulary retention—short term and long term—of EFL learners. Their result showed that strategy instruction in short term retention far outweighs than in long term retention.

  Related to studies done in this field of study, Laufer and Shmueli (1997) also conducted a research. They examined the relationship between different teaching techniques and memorization of new words in terms of both short-term and long-term. In the same line of research, findings of the present study also investigated the effect of two types of metacognitive vocabulary learning strategy, namely selective attention and self-initiation, on the recall and retention of idioms. The results showed

that metacognitive strategy instruction had significant effect on both types of remembering (recall, and retention). Strategy instruction not only had significant effect on recall and retention, but also as the instruction proceeded, the effect became more and more significant. This latter effect was shown by the result of the ANOVA procedure. This means that strategy instruction is a business that requires long term investment from all participants including teachers, students, course designer, and so forth.

Based on the findings of previous studies and considering the fact that research which has directly addressed the issue of prefabricated chunks or formulaic expressions in terms of recall or retention, it can be claimed that this study is a new version of research done in relation to language learning strategies. It is new because rather than engaging itself with providing evidences regarding the effectiveness of strategies, it takes the effectiveness of strategy instruction for granted, and instead wanted to show the depth of that effectiveness on retention (that is, short-term, or long-term) of idioms.

Conclusion

The general conclusion that can be drawn from the results of the present study is that tasks and activities that encourage learners to reflect on their own learning should be incorporated into instructional materials, as these help learners become more strategic and independent learners. Returning to the same distinction between the role of memory-based and a lexical-based function of language, and interpreting it in terms of theoretical implications, we see that what we have been trying to say in this study is that much of the things we say or produce are likely to have been produced before. In pedagogical terms, we are emphasizing a reorientation to the teaching of vocabulary (as an umbrella term), and its subcomponents such as idioms, collocations, phrasal verbs, and so forth. Based on the findings of this study, it can be concluded that if regular instruction of vocabulary learning strategies is provided, and if learners are consciously aware of their own learning, learning will be dramatically enhanced. In this way, teachers can have a stronger effect on the learning process. But it should be kept in mind that sheer presentation of these strategies is not enough. What is needed, in addition to instruction, is that teachers should routinely conduct research in their own classroom to better understand the nature of these learning strategies. Related to the findings of this study, the following implications are presented:

- Closely related to the first implication is the fact that learners need to be provided with tasks to help them organize their mental lexicon.

- Learners need to be provided with opportunities to be actively involved in the learning of words.

- Learners have to be encouraged to take the responsibility for vocabulary expansion

- Teachers should raise students' consciousness of the importance of these strategies by explaining the nature and significance of learning strategies to their students.

- EFL teachers should try to discourage application of non-communicative learning strategies such as translation, rote memorization, and repetition, and instead try to raise students' awareness of topics such as metacognitive (thinking about thinking) strategies, which are profitable means for overcoming the difficulties, dilemmas, embarrassments, and potholes facing EFL students.

Suggestions for Further Research

The following research avenues, which are related to the issue investigated in the present study, are recommended for further research:

1. Vocabulary learning strategies are so varied that including all of them in this study was next to impossible. Other studies can be conducted to find other vocabulary learning strategies and measure their effect on EFL learners

2. The subjects of this study were all female advanced students. Similar studies can be conducted with different levels of language proficiency, gender, age, and social classes.

3. The variables under investigation were vocabulary learning strategies and recall of idioms. Other studies can be conducted to consider the effect of metacognitive strategy training on other types of multi-word utterances such as collocation, phrasal verbs, and the like.

References

Atai, M. R., Akbari, I., & Afzali Shahri, M. (2004). The Effect of sentence writing versus multiple exposures in different contexts on EFL learners' acquisition of idioms. *IJAL, 7, (1).* 46-47.

Brown, D. (2000). *Principle of language teaching and learning*. (Fourth Edition). White Plains, NY: Pearson Education.

Carter, R., & McCarthy, M. (Eds.). (1988). *Vocabulary and language teaching.* London: Longman.

Chamot, A. U. & O'Malley, J. M. (1994). Language learner and learning strategies. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 371-392). London: Academic Press.

Coady, J. & T. Huckin (Eds). (1997). *Second language vocabulary acquisition.* Cambridge: Cambridge University Press.

Cohen A. D. and Aphek, E. (1980). Retention of second language vocabulary over time: Investigating the role of mnemonic association. *System, 8 (2),* 221-35.

Ellis, R. (1999). *Understanding second language acquisition.* Oxford: Oxford University Press.

Eslami Rasekh, Z. and Ranjbary, R. (2003). Vocabulary learning strategy training for vocabulary learning. *TESL-EJ, 7, (2) p, 23-38.*

Flevel, J. (1976). Vocabulary learning aspects of problem solving. In Resnick, L. (Ed). *The nature of intelligence.* Hillsdale, New Jersey: Lawrence Erlbaum Association.

Gu, P. Y. (2003). Vocabulary learning in a second language: Person, task, context and strategies. *TESOL, (7), 2*-23

Gu, Y., & Johnson, R. K. (1997). Vocabulary learning strategies and language learning outcomes. *Language Learning, (46),* 643-679.

Hall, H. (2004). *Metacognition.* Unpublished Ph. D Thesis. University of Missouri.

Hulstijn, H. (1997). Mnemonic Method in Foreign Language Vocabulary Teaching. In Coady, J. & Huckin, T. (1997). *Second vocabulary acquisition; a rationale for pedagogy (pp. 166-184).* Cambridge: Cambridge University Press.

Knowles, M. (1997). *The Modern Practice of Adult Education.* New York: Association Press.

Laufer, B. & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it. *RELC JOURNAL Vol (28). No.1.* 89-108.

Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward.* Hove: Language Teaching Publications.

Marefat, H. and Shirazi, M. A. (2003). The Impact of teaching direct learning strategies in the retention of vocabulary by EFL Learners. *The Reading Matrix,* 3, 2.

O'Malley, J. M. & A. U. Chamot. (1990). *Learning strategies in second language acquisition.* Cambridge: Cambridge University Press.

O'Malley, J. M. & A. U. Chamot. (1995). *Learning strategies in second language acquisition.* UK: Cambridge University Press.

Oxford, R. (1990). *Language learning strategies*: What every teacher should know. New York: Newbury House.

O'Malley, J. M. & A. U. Chamot. (2003). Language learning strategies: Concepts and relationship*, IRAL, 41, PP.* 271-278.

Riazi, Sadeghi. & Zare. (2005). The effect of EFL student's proficiency and age, their overall pattern of vocabulary strategy. *IJAL, 6, (2),* 87-106.

Richards, J. C. and T. S. Rodgers. (1986). *Approaches and methods in language teaching*: A description and analysis. New York: Cambridge University Press.

Royer, J. M. (1973). Memory effects for test-like events during acquisition of foreign language vocabulary. *Psychological Reports, 32 (2),* 195-8.

Rubin, J. (1975). What the "good language learner" can teach us. *TESOL Quarterly* 9 (1): 41-51.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Stern, H.H. (1992)*. Issues and options in language teaching*. Oxford: OUP.

Thornbury, S. (2002). *How to teach vocabulary*. Malaysia: Longman.

Wenden, A. (1991).   Learning strategies for learner autonomy: Planning and implementing learner training for language learners. Englewood Cliffs, NJ: Prentice Hall.

Williams, M. & Burden, R. (1997). *Psychology for language teachers*: A social constructivist approach. Cambridge: Cambridge University Press.

Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition*: *A rationale for pedagogy* (pp. 5-19). Cambridge: Cambridge University Press.

Appendix: List of all idioms used in this study

1. *Shake a leg*
2. *paid through the nose*
3. *keep a stiff upper lip*
4. *make it from scratch.*
5. *kicked the bucket*
6. *I get hot under the collar*
7. *face the music.*

8. *mind your P's and Q's.*
9. *bury the hatchet.*
10. go *to bat for* sb.
11. *talking through his hat.*
12. *bite the bullet*
13. *out of the woods*
14. *smell a rat*
15. *a lemon.*
16. *stuck to his guns.*
17. *fed him a line*
18. *to be like Caesar's wife*
19. *you would make a bundle.*
20. *take the world by storm.*
21. *twisted her arm.*
22. *get the show on the road,*
23. *Way to go*!
24. *tongue-in-cheek.*
25. *bent over backwards*
26. *got the sack*
27. *get the ball rolling,*
28. *threw the gauntlet,*
29. *take the bull by the horns*
30. *Shooting his mouth off.*
31. *green with envy.*
32. *as broad as it long*
33. *something up my slaves*
34. *pull his own weight.*
35. *You have got a heart of gold.*
36. *Why don't you pig out?*
37. *play it by ear.*
38. *left me high and dry.*
39. *spilled the beans.*
40. *bite the dust*

Title

Critical Discourse Analysis in an EFL Classroom

Authors

Aram R. Sadeghi (Ph.D.)

Semnan University, Semnan, Iran

Ali Asadi (M.A. Student)

Semnan University, Semnan Iran

Bio data

Aram Reza Sadeghi is an associate professor and the head of English Department at Semnan University, Iran. He has been teaching English at BA and MA levels for more than 14 years. His research interests include culture studies, material preparation, English as an International Language (EIL), and Critical Discourse Analysis (CDA).

Ali Asadi is an MA student in TEFL at Semnan University, Iran. He got his B.A from Payam-e Noor University of Qom. He has been teaching English for more than five years in different institutes in Tehran. He is currently teaching at KISH Institute of Science and Technology. His research interests lie mostly in the realm of Critical Discourse Analysis.

Abstract

The purpose of this study was to examine the applicability of critical discourse analytical tools in an EFL classroom. To do so, 20 upper-intermediate EFL learners were chosen via a standardized TOEFL test. Then, a pre-test was given to the learners to evaluate their current level of writing skills. After that, in the treatment, the learners were trained how to apply the critical discourse analytical tools into their writings for sixteen sessions. Comparing the scores of the learners on pre-test and post-test via t-test showed that the treatment significantly improved the learners' writing skills. Finally, it was concluded that teaching critical discourse analytical tools can lead to improvements in writings of upper-intermediate EFL learners.

Introduction

Since its emergence, critical discourse analysis (henceforward CDA) has attracted the attentions of many linguists and scholars. Advocates of this approach have tried to track down discriminations and language abuses in texts and talks of the people, such as racists, feminists, etc., who try to manipulate thoughts with their words. It is believed that CDA helps people to be able to identify these biases and abuses. As van Dijk (1985) puts it, "CDA is a type of discourse analytical research that primarily studies the way social power abuse, dominance, and inequality are enacted, reproduced, and resisted by text and talk in the social and political context." In addition, several scholars have shown their interests in applying CDA in classroom. One such scholar is Betsy Rymes (2008) with her book *Classroom Discourse Analysis: A Tool for Critical Reflection*. She brings examples in which by analyzing talk in classrooms "teachers were able to use their knowledge of these different language practices as a resource to build mutual, collaborative understandings of the ways stories can be told, questions can be responded to, and problems can be solved" (p. 7). Moreover, *Navas Brenes* (2005) implements some discourse analytical tools which he believes are "useful in analyzing and understanding different segments of an oral narrative". Besides, he concludes that "language teachers can apply different concepts from the area of discourse analysis (e.g., foreground, background, centers of interest, and background knowledge) on the teaching of certain skills such as reading or listening comprehension".

CDA, according to Jorgensen and Phillips (2002), "provides theories and methods for the empirical studies of the relations between discourse and social and cultural developments in different social domains" (p. 60). As a multidisciplinary approach, CDA aims at identifying "ideological biases" (Widdowson, 2007) in discourses and also considering how people might be "manipulated" by the "abuse of power" (Huckin, 2002). To do so, CDA makes use of "discursive and text-level" relations and, in addition, introduces some tools by which "higher-level" and social relations of the discourse can be analyzed (Huckin, 2002).

According to Bloor and Bloor (2007, p.12), a critical discourse analyst tries to find out "the origins of social problems" and find "ways to analyze them productively".

They believe that critical discourse analysts "see discourse both as a product of society and also as a dynamic and changing force that is constantly influencing and re-constructing social practices and values, either positively or negatively". They further explain that a critical discourse analyst may face "macro" or "micro" issues. Put simply, macro issues are the problems that are of "major international importance", whereas micro issues are smaller problems related to "single individuals" (p. 12). Accordingly, one can start to deal with international issues of feminism (Lazar, 2005) and racism (van Dijk, 1986 and 2005; Wodak, 2000a) or investigate a simple case of abuse of language by a lawyer.

CDA has various aspects, each of which can be studied and investigated in details. Bloor and Bloor summarize some of the common goals of the approach as follows (2007, pp. 12-13):

- "To analyze discourse practices that reflect or construct social problems;

- To investigate how ideologies can become frozen in language and find ways to break the ice;

- To increase awareness of how to apply these objectives to specific cases of injustice, prejudice, and misuse of power; …

- To demonstrate the significance of language in the social relations of power;

- To investigate how meaning is created in context;

- To investigate the role of speaker/writer *purpose* and authorial *stance* in the construction of discourse" [original italics].

The present study aims to examine the extent to which teaching critical discourse analytical tools (henceforward CDA tools) helps learners improve their writing skills. However, in this study, in contrast with many other researches in the field of writing, the purpose is not to teach structure, paragraph writing, or essay writing. Besides, it was not attempted to familiarize learners with the processes of writing per se. In fact, researchers tried to exercise a rather new approach to developing learners' writing skills, an approach which is missing in almost all EFL classes in Tehran that deal with teaching writing. The aim of this work, in fact, is adding CDA tools to the process of teaching writing. In so doing, some upper-intermediate participants were selected, and during the period of experiment, the intended CDA tools were taught, practiced, and

reviewed to see whether this treatment could help learners improve their writing skills in English.

Purpose of the study

There are several L2 users or learners who are able to write very professionally and interestingly. These writers usually make the reader go "*wow*" while nodding their heads in approval; we believe they are "*naturals*". Yet, there are other writers who, one might think, *definitely* need to improve their writing skills; their writings just "*do not feel right*". So how can they improve their writings?

   When it comes to teaching writing, a lot of attention is drawn to concepts such as accuracy, choice of words, process and product, free and controlled writing, sentence construction, etc. Different methods have different perspectives towards these concepts, and, very often, teaching these concepts becomes the main theme of writing classes. The main focus of this study, however, is to investigate a different way of improving writing skills of English learners. In fact, the present study aims at quantitatively examining the applicability of CDA in an EFL context. That is, it was assumed that by teaching learners how to implement CDA tools, the researchers might help them improve their writing skills. Thus, the following question was posed as the research question:

Q: Does the teaching of CDA tools to EFL learners have any significant effects on improving their writing skills?

Based on this question, the following null hypothesis was developed:

H: Teaching CDA tools to EFL learners has no significant effects on improving their writing skills.

Theoretical framework

The emergence of discourse analysis dates back to 1960s and early 1970s, when Zelling Harris published the paper "Discourse analysis" (Harris, 1952, cited in McCarthy, 1991):

   "Harris was interested in the distribution of linguistic elements in extended texts, and the links between the text and its social situation, though this paper is a far cry from the discourse analysis we are used to nowadays". (p. 5)

McCarthy also refers to some figures such as Dell Hymes (1960s), Austin (1962), Searle (1969), and Grice (1975; 1989) as the ones who contributed to the "study of language as social action" by their theories, for example, "speech-act theory", "conversational maxims", and "pragmatics" (McCarthy, 1991, pp. 5-6). Wodak (2001) also mentions that in 1070s a new form of discourse started which considered the role of language as a way of developing power-relations in society (Anthonissen, 2001; Wodak, 2001). At that time, several researchers paid attention to "language change" and "communicative interaction", while the notion of "social power" did not seem to have attracted many attentions (Labov, 1972; Hymes, 1972; Wodak, 2001). Later, thanks to the works of the scholars such as Kress and Hodge (1979), Fowler *et al.* (1979), van Dijk (1985), Fairclough (1989), and Wodak (ed.) (1989), critical linguistics was introduced (Wodak, 2001, p. 5). According to Kress (1990), a team of scholars at the University of East Anglia in the 1970s agreed to choose the term *Critical Linguistics*, adapted from social philosophy (Kress, 1990, p. 88, cited in Wodak, 2001, p. 5).

Wodak (2001) states that in January 1991, in a small meeting in Amsterdam and by the support of the University of Amsterdam, some scholars including Teun van Dijk, Norman Fairclough, Gunther Kress, Theo van Leeuwen, and Ruth Wodak had the chance to spend two days together to talk about discourse analysis and CDA. Thus, they had the opportunity to share and discuss their views and theories, an event that can be considered as a turning point in the history of CDA in early 1990s (p. 4). In a search for the contributions to the emergence of CDA, works such as the journal *Discourse and Society* (1990) by van Dijk, some books by Fairclough such as *Language and Power* (1989), a book by van Dijk, *Prejudice in Discourse* (1984), or *Language, Power, and Ideology* by Wodak (1989) can be referred to (Wodak, 2001, p. 4). The symposium held in Amsterdam was the beginning of further "collaborations" between these figures and since then, many other scholars and researchers started to explore and investigate this approach (p. 4). In Wodak's own words, "CDA and CL had existed before, but not as such an international, heterogeneous, closely knit group of scholars" (p. 4). It should also be mentioned that the development of CDA was to a great extent influenced by the works of the scholar Norman Fairclough, especially after publishing his 1995 book by the title *Critical Discourse Analysis* with the subtitle *the critical study of language* (Billing, 2003, p. 35).

Widdowson believes that the main purpose of CDA is to study language for achieving "socio-political power" (2007, p. 70). Discourse analysts usually relate a text to broader issues of "ideology" and "social belief" (p. 70). As Fairclough and Wodak (1997) put it, CDA considers language as "social practice". In this approach, the term *discourse* has a specific meaning which might be a bit different from other genres. According to Widdowson (2007), for the proponents of CDA, "discourses are kinds of genre, institutionalized modes of thinking and social practice, and those who compose texts are taken to be not so much individuals as socially constructed spokespersons or representatives of discourse communities" (p. 70). Therefore, for a critical discourse analyst, the notions of sociological and political relations are considered to be significantly crucial.

Since CDA is a multidisciplinary approach, it can be used in various contexts for various purposes. One context in which CDA can be applied is the context of classroom. Betsy Rymes (2008) refers to four reasons why CDA can be applied to classrooms (p. 5):

1. "Insights gained from classroom discourse analysis have enhanced mutual understanding between teachers and students;
2. by analyzing classroom discourse themselves, teachers have been able to understand local differences in classroom talk—going beyond stereotypes or other cultural generalizations;
3. when teachers analyze discourse in their own classrooms, academic achievement improves; and
4. the process of doing classroom discourse analysis can itself foster an intrinsic and lifelong love for the practice of teaching and its general life-affirming potential."

Rymes believes that by "recording, viewing, reviewing, and analyzing" discourse, the (linguistic) relations between teachers and students will improve. She brings several examples of misunderstandings in classroom discourse between teachers and students and proves how, by analyzing discourse, they can be removed and thus better communications and relations are established. In fact, one possible misunderstandings between teachers and students is that each of them might use different "discourse patterns" which may seem problematic at first place, but can be resolved by analyzing

discourse (Rymes, 2008, p. 12). Rymes also mentions that "in the classroom, context can range from the talk within a lesson, to a student's lifetime of socialization, to the history of the institution of schooling" (p. 12).

CDA tools

CDA introduces several tools which make it possible to analyze discourse, either in written or spoken forms. By analyzing discourse, one can gain a better understanding of social and cultural factors, ideological biases, social and cultural factors, etc. CDA tools include a diverse number of tools and this gives a researcher a choice to choose those tools which are most appropriate for his/her purpose of analysis. For example, these tools can be employed in order to trace embedded ideologies or abuses of language in the intended discourse. Huckin (2002, pp. 7-12) categorizes these tools into the following:

(a) "Word/phrase level concepts" including "classification, connotation, metaphor, lexical presupposition, modality, and register";

(b) "Sentence/utterance level concepts" including "transitivity, deletion, topicalization, register, politeness, presupposition, insinuation, and intertextuality;

(c) "Text level concepts" including "genre, heteroglossia, coherence, framing, extended metaphor, foregrounding/backgrounding, omission, and auxiliary embellishments"; and

(d) "Higher level concepts" including "central processing, peripheral processing, heuristics, reading position, naturalization, cultural models and myths, resistance, hegemony, and ideology".

In this classification, while the first three levels are assumed to be text/discourse-based, the fourth one deals with the way "particular discourses can lead to abuses of power" (p. 11). Below are a number of CDA tools which were chosen for this research:

Connotation: According to *Oxford Advanced Learner's Dictionary* (2000), connotation is "an idea suggested by a word in addition to its main meaning". For example, the word "golf" can connote being rich.

  Metaphor: Oxford Advanced Learner's Dictionary (2000) defines metaphor as "a word or phrase used in an imaginative way to describe sb/sth else, in order to show

that the two things have the same qualities and to make the description more powerful, for example *she has a heart of stone*".

Simile: Oxford Advanced Learner's Dictionary (2000) defines simile as "a word or phrase that compares sth to sth else, using the words *like* or *as*, for example *a face like a mask* or *as white as snow*".

Exclusion: Exclusion refers to the condition in which one piece of information (i.e., the social actor) in a statement is deleted. Through exclusion, one can change the focus of the sentence and thus distract the attention of the reader from the intended part of the sentences. For example, in the sentence "*in Iran concerns are being expressed about the drug dealings occurring near eastern borders",* the "passive agent" of the sentence has been deleted. That is, it tells us that *concerns are being made*, but it is not known *who* expresses them (van Leeuwen, 1996).

Backgrounding is a particular form of exclusion. According to van Leeuwen (1996, p. 41), "backgrounding can result from simple ellipses in non-finite clauses with *-ing* and *-ed* participles, in infinitival clauses with *to,* and in paratactic clauses. In all these cases the excluded social actor is *included* elsewhere in the same clause or clause complex" [original italics].

Nomination: As van Leeuwen (1996, p. 53) puts it, nomination can be identified by "proper nouns", for example by "surname only, with or without honorifics", "given name and surname", and "given name only" (p. 53), such as "*Paul Richards, aged 57, and his wife Susan, 41".* The point, according to van Leeuwen, is that "nameless characters fulfill only passing, functional roles, and do not become points of identification for the reader or listener" (p. 53). Thus, in order to emphasize or de-emphasize the role of one social actor, one can decide to use or deliberately avoid nomination.

van Leeuwen (1996) introduces "pseudo titles", for example, "controversial cancer therapist Milan Brych" (p. 54). In this example, the pseudo title *controversial* is given to Milan Brych by the writer/speaker and does reflect his/her opinion toward that cancer therapist. It should be mentioned that for those who do not know Milan Brych, the pseudo-title "controversial" can affect the way they think of this person.

Functionalization: Functionalization, according to van Leeuwen (1996), "occurs when social actors are referred to in terms of something they do, for instance, an occupation or role". He continues that functionalization can be structured in three ways (p. 54):

(a) By adding suffixes such as "*-er, -ant, -ent, -ian,* and *–ee*" to a verb in order to form a noun. Examples of this kind of functionalization can be: *employer*, *contestant*, "*correspondent*", "*guardian*", and *employee*.

(b) By adding suffixes such as "*-ist* and *-eer*" to another noun "which denotes a place or tool closely associated with an activity (a noun which forms the 'Range' of that activity)" such as *violinist* and "*mountaineer*".

(c) By adding nouns such as "*man*, *woman*, *person*, and *people*" to another noun and make a "compound" noun, for example *policeman*, *salesperson*, and *businesswoman*.

Physical identification: This is how "physical" features of a social actor become a way of referring to him/her, for example "brunette", "cripple", or "fat" (van Leeuwen, 1996, p. 57). In fact, physical identification gives a "unique identity" to "social actors" when there is no nomination in the discourse (p. 57). Moreover, it is a way of "selectively" attracting the readers/listeners' attention to some particular features of the "social actors" (p. 57).

Metonymy: According to George Yule (1996), metonymy is a "type of relationship between words based simply on a close connection in every day experience", for example, "bottle" can represent "beer" or "roof" can be a symbol of "house" (p. 122). Metonymy can be used to *de*-emphasize the role of the social actor. For example, in "Tehran accepted to apologize", it may not be easy to understand whether by *Tehran* the author meant *the President*, *the Leader*, or *the Senate*.

Indetermination and differentiation: van Leeuwen (1996, pp. 51-52) explains that indetermination is the act of making a social actor "anonymous" or "unspecified". In fact, indetermination backgrounds the role of the social actor. Indetermination is usually formed "by indefinite pronouns ('somebody', 'someone', 'some', 'some people') used in nominal functions", such as "Some people believe that UFOs really exist" (p. 51).

Differentiation helps us make a distinction between "an individual social actor or group of social actors from a similar actor or group, creating the difference between the 'self' and the 'other, or between 'us' and 'them', as with 'others' in:
And though many of the new migrants are educated high-achievers from places like Singapore and Hong Kong – 'uptown' people in American terminology – others are 'downtown' people from places like Vietnam, the Philippines, and Lebanon" (van Leeuwen, 1996, p. 52).

Anaphora: Anaphora is a rhetorical technique through which one can strongly attract the attention of the reader/listener to some intended parts of the discourse. Harris (2008, online) defines anaphora as "the repetition of the same word or words at the beginning of successive phrases, clauses, or sentences…", for example, "Slowly and grimly they advanced, not knowing what lay ahead, not knowing what they would find at the top of the hill, not knowing that they were so near to Disneyland". Anaphora is a way of emphasizing one part of discourse. Another example can be "not for money, not for promotion, not for public admiration, and not for anyone, he did it for the sake of love" (Harris, 2008).

Epistrophe: Epistrophe is the reverse of anaphora. That is, in anaphora a word of phrase is repeated at the beginning of a sentence, while in epistrophe a word or a phrase is repeated "at the end of successive phrases, clauses, or sentences", for example, "You will find washing beakers helpful in passing this course, using the gas chromatograph desirable for passing this course, and studying hours on end essential to passing this course" (Harris, 2008).

Epanalepsis: Epanalepsis is a combination of anaphora and epistrophe. Here you should end the sentence by the same term the sentence was started. For example, "Water alone dug this giant canyon; yes, just plain water" or "Our eyes saw it, but we could not believe our eyes" (Harris, 2008).

Hypophora: In order to draw the attention of the readers to a specific point, it is possible to "raise one or more questions" and then try to answer them (Harris, 2008).This technique is called hypophora. It is usually used to start a paragraph where the rest of the paragraph is spent providing the answer to the raised question(s).

Rhetorical question: Rhetorical question is another way of raising questions in writing. However, Harris (2008) explains that it is different from hypophora in a way that the rhetorical question does not need an answer by the writer because "the answer is obvious or obviously desired". One example of rhetorical question can be "But how can we expect to enjoy the scenery when the scenery consists entirely of garish billboards?" (Harris 2008).

Writing

When it comes to teaching writing, a lot of attention is drawn to concepts such as organization, structure, cohesion, coherence, vocabulary, etc. Different methods have different perspectives towards these concepts, and, so far, a lot of researches have

been conducted in these areas and several books and papers have been published (see Howatt 1984; Robb *et. al.* 1986; Ishikawa 1995; Victori 1999; Sasaki 2000; Carter and Nunan 2001; Richards and Rodgers 2001; Huckin 2002; Pan 2002).

In this research, however, it has been attempted to test the effect of CDA tools on the improvement of writing skills of upper-intermediate EFL learners. The reason for choosing upper-intermediate learners was that they were assumed to be able to communicate in L2 and to have an acceptable level of accuracy and perhaps it was time to implement some CDA tools in their writing so that they were able to make a better impression on the readers than before. Therefore, in this study, the concept of writing was viewed from a different perspective, that is, from the eyes of a critical discourse analyst. In fact, some tools extracted from CDA were taken, hoping to be able to help the learners practice them and thus, improve their writing skills in L2.

Methods

Participants

The participants of this study included 20 upper-intermediate EFL learners at Valiasr English Institute in Tehran which were chosen by means of a standardized TOEFL test (see Phillips, 2001, p. 538). They included 17 female and three male learners, all of whom were between 17 and 25 years old. All the participants had at least two years of English learning background at different English institutes and two of them were also studying English in college as freshmen. In order to select upper-intermediate participants, a TOEFL test was given to 32 learners who were willing to take part in the study, but only the learners whose scores were above 450 (out of 677) in the TOEFL test were qualified for the treatment and others were ruled out. The purpose of giving this test was to identify the learners with the intended level of proficiency, namely upper-intermediate. In addition, the homogeneity of the subjects was assured by excluding those learners whose scores were not within one standard deviation below or above the mean. Thus, 12 learners were excluded from the study and the experiment began with 20 participants.

Procedure

The present study was conducted in Valiasr English Institute in Tehran, Iran. First, a standardized TOEFL test was given to a group of learners in order to pick those with

gain scores of more than 450. Then, from the learners whose scores were above 450, 20 participants with scores falling within one standard deviation above or below the mean were selected as the sample of the study. Next, a pre-test (see Appendix), taken from the "Longman Complete Course for the TOEFL test: Preparation for the computer and paper tests" by Phillips (2001, p. 538), was given to the sample in order to pinpoint their current levels of writing skills. Two raters evaluated the performances of the participants on the pre-test. The scoring criteria based on which the writing performances of the participants were evaluated were taken, with a little change, from the scoring criteria of the TOEFL iBT writing test of the Longman Preparation Course for the TOEFL Test: iBT, Second Edition (2008, pp. 101-102). That is, in scoring a TOEFL iBT writing test, the score for each task ranges from 0 to 5, but in this study, the scoring procedure ranges from 10 to 50 (see below) for the writing. Another difference between the current scoring criteria and TOEFL iBT is the addition of four more criteria, namely cohesion/coherence, persuasiveness, reasoning, and critical discourse analytical tools. The following table shows the criteria used to score the writing performances of the participants on the pre-test and post-test.

Table1. Scoring criteria of pre-test and post-test

	Excellent	Good	Not bad	Poor	Awful
Answer to question (relevance)	5	4	3	2	1
Comprehensibility	5	4	3	2	1
Organization	5	4	3	2	1
Flow of ideas	5	4	3	2	1
Cohesion and coherence	5	4	3	2	1
Grammar	5	4	3	2	1
Vocabulary	5	4	3	2	1
Persuasiveness	5	4	3	2	1
Reasoning	5	4	3	2	1
Discourse Analytical tools	5	4	3	2	1

After the pre-test, the experiment began. The participants were asked to attend 90-minute classes in the afternoon on even days. In the classes, CDA tools were taught to the participants and they were trained how to use them in their writings. During the

experiment, there was no training in writing skills of the participants except for the teaching of the intended CDA tools. After sixteen sessions of training, the post-test (see Appendix) was given to the participants to determine their probable gains. As with the pre-test, the post-test was also taken from the "Longman Complete Course for the TOEFL test: Preparation for the computer and paper tests" by Phillips (2001, p. 486). Again, two raters assessed the performances of the participants on the post-test.

During the experiment, the validity of the study was attentively being taken into consideration to make sure that the results are valid. In so doing, the researchers attempted to control and remove the possibilities of intervening variables, those that were potentially a threat to the truthfulness of the results based on the treatment. Thus, the experiment was extremely limited to the teaching of CDA tools and the researchers fully attempted to avoid teaching other skills or sub-skills of writing such as grammar, vocabulary, sentence structure, etc. Moreover, the participants were asked not to attend any other English learning courses at the time of the experiment so that the role of out-of-classroom learning was kept low. In addition, the reliability of the scores of the participants on both pre- and post-tests were calculated and reported in the next section.

Results

The first instrument used in this study was a TOEFL test, resulting in the selection of 20 participants. The descriptive data of this proficiency test for the 32 learners who took the test were calculated and summarized in the following table:

Table 2. Descriptive statistics of participants' scores on TOEFL test

Descriptive Statistics

	N	Range	Mean	Std. Deviation	Variance
Participants	32	65.00	476.8750	15.17373	230.242
Valid N (listwise)	32				

After the selection of the participants, the pre-test was administered. The following table shows the descriptive statistics of the participants' performances on the pre-test:

Table 3. Descriptive statistics of the pre-test

Descriptive Statistics

	N	Range	Mean	Std. Deviation	Variance
Pretest	20	16.00	28.4500	3.84537	14.787
Valid N (listwise)	20				

As it is shown in the above table, the mean score of the participants on the pre-test is 28.45 (out of 50). Besides, the correlation coefficients of the ratings were also calculated and determined to ensure the consistency of the scores:

Table 4. Correlation and reliability of scores on pre-test

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.864	.864	2

Inter-item correlation matrix

		Rater1	Rater2
Rater1	Pearson Correlation	1	.761**
	Sig. (2-tailed)		.000
	N	20	20
Rater2	Pearson Correlation	.761**	1
	Sig. (2-tailed)	.000	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Then, the descriptive statistics of the post-test was calculated and summarized in the following table:

Table 5. Descriptive statistics of the post-test

Descriptive Statistics

	N	Range	Mean	Std. Deviation	Variance
Posttest	20	16.00	39.0000	3.83886	14.737
Valid N (listwise)	20				

The descriptive statistics of the post-test shows that the mean score of the participants on the post-test is 39.00, which shows an improvement of the scores on the post-test, in comparison with the pre-test. In addition, the correlation coefficient of the ratings of the two raters of the post-test was calculated and summarized in the following table:

Table 6. Correlation and reliability of scores on post-test

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.923	.926	2

Correlations

		Rater1	Rater2
Rater1	Pearson Correlation	1	.861**
	Sig. (2-tailed)		.000
	N	20	20
Rater2	Pearson Correlation	.861**	1
	Sig. (2-tailed)	.000	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Comparing the means of the scores on the pre-test and post-test, it is easy to subjectively conclude that teaching critical discourse analytical tools to the learners helped them improve their writing skills. However, in order to quantify the findings of the study, a paired-sample t-test was administered to compare the performances of the participants on the pre-test and the post-test. The significance level was set at 0.05.

Table7. Paired-sample t-test of pre- and post-tests

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair	Pretest	28.4500	20	3.84537	.85985
	Posttest	39.0000	20	3.83886	.85840

Paired Samples Correlations

		N	Correlation	Sig.
Pair	Pretest & Posttest	20	.980	.000

Paired Samples Test

		Paired Differences			95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair	Posttest - Pretest	10.55000	.75915	.16975	10.90530	10.19470	62.149	19	.000

As it was shown in the table7, the mean score of the participants on the pre-test was 28.45, while in the post-test this score was significantly higher, i.e., 39.00. According to the t-test, the difference between the mean scores of the pre-test and post-test in the above-mentioned analyses is significant. In other words, it can be concluded from the analyses that the treatment was quite efficient and that CDA tools practically improved the writing skills of the participants. Therefore, the increase in the mean score of the participants can be attributed to the independent variable of the research, i.e., CDA tools. Consequently, the null hypothesis of the research was rejected.

Discussion

As mentioned earlier, the null hypothesis claimed that teaching CDA tools to EFL learners has no significant effects on improving their writing skills. However, this null

hypothesis was rejected at 0.05 level of significance. Therefore, it can be concluded that teaching CDA tools *did* improve the writing skills of the learners. In other words, the analyses of the result showed that the treatment was quite efficient.

This study aimed at proposing a way to improve the writing skills of upper-intermediate EFL learners. However, in this study, the participants of the study had a more or less good command of English, since they were able to gain scores higher than 450 in a TOEFL test. Besides, they were able to comprehensibly communicate in L2 and they could also write about different issues, using English language. Thus, while the research was being conducted, the only focus of the class was on learning and practicing CDA tools and on how to use them in different discourses.

The reason why CDA tools were selected was that the learners seemed to be unfamiliar with the role of persuasive discourse and rhetoric and the importance of CDA in various contexts. Put simply, in their writings, learners mostly used the simplest forms of language in order to convey their messages. Therefore, it was tried to make them familiar with CDA tools and teach them how to use these tools in their writings in order to be able to make a better impression on the readers. For example, they learned how to hide some pieces of information in the context, and how to exclude the social actor(s) of a sentence without violating grammatical rules. They also understood how to direct the attention of the readers to specific pieces of information, while distracting them from the backgrounded parts. Moreover, they were taught how to use some figures of speech such as simile, metaphor, and metonymy, they studied and practiced some rhetorical devices such as anaphora, epistrophe, metabasis, hypophora, and procatalepsis and they became familiar with persuasive discourse.

After the participants began to use CDA tools, they seemed to have improved in their writing skills. According to the results of the post-test, the participants performed better in comparison with their performances on the pre-test. Consequently, it might be a good idea to start to teach these tools to upper-intermediate and advanced learners of English in order to help them improve their writing skills.

Another advantage of familiarizing learners with critical discourse analytical tools is that as learn how to use the techniques of this approach, their awareness of language and ideological bias increases (Widdowson, 2007) and they learn how easy it might be to be manipulated by language. In other words, when a learner learns how

to use language to influence the readers or how to background, foreground, or exclude information in a discourse, he or she is less likely to be under the control of discourse. That is, he or she learns to be critical. In addition, as Jorgensen and Phillips (2002, p. 60) point out, CDA enables researchers to study "the relations between discourse and social and cultural developments in different social domains". This research, too, attempted to bring CDA to an academic context in order to help the learners improve their writing skills in L2. Moreover, the tools introduced and categorized by Huckin (2002) which refer to "discursive and text level" and also "higher-level" relations of discourse were implemented in the study and it was noticed that they can also be used in an EFL classroom to help learners be better writers.

Conclusion

In summary, it was shown that the writing skills of upper-intermediate EFL learners improved after they practiced CDA tools. Huckin (2002) divides CDA tools into four categories of "word/phrase level", "sentence/utterance level", "text level", and "higher level" concepts. Each level contains several tools with different functions. In this study, some of these tools were selected to be practiced and implemented by EFL learners which helped them perform better in their writings. The present findings suggest some implications for researchers, teachers, learners, and curriculum developers in the realm of EFL teaching and learning. In fact, the purpose of the study was to introduce an alternative way of improving writing skills of learners in writing classes. It means that while, in a typical English class in Iran, teachers and learners work on structure, vocabulary, style, cohesion, coherence, etc., it might be a good idea to start to teach CDA tools to learners to help them enhance their writing skills. In addition, the results of this study can be used by curriculum developers in order to provide teachers and learners with information about CDA and CDA tools. It goes without saying that more research in this area is needed to be conducted to help introduce and elaborate the theoretical implications and practical applications of CDA.

References

Anthonissen, C. (2001). *On the effectiveness of media censorship: linguistic, paralinguistic and other communicative devices of media regulation*. PhD thesis, University of Vienna.

Austin, J. L. (1962). *How to Do Things with Words*. Oxford: Oxford University Press. Billing, M. (2003). 'Critical Discourse Analysis and the Rhetoric of Critique', in Weiss, G. and Wodak, R. (2003). *Critical discourse analysis: theory and interdisciplinarity*. Hampshire: Palgrave Macmillan.

Bloor, M. and Bloor, T. (2007). *The Practice of Critical Discourse Analysis: An Introduction.* London: Hodder Education.

Carter, R. and Nunan, D. (2001). *The Cambridge Guide to Teaching English to Speakers of Other Languages*. Cambridge: Cambridge University Press.

Fairclough, N. (1989). *Language and Power*. London: Longman.

Fairclough, N. (1995). *Critical Discourse Analysis*. London: Longman.

Fairclough, N. and Wodak, R. (1997). 'Critical discourse analysis', in T. van Dijk (ed.), *Discourse Studies: A Multidisciplinary Introduction*. Volume 2. London: Sage, pp. 258-84.

Fowler, R., Hodge, G., Kress, G. and Trew, T. (eds) (1979). *Language and Control*. London: Routledge and Kegan Paul.

Grice, H. P. (1975). 'Logic and conversation', in Cole, P. and Morgan, J. (Eds.) *Syntax and Semantics, Volume 9: Pragmatics*. New York: Academic Press.

Grice, H. P. (1989). *Studies in the Way of Words*. Harvard University Press.

Harris, R. (2008). 'A Handbook of Rhetorical Devices.' *VirtualSalt* (online). http://www.virtualsalt.com/rhetoric.htm

Harris, Z. (1952). *Discourse Analysis*. Language, 28, 1-30.

Howatt, A. P. R. (1984). *A History of English Language teaching*. Oxford: Oxford University Press.

Huckin, T. (2002). 'Critical Discourse Analysis and the Discourse of Condescension'. Chapter 7 in *Discourse Studies in Composition*, ed. E. Barton and G. Stygall. Hampton, (2002).

Hymes, D. (1964). *Towards ethnographies of communication*. In J. J. Gumperz and D. Hymes (Eds), *The Ethnography of Communication. American Anthropologist*, 66(6), 1-34.

Hymes, D. (1972). 'Models of interaction of language and social life', in J.J. Gumperz and D. Hymes (eds), *Directions in Sociolinguistics - The Ethnography of Communication*. New York: Holt, Rinehart and Winston, pp. 35-71.

Ishikawa, S. (1995). 'Objective measurement of low-proficiency EFL narrative writing'. *Journal of Second Language Writing*, Volume 4, Issue 1, January 1995, Pages 51-69. Published by Elsevier Science Inc.

Jorgensen, M. and Phillips, L. (2002). *Discourse Analysis as Theory and Method*. London: SAGE Publications Ltd.

Kress, G. (1990). *Critical discourse analysis*, Annual Review of Applied Linguistics, 11: 84-97.

Kress, G. and Hodge, B. (1979). *Language as Ideology*. London: Routledge.

Labov, W. (1972). *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.

Lazar, M. (ed.) (2005). *Feminist Critical Discourse Analysis. Gender, Power and Ideology in Discourse.* Houndsmills, UK: Palgrave MacMillan.

McCarthy, M. (1991). *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.

Navas Brenes, C. A. (2005). Analyzing an Oral Narrative Using Discourse Analysis Tools: Observing How Spoken Language Works. *Revista Electrónica "Actualidades Investigativas en Educación".* Volumen 5, Número 1, Año 2005.

Oxford Advanced Learner's Dictionary, (2000). Oxford: Oxford University Press.

Pan, S. H. (2002). 'A Study of the Effects of Summary Writing with Structure Guidelines on the Writing of EFL Beginning Writers'. (Master's Thesis). http://ethesys.lib.ncku.edu.tw/ETD-db/ETD-search/view_etd?URN=etd-0709102-171214

Phillips, D. (2001). *Longman Complete Course for the TOEFL test: Preparation for the computer and paper tests*. New York: Pearson Education, Inc.

Phillips, D. (2008). *Longman Preparation Course for the TOEFL Test: iBT*. New York: Pearson Education, Inc.

Reid, J. (2001). Writing. In Carter, R. and Nunan, D. 2001. *The Cambridge Guide to Teaching English to Speakers of Other Languages*. Cambridge: Cambridge University Press.

Richards, J. C. and Rodgers, T. S. (2001). *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press.

Robb, T., Ross, S., and Shortreed, I. (1986). 'Salience of feedback on error and its effect on EFL writing quality'. *TESOL quarterly*, Vol. 20, No. 1.

Rymes, B. (in press) (2008). *Classroom Discourse Analysis: A Tool for Critical Reflection.* Cresskill, NJ: Hampton Press.

Sasaki, M. (2000).'Toward an Empirical Model of EFL Writing Processes: An Exploratory Study'. *Journal of Second Language Writing*, Volume 9, Issue 3, September 2000, Pages 259-291. Published by Elsevier Science Inc.

Searle, J. R. (1969). *Speech Acts*. Cambridge: Cambridge University Press.

van Dijk, T. A. (1984). *Prejudice in Discourse: An Analysis of Ethnic Prejudice in Cognition and Conversation*. Amsterdam: Benjamins.

van Dijk, T. A. (ed.) (1985). *Handbook of Discourse Analysis* (4 vols.). New York: Academic Press.

van Dijk, T. A. (1986). *Racism in the Press*. London: Arnold.

van Dijk, T. A. (2005). *Discourse and racism in Spain and Latin America*. Amsterdam: Benjamins.

van Leeuwen, T. (1996). 'The representation of social actors', in Caldas-Coulthard, C.

Victori, M. (1999). 'An analysis of writing knowledge in EFL composing: a case study of two effective and two less effective writers'. *System*, Volume 27, Issue 4, December 1999, Pages 537-555 . Published by Elsevier Science Ltd.

Widdowson, H. G. (2007). *Discourse Analysis*. Oxford: Oxford University Press.

Wodak, R. (1989). 'Introduction', in R. Wodak (ed.), *Language, Power and Ideology*. Amsterdam: Benjamins, pp. i-ix.

Wodak, R. (2000a). The rise of racism - an Austrian or a European phenomenon?. *Discourse and Society*, 11 (1): 5-6.

Wodak, R. (2001). 'The discourse -- historical approach', in Wodak, R. and Meyer, M. 2001. *Methods of Critical Discourse Analysis*. London: SAGE Publications Ltd.

Wodak, R. (2001). 'What CDA is about -- a summary of its history, important concepts and its developments', in Wodak, R. and Meyer, M. 2001. *Methods of Critical Discourse Analysis*. London: SAGE Publications Ltd.

Yule, G. (ed.) (1996). *The Study of Language*. Cambridge: Cambridge University Press.

APPENDIX

Appendix: Copy of the Pre-test and the Post-test
Pre-test
Do you agree or disagree with the following statement?
   People should never live at home with their parents after the age of twenty-five.
Use specific reasons and details to support your answer.

Post-test
Do you agree or disagree with the following statement?
   It is better to save your money for the future than to enjoy it now.
Use specific reasons and examples to support your answer.

Title

Foreign Language Learners' Processing of Relative Clause Ambiguity: A Relevance Theory Perspective

Authors

Sima Khezrlou (M.A.)
University of Tehran, Iran

Ali Akbar Khomeijani Farahani (Ph.D.)
University of Tehran, Iran

Fateme Layeghi (M.A.)
University of Tehran, Iran

Bio Data

Sima Khezrlou, M.A. in TEFL from the University of Tehran, Tehran, Iran. Her major academic interests include SLA issues and psycholinguistics. She is currently involved in teaching English at foreign language teaching centers.

Ali Akbar Khomeijani Farahani, assistant professor of English and Linguistics in the English department of University of Tehran, Tehran, Iran. His research interests include Discourse Analysis and Systematic Functional Linguistics. He has taught extensively in these areas in MA and PhD levels.

Fateme Layeghi, M.A. in TEFL from the University of Tehran, Tehran, Iran. Her major academic interests include ELT and SLA.

Abstract

This study explored the way adult Persian-speaking foreign language learners of English attempted to resolve ambiguities of relative clause type. Two groups of advanced (n=100) and intermediate (n=90) English female and male learners with Persian as their first language tried to resolve ambiguities in an off-line task. The study aimed to explore the learners' disambiguation of relative clause sentences linked by either *of* or *with* prepositions from relevance theory perspective. The results show that learners' attachment preferences were largely influenced by the contextual effect and processing effort principles of relevance theory that have been claimed to influence sentence processing. The article also demonstrates that although the learners' disambiguation

preferences were influenced by the phrase-structure based locality principles (recency or predicate proximity) for the ambiguous sentences linked by *of* preposition, there was no evidence that they applied any phrase structure–based ambiguity resolution strategies of the same type to the interpretation of complex genitives linked by *with* antecedent. The Persian-speaking participants of different level of proficiency and gender differed significantly in terms of their disambiguation preferences. Results are further discussed.

Keywords: Relevance theory, Disambiguation, Relative clause ambiguity, Contextual condition, Non-contextual condition.

Introduction
Relevance theory, originally proposed by Sperber and Wilson (1986) as a cognitive approach to communication, is based on the assumption that human cognition is relevance-based: "we pay attention to information that appears relevant to us, construct relevant representations of such information, and process these representations in a context that maximize its relevance" (Ying, 2005, p. 556). The theory has been studied by researches in the areas of interlanguage pragmatics of speech-act production and comprehension and the psycholinguistic studies of language understanding and production (Foster-Cohen, 2000). Foster-Cohen (2000) refers to the scarcity of research in relevance theory with only few attempts to investigate this area (e.g., Ying, 2004; Ying, 2005). Although numerous studies exist that have examined the way L2 learners disambiguate relative clause attachment ambiguities, there has been no study to investigate the relative clause attachment disambiguation from the perspective of relevance theory. Therefore, the current study has been undertaken in an effort to bridge this gap. The study aims to investigate the foreign language learners' interpretation of the sentences with relative clause attachment ambiguities in which the ambiguity concerns the attachment of the relative clause. There are two types of ambiguous relative clause sentences, with the two noun phrases in the NP complex being joined either by functional *of* or by attributive *with*, as illustrated by examples 1a and 1b, respectively.
1. a. The dean liked the secretary *of* the professor who was reading a letter.
    b. The dean liked the professor *with* the secretary who was reading a letter.
 Each sentence allows two possible interpretations with a relative clause modifying either the head of the overall object NP (= NP1attachment) or the embedded noun phrase (= NP2 attachment).
  This type of ambiguity raises significant issues in sentence processing including the relationship between syntax and semantics and the psycholinguistically-based parsing tendencies in the disambiguation of ambiguous sentences. These issues are discussed as the background of the present study.

The relationship between syntax and semantics in sentence interpretation

A significant issue related to the parsing of ambiguous sentences is the interaction of the effects of syntax and semantics. Some buttress the relevance of semantics and syntax during sentence interpretation and some assume the autonomy of these two. Rayner, Carlson, and Frazier (1983) touched the issue in this way:

The central issue in the debate between autonomy and interaction in processing is not whether all the relevant types of information can be exploited at some point in the comprehension of a linguistically conveyed message. Rather, it concerns how the distinct types of information are utilized. (p. 359)

However, intermediate positions in between these two extremes of strong autonomy and interaction or integration are also possible which were tested by Rayner, Carlson, and Frazier (1983) in two experiments.

MacDonald (1993) conducted two experiments in order to investigate the interaction of lexical and syntactic ambiguity in the resolution of lexical category ambiguities. The goal of the first experiment was to investigate the role of delay strategy in which semantic bias is delayed in its impacts on the ambiguity resolution. The aim of the second experiment was to investigate the effect of semantic information on ambiguity resolution. Using the self-paced reading task in experiment 1, reading times for sentences containing lexical category ambiguities and their unambiguous counterparts were measured. Experiment 2 had two parts: first, reading times in two bias and two ambiguity conditions were measured, and then reading times were evaluated in the light of normative data (MacDonald, 1993).The results of the first experiment were against the delay strategy in the lexical category ambiguity resolution. In fact, the pattern of reading times taken to support the delay strategy had other explanation; they were due to added differences across ambiguity conditions. The findings of the second experiment indicated the influence of semantic information which appeared early in the lexical category ambiguity resolution. In other words, findings supported the constraint-based model in which alternative interpretations are "partially activated and rapidly constrained by probabilistic information such as the relative frequency of the alternative interpretations, combinatorial semantic information and word concurrence information"(MacDonald, 1993, p.711). As it is evident from the results of this study, constraint based models provide an opportunity to unify accounts of lexical and syntactic processing and to explore a number of subtle interacting constraints such as lexical, syntactic, and discourse level constraints.

Psycholinguistic theories of ambiguity interpretation
There have been several attempts to describe the parsing tendencies in the disambiguation of ambiguous sentences. One approach has been termed 'the minimal attachment principle" (Frazier & Fodor, 1978). According to Frazier and Fodor (1978) "each lexical item is to be attached into the phrase marker with the fewest possible number of nonterminal nodes linking it with the nodes which are already present" (p. 320).

The second approach to the parsing of ambiguous sentences is the theory of referential context proposed by Altmann and Steedman (1988). According to Altmann

and Steedman (1988), the principle of referential support can be considered as an interpretation which is referentially supported and would be favored over one that is not. Irrespective of a psycholinguistics tradition of accounting for the resolution of syntactic ambiguities by purely structural criteria rather than by any appeal to semantics, Altmann and Steedman (1988) argue that contextual cues can influence the parsing decisions. In other words:

If interpretation and even referents are available at every turn in sentence processing, then it is clearly possible that they play some part in resolving the local syntactic ambiguities that are so surprisingly abundant in natural language sentences, by selecting among analyses according to their appropriateness to the context of utterance. (p.192)

Relevance theory

Relevance theory is a theory of mind and cognition which considers the nature of human communication from two perspectives: the contextual effect and the processing effort. The more the contextual support for any particular sentence, the better the understanding of that sentence. The notion of processing effort rests on the assumption that humans perceive the first meaning of the sentence that comes to their mind and they are reluctant to spend more effort investigating other meanings.

Foster-Cohen (2000) in a review article has focused on Sperber and Wilson's (1995) ideas on the issue of relevance theory. Foster-Cohen (2000) considered angels of relevance theory in terms of communication, inference, relevance, and aspects of verbal communication. The role of inferential processes on contextualized linguistic expressions that result in a likely interpretation has been emphasized (Foster-Cohen, 2000). This view toward the role of inferential processes stems from Sperber and Wilson's (1995) claim that human communication is characterized by its being prone to failure, what they called "non-demonstrativeness".

Based on Sperber and Wilson's argumentation that inferencing is not merely a heuristic activity but an activity based on a deductive operation, Foster-Cohen defines the rudimentary operation of the deductive device as:

The assumptions (defined as structured conceptual representations) that have been formed on the basis of perception, of linguistic decoding, from encyclopedic memory or as a result of a prior deductive process and are used as premises in the deduction of new assumptions. (p. 81)

Therefore, what the deductive device, according to Foster-Cohen (2000), does is "to improve the individual's representation of the world" (p. 81).

The third perspective, i.e. the relative nature of the relevance theory, has been best approached by Foster-Cohen (2000):

Relevance is argued to be a relative notion (an utterance may be more or less relevant) in proportion to two factors. First, in relation to the contextual effects it has. An utterance may, therefore, be irrelevant (and communication will thereby fail) if the new information conveyed by the utterance fails to connect with anything in the (cognitive) context, is simply repetitive of existing assumptions, or is inconsistent with existing assumptions but is too weak to overthrow them. Second, the relevance

of an utterance is in proportion to the effort required to derive the contextual effects. An utterance may be irrelevant, even if it has significant contextual effects, if it takes too much effort to derive them. (pp. 82-83)

The above explanation of relevance by Foster-Cohen (2000) is in parallel with the psycholinguistic account of the disambiguation of the ambiguity because of the fact that "your system will not put in any more effort than it can help, you will then stop at that first interpretation, unless there is something else urging you on" (p. 83).

Finally, Foster-Cohen (2000) includes the last part of Sperber and Wilson's (1995) book on relevance theory which is centers on the issue of utterance interpretation, a process of top-down nature.

Yus (2002) refers to relevance theory as a movement toward the revision of some of the features of Grice's (1975) conversational theory. According to Yus (2002):
Sperber and Wilson also inherited certain aspects of Grice's theory such as the importance of manifest underlying intentions in communication and the difference between *what is said* and *implicatures* (developed into the not quite equivalent dichotomy *explicatures-implicatures*, see Carston, 2001). However, unlike Grice's (1975) *cooperative principle*, which speakers voluntarily follow or disobey Sperber and Wilson's *principle of relevance* (the fact that utterances communicate a presumption of being relevant to the hearer) is spontaneous and biologically rooted in human cognition. (p. 2)

Relevance theory and pragmatics
According to Rumos (1998), Sperber and Wilson's (1986) relevance theory was first introduced to the field of pragmatics in 1986. That was an attempt to analyze the media text in which there was not a face-to-face interaction between both sides of the communication. Later on, Rumos (1998) used this theory to provide a model in interpretation of media discourse. Despite some opposition (e.g., Clark, 1987) that doubts the applicability of this theory to media, Rumos attempted not only to apply it, but also to provide a model. He claims that interpretation from the pragmatic perspective requires considering two kinds of situations: the one in which the speaker's intention accords with the hearer's interpretation, the desired one, and the one in which the needed interpretation is not possible, due to interruption. Randall (1988) asserts that in contexts where the lack of face-to-face interaction results in a great degree of misunderstanding, the main reliable source is the text and its linguistic form. Rumos believes that the hearer should make a balance between the contextual effects and the processing effort and select the first proposition that seems to adapt these two necessary conditions. He believes in order to propose a pragmatic model for RT to be used in discourse four factors should be taken into account: (1) whether the speaker and hearer's communication is direct, (2) whether this communication is direct or implied, (3) whether this is deliberate, and (4) whether the hearer has chosen the intended intention. Sperber, Cara, and Girotto (1995) signify the value of pragmatic process involved in recognizing where the relevance exists.

Wilson (1994) belies that relevance theory has four assumptions with regard to the possible interpretations of the text: (1) the linguistic context allows various

interpretations of the text, (2) all of these interpretations are not within the access of the hearer at the same time, (3) hearers have a criterion for evaluating the incoming interpretations, and (4) this criterion helps the hearer to expunge other possible interpretations in favor of the most desired one. The essence of relevance theory is that the addressee should decide on the best interpretation of what s/he receives during communication. This interpretation should be in accordance with the speaker's intention. That is the proposition the speaker tries to convey through the linguistic codes. Besides, the hearer should apply the least processing effort in doing so. This processing effort is exerted in understanding the linguistic codes and comparing the likely and possible interpretations and finally choosing the closest to the real intention. Sperber and Wilson (1986) defined the conditions for relevance theory: maximal amount of contextual effect should be supplied with the text and it should also have the text with less processing challenge as possible. The first interpretation that accords with these two is accepted at the expense of expunging the others. Evans (1994) uses the relevance theory to name his Selection Task. In this regard, he considers relevance as both linguistic and contextual clues. According to Sperber et al. (1995), the learners' performances in the Selection Tasks are affected by expectations of relevance raised by both the context and content of the rule. They claim that relevance theory, as a mental process, acts as a guide for learners to select relevant information in doing these tasks. What they mean by content relevance is related to the interpretation of tasks. Sperber et al. (1995) summarize Sperber and Wilson's (1986) relevance theory as a process in which the linguistic codes and context alone cannot lead to a meaningful interpretation. Instead, it is a cognitive process in which the previously confirmed information in mind helps the learner to relate the new coming information to it and figure out the possibility of some relations that lead to the proper proposition. This cognitive process engages processing effort which itself leads to a negative effect on the interpretation. Human beings pay attention to the relevant information and attempt to make a context in which this relevance is increased. Therefore, those stimuli that provide this relevance are chosen first to be processed. Sperber et al. (1995) make a distinction between communicated and ordinary information. They claim that the former bears a relative amount of relevance within its context as it is this context that makes the communication possible and the latter may not be relevant. According to Sperber et al. (1995) the relative amount of relevance has two aspects: side and effort. What side presumes is the sufficient amount of relevance that can be a good reason for the required effort. The effort aspect requires the least amount of effort on both communicators' sides.

Prior studies of ambiguity
In a study, Ying (1996) investigated the interpretation of the ambiguously attached PPs by adult L2 learners. Ying's study investigated the interpretation of ambiguous sentences in terms of minimal attachment principle (Frazier & Fodor, 1978) and referential context (Altmann & Steedman, 1988). Ying conducted four experiments in which the first experiment tested the effect of minimal attachment principle, the second tested the effect of referential context, and the third and fourth experiments

tested the effect of lexical preference in the interpretation of ambiguous sentences with a difference that experiment 3 utilized "action verbs" but experiment 4 utilized "psych and perception verbs". Ying reached the conclusion that the lexical preferences were the same for experiments 3 and 4 except for the fact that the effect of lexical information was not so much strong for the psych and perception verbs as for the action verbs.

Harley et al. (1995) investigated the interpretation of ambiguous sentences by participants of different ages. Harley et al. (1995) presented sentences with prosodic contours that were in conflict with the syntax of the sentence. The examples below show these sentences:

1. Our dogs bark at the neighborhood cat.
2. Our dogs bark sometimes frightens people.

The results of their study revealed that despite their expectations of the superior performance by younger learners, the older participants had the same performance in following the prosodic cues in the ambiguous sentences. The findings of the study by Harley et al. (1995) have been subjected to criticism. One major shortcoming of the study, according to Ying (1996), was that the participants were surprised to hear after the experiment that the experimental sentences were ambiguous. Yet another drawback of Harley et al.'s study was the vagueness of the relation between grammar and parsing (pp. 687-688).

Ying (2004) investigated the interpretation of structurally ambiguous sentences from the perspective of relevance theory. Ying tested these two interpretations in two types of experiments: 1) the ambiguous sentences were provided to the learners, and 2) the same ambiguous sentences were presented with a referential context favoring the relative clause interpretation. The results of the experiment showed that the participants selected the complement interpretation of the ambiguous sentences in the first experiment which confirmed the processing effort assumption of the relevance theory because they selected the complement clause interpretation that had less syntactic nodes compared to the relative clause. In the second experiment, learners tended to select the contextualized relative clause, but in the non-referential condition they had no preference indicating the negative effect of context.

In another study, Ying (2005) investigated the relevance theory in relation to the interpretation of reflexive anaphora in VP-ellipsis structures. These structures are ambiguous because they yield two types of interpretations:

- John defended himself and Bill did too.

The example above demonstrates two meanings:

a. John defended himself+ and Bill× did [defend himself×] too. (i.e., Bill).
b. John defended himself+ and Bill× did [defend him+] too.   (i.e., John).

The first interpretation is known as the sloppy reading and the second as the strict reading. The learners were provided with these ambiguous sentences with and without context in two experiments. The results indicated that in the first experiment, learners preferred the sloppy reading because of these reasons:

a) the pronoun reading, which involves a change to the pronoun from the reflexive, is not immediately available in the linguistic information of the sentence, b) the distance

between the reflexive and its higher antecedent is greater than the distance between the reflexive and its lower antecedent, and c) the binding relationship of the reflexive in the elided VP with its higher subject is not within a single bounding node. (p. 563)

The results of the second experiment showed that the learners have chosen the strict reading because the preceding context in the referential form has supported this interpretation, while the negative context in non-referential form caused the learners to have no certain preference.

In 1972, Hakes carried out two experiments to recognize the effect of reducing the complementizer on the comprehension difficulty of native speakers of English. In doing so, he considered Fodor and Garrett's (1967) sentence comprehension theory which is based on the assumption that when a prominent clue is omitted from the linguistic codes of a sentence, the process of comprehension encounters some difficulty. They constructed twenty experimental sentences with the complementizer employed in all the experimental sentences along with thirty fillers. The structure of experimental sentences was [Det + (Adj) + N], with the prenominal adjective being present in some sentences but not in others. The speed of reading for these sentences was calculated to determine the differences in the presence and absence of the main clue. The analysis of the results indicated that the presence of the complimentizer helped native speakers not only in comprehension success but also in speed.

Purpose of the Study

The purpose of the present study was to investigate the manner in which advanced and intermediate Persian-speaking L2 learners of English resolved relative clause attachment ambiguities in the target language. The research questions explored in this study were as follows:

1. Are L2 learners constrained by a minimal effort principle in processing relative clause attachment ambiguities?
2. Does a preceding referential context impose procedural constraints on the comprehension of ambiguous sentences by L2 learners?
3. Are there any differences between learners' sex and their disambiguation of relative clause attachment ambiguities?
4. Are there any differences between learners' level of proficiency and their disambiguation of relative clause attachment ambiguities?

Method

Participants
The subjects consisted of 90 (40 male and 50 female) intermediate and 100 (35 male and 65 female) advanced Persian-speaking learners of English. The intermediate participants have been selected out of 100 learners based on their proficiency scores. The advanced participants have been chosen based on their proficiency scores out of 120 learners. Both the intermediate and advanced participants consisted of female and male learners to investigate the effect of sex on their disambiguation preferences.

The intermediate participants were learners in a language center in Tehran. The advanced learners were second-year university students of engineering at the BA level in Tousi University.

Instrumentation

TOEFL language proficiency test

Participants were divided into two groups of intermediate and advanced, based on their scores on a language proficiency test. The language proficiency test used for this purpose was the paper-based version of the Test of English as a Foreign Language (TOEFL) Preparation Kit (2003). The test did not include the listening comprehension questions, because the study did not aim at testing the participants' level of listening comprehension. Based on their scores on the TOEFL test, the intermediate and the advanced subjects have been selected to respond to the questionnaire. The results of the test at each level are presented at in Table 1.

Table 1 TOEFL test results

	Mean	Standard Deviation
Intermediate	31.95	2.22
Advanced	46.12	2.20

Questionnaire materials

The questionnaire materials included 80 experimental sentences and 30 fillers (see Felser, Roberts, Marinis, & Gross, 2003). All experimental sentences were ambiguous and of the form (NP–V–[NP1–P–NP2]–RC), where V is the verb, P is preposition, and RC is relative clause. There were two versions of each sentence, with the two noun phrases in the complex NP being joined either by functional *of* or by attributive *with*, as illustrated by examples 2a and 2b, respectively.

2. a) The dean liked the secretary *of* the professor who was reading a letter.

   b) The dean liked the professor *with* the secretary who was reading a letter.

The fillers were sentences with structural ambiguity as illustrated by examples 3.

3. The young woman bought an antique desk with long legs at the sale.

a) The desk had antique legs.

b) The woman had long legs.

In order to identify the learners' tendency in attaching the relative clause to one of the NPs, the experimental sentences consisted of 20 relative clause ambiguous sentences linked by *of* and 20 ambiguous sentences linked by *with.* For the second question, 20 sentences linked by *with* favoring an NP2 interpretation, and 20 sentences linked by *of* favoring an NP2 interpretation were employed. The ambiguous sentences presented with contextual cues were different from the non-contextual sentences.

Research procedure

Participants were instructed that they could refuse to participate without penalty, that participation was anonymous, and that completion of the test battery would

constitute informed consent. Before the experiment, the Persian-speaking learners of English were first asked to take the TOEFL Preparation Kit (2003). Therefore, only the intermediate and advanced level learners were selected for the study. After taking the TOEFL test, participants were instructed to read the sentences in the questionnaire carefully and to indicate one interpretation which seemed more appropriate for each question. Administered in their respective classrooms, the students took approximately 90 minutes to complete the questionnaire. Example 4 shows the procedure.

4. The dean liked the secretary of the professor who was reading a letter.

i. the secretary was reading a letter.

ii. the professor was reading a letter.

   In half of the choices, the NP1 in the complex appeared first, and in the other half, it appeared second, to avoid the subjects developing a strategy for answering the questions. Participants were encouraged to make their choices as immediately as possible.

The psychometric properties of the questionnaire

   Factor analytic procedure was used to test the validity of the questionnaire. Although the scale had high enough factor pattern/structure coefficients to qualify the respective items as marker variables (near pure representations of the factors), a few items were problematic and some refinements were considered necessary. This was accomplished by replacing the suspect items identified with appropriate ones. These changes were implemented and validated.

   The reliability of the questionnaire was computed using the Cronbach Alpha which turned out to be 0.78. The reliability estimate of this questionnaire was acceptable.

Results

Table 2 reports the performance on relative clause sentences containing complex *with* antecedents in the contextual and non-contextual conditions. In the non-contextual condition, as table 2 indicates, the participants provided more NP2 than NP1 responses. The disambiguation preferences of the participants have been affected by their level of proficiency and sex.

Table 2 Descriptive statistics for disambiguation preferences of complex NPs linked by *with*

	NP1(non-contextual)	NP2(non-contextual)	NP1 (contextual)	NP2 (contextual)
Mean	2.53	3.88	2.30	3.80
Std. Deviation	1.664	1.757	1.713	1.814
Variance	2.769	3.087	2.933	3.292

   Three statistical tests were performed to examine whether the statistical assumptions underlying the use of Multivariate Analysis of Variance (MANOVA) were violated in the data set. First, Box's M test of equality of covariance matrices

indicated that there were no significant differences between the covariance matrices. Therefore, the assumption of homogeneity of covariances across groups was not violated. Secondly, Levene's test of equality of error variances indicated that the homogeneity of variance for each of the dependent measures was not violated in the data set (p>0.05). The third test used was the multivariate test of significance, Wilks' Lambda criterion variance indicated that there was a statistically significant multivariate effect for gender levels (F= 2.615, p<0.05) and proficiency levels (F= 2.049, p<0.05). Having determined that the results met the statistical criteria set out above, the next step was to conduct MANOVA.

Table 3 Tests of between subjects-effects for complex NPs linked by *with*

Source	Dependent Variable	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected	withNP1	22.313^a	3	7.438	3.126	.038	.207
Model	withNP2	25.531^b	3	8.510	3.230	.034	.212
	conxwithNP1	24.581^c	3	8.194	3.284	.032	.215
	contxwithNP2	10.318^d	3	3.439	1.049	.383	.080
		248.090	1	248.090	104.261	.000	.743
Intercept	withNP1						
	withNP2	515.552	1	515.552	195.688	.000	.845
	conxwithNP1	223.502	1	223.502	89.581	.000	.713
	contxwithNP2	475.591	1	475.591	144.995	.000	.801
		2.452	1	2.452	.747	.393	.020
proficiency	withNP1						
	withNP2	19.629	1	19.629	7.451	.010	.171
	conxwithNP1	1.845	1	1.845	.740	.396	.020
	contxwithNP2	9.903	1	9.903	4.162	.049	.104
		7.722	1	7.722	3.245	.080	.083
Sex	withNP1						
	withNP2	7.646	1	7.646	2.902	.097	.075
	conxwithNP1	5.034	1	5.034	1.535	.223	.041
	contxwithNP2	13.845	1	13.845	5.549	.024	.134
		8.254	1	8.254	3.469	.071	.088
proficiency *	withNP1						
sex	withNP2	3.863	1	3.863	1.466	.234	.039
	conxwithNP1	7.994	1	7.994	3.204	.082	.082
	contxwithNP2	3.665	1	3.665	1.117	.298	.030
		85.662	186	2.379			
Error	withNP1						
	withNP2	94.844	186	2.635			
	conxwithNP1	89.819	186	2.495			
	contxwithNP2	118.082	186	3.280			
		363.000	190				
Total	withNP1						
	withNP2	721.000	190				
	conxwithNP1	326.000	190				
	contxwithNP2	706.000	190				
		107.975	189				
Corrected	withNP1						
Total	withNP2	120.375	189				
	conxwithNP1	114.400	189				
	contxwithNP2	128.400	189				

a. R Squared = .207 (Adjusted R Squared = .141)
b. R Squared = .650 (Adjusted R Squared = .567)
c. R Squared = .215 (Adjusted R Squared = .149)
d. R Squared = .180 (Adjusted R Squared = .154)

Multivariate Analysis of Variance (MANOVA) showed that the number of NP2 responses was significantly above chance level for *with* condition. The disambiguation preferences of the participants were affected by their level of proficiency and sex. The effect of level of proficiency on the selection of NP2 responses was significant in both the non-contextual (F= 7.451, p< 0.05) and contextual conditions (F= 4.162, p< 0.05). This positive effect of proficiency on NP2 selection signifies that the higher the level of proficiency, the greater the NP2 selection. Moreover, NP2 preference in the contextual condition was the only variable that differed significantly between genders. Based on pairwise comparisons using estimated marginal means, females reported significantly higher NP2 selection than males (F= 5.549, p< 0.05). In this analysis, there was no interaction effect (p>0.05).

Table 4 presents the mean percentages of responses provided for complex genitive NPs. In the non-contextual condition, the Persian-speaking L2 learners showed a marked preference for NP1 attachment for complex NPs linked by *of*. Regarding the fact that the participants provided more NP2 responses than NP1 responses in the contextual condition, it can be argued that their attachment decisions were influenced by the contextual effect.

Table 4 Descriptive statistics for disambiguation preferences of complex NPs linked by *of*

	NP1(non-contextual)	NP2(non-contextual)	NP1 (contextual)	NP2 (contextual)
Mean	6.00	2.95	1.50	3.63
Std. Deviation	2.230	2.207	1.109	1.807
Variance	4.974	4.869	1.231	3.266

In response to the learners' disambiguation preferences of complex NPs linked by *of* antecedent, MANOVA was used. Box's M test of equality of covariance matrices demonstrated that the homogeneity of covariance matrices was met. Levene's test of equality of error variances indicated that the homogeneity of variance for the dependent measures was not violated in the data set. The multivariate tests of significance, Wilks' Lambda criterion variance, indicated that there was a statistically significant multivariate effect for gender levels (F= 2.521, p<0.05) and proficiency levels (F= 3.170, p< 0.05). Table 5 presents the results of the MANOVA.

Table 5 Tests of between subjects-effects for complex NPs linked by *of*

Source	Dependent Variable	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	ofNP1	35.302[a]	3	11.767	2.669	.000	.883
	ofNP2	45.667[b]	3	15.222	3.799	.018	.240
	contxofNP1	7.565[c]	3	2.522	2.245	.100	.158
	contxofNP2	13.291[d]	3	4.430	1.398	.015	.680
Intercept	ofNP1	1270.829	1	1270.829	288.282	.000	.889
	ofNP2	317.017	1	317.017	79.126	.000	.687
	contxofNP1	78.615	1	78.615	69.992	.000	.660
	contxofNP2	422.104	1	422.104	133.198	.000	.787
Proficiency	ofNP1	18.721	1	18.721	4.247	.047	.106
	ofNP2	22.826	1	22.826	5.697	.072	.137
	contxofNP1	6.179	1	6.179	1.950	.171	.051
	contxofNP2	6.820	1	6.820	6.072	.019	.144
Sex	ofNP1	24.852	1	24.852	5.638	.043	.135
	ofNP2	33.687	1	33.687	8.408	.116	.189
	contxofNP1	1.435	1	1.435	1.278	.266	.034
	contxofNP2	1.752	1	1.752	.553	.462	.015
proficiency * sex	ofNP1	.566	1	.566	.128	.722	.004
	ofNP2	.157	1	.157	.039	.844	.001
	contxofNP1	.755	1	.755	.672	.418	.018
	contxofNP2	5.497	1	5.497	1.735	.196	.046
Error	ofNP1	158.698	186	4.408			
	ofNP2	144.233	186	4.006			
	contxofNP1	40.435	186	1.123			
	contxofNP2	114.084	186	3.169			
Total	ofNP1	1634.000	190				
	ofNP2	538.000	190				
	contxofNP1	138.000	190				
	contxofNP2	653.000	190				
Corrected Total	ofNP1	194.000	189				
	ofNP2	189.900	189				
	contxofNP1	48.000	189				
	contxofNP2	127.375	189				

a. R Squared = .883 (Adjusted R Squared = .816)
b. R Squared = .240 (Adjusted R Squared = .177)
c. R Squared = .158 (Adjusted R Squared = .087)
d. R Squared = .680 (Adjusted R Squared = .498)

Multivariate Analysis of Variance (MANOVA) revealed that the disambiguation preferences for NP1 selection in the non-contextual and NP2 selection in the contextual conditions were significantly above chance level. The advanced Persian-speaking L2 learners showed a greater preference for NP1 selection in the non-contextual condition (F= 4.247, p< 0.05) and NP2 selection in the contextual condition (F= 6.072, p< 0.05). Moreover, the female and male participants have differed in NP1 attachment preferences in the non-contextual condition. No interaction effects were found for the level of proficiency and sex variables.

Discussion

The present study investigated the way adult foreign language (FL) learners of English disambiguate relative clause attachment ambiguities in relation to the principles of relevance theory. It has provided empirical evidence in support of relevance theory. First, the results of the non-contextual disambiguation preferences of learners for the complex genitive NPs linked by *of* preposition indicated that both the intermediate and advanced participants interpreted the relative clauses with a strong preference for NP1 selection. This finding supports the processing effort principle of relevance theory because learners have applied the phrase-structure based locality principles (recency or predicate proximity). By attaching the preposition *of* to the most recently processed phrase (recency) and/or by attaching the preposition as structurally close as possible to the head of the phrase (predicate proximity), learners had fewer effort in attaching the preposition *of* to NP1. In other words, when learners are not aided with lexical-semantic information in genitive complex nouns, they tend to use locality principle of predicate proximity. Persian is a language in which the verb is more active during the processing, so verb attracts the closest NP. The results of MANOVA demonstrated that the difference between the intermediate and advanced participants was statistically significant with the advanced group showing a stronger preference for NP1 disambiguation. In the contextual condition, the participants' disambiguation preferences have been influenced by the contextual cue favoring a strict interpretation of the relative clause ambiguity. In the contextual condition, too, the advanced participants showed a clear preference for NP2 selection of NPs linked by the preposition *of*. Therefore, involving contextual information constrained the inferred conclusion about encoding the relative clause. In this way, the contextual information also reduced the overall computational effort required of the learner and guided him/her towards the strict interpretation. The results of the empirical analyses have also indicated a statistically significant difference between female and male participants in the disambiguation of NPs, indicating that the gender factor was responsible for the difference. NP1 selection in the non-contextual condition was the only variable that differed significantly between genders. Females reported significantly greater preference for the selection of NP1 in genitive complex conditions.

By and large, the results of MANOVA for the relative clause sentences containing complex *with* antecedent revealed similar patterns of disambiguation. It is widely observed that in complex NPs linked by the thematic prepositions such as *with*, NP2 disambiguation is universally preferred over NP1 disambiguation even in languages for which an NP1 selection has been attested. This suggests that the lexical bias is strong enough to expunge the effect of any phrase-structure-based locality principle (Felser, Robert, Marinis, & Gross, 2003).Therefore, the lexical bias helps learners select the NP2 interpretation of the ambiguous relative clauses with less processing effort. This assumption can be supported by construal theory according to which attaching a relative clause to constituents outside the thematic domain is dispreferred. Previous studies have revealed that English native speakers in the same conditions, also, prefer NP2s (e.g., Dussias, 2001; Fernandez, 1999) and Persian-speaking

learners of English seem to be in line with this universal preference. The advanced and intermediate participants differed significantly in their NP2 selection preference, with advanced learners having a stronger preference. However, female and male participants were not different in their attachment ambiguity resolution in the non-contextual condition.

In the contextual condition, both the advanced and intermediate participants provided more NP2 responses. It is clear that contextual condition along with the lexical bias helped learners chose the cued NP. The results indicated that the differences between the intermediate and advanced learners' interpretation of ambiguous sentences were statistically significant, although both showed a preference for the NP2 selection. The only variable that has been found to differ across female and male learners was the contextual NP2 preference.

Conclusion

Taken together, the findings signify that learners follow the principles of relevance theory in that they tend to construe and represent the relevant information and they also spend the least effort in their interpretations.

The state of the art of ELL is definitely a colorful tapestry. As more and more learners start learning FLs in different educational contexts and under so many varying conditions, many of the issues identified in previous studies are still on the agenda, but new ones have also surfaced. More studies are needed on how learners deal with the ambiguities in the target language, how several factors affect their interpretations, and how they arrive at a conclusion. Further research is also needed in conducting similar studies with other groups of learners and teachers both in Iran and with students of different first languages studying other target languages.

The models that predict L1 transfer of processing strategies, whether universalist or purely exposure based, account for the findings of the presents study. The participants of the study positively transferred the L1 disambiguation preferences to L2. This is in line with VanPatten's input processing theory which states that learners process input for meaning before they process it for form. Therefore, the implications suggest that teachers who integrate tasks that involve meaningful context and that encourage learners to attend to the meaning before form ensure successful language acquisition.

References

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30,* 191-238.

Clark, H. H. (1987a). Relevance to what? *Behavioral and Brain Sciences, 10*, 714-715.

Dussias, P. (2001). Sentence parsing in fluent Spanish-English bilinguals. In J. Nicole (Ed.), *One mind, two languages: Bilingual language processing* (pp. 159-176). Malden, MA: Blackwell.

Evans, J.St.B.T. (1994). Relevance and reasoning. In S. Newstead & J.St.B.T. Evans (Eds.), *Current directions in thinking and reasoning*. Hillsdale, NJ: Erlbaum.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453-489.

Fernandez, E. (1999). Processing strategies in second language acquisition: Some preliminary results. In E. Klein & G. Martodardjono (Eds.), *The development of second language grammars: A generative approach* (pp. 217-240). Amsterdam: Benjamins.

Fodor, J. A., & Garrett, M. F. (1967). Some syntactic determinants of sentential complexity. *Perception and Psychophysics, 2*, 289-296

Foster-Cohen, S. (2000). Relevance: Communication & cognition. *Second Language Research, 16*, 77-92.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6,* 291–326.

Hakes, D. T. (1972). The effect of reducing complement constructions on sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, *11*, 278-286.

Harley, B., Howard, J., & Hart, G. (1995). Second language processing at different ages: Do younger learners pay more attention to prosodic cues to sentence structure? *Language Learning*, *45,* 43-71.

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language, 32*, 695-715.

Randall, M. (1988). The context of literary communication: Convention and presupposition. *Journal of Literary Semantics, 17*, 46-53.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning & Verbal Behavior, 22,* 358-374.

Rumos, F. Y. (1998). Relevance theory and media discourse: A verbal-visual model of communication. *Poetics, 25*, 293-309.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57*, 31-95.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.

Wilson, D., (1994). Relevance and understanding. In G. Brown, K. Malmkjar, A. Pollit & J. Williams (Eds.), *Language and Understanding*, (pp. 35-58). Oxford: Oxford University Press.

Ying, H. (1996). Multiple constraints on processing ambiguous sentences: Evidence from adult L2 learners. *Language Learning*, *46*, 681-711.

Ying, H. (2004). Relevance mapping: a study of L2 learners' processing of syntactically   ambiguous sentences in English. *Second Language Research, 20*, 232-255.

Ying, H. (2005). Relevance and L2 learners' interpretation of reflexive anaphora in VP-ellipsis: An exploration of relationship between relevance theory and typological universal. *Journal of Universal Language, 6*, 159-194.

Yus, F. (2002). Humor and the search for relevance. *Journal of Pragmatics, 2*, 1-37.

Title

Overt Pronoun Constraint in Persian

Authors

Seyyed Mohammad Reza Amirian (Ph.D. Candidate)

University of Tehran, Iran.

Reza Bagheri Nevisi (Ph.D. Candidate)

University of Tehran, Iran.

Bio Data

Seyed Mohammad Reza Amirian is a Ph.D. candidate of TEFL at the University of Tehran. He is a lecturer in applied linguistics at Sabzevar Teacher Training University, Iran. He has published books and articles and presented in a number of national and international conferences. His primary research interest includes test fairness, differential item functioning, reading assessment and UG.

Reza Bagheri Nevisi is a Ph.D. candidate of TEFL at the University of Tehran, Iran. He is a lecturer in applied linguistics at Qom University. He has published articles and presented in many national and international conferences. His research interest includes task based language teaching and learning, communicative language teaching, language assessment, and second language acquisition.

Abstract

This study investigated the status of Overt Pronoun Constraint (OPC) in Persian. OPC asserts that in a language with overt/null pronominal alternation, an overt pronoun cannot receive a quantified antecedent. A group of 36 adult Persian learners with different first language backgrounds participated in the study. For data collection purposes, a written questionnaire consisting of four different patterns was developed. There were four tokens for each pattern: (1) the biclausal sentences with quantified antecedent and null subject, (2) quantified antecedent and overt pronoun, (3) referential antecedents and null subjects, and (4) referential antecedents and overt pronouns. In order to have a baseline for comparison, a control group of 19 adult

Persian native speakers took part in the endeavor. The results of the study showed that the difference in the interpretation of null and overt pronominals by this group of Persian learners was meaningful. Moreover, it was found that Persian learners demonstrated a native-like mastery of Overt Pronoun Constraint in Persian.

Introduction

The question of how to account for the use and interpretation of overt versus null pronouns in pro-drop languages like Persian has been addressed within various fields of linguistics, theoretical syntax, sociolinguistics, and discourse pragmatics (Lubbers Quesada & Blackwell, 2009). Overt Pronoun Constraint (OPC) principle states that in a language with overt/null pronominal alternation, an overt pronoun cannot receive a quantified antecedent. Some researchers consider it as a UG principle while others draw on non UG explanations.

OPC involves the contrast between overt and null pronouns with regard to their ability to take quantified NPs as antecedents. In languages with both overt and null pronouns (e.g., Persian), only null pronouns can be bound by a universal quantifier such as *everyone* or a question word such as *who.*

It is argued that the knowledge of OPC learned by native speakers goes beyond the input that they receive as young children. In L2 acquisition, learners come across a similar task to that of L1 learners, namely the need to arrive at a system accounting for L2 input. There has been an ongoing debate in SLA research about whether L2 learners' mental grammars derive from UG or not for over twenty five years. The strongest case for the operation of principles of UG in interlanguage grammar can be made if L2 learners demonstrate knowledge of the subtle and abstract linguistic property of OPC.

A seminal study by Kanno (1997) shows that English speakers are sensitive to overt pronoun constraint in L2 Japanese, although neither English nor Japanese provide evidence for it. She investigated whether the OPC is accessible in the L2 acquisition of Japanese. She wanted to assess subjects' knowledge of the contrast between null and overt pronominals in terms of their co-occurrence with a quantified noun phrase antecedent. She concluded that subjects' compliance with OPC constitutes evidence for access to UG.

However, Yamada (2002) points to the problems for the OPC in Kanno's study. The first problem is whether overt embedded subject pronouns can really refer to the referential NPs in matrix subject position, because of the growing evidence that in languages that permit null pronouns, they are not in free variation with the overt pronouns. A second problem is whether there are phenomena in other languages of a

similar type to the null/overt alternation in Japanese. Marsden (1998, cited in Yamada, 2002) found that null and overt pronouns do not refer to the same antecedent in Japanese.

In another study, Lozano (2002) studied whether overt and null pronouns are in free variation in Spanish. He states that the occurrence of overt/null pronouns in OPC contexts is configurationally-bound. If the relevant constituent is informationally-focused, overt pronouns will appear and if the constituent has neutral focus, null pronouns will pop up. In a similar study, Gurel (2003) investigated the effects of OPC in Turkish. Despite arguments for the universality of the OPC, he presented evidence for the similar distribution of null and overt pronouns.

What triggers OPC?
Yamada (2002) considers the status of the OPC in the grammatical theory. He asks the following questions: why can null pronouns have a bound variable interpretation while overt pronouns cannot? What actually causes OPC effects? He discusses two different hypotheses. 1) Distributional Difference Hypothesis: OPC effects result from a difference in the distribution of null/overt pronominals in syntactic structures. 2) Morphological Under -specification Hypothesis: the morphological component, i.e. feature specification is responsible for the effects. The first hypothesis states that OPC effects might be due to a difference in the syntactic structure of overt pronouns and null pronouns which is linked to differences in binding properties. The second hypothesis is that OPC might be located in the morphological component. Yamada's first explanation is more acceptable and it is adapted as the possible explanation for OPC in the present study.

Universality of OPC
Montalbetti (1984) formulated the binding contrast between overt and null pronouns in the context of quantified antecedents under OPC and postulated it as a property of Universal Grammar. After Montalbetti's claim on the universality of the OPC in pro-drop languages, Perez-Leroux & Glass (1999) and Kanno (1997) considered this phenomenon in Spanish and Japanese respectively. Native-like performance in adult L2 acquisition of the OPC was found and the presence of UG-constrained L2 grammars in adult learners was used to account for such performance.

Due to the fact that there seems to be a lack of research on OPC in the literature considering its significance as a UG principle, there are some good reasons why OPC is the right venue for a good amount of research. The phenomenon under investigation is not discussed in any of the second language books or instructional materials. It has no direct counterpart in English. Instruction and transfer cannot be regarded as possible explanations for OPC influence (Kanno, 1998). Therefore, to bridge the gap, the present study could be illuminating on the status of OPC in Persian and universality of this principle

Purpose of the study

Although there is enough reason to believe that Overt Pronoun Constraint is part of UG since this knowledge is underdetermined by L2 input and also cannot be derived from the grammar of the mother tongue, alternative ways of analyses have recently appeared and possible pragmatic explanations and non-UG accounts have been proposed.

Therefore, in order to shed more light on this issue and examine its availability to L2 Persian learners as a UG principle, the present study aimed at finding out whether Persian interlanguage grammar is constrained by this UG principle. To date, almost no study has been conducted to examine the status of OPC in Persian. Hopefully the results of the present study will throw more light on the operation of this UG principle in Persian.

Method
Participants
The participants in the present study consisted of a group of 35 male and female Persian learners with different first language backgrounds including English, Arabic, Russian, Hindi, Turkish, Chinese, Korean, and Japanese. Their proficiency levels ranged from lower- intermediate to upper-intermediate and they were taking Persian classes at Dehkhoda institute in Tehran.Their length of residence in Iran varied from 6 months to 4 years. Some of them have been familiar with Persian before coming to Iran. A control group of 19 adult native speakers of Persian was also selected to provide a baseline against which to measure L2 learners' performance.

Procedures
A written questionnaire was developed to examine the participants' knowledge of the contrast between null and overt pronominals with respect to co-occurrence with a quantified antecedent. All the instructions on the questionnaire were written in Persian. It was intended to encourage the students to approach the questionnaire from the perspective of Persian grammar. From time to time, some explanations were also made in English whenever required. The test materials consisted of four sets of biclausal sentences with each set containing four tokens. There were also four filler sentences intended to distract the learner's attention from the point being tested. The sentences in Examples 1 and 2 show the contrast between overt and null pronouns with respect to a quantified antecedent.

1. Null pronoun with a quantified NP as antecedent
   کی دلیل آورده که بی گناه است؟
   Who$_i$ has brought evidence that (he$_i$) is innocent?

2. Overt Pronoun with a quantified NP as antecedent.
   کی باور میکرد که او پیروز شود؟
   Who$_i$ could believe that (he$_i$) won the game?

Participants had to indicate the interpretation of the subject argument in the embedded clause by responding to the options (a) "the same as کی or (b) "another person". By selecting choice (a) for the first sentence and (b) for the second sentence learners would indicate that they had learned the OPC in Persian.

To ensure that the subjects know that "او (he/she), آنها (they)" can take an intrasentential antecedent that is a referring NP, another set of four sentences of the following type was also included.

1. Zero pronoun with a referring NP as antecedent.
آقای احمدی گفت که میخواهد به مشهد برود
Mr Ahmadi$_i$ said that (he$_i$) wanted to go to Mashhad.

2. Overt pronoun with a referring NP as antecedent.
حامد باور نمی کرد که او برنده شود
Hamed$_i$ couldn't   believe that (he$_i$) won.

By selecting Hamed as antecedent for " او " in such sentences, subjects demonstrate that they do not have any general prohibition against the use of the overt pronoun with an intrasentential antecedent. In order to establish knowledge of the OPC, subjects must accept the co-referential interpretation in sentences like 4.

Results

Results from Persian native speakers
 Table 1 presents the results from the control group of Persian native speakers. These sentences were designed to examine knowledge of the OPC.

Table 1.   Interpretation of Null and Overt Pronoun with respect to a quantified antecedent
(Control group) (n=19)

Type of argument in the embedded clause	Permits matrix subject antecedent		Does not permit matrix subject antecedent	
	N	Percentage	N	Percentage
Null	27	72 %	11	(28%)
Overt	13	34%	25	(66%)

As predicted by Overt Pronoun Constraint, the native speakers of Persian made a distinction between Null and Overt Pronoun in terms of their compatibility with a

101

quantified antecedent. The native speakers selected the Null subject 72% of the time. With Overt Pronoun, this interpretation was made 28% of the time which is negligible.

Now we consider the results from the sentences with a referring NP as antecedent. As you can see in table 2 below, native speakers accepted the matrix subject as the antecedent of Null Pronoun 100% of the time. It shows that when the null pronominal was used in the embedded clause, there was a very strong preference of intrasentential antecedent.

Table 2. Interpretation of Null and Overt pronoun with respect to a referring antecedent (Control group) (n= 19)

Type of argument in the embedded clause	Permits matrix subject antecedent		Does not permit matrix subject antecedent	
	N	Percentage	N	Percentage
Null	38	(100%)	0	
Overt	20	(53%)	18	(47%)

In the case of Overt Pronoun, however, there is not such a preference. As indicated by the table, the referring NP in the matrix subject position was chosen as the antecedent 53% of the time. It suggests that co-reference is permitted but there is no special preference for it. Therefore we can conclude that co-referential interpretation is not prohibited for referring NPs while for quantified NPs co-referential interpretation is rejected.

Results from the L2 learners

Now we are going to compare these results with those of the L2 learners. The results of sentences with a quantified NP are shown in table 3.

Table 3. Interpretation of Null and Overt Pronoun with respect to a quantified antecedent

(L2 learners, n=36)

Type of argument in the embedded clause	Permits matrix subject antecedent		Doesn't permit matrix subject antecedent	
	N	percentage	N	percentage
Null	52	(72%)	20	(28%)
Overt	32	(44%)	40	(56%)

As the findings indicate Persian learners selected the matrix subject as the antecedent for a null pronoun 72% of the time. Native speakers had also chosen the matrix subject as the antecedent of null pronoun 72% of the time which is exactly repeated by Persian L2 learners. Thus, there is no difference between the two groups. The matrix subject was also selected as the antecedent of overt pronoun 44% of the time. It suggests that Persian learners prefer to choose the matrix subject as the antecedent of null pronoun to choosing it as the subject of overt pronoun. Table 4 reports on the results of sentences with a referring NP as the matrix subject.

Table 4. Interpretation of null and Overt Pronoun with respect to a referring antecedent
(Persian learners, n= 36)

Type of argument in the embedded clause	Permits matrix subject antecedent		Doesn't permit matrix subject antecedent	
	N	percentage	N	percentage
Null	61	(85%)	11	(15%)
Overt	24	(33%)	48	(67%)

As it can be inferred from the table, in patterns with a null subject the referring NP in subject position was obviously the preferred antecedent (it was chosen 85% of the time). Once again we found out that the responses from Persian learners were very similar to those of control group. Therefore, it is concluded that since there is no obvious positive evidence in Persian that would allow speakers to infer this, they must have access to knowledge coming from UG.

Discussion

This article concentrated on the acquisition of OPC in Persian. It was found that OPC, as a principle of UG, is observed by L2 learners of Persian indicating that their interlanguage grammar is UG–bound because this knowledge of OPC could not have been learned from L2 input nor derived from the grammar of mother tongue. The participants received no instruction on the interpretation of null and/or overt pronouns, and it is extremely unlikely that Persian learners could have received previous exposure to any of the sentences or judgments that illustrate the operation of this principle. There is nothing in experience to tell language learners overt pronouns from null pronouns with regard to the permissibility of a quantified antecedent. This finding supports the results of Kanno's (1997) study where he found that OPC can be found in a wide range of unrelated languages such as Spanish, Korean, and Chinese. He concludes that since instruction and transfer has no role in the acquisition of OPC it could be considered as one of UG principles.

   For many years there has been an ongoing debate in SLA research about whether L2 learners' mental grammars derive from UG or not. Although there have been proponents of a `no access to UG' view (Clahsen and Muysken 1986,Bley-Vroman, 1990 & Newmeyer, 1998), many researchers have shown that L2 learners can acquire grammatical distinctions proposed by linguists working within a UG framework (White 1990, Schwartz and Sprouse 1996 & Grondin and White 1996). This study is further evidence for the availability of UG to L2 learners. That is if L2 learners can acquire such distinctions as OPC, the implication is that they have access to UG.

   However, since we are unaware of the status of OPC in some of the learners' first languages such as Russian and Hindi, we cannot strongly argue that transfer plays no role in the acquisition of this principle and it could be regarded as a possible explanation for OPC influence. Therefore, a more comprehensive study could address this topic by studying the status of OPC in relation to learners' mother tongue or by investigating participants that all share the same first language.

References

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, *20*,3-49.

Clahsen, H. & Muysken, P. (1986) The availability of Universal Grammar to adult and child learners: a study of the acquisition of German word order. *Second Language Research*, *2*,93-119.

Grondin, N. & White, L. (1996) Functional categories in child L2 acquisition of French. *Language Acquisition*, *5*, 1-34.

Gurel, A. (2003). Is the Overt Pronoun Constraint Universal? Evidence from L2 Turkish. Proceedings of the 6th Generative approaches to Second Language Acquisition Conference.

Kanno, K. (1997). The acquisition of null and overt pronominals in Japanese by English Speakers, *Second Language Research*, 13(3), 265-287.

Kanno, K. (1998). Consistency and variation in second language acquisition. *Second Language Research,* 14 (4), 376-388.

Kanno, K. (1998). The stability of UG principles in second language acquisition: Evidence from Japanese, *Linguistics*, 36 (6), 1125-1146

Lozanov, C. (2002). The interpretation of overt and null pronouns in non-native Spanish. *Durham Working Papers in Linguistics,* 8, 53-66.

Lubbers Quesada, M. & Blackwell S. E. (2009). The L2 acquisition of null and overt Spanish subject pronouns: A pragmatic approach. In *Selected Proceedings of the 11th Hispanic Linguistics Symposium*,ed. Joseph Collentine et al., 117-130. Somerville, MA: Cascadilla Proceedings Project.

Montalbetti, M. (1984). *After binding*. Unpublished PhD dissertation, Massachusetts Institute of Technology.

Montrul, S. & Bruhn de Garavito, J (1999). The L2 acquisition of Spanish: generative perspctives. *Second Language Research,* 15(2), 111-114

Preze-Leroux, A. & Glass, w. (1999). Null anaphora in Spanish second language acquisition: probabilistic versus generative approaches. *Second language Research,* 15 (5), 220-249.

Schwartz, B. D. & Sprouse, R. (1996). L2 cognitive states and the 'full transfer/full access' model. *Second Language Research*, *12*, 40-72.

White, L. (1990). Second language acquisition and universal grammar. *Studies in Second Language Acquisition*, *12*, 121-133.

Yamada, K (2002). The status of the Overt Pronoun Constraint in grammatical Theory and SLA of Japanese. Paper presented in Second Language Discussion Group at the University of Essex.

Title

The Effect of Intrinsic and Extrinsic Motivation on Iranian EFL
Learners' Language Learning

Authors

Abbas Pourhosein Gilakjani (Ph.D. candidate)

Islamic Azad University, Lahijan Branch, Iran.

Seyedeh Masoumeh Ahmadi (B.A.)

Islamic Azad University, Lahijan Branch, Iran.

Bio Data

Abbas Pourhosein Gilakjani, Ph.D. candidate of TEFL at the Universiti Sains Malaysia, School of Educational Studies, Pinang, Malaysia.. He is the faculty member of the Islamic Azad University, Lahijan Branch from 1999 to 2010. His main research interests are pronunciation listening comprehension, motivation. He has published five articles in international journals and presented five papers in international conferences.

Seyedeh Masoumeh Ahmadi, B.A. graduate of Computer Engineering from the Islamic Azad University, Lahijan, Iran. She will start her education in M.A. in Computer Engineering at the Universiti Sains Malaysia in September 2010. Her main research interests are multimedia, listening comprehension, motivation. She has published five articles in international journals and presented five papers in international conferences.

Abstract

The study investigates the effect of intrinsic and extrinsic motivation on Iranian EFL learners' language learning. The research examined both male and female learners using Gardner and Lambert's (1972) "extrinsic and intrinsic motivation" model in order to achieve the desired results. Sixteen items were selected from Gardner and Lambert's model. They were distributed among students for testing. Results indicated that female students have stronger intrinsic motivation than extrinsic motivation and

male students have stronger extrinsic motivation than intrinsic motivation. Findings provide a better understanding of the theoretical and practical facets for teachers.

Introduction

Motivation is one of the most appealing, complex variables used to explain individual differences in language learning. By the 1990s Gardner's motivation had overwhelming dominance in second language motivation research (Dornyei, 2001). Gardner and Lambert (1972) focus on both intrinsic and extrinsic motivations in second language learning. Student motivation affects how much a student desires to participate in the learning process. Intrinsic motivation refers to "motivation to engage in an activity for its own sake", and extrinsic motivation refers to "motivation to engage in an activity as a means to an end" (Pintrich & Schunk, 2002, p. 245). Extrinsic motivation refers to outside sources or values that influence a person to act or learn. Examples of these outside sources are rewards; positive or negative outcomes; and comfort or discomfort. As long as this external source provides the sufficient incentives or conditions, learning can take place. However, once the external input stops or no longer provides sufficient value to the student, then the willingness and effort to learn will also stop (Bomia et. al., 1997, p. 4). Intrinsic motivation refers to influences that originate from within a person which cause a person to act or learn. Examples of these influences are one´s self-concept, self-esteem, self-satisfaction, personal values, and personal/emotional needs and drives. Self-motivation can lead the student to go beyond the scope and requirements of an educational course because they are seeking to learn about the subject, not just fulfil a limited set of requirements (Bomia et. al., 1997, pp. 3-4).

It seems that females have a better talent for language than males. There are contradictory beliefs about the tendency of females towards more language learning. Hagborg (1995) indicated that there were no significant gender differences on 'mastery motivation' intrinsic motivational components. Boggiano and Barrett (1992) found that females are more extrinsically motivated than males, and that males are intrinsically motivated, have fewer incidents of depression than females. Therefore, it is necessary to investigate EFL student motivations for learning English. This research will investigate whether Iranian EFL students tend to have a stronger extrinsic motivation or intrinsic motivation.

Theoretical framework

Motivation has strong effect on students' achievement in numerous studies. Studies have confirmed the relationship between intrinsic motivation and course material and higher academic performance (Noels, et. al., 1999), indicating that intrinsic motivation may be critical predictor of learners' academic performance. According to the most researchers, individuals who are intrinsically motivated, compared to these

who are controlled by others to perform an activity(extrinsically motivated) have more interest, excitement, fun, confidence, which leads to enhanced performance, creativity, persistence, vigour, well-being, and self-esteem (Ryan & Deci, 2000). Noels, Clement, Pelletier (2001) investigated French Canadian students' intrinsic and extrinsic motivation for language learning. The results showed that integrative motivation has strongly correlation with intrinsic motivation. In other studies, female students were also found to have stronger intrinsic motivation. Nilsson and Stomberg (2008) also revealed that female nursing students have higher intrinsic motivation than male students. Jacobsson (2000) found that female students had higher intrinsic motivation than males. It was also found that female students work harder, having more thought-out strategies for their studies. Male students were found to have stronger extrinsic motivation than girls.

Learners' use of strategy may reflect their motivational motivation. Learning motivation was confirmed to have significant correlation with learning strategies. Chang and Huang (1999) investigated English learners' learning motivation and learning strategies. It was found that intrinsic motivation was significantly related to motivation level as well as the deep processing strategies: cognitive and meta-cognitive strategies. They suggested that intrinsic motivation may be powerful predictors for language learning and should be put emphasis in the EFL classroom. Liao (2000) examined Taiwanese high school students' learning motivation and use of strategies. It was found that learners showed stronger extrinsic motivation than intrinsic motivation and that social and meta-cognitive strategies were found to be used most frequently.

The teaching and learning of English has long been a difficult task for both EFL learners and teachers in Iran due to reasons such as lack ok of resources and little contact with the target language (Sadeghi, 2005). Iranian learners are highly motivated to study English. It will be interesting to investigate Iranian learners' motivation to learn English. The present study intends to emphasize the extent of Iranian learners' motivation in learning English as a foreign language and the differences in intrinsic and extrinsic motivation. Through this study, we hope to better understand their needs and help them develop appropriate learning strategies that may enhance their language learning. Furthermore, the findings of this study maybe provide some improvements and changes concerning teaching and learning L2 in Iran.

This study wants to provide a better understanding of motivation for Iranian EFL students.

There are two questions for this study:

1: Are there any motivational differences among male and female students?

2: Are Iranian EFL students more intrinsically or extrinsically motivated?

The hypotheses of this study, on the basis of its research question, are as follows:

1. Iranian EFL students are both intrinsically and extrinsically motivated.

2. There are motivational differences between Iranian EFL male and female students.

Methodology

Subjects

The researchers conducted a survey of 60 Iranian EFL students from 2 different English Translation classes at the Islamic Azad University of Lahijan, Iran. They were 30 female 30 male students between 21 and 25 years of age. They have been studying English since 1999.

Instruments

The students were required to complete a questionnaire showing their motivation concerning learning English. 16 items were selected for test administration: 8 intrinsic motivation and 8 extrinsic motivation. The questionnaire put emphasis on the motivational differences among Iranian EFL students. The questionnaire was translated into Persian language so that every student could understand the questions.

Procedure

The questionnaire consisted of 16 different items, each with a five-point scale ranging from Strongly Agree (1) to Strongly Disagree (5). 8 measured intrinsic motivation and 8 measured extrinsic motivation. The instructions for completing the questionnaire were explained by the researchers. Respondents were given 20 minutes to answer the questionnaire.

Data analysis and results

Descriptive analyses were used to describe the data in order to get the results for the motivation orientation of Iranian EFL students learning English. Each 8-item questionnaire was subjected to two reliability tests to check for internal consistency of both intrinsic and extrinsic motivation. The reliability coefficients were high, confirming that the internal consistency of 16 items in the questionnaire was high.

Hypothesis 1: Iranian EFL students are both intrinsically and extrinsically motivated

In order to test the research hypothesis, the X and SD of the items were compared. Then, two separate t-tests between pairs of the means were computed for items to examine whether there were significant differences between pairs of intrinsic and extrinsic items (tables 1 and 2).

Table1 Students' Scores on Intrinsic Items

Male	Female
X = 4.11	X = 3.55
$\sum X = 247$	$\sum X = 213$
S = 1.47	S = 1.94

T obs =1.86   S=Standard Deviation

Table one indicates the students' scores on intrinsic items. The mean and standard deviation of the intrinsic items for male are 4.11, 1.47 respectively. The mean and standard deviation of the intrinsic items for female are 3.55, 1.94 respectively. Then the t-observed of the study for the participants on intrinsic items was calculated. It was about 1.86.

Table 2 Students' Scores on Extrinsic Items

Male	*Female*
X = 4.15	*X = 3.85*
$\sum X = 268$	*$\sum X = 261$*
S = 1.05	*S = 1.15*

T obs = 1.76   S=Standard Deviation

Table two indicates the students' scores on extrinsic items. The mean and standard deviation of the intrinsic items for male are 4.15, 1.07 respectively. The mean and standard deviation of the intrinsic items for female are 3.85, 1.15 respectively. Then the t-observed of the study for the participants on intrinsic items was calculated. It was about 1.76. By applying the T-test, it was found that there is a difference between the two scores on intrinsic and extrinsic motivation.

Table 3 Summary of statistics: The Mean, Standard Deviation, and reliability for the questionnaire according to gender (N=60)

Variables	Gender	Subjects	Mean	Standard Deviation	Reliability
Intrinsic	Female	30	4.11	1.47	0.88
motivation	Male	30	3.55	1.94	0.81
Extrinsic	Female	30	3.85	1.15	0.82
motivation	Male	30	4.15	1.05	0.92

Table 3 shows the subjects' mean, standard deviation, and reliability for the questionnaire. The mean, standard deviation, and reliability of questions related to intrinsic motivation are 4.11, 1.47, 0.88 for female and 3.55, 1.94, 0.81 for male respectively. The mean, standard deviation, and reliability of questions related to extrinsic motivation are 3.85, 1.15, 0.82 for female and 4.15, 1.05, 0.92 for male respectively. The reliability for intrinsic and extrinsic items was 0.88, 0.81 and 0.82, 0.92 respectively. This indicated that the scorers' scores are correlated. So they are reliable enough to be able to show the results of the study. The statistical computation of the study showed the results regarding the hypotheses. This means that the hypotheses of the study have been supported, which means that there are motivational

differences between Iranian EFL male and female students. The results of the study showed a relationship between motivational orientations and gender.

Hypothesis 2: There are motivational differences between Iranian EFL male and female students
Intrinsic motivation among male and female
Comparing the overall mean scores of the items in intrinsic motivation among male [M=3.55] to the female [M=4.11], it can conclude that the female learners are to a certain extent intrinsically motivated but they still have an extrinsic motivation towards foreign language learning (see figure 1).
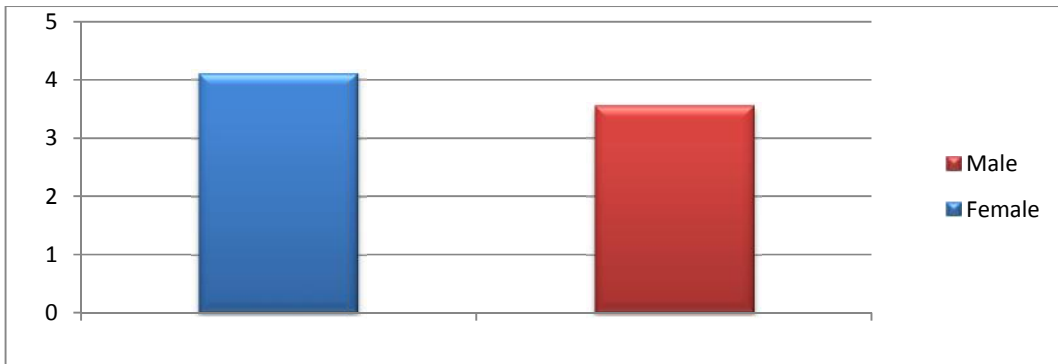
Figure 1 The overall means of learners' intrinsic motivation among male and female

Extrinsic motivation among male and female
Comparing the overall mean scores of the items in intrinsic motivation among male [M=4.15] to the female [M=3.85], it can conclude that the male learners are to a certain extent extrinsically motivated but they still have an intrinsic motivation towards foreign language learning (see figure 2).

Figure 2 The overall means of learners' extrinsic motivation among male and female

Discussion

The study showed that Iranian students had a stronger extrinsic motivation to learn English. The following table shows the gender differences between extrinsic and intrinsic motivation.

Based on the results of this study, it was indicated that students with stronger motivation tend to use more learning strategies than learners with less strong motivation. The results confirmed that learners who learn language for intrinsic reasons tend to be more willing to use various kinds of language learning strategies. In this study, it was shown that gender has significant effects on motivation, especially on intrinsic motivation, indicating that gender may be a critical factor influencing learners' motivation. In addition, it was shown in the previous studies that learning achievement has an important effect on learners' motivation.

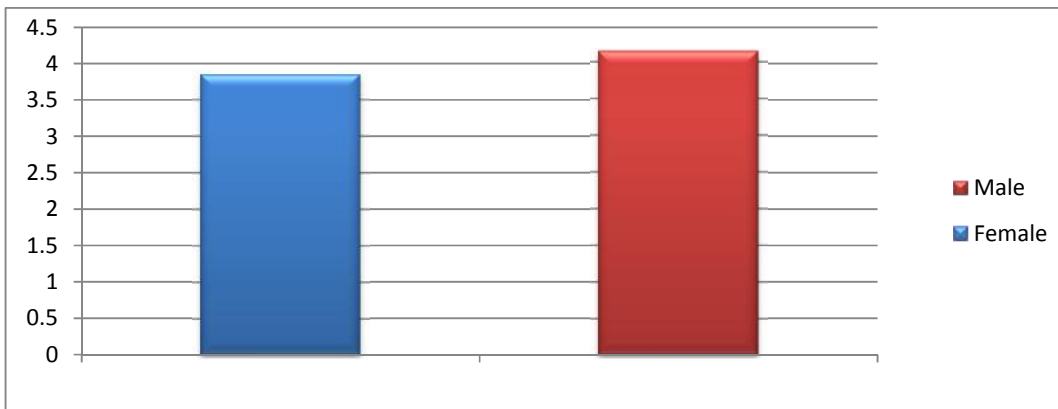This study showed that Iranian EFL students have a stronger extrinsic motivation to learn English. *Female students have stronger intrinsic motivation orientation than extrinsic motivation.* The reasons are as follows: females tend to be more socially dependent and they are more eager to integrate with a social norm, because they see other people's criticism more important than do males. *Male students have stronger extrinsic motivation orientation than intrinsic motivation.* The reason could be their thinking is more career-oriented than females, thus learning English is mainly for vocational purposes. On the whole, this study found that female students had stronger intrinsic motivation while male students had stronger extrinsic motivation.

Conclusion

This study indicated that Iranian EFL students have stronger extrinsic motivation than intrinsic motivation toward language learning. The most important factors affect students' motivation are parents, teachers, social personality, university environment, examinations. Females were found to have stronger intrinsic motivation than males to learn English. Findings of this study have implications for teaching pedagogy. Teachers should put emphasis on increasing students' intrinsic and extrinsic motivation, although findings of the study indicated that males have stronger extrinsic motivation than intrinsic motivation on language learning. Teachers should find out appropriate activities that enhance students' motivation. Teachers should use effective strategies to motivate students of different groups.

They can inform their students the significance and usefulness of these strategies. In order to motivate male and female students to learn English, teachers should adjust their teaching techniques based on their students' needs to enable them to learn English easily. In conclusion, motivating students is necessary to ensure learners' academic growth. Encouragement can make learning more efficient and improve the classroom atmosphere. Recognizing learners' extrinsic motives can be obtained by preparing learners for examinations and focusing more on practical skills like how to communicate with other people when travelling abroad. On the other hand, raising their interests towards the culture of the target language can be done by activities like giving information on the lifestyle, literature of the English-speaking countries through visual, written and audio forms. Successful learners are motivated by both

internal and external factors and both types should be promoted in the classroom by teachers.

Suggestions for motivating language learners

1. Teachers should be sensitive to learners' motives to raise their motivation because both kinds of motivation are required to promote people to learn.

2. Teachers should encourage a balanced development of both kinds of motivation for learners.

3. Teachers can raise learners' intrinsic motivation by enhancing their positive attitudes and correcting their negative stereotypes towards English-speaking countries and people as well as the English language itself.

4. Teachers can improve the contents, teaching methods, classroom activities to raise learners' interests and motivation in language learning.

5. Teachers' feedback can affect learners' motivation. The feedback should be encouraging rather than controlling; otherwise, learners may lose their motivation.

6. Teachers should develop learners' confidence with praise and encouragement in their feedback during or after class.

7. Teachers should encourage positive classroom behaviours to enhance learners' positive attitudes which raise motivation.

8. Teachers should create a pleasant, relaxed atmosphere in the classroom.

9. Teachers should develop a good relationship with the learners, increase the learners' self-confidence, make the language class interesting.

10. Teachers should promote learner autonomy, increase the learners' goal-orientedness, and familiarize learners with the target language culture.

11. Teachers can encourage learners to develop their own intrinsic rewards through positive self-talk, guided self-evaluation, and mastery of specific goals, rather than comparison with other learners. Teachers can promote a sense of greater self-efficacy, increasing motivation to continue learning language.

12. Teachers can learners improve motivation by showing that L2 learning can be an exciting mental challenge, a means to cultural awareness and friendship, and a key to world peace.

References

Boggiano, A. K., & Barrett, M. (1992). Gender differences in depression in children as a function of motivational orientation. *Sex Roles*, 26(1-2), 11-17.

Bomia, L., Beluzo, L., Demeester, D., Elander, K., Johnson, M., & Sheldon, B. (1997). *The impact of teaching strategies on intrinsic motivation*. 1st edition. Champaign, IL: ERIC Clearinghouse on Elementary and Early Childhood Education. 28s ISBN not stated.

Chang, S. M., & Huang, S. H. (1999). Taiwanese English learners' learning motivation and language learning strategies. *Proceedings of the Sixteen Conferences on English Teaching and Learning in the Republic of China, 111-128.*

Dornyei, Z. (2001). *Teaching and Researching Motivation*. London: Longman.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and Motivation in Second Language Learning.* Rowley, Mass: Newbury House.

Hagborg, W. J. (1995). *Gender and Motivational Orientation among High School Students.* Paper presented at the Annual Meeting of the American Psychological Association. New York, NY.

Jakobsson, A. K. (2000). *Motivation and Learning from a Gender Perspective. A study of students at gymnasium in theoretical program.* An unpublished thesis of Göteborg, Göteborgs Universitet: Acta Universitatis Gothoburgensis.

Liao, Y. F. (2000). Taiwanese secondary school students' learning motivation and language learning strategies in EFL learning, 155-183.

Nilsson, K., & Stomberg, M. I. W. (2008). Nursing students' motivation toward their studies – a survey study. *BMC Nursing* 7(6).

Noels, K. A., Pelletier, L., & Pelletier, L. G. (1999). Perceptions of teacher communicative style and students' intrinsic and extrinsic motivation. *Modern Language Journal, 83, 23-34.*

Noels, K. A., Clement, R., & Pelletier, L. G. (2001). Intrinsic, extrinsic, and integrative orientations of French Canadian learners of English. *The Canadian Modern Language Journal, 57(3), 425-442.*

Pintrich, P, R., & Schunke, D. H. (2002).Motivation in education: Theory, research, and application. (2nd, ed). New Jersey: Merill Prentice Hall.

Ryan, M. R., & Deci, E. L. (2000). Intrinsic and extrinsic motivation: classic definitions and new directions. Contemporary Educational Psychology, 25, 54-67.

Sadeghi, A. R. (2005). ESP in Iran: A transition from present state. In: G. R Kiany, & M. Khayyamdar, (Eds.), *Proceedings of the First National ESP/EAP Conference,* 2, Tehran.

Appendix
Motivations for Learning English
I would like to find out what motivates you to learn English. Please look at the statements below and indicate how much you agree or disagree with them. Circle the number that corresponds to your own opinions.

I want to learn English, because:	*Strongly agree*	*Agree*	*Neutral*	*Disagree*	*Strongly disagree*
1. I want to travel overseas in the future.	5	4	3	2	1
2. I want to further my studies in the future	5	4	3	2	1
3. English can help me to make	5	4	3	2	1

friends with people of different nationalities and background.					
4. English can enhance my chances of emigrating to other countries in the future.	5	4	3	2	1
5. English is the way to gain more knowledge	5	4	3	2	1
6. English can raise my social status.	5	4	3	2	1
7. A good standard of English can help me do	5	4	3	2	1
8. English can help me to find a better job in the future.	5	4	3	2	1
9. English can help me understand Western culture.	5	4	3	2	1
10. It can help me to broaden my horizons.	5	4	3	2	1
11. English can enable me to appreciate Western films and music.	5	4	3	2	1
12. It can satisfy my interests and curiosity.	5	4	3	2	1
13. English is the mark of an educated person	5	4	3	2	1
14. It is part of my schoolwork.	5	4	3	2	1
15. I like conversing with foreigners	5	4	3	2	1
16. I want to pass my public examinations.	5	4	3	2	1

Thank you very much for your

help.

Title

A Study on Iranian University Students' and Teachers' Beliefs about Language
Learning

Authors

Sanaz Ghobadi Mohebi (M.A.)

Islamic Azad University, Mashhad Branch, Mashhad, Iran

Ebrahim Khodadady (Ph.D.)

Ferdowsi University of Mashhad, Iran

Bio Data

Sanaz Ghobadi Mohebi, Graduated in Teaching English as a foreign language (M.A)
from Ferdowsi University of Mashhad. She is a lecturer at the department of English
language, Islamic Azad University, Mashhad Branch, Iran. Her major interest are
Teaching English, Language learning, psychology of Language, and Individual
Differences.

Ebrahim Khodadady is assistant professor of Applied Linguistics at Ferdowsi
University of Mashhad, Iran. He has offered undergraduate and graduate courses such
as Advanced Writing, Advanced Speaking, Teaching Language Skills, Testing,
Research Methods and Principles and Syllabus Design in Canada and Iran. He has
published widely in the field.

Abstract

The aim of the study is to investigate beliefs students usually held about language
learning, based on the Beliefs About Language Learning Inventory (BALLI)
questionnaire (Horwitz, 1988) and comparing them with teachers. For this purpose,
423 Iranian learners of English were selected. Running descriptive statistics and the
scree plot test, five factors were extracted: Formal Learning and Motivation for
English Learning, Learner's Confidence in Learning English, Strategy and Attitude in
Learning English, Aptitude and Strategy in Learning English, and the Importance of
English and Formal Learning. Evidence of the relative similarities and significant
differences between teachers' and learners' beliefs about language learning on nine
items was provided. The findings of the current study suggest that teachers should be

aware of learners' beliefs as well as their own in order to assist less successful language learners.

Introduction

Beliefs about language learning refer to learners' notions, perceived ideas, insights, concepts, opinions, representations, assumptions, expectations or mini-theories of the nature of language or language learning (Hong, 2006). It is generally agreed that individual language learners hold different beliefs about how language is learned. The belief systems learners hold or develop help them to understand and adapt to new environments, to define what is expected of them and to act in accordance with those understandings (White, 1999). Influenced by previous experiences of language learners, or shaped by their own cultural backgrounds, second language learners often hold different beliefs or expectations about language learning that would likely affect the way they use their learning strategies and learn a second language. In case where these beliefs differ significantly from teacher ideas, they would influence learner satisfaction with the course (Yang, 1999).

Since both language students and teachers bring their unique sets of beliefs to bear in situations and decisions related to language learning and teaching, understanding students' beliefs about English learning is essential for the teacher to provide appropriate English instruction (Frugé, 2007). Liao and Chiang (2003) pointed out that the learning beliefs were often based on the previous learning experiences and cultural backgrounds, and would further influence strategies people would use to enhance their English learning and teaching. Learners take different approaches to language learning because they have different beliefs about language learning and also rich learning experience seems to be associated with more refined beliefs (Mori, 1999; King, 2000). As Horwitz (1985, 1987, & 1999) insisted, understanding the beliefs of learners is important because it helps teachers to understand learners' approaches to language learning and learners' use of learning strategies better, so that they can plan language instruction appropriately.

Teachers who have access to their learners' beliefs may choose to reinforce or to challenge certain beliefs. In either case, they will be better equipped to engage in meaningful dialogue about learning with their learners (Cotteral, 1999). In recent years, researchers have realized the important influence of language beliefs in foreign language learning process and have increasingly focused on students' beliefs about language learning and their effect on students' motivation, anxiety and the use of strategy (Horwitz, Horwitz & Cope, 1986; Kuntz, 1996; Banya and Chea, 1997; Mori, 1999; Sakui & Gaies, 1999; Yang, 1999; Horwitz, 2001; Bernat, 2002; Gregersen & Horwitz, 2002; Lee, 2002; Liao & Chiang, 2003; Le, 2004; Diab, 2006; Hong, 2006; Huang, 2006; Camille Bakker, 2008). Horwitz (1985) developed the Beliefs About

Language Learning Inventory (BALLI) to assess students' opinions on a variety of issues and controversies related to language learning via free-recall protocols and group discussions with both foreign language and ESL learners and teachers. The BALLI has become a popular instrument for investigating beliefs about language learning. This questionnaire has proven very effective in increasing student learning as well as student satisfaction with the course. Altan (2006) asserted that BALLI can be helpful to language teacher educators both by determining popular beliefs of their students who are going to be teachers in future as well as in identifying minority groups with different opinions. Research on the topic, since Horwitz's pioneering study in 1985, has indicated that some of these beliefs are detrimental to learning.

Mantle-Bromley (1995) used the BALLI to investigate the beliefs of 208 seventh grade middle school students. Similar to Horwitz's study (1988), the students underestimated the difficulty of language learning to a certain extent. She found that positive attitudes and realistic beliefs had "links to proficiency" and recommended teachers to design and implement lessons on the language learning process that incorporated attitude-change methods. Banya and Chea (1997) found that students with positive beliefs about foreign language learning tended to have stronger motivation hold favorable attitude and higher motivational intensity, use more strategies, are less anxious, have better language achievement and are more proficient. Yang's (1999) study on Taiwanese students concluded as beliefs about language learning can affect the use of strategies, learning strategies may also influence learners' beliefs about language learning. Peacock (2001) investigated the beliefs of 146 undergraduate trainee teachers in a three-year study in Hong Kong and found a positive association between their beliefs on vocabulary and grammar and proficiency. Siebert's (2003) findings revealed that male were much more optimistic about their own abilities and the length of time it would take them to learn English, and they were almost twice as likely as females to endorse excellent pronunciation, and more than twice as likely to view grammar learning as the most important part of language acquisition. To identify the relationship between students' and teachers' conceptions on English learning Liao and Chiang (2003) did a study on a total of 143 students and 15 teachers in Taiwan. Their findings showed that students' and teachers' responses to the BALLI were on the whole rather consistent and strikingly similar with one another; but not in the areas of the nature of language learning and learning and communication strategies. Students' preference of learning grammar and translation did not go well with their teachers' communicative teaching methods, leading to a possible lack of learner satisfaction or self-confidence in the classroom.

In an extensive study on 428 monolingual Korean and 420 bilingual Korean-Chinese university students, Hong (2006) investigated the relationship between the learners' beliefs and their learning strategy use and also the influence of background variables. Significant influences of the individual variables of academic major and self-rated English proficiency on strategy use and beliefs concerning language learning were found. Diab (2006) supported the general contention that different cultural backgrounds, background variables within group and variation in a particular group's belief about learning different target languages are influential factors on

learner belief. His findings indicated that learning a foreign language seemed to be related to the political and socio-cultural context. Bernat (2006) found that the beliefs held by participants in the Australian and American context were similar in all categories. Statistically significant differences were also claimed with respect to gender.

Although the BALLI and its modified or enlarged versions have been used in eliciting learners' beliefs about language learning in studies of foreign language learners in the U.S., ESL learners in English speaking countries and EFL learners in foreign countries, to date there is no comparable study on learners' beliefs about language learning done in Iran, especially regarding the relation between belief system of students and teachers. Therefore, this article explores the beliefs of both teachers and students and argues that teachers should try to eliminate any detrimental beliefs in their trainees before they start to teach ESL, to ensure that they do not form their trainees' teaching.

Method
Participants
To investigate the beliefs about language learning, 423 students from four universities and three Teacher Training Centers in Iran were selected. Among the 423 participants, 299 students had not taught English before (70.7 %) and 124 students had (29.3%). Most of the students who had the experience of teaching English had taught English for about one year (about 10.0% of the 299 students). In terms of academic degree, 419 of the participants were undergraduate and 4 participants were students of PhD. Most of the students were juniors (52.3%), 23.1% were sophomores and 24.7% were seniors. In terms of gender, the students were not balanced with 108 males and 315 females. Their age ranged from 19 to 49 with an average age of 22.41. Less than 4% of the subjects had the experience of living abroad.

Instruments
In order to collect additional information on the individual background, the researcher designed the Personal Information Questionnaire. It contained items related to demographic information of the participants (e.g., age, gender, institution, academic major, degree of study, and year of study).

The EFL BALLI version was used to identify the beliefs held by EFL learners about language learning. The version of the ESL/EFL BALLI used for this study contained 34 items. The items on the BALLI assess learners' beliefs in five areas: 1) the difficulty of language learning (six items), 2) foreign language aptitude (nine items), 3) the nature of language learning (six items), 4) learning and communication strategies (eight items), and 5) motivation and expectations (five items). BALLI was designed based on free recall tasks asking respondents to list beliefs about language learning. Horwitz (1985) firstly developed the BALLI to assess beliefs about language learning. Horwitz originally designed three different versions of BALLI; one for foreign language teachers (1985) with 27 items, another one for ESL students (1987) with 27 items, and a third for U.S. students learning a foreign language (1988) with 34

items. Items were scored on a five-point Likert scale: A=strongly disagree, B=disagree, C=neither agree nor disagree, D=agree, E=strongly agree.

Data collection and analysis

The instruments were distributed during class time preceded by a brief explanation of the purpose and nature of the study. After the completion of the instrument, the questionnaires were collected by the researcher for data analysis. First, descriptive analysis was run to summarize the learners' beliefs about language learning item by item. The principal component analysis and a scree test were used to extract the set of variables using a criterion of eigen values equal to or greater than 1.0. The varimax rotation (uncorrelated factor rotation) was run to yield the final factor loadings and to increase interpretability of the underlying factors (Kim & Mueller, 1978). Then students were divided into two groups of students and teachers and descriptive analysis was used to compare their beliefs.

Results

Results of descriptive analyses and factor analyses on the BALLI items

The BALLI was used to examine the learners' beliefs about language learning. Descriptive statistics were computed on the students' responses to the BALLI items. In order to investigate patterns in the subjects' responses to the 34 questionnaire items, a factor analysis was performed. Based on the results of principal component analyses on this study, 14 factors were found as an initial solution on the BALLI. However, after the application of the scree plot test, again to refine the factor dimensions, five factors were extracted: 1) Formal Learning and Motivation for English Learning, 2) Learner's Confidence in Learning English, 3) Strategy and Attitude in Learning English, 4) Aptitude and Strategy in Learning English, and 5) the Importance of English and Formal Learning. These five factors accounted for 28.13% of the total variance for the students and included 28 of the 34 items. A varimax rotation test was applied to make the factors more interpretable. Table 1 presents the final factor loading of the BALLI items. The scree plot is shown below.
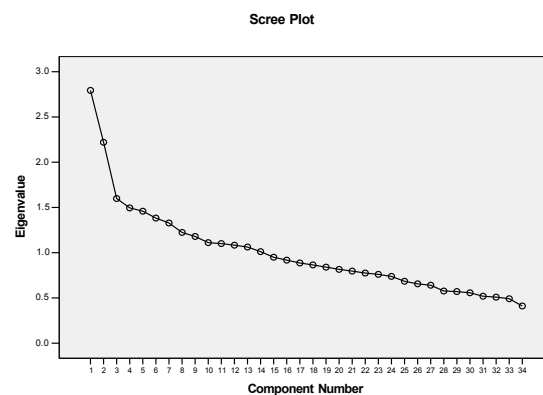
Table 1. Rotated Factor Structure of the BALLI Variables

Items	F 1	F 2	F 3	F 4	F 5
I19	0.64				

I15	0.63				
I31	0.49				
I25	0.48				
I07	0.41				
I20	0.35				
I26	0.32				
I30	0.31				
I22	0.31				
I14		0.60			
I06		0.57			
I04		-0.57			
I17		0.51			
I10		0.30			
I24			0.53		
I11			0.49		
I08			0.44		
I13			0.39		
I01			0.37		
I12			0.33		
I28				0.63	
I23				0.56	
I09				0.53	
I02				0.40	
I32					0.47
I33					0.38
I29					0.32
I05					0.32

Note: Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
A. Rotation converged in 10 item rations

Tables 2 through 6 present the items that loaded with a value equal to or greater than ± .30 (Hatch and Lazaraton, 1991), overall frequency (%) on these items, the means and the standard deviation, as well as the name of each factor and the number and content of the items[i].

Table 2. BALLI Factor 1. Formal Learning and Motivation for English Learning

Item Description	1*	2	3	4	5	Load	M	SD
Learning English is mostly a matter of learning many of grammar rules.	5.2	25.8	17.0	42.3	9.7	0.64	3.26	1.10
Learning English is mostly a matter of learning many new vocabulary words.	14.2	38.8	12.5	29.3	4.0	0.63	2.67	1.18
People who speak more than one language well are very intelligent.	24.3	44.0	16.1	12.8	2.6	0.49	2.25	1.04
Learning English is mostly a matter of translating from English into Persian.	2.8	9.7	14.9	43.5	28.8	0.48	3.85	1.04
It is important to speak English with an excellent accent.	37.8	38.1	10.9	11.6	1.7	0.41	2.01	1.05
It is important to practice in the language laboratory.	24.1	46.3	17.0	11.1	1.4	0.41	2.19	0.97
If I learn to speak English very well, it will help me to get a good job.	34.0	41.4	13.7	9.2	0.7	0.35	1.98	0.97
I would like to learn English so that I can get to know its speakers better.	21.0	47.5	22.0	6.4	2.6	0.32	2.21	0.94
If I get to speak English very well, I will have many opportunities to use it.	44.2	40.7	8.0	4.5	2.1	0.31	1.78	0.93

Note. * 1= Strongly agree; 2= Agree; 3= Neither agree nor disagree; 4= Disagree; 5= Strongly disagree
+ The remained percentage in some items belongs to the blank answers.

As shown in Table 2, Factor 1 was labeled as formal learning and motivation for English learning because many items (19, 15, 31, 25, 7, 20, 26, 30 and 22) related to motivational beliefs and formal learning of English were identically loaded on Factor 1.

More than half of the students (52%) disagreed that grammar plays an important role in language learning. Regarding the importance of vocabulary in language learning, fewer students (33.3%) disagreed with the importance of learning new vocabulary than agreed with (53%). A great number of students (72.3%) disagreed that translating from English into Persian played the most important role in English learning. Many of the students (75.9%) believed excellent pronunciation was important. A great number of participants (70.4%) felt that "It is important to practice in the language laboratory". This case also happened for getting a good job; about 75% agreed that "If I learn to speak English very well, it will help me to get a good job". The majority of them agreed (85%) that "If I spoke English well I will have many opportunities to use it". A larger number of students (68.5%) believed that they would like to learn English so that they could get to know English speakers better.

Factor 2 had five items (14, 6, 4, 17, and 10) which dealt with beliefs about Learner's Confidence in Learning English (see Table 3). About three quarter (71%) disagreed that they had an aptitude for learning English. The majority of the participants (82.7%) believed that "I will ultimately learn to speak English very well". When asked about the difficulty involved in learning English, most of the students did not consider English as a difficult language (81.8%). Also, 67% of students respectively agreed that it was easier for someone who already spoke a foreign language to learn another one.

Table 3. BALLI Factor 2. Learner's Confidence in Learning English

Item Description	1*	2	3	4	5	Load	M	SD
. I have an English aptitude. i.e., have the ability to learn it.	21.3	49.6	21.5	5.4	0.9	0.60	2.12	0.87
I believe I will ultimately learn to speak English very well.	36.6	46.1	9.9	5.4	1.7	0.57	1.89	0.91
Learning English is very difficult.	2.1	5.7	9.5	56.7	25.1	-0.57	3.94	0.94
. I feel self-conscious speaking English in front of other people.	15.1	32.2	26.2	19.6	5.9	0.51	2.66	1.15
. It is easier for someone who already speaks a foreign language to learn another one.	21.0	46.1	23.4	6.9	2.4	0.30	2.23	0.94

Note. * 1= Strongly agree; 2= Agree; 3= Neither agree nor disagree; 4= Disagree; 5= Strongly disagree
+ The remained percentage in some items belongs to the blank answers.

Table 4. BALLI Factor 3. Strategy and Attitude in Learning English

Item Description	1*	2	3	4	5	Load	M	SD
. Learning English is different from learning other school subjects.	17.0	43.7	21.5	13.2	4.3	0.53	2.43	1.06
. It is better to learn English in an English speaking country.	61.9	27.0	4.7	3.5	2.8	0.49	1.58	0.94
It is necessary to know English culture in order to speak it.	22.7	38.3	19.4	17.0	2.6	0.44	2.39	1.09
. It is okay to guess if you do not know a word in English.	32.6	48.7	10.4	5.7	2.1	0.39	1.95	0.93
It is easier for children than adults to learn English.	53.2	37.8	3.5	4.5	0.7	0.37	1.61	0.81
. If I heard some people speaking English, I would go up to them so that I could practice speaking the language.	27.4	41.6	25.1	5.0	0.9	0.33	2.10	0.89

Note. * 1= Strongly agree; 2= Agree; 3= Neither agree nor disagree; 4= Disagree; 5= Strongly disagree
+ The remained percentage in some items belongs to the blank answers.

As shown in Table 4, factor 3 included six items (24, 11, 8, 13, 1, and 12) related to strategy and attitude in learning English. More than half of the students (60.7%) agreed that learning English was different from learning other academic subjects. The majority of the participants (88.9%) felt that it would be better to learn English in an English speaking country. Many of the students (61%) believed that "it is necessary to know English culture in order to speak English". The students were greatly disposed (81.3% agreement) toward guessing unknown words in English. About 90% of the students agreed that it was easier for children than for adults to learn English. Only 5.9% of the participants were timid in going up and practicing English if they heard some people spoke English.

Table 5. BALLI Factor 4. Aptitude and Strategy in Learning English

Item Description	1*	2	3	4	5	Load	M	SD
People who are good at math and science are not good at learning English.	1.9	5.7	33.3	32.4	25.8	0.63	3.72	1.03
It is easier to speak than understand English.	5.4	14.2	17.5	44.9	17.7	0.56	3.55	1.11
You should not say anything in English until you can say it correctly.	4.0	10.4	5.9	42.6	37.1	0.53	3.98	1.10
Some people are born with a special ability which helps them learn English.	11.3	44.9	19.6	19.4	4.7	0.40	2.61	1.06

Note. * 1= Strongly agree; 2= Agree; 3= Neither agree nor disagree; 4= Disagree; 5= Strongly disagree
+ The remained percentage in some items belongs to the blank answers.

As shown in Table 5, items 28, 23, 9, and 2 loaded on Factor 4, aptitude and strategy in learning English. When asked about the item "People who are good at mathematics or science are not good at learning foreign languages," only 7.6% of students agreed. With regard to ease of skills, only 19.6% felt that speaking was easier than understanding. However, 79% students were much more likely to disagree that they should not say anything in English until they could say it correctly, indicating their greater belief in the importance of accuracy. More than half of the participants (55%) believed that some people had a special ability for learning English.

The last factor, factor 5, had four items (32, 33, 29, and 5) which dealt with the importance of English and formal learning of English (Table 6). An interesting number of participants (78.9%) strongly agreed or agreed that everyone could learn to speak English. Only 12.5% did not feel that "Iranians are good at learning foreign languages". Most of the participants (68.4%) did not believe English was structured in the same way as Persian. When asking about "Iranians think that it is important to speak English", 63.6% agreed.

Table 6. BALLI Factor 5. The Importance of English and Formal Learning

Item Description	1*	2	3	4	5	Load	M	SD
. Iranians are good at learning English.	10.6	39.2	36.9	10.4	2.1	0.47	2.52	0.91
	31.9	46.8	11.1	7.6	0.9	0.38	1.94	0.93
. Everyone can learn to speak English.	16.3	47.3	21.3	10.9	3.3	0.32	2.35	2.35
. Iranians think that it is important to speak English.	2.1	8.5	21.3	49.9	17.5	0.32	3.70	0.97
English is structured in the same way as Persian.								

Note. * 1= Strongly agree; 2= Agree; 3= Neither agree nor disagree; 4= Disagree; 5= Strongly disagree
+ The remained percentage in some items belongs to the blank answers.

Beliefs of Students and teachers

In order to see what influence teaching of English can have on a person, the participants were divided into two groups. The first group was the participants who had not taught English previously. The second group was the ones who had the experience of teaching English. The results come below.

Table 7. Students' and Teachers' Beliefs about Language Learning

Item		1		3
It is easier for children than adults to learn English.				
	Students	94.3	2.7	3.0
	Teachers	83.9	5.6	10.5
8. It is necessary to know English culture in order to speak it.				
	Students	57.9	20.7	21.4
	Teachers	68.5	16.1	15.3
. Learning English is mostly a matter of learning many new vocabulary words.				
	Students	57.5	11.6	31.0
	Teachers	44.4	15.3	40.3
. I feel self-conscious speaking English in front of other people.				
	Students			
	Teachers	43.1	27.9	29.0
		59.0	23.0	18.0
. Learning English is mostly a matter of learning many of grammar rules.				
	Students	33.8	17.4	48.8
	Teachers	24.2	16.1	59.7
. Learning English is mostly a matter of translating from English into Persian.				
	Students	16.1	14.8	69.1
	Teachers	4.0	15.3	80.6
. If I learn to speak English very well, it will help me to get a good job.				
	Students	79.4	10.8	9.8
	Teachers	68.3	21.1	10.6
. I would like to learn English so that I can get to know its speakers better.				
	Students	72.1	20.5	7.4
	Teachers	61.3	25.8	12.9

On eight items (1, 8, 15, 17, 19, 25, 26, and 30), the answers may have implications for the learning and teaching of EFL. Since the teachers were mostly inexperienced student teachers, the reported differences were not so considerable. Therefore, difference of 10% or more was chosen as contrast. In order to have easier interpretation of the results and for more clarity, strongly agree and agree (1) and strongly disagree and disagree (3) are grouped together. Number 2 shows the percentage of neutrality. All results are expressed as percentages:

On item 1, 94.3% of students and 83.9% of teachers agreed that "It is easier for children than adults to learn English". On item 8, about 58% of students believed that it was necessary to know English culture in order to speak it; whereas, 68.5% of teachers believed in it. On item 15, "Learning English is mostly a matter of learning many new vocabulary words", 57.5% and 44.4% of students and teachers agreed; 31.0% and 40.3% disagreed respectively.

On item 17, "I feel self-conscious speaking English in front of other people", a difference of 16% was found. On item 19, only 24.2% of teachers believed grammar

rules were important, about 60.0% did not believe so. These percents were 33.8% and 48.8% respectively for students. On item 25, about the importance of translating from English to Farsi for learning English, only 4.0% of teachers agreed; about 80.0% disagreed. The percent of students' agreement and disagreement were 16.1% and 69.1% respectively. Among students, 79.4% believed in item 26, "If I learn to speak English very well, it will help me to get a good job"; 68.3% of teachers agreed with this notion. As for item 30, about 72.0% of students and 61.0% of teachers felt that they liked to learn English so that they could get to know English speakers better.

Discussion

In the present study, the beliefs of English students about language learning were explored, as measured by the BALLI (Horwitz, 1988). It was reported that students hold a variety of beliefs about language learning.

Ninety-five percent of the subjects agreed with item 7 ("It is important to speak English with an excellent accent"). People in Iran usually place a high value on English proficiency. Perhaps, since most Iranian students have perfectionism view toward learning English, speaking English with an excellent pronunciation is particularly important for them to attain this goal. Sixty-nine percent of the subjects agreed with item 12 ("If I heard some people speaking English, I would go up to them so that I could practice speaking the language") (n= 292). Wanting to speak English with native speakers is probably one of the main reasons why these students came to study English. Eighty-one percent of the subjects overall agreed with item 13 ("It is O.K. to guess if you don't know a word in English") (n= 344). Findings may show that Iranians are risk taker and ready to guess (81.0%). Iranian alphabetical system consists of many components which are related to the sound or the meaning of words and therefore, makes guessing strategy very useful in memorizing, recognizing and recalling a word. The agreement rate for item 11 ("It is best to learn English in an English speaking country") is so high, i.e. 89% (n= 376). This might imply that ethnic backgrounds might influence beliefs about learning target languages. It may also show that the beliefs and attitudes on the importance of knowing culture in target language learning are almost important for Iranians.

The responses to item 14 ("I have a foreign language aptitude") would appear that many of the subjects see themselves as gifted language learners, with an adequate ability to the task of language learning. Iranian students have a very high estimate of their aptitude in learning English. Half of the subjects agreed with item 17 "Learning English is mostly a matter of learning many new vocabulary words" (53% agreement and 33.3% disagreement). The findings on item 15 show that learning vocabulary words has great value among Iranian (53.0%). One of the likely reasons is culture, learning environments and teaching methods in Iran. It seems that in Iran texts and readings are usually emphasized. In such a learning environment, it is understandable why students studying English attach more importance to learning vocabulary. These beliefs may lead students to invest their time memorizing vocabulary lists at the expense of other language learning tasks. About 70% of students agreed on item 8 ("It is necessary to know the foreign language culture in order to speak the foreign

language."). Also, 69% of the subjects agreed with item 30 "I would like to learn English so that they know English speakers better". The data demonstrate that the Iranian students have a strong integrative motivation. These findings imply that ethnic backgrounds might influence beliefs about learning target languages. Iranians consider culture of the target language so significant; they think that foreign languages are so important and knowing their speakers will yield so many benefits. They attach so importance to and are so enthusiastic toward culture. They show more integrative motivation for learning a language.

The factors found in this study include Motivation and Formal English Learning, Learner's Confidence in Learning English, Strategy and Attitude in Learning English, Aptitude and Strategy in Learning English, and The Importance of English and Formal Learning. The special combination of formal learning and motivation as the first factor may suggest the attention participants in this study pay to the formal language learning and their motivation for learning it. It seems likely that they are motivated both instrumentally and integrativelly to learn formal English rules and any related subjects to get better situations. Contrary to the expectation, having confidence as the second factor and the high percentage of students' agreement on the related items may show a new picture of the Iranians' self-efficacy in learning English and, therefore, more attention on the part of teachers to employ the learners' confidence in teaching is needed. The third and fourth factors might show the distinct characteristic and beliefs of Iranian students toward learning strategies. The last factor, the importance of English and formal learning, mainly shows Iranian beliefs about themselves, the importance of English and the structural differences between the two languages.

Conclusion

The present study has identified some unique and important beliefs of Iranian students studying English. It can be pointed out that the items with which respondents agreed most strongly (items which had a percentage of 80 or higher) cover a variety of beliefs. It was shown that students reported holding various opinions about language learning. For instance, the majority of students endorsed the easiness of some languages, superiority of children in learning English, their ultimate success in language learning, guessing the meaning of unknown words, and the importance of repetition and practice in learning English. A great number of them also stated that "it is better to learn English in an English speaking country" and "everyone can learn to speak English".

Furthermore, the participant in this study held not only similar but also different beliefs concerning language learning from those of American foreign language learners (Horwitz, 1988). Some findings of the present study contrast with those conducted in various learning and cultural contexts to support the argument that learners' beliefs are influenced by the different language learning contexts (ESL, EFL, or FL), educational or cultural background, and stages of language learning. Students in the present study, for example, held strong instrumental motivation for learning English, possibly, because of self imposed or other-imposed pressures at

home. They were also motivated to learn English for academic purposes and better job opportunities, more so than the American students in Horwitz's (1988) work. Interestingly, they expressed a high agreement on the area of aptitude and self-efficacy and believed that they would ultimately be successful in learning and speaking English. Their percentage of agreement was more than Horwitz's. Furthermore, the students saw English as an easy language and believed "Iranians are good at Learning English". They were reported having more self-confidence, comparing with Americans.

In addition to differences among students' beliefs, some differences were found between students' and teachers' beliefs on items 1, 8, 15, 17, 19, 25, 26, 30, and 33. Compared with teachers, more students believed the superiority of children in language learning, importance of learning new vocabulary, importance of learning grammar rules, and importance of translation. More students also liked to learn English so that they could get to know English speakers better. On the other hand, more teachers endorsed guessing the meaning of unknown vocabulary and having self-consciousness in speaking English in front of people. They also stated that if they learned to speak English very well, it would help them to get a good job.

In terms of pedagogical implications of learner-beliefs research, identification of beliefs and reflection on their potential impact on language learning and teaching in general, as well as in more specific areas such as learners' expectations and strategy use, can inform future syllabus design and teacher practice in an EFL course as well as increased awareness and adjustment of expectations concerning language learning (Bernat & Llyod, 2007). More importantly, teachers should be more aware of the conflicting beliefs between them and students, and try to integrate these conflicting beliefs into teaching through strategy training in order to develop greater flexibility in learners' ways of approaching English.

References

Altan, M. Z. (2006). Beliefs about language learning of foreign language-major university students. Erciyes University. *Australian Journal of Teacher Education, 31 (2),* 45-51.

Bakker, S. C. (2008). Byu students' beliefs about language learning and communicative language teaching activities. Unpublished PhD dissertation, Brigham Young University.

Banya, K., & Chen, M. (1997). Beliefs about foreign language learning: A study of beliefs of teachers' and students' cross cultural settings. Paper presented at the 31st Annual Meeting of the Speakers of Other Languages, Florida. (*ERIC Document Reproduction Service* No. ED411691).

Bernat, E. (2002). Investigating Vietnamese ESL Learners' beliefs about language learning. EA Journal, 21 (2), 40-54.

Bernat, E. (2006). Assessing EAP learners' beliefs about language learning in the Australian context. *Asian EFL Journal*, 8 (2).

Bernat, E., & Lloyd, R. (2007). Exploring the Gender Effect on EFL Learners' Beliefs about Language Learning. *Australian Journal of Educational & Developmental Psychology. 7,* 79-91.

Cotteral, S. (1999). Key variables in language learning: what do learners believe about them? *System, 27 (1999),* 493-513.

Diab, R. L. (2006). University students' beliefs about learning English and French in Lebanon. *System, 34 (1),* 80-96.

Frugé, Ch. (2007). Epistemological congruency in community college classrooms: Effects of epistemological beliefs on students' experiences. Unpublished doctoral dissertation. Louisiana State University and Agricultural and Mechanical College. (*ERIC Document Reproduction Service* No. EJ794367).

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston, Mas: Heinle & Heinle.

Hong, K. (2006). Beliefs about language learning and language learning strategy use in an EFL context: a comparison study of monolingual Korean and bilingual Korean-Chinese university students. Unpublished doctoral dissertation. University of North Texas.

Horwitz, E. K. (1985). Using student beliefs about language learning and teaching in the foreign language methods course. *Foreign Language Annals, 18 (4),* 333-340.

Horwitz, E. K. (1987). Surveying student beliefs about language learning. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 119-129). Englewood Cliffs, NJ: Prentice/Hall International.

Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal, 72 (3),* 283-294.

Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70 (2),* 125-132.

Horwitz, E. K. (1999). Cultural and situational influences on foreign language learners' beliefs about language learning: A review of BALLI studies. *System, 27 (4),* 557- 576.

Horwitz. E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics, 21,* 112-126.

Horwitz, E. K., & Gregersen, T. (2002). Language learning and perfectionism: Anxious and non-anxious language learners' reactions to their own oral performance. *The Modern Language Journal, 86 (4),* 562-570.

Huang, Z. W. (2006). Learner Beliefs of Language Learning Revisited. *Sino-US English Teaching, 3 (3),* 62-67.

Kim, J. O., & Mueller, C. W. (1978). *Introduction to factor analysis*. Beverly Hill, CA: Sage Publications.

King, P. M. (2000). Learning to make reflective judgments. *New Directions for Teaching and Learning, 82,* 15-26.

Kuntz, P. S. (1996). Beliefs about language learning: The Horwitz model. HEA Title VI African Studies Centers. Madison, WI: University of Wisconsin,

Department of Curriculum and Instruction. (*ERIC Document Reproduction Service* No. ED 397649).

Le, J. (2004). Affective characteristics of American students studying Chinese in China: A study of heritage and non-heritage learners' beliefs and foreign language anxiety. Unpublished doctoral dissertation. Texas University, Austin.

Lee, S. I. (2002). The beliefs and values of Japanese of low language proficiency. Center for Comparative Studies in Race and Ethnicity and Ryukoku.

Liao, P. S., & Chiang, M. Y. (2003). The study of students' and their teachers' beliefs about English learning. International Conference on English Teaching and Learning in the Republic of China.

Mantle-Bromley, C. (1995). Positive attitudes and realistic beliefs: Links to proficiency. *Modern Language Journal*, *79 (3),* 372-386.

Mori, Y. (1999). Epistemological beliefs and language learning beliefs: What do language learners believe about their learning? *Language Learning, 49 (3),* 377-415. (*ERIC Document Reproduction Service* No. ED406657).

Peacock, M. (1999). The links between learner beliefs, teacher beliefs, and EFL proficiency. *Perspective, 10 (1),* 125-159.

Peacock, M. (2001). Pre-service ESL teachers' beliefs about second language learning: a longitudinal study. *System, 29(2),* 177-195.

Sakui. K. & Gaies, S. J. (1999). Investigating Japanese Learners' beliefs about language learning. *System, 27(4),* 473-492.

Siebert, L. L. (2003). Student and teacher beliefs about language learning. *The ORTESOL Journal, 21,* 7-39.

White, C. (1999). Expectations and emergent beliefs of self-instructed language learners. *System, 27 (4),* 443-457.

Yang, N. D. (1999). The relationship between EFL learners' beliefs and learning strategy use. System, 27 (4), 515-535.

Title

Refusals in English and Persian: A Pragmalinguistic Investigation[1]

Authors

Naser Atasheneh (M.A.)

Islamic Azad University, Abadan Branch, Iran.

Ahmad Izadi (Ph.D. candidate)

Islamic Azad University, Abadan Branch, Iran.

Bio Data

Naser Atasheneh is a full-time academic member of the Department of English, Islamic Azad University, Abadan Branch, Abadan, Iran. He received his B.A. from Allame tabatabaie University and M.A. from Tarbiat Moallem University. His research interests include Syllabus Design, Pragmatics and SLA.

Ahmad Izadi is faculty member of the Department of English, Islamic Azad University, Abadan Branch, Abadan, Iran. He is also a Ph.D. candidate at the faculty of Languages and Linguistics, University of Malaya, Malaysia. He received his B.A. and M.A. in TEFL from Shiraz University and Shahid Chamran University of Ahvaz, respectively. He has also presented articles in three international conferences. His research interests include pragmatics, politeness theory and academic discourse.

Abstract

This study investigated the similarities and differences between Iranian EFL learners' use of English and Persian refusals, using role play scenarios. It also examined the influence of participants' native language on the production of refusals in English as a foreign language. How the refuser-refusees' social distance and power relations affected the refusers' choice of strategies was another question at issue. The informants of the study were 30 Iranian senior level students majoring in English. The study was done in two phases with an interval time distance of about three months to reduce the probable effects of the first phase on the second one. The same subjects participated in both phases. Once they acted out the role play situations in English and once in Persian. The role plays' provided data were recorded, transcribed and analyzed to show the average frequencies and length of direct and indirect strategies, the types of employed strategies, and the effects of addressees' power and gender on the responses. The results showed that role play interactants used more indirect strategies in Persian in comparison to English. Also, both the interlocutors' (refusee's)

power and social distance affected the type, frequency and length of the strategies used by these EFL learners.

Introduction

Research in the field of 'interlanguage pragmatics' is increasingly broadening its scope to find ways to help foreign/second language learners develop communicative competence; " the ability [they need] to use language effectively in order to achieve a specific purpose and to understand language in context" (Thomas, 1983, p. 94). It has been frequently observed that EFL/ESL learners often appear to lack the pragmatic knowledge of the target language, and hence fail to communicate effectively; their attempt to communicate with the native speakers of the target language is more likely to result in intercultural miscommunication. One of the main sources of miscommunication, is these learner's inability to perceive and produce 'speech acts' (Searl, 1969) appropriately in the context.

According to Brown and Levinson (1978/87) some speech acts are intrinsically 'face threatening'. Thus production of these kinds of acts brings about more challenge for language learners in cross-cultural settings. Refusals are postulated as belonging to such category. They are known as dual face-threatening acts since they endanger the speaker

to threaten both his own face and his/her interlocutor(s). Moreover, their perception and

production requires a high level of awareness of the socio-cultural factors of a given target community (Chen, 1995).

Previous research, however, has shown that refusals are the "sticking point" for FL learners at various levels of proficiency (Beebe, Takahashi & Uliss-Weltz, 1990, p. 56). These learners are often observed to lack the pragmatic knowledge necessary to mitigate such a face-threatening act (Felix-Brasdefer, 2004, Kwon, 2003). These are the motives for the current study to bring this face threatening act into scrutiny.

An important reason for FL learners' failure to act appropriately in the realization of speech acts and refusals in particular is their resort to the norms of mother tongue in the production of these acts (Beebe, et al., 1990), as their realizations vary across cultures (Eslami, 2010). This is technically known as 'pragmatic transfer' Beebe, et al., 1990). Of the main sources of the pragmatic errors, committed by EFL/ESL learners, is negative pragmatic transfer, as evident in some comparative studies of native and nonnative refusals (Beebe et al., 1990; Kwon, 2003, Wannaruk, 2008). Negative pragmatic transfer is the use of native language pragmatic feature which leads to an inappropriate form in the target language, and hence miscommunication (Kasper, 1992).

Previous research on refusals

Research on refusals, in line with other speech acts, has sought different purposes, using different methodology for data collection. A large proportion of the previous studies have explored speech acts cross-culturally, while others have focused on interlanguage refusals and examined the ESL/EFL learners' perception and production of speech acts.

Interlanguage pragmatics (ILP), (the concern of this research) has received the attention of scholars for about three decades. A crucial work in this area is the study of Kasper and Blum-Kulka (1993), who made an attempt to present a comprehensive scope of the field. They extended the area within interlanguage pragmatics to five categories: (a) pragmatic comprehension, (b) production of linguistic action, (c) development of pragmatic competence, (d) pragmatic transfer, and (e) communicative effect. The present study focuses on the last category with respect to L2 refusals.

An influential study on refusals is Beebe, Takahashi and Uliss-Wells' (1990). They created a Discourse Completion Test (DCT) and a comprehensive categorization of refusals which have been used in many later studies carried out on refusals. The results of the study revealed that there is an interaction between the status of the participants and the directness of refusals. Americans usually employ indirect strategies in refusing regardless of the status of the interlocutor, whereas Japanese tend to use more direct strategies when addressing a lower status person and more indirect strategies when refusing persons of higher ranks.

Beebe and Cummings (1996) studied refusal responses collected with two different data collection procedures. Eleven ESL teachers completed a questionnaire while another eleven teachers were asked the same questions on the phone to be answered verbally. The comparison of these naturally occurring talks and DCTs indicated that DCTs are an efficient data-eliciting device for quick collection of a large amount of data. The oral data, however, resulted in more lengthy responses that were also deeper emotionally and psychosocially. For example, role plays allow the researcher to extend the turns of the interaction so that a speech act is finally accomplished. "The conversational activity of the role play will result in more negotiation, and as a result of the negotiation each party will present more pragmatic features like hesitation, turn taking, use of discourse markers, etc which will be found in natural conversation "(Kasper 1990). This is the motive for the present research to draw on role play as a method of data collection.

Nelson et al. (2002) applied a modified version of the DCT for studying similarities and differences between Americans and Egyptians in making refusals. Arabs tend to show more awareness of status differences in refusing of a person in higher status than Americans do. They found refusals in Arabic language and culture even more 'face' threatening than what is perceived in American culture.

Kwon (2003) has examined the use of the refusal strategies among Korean speakers and American English speakers. Kwon has also used Beebe et, al classification of refusals. The data were gathered by using Discourse Completion Test (DCT). The results revealed that despite the fact that both groups have used similar refusal strategies in their languages, cross cultural differences were obvious. Kwon suggested

that negative pragmatic transfer is indeed a potential source of miscommunication due to the inadequacy of ESL learners' knowledge of sociolinguistic rules.

Felix- Brasdefer's (2006) study is relevant to the present research in that it used open role play as a method, Beebe et al.'s classification of refusal strategies as an analytical tool, and focused on 'social distance' and 'power'. In her study, too, social power and social distance played an important role in selecting the linguistic strategies by Mexican speakers of Spanish.

Purpose of the study

The present research aims at focusing on two issues: First, it is an attempt to provide a baseline for refusal strategies in Persian as a First Language in order to contribute to a more fruitful cross-cultural communication with reference to the two important social variables of 'power' and 'distance' in Persian language. Second, once more, it brings learner language into focus to investigate the possible amount of pragmatic transfer from L1 in the learner language (L2), as far as the production of refusals is concerned. In short, the study addresses the following research questions:

1. How do Iranian EFL learners realize the speech act of refusals in Persian and English?

2. How do social variables "relative power" and "social distance" between the interactants affect the production of refusals in both L1 and L2 of Iranian EFL learners?

3. To what extent Iranian EFL learners use "pragmatic transfer" strategy in their performance of refusals in L2?

A refusal is a non-compliant act on the part of the speaker in response to an invitation, offer or request. It is high-risk and can easily cause offence to the hearer. According to Beebe, Takahashi, & Uliss-Weltz "refusals are a major cross-cultural sticking point for many non-native speakers" (1990, p. 56). Given the face–threatening nature of refusal, its performance thus calls for considerable linguistic and cultural expertise on the part of the speaker. Thus to refuse without causing offence requires not only recognizing the linguistic forms necessary for its successful performance but also being aware of the socio-cultural values of the target community. This requirement of a high level of pragmatic competence can constitute a major cross-cultural challenge for EFL learners.

To date, little attempt has been made to conduct an Interlanguage Pragmatics study that investigates the speech act of refusal produced by Iranian EFL learners (Izadi and Zuraidah, 2010; Eslami 2010; Shokouhi and Khalili, 2008). This kind of research can be particularly beneficial, considering its pertinent pedagogical implications for future research. Findings from the analysis of data can provide useful background information for the development of learning materials and teachings methods for the enhancement of Iranian EFL learners' sociopragmatic competence.

Method

Participants

30 students of English (Persian as L1) at Azad University of Abadan, Iran participated in the study. Their age ranges between 20 to 29, and of both genders. They had

enrolled in "Conversation Two" course and were at the beginner to intermediate level of English language proficiency.

Data Collection

The data collection instrument in this study was the role play scenarios (Appendix A). After giving instructions of how the scenario works, the students were asked to play the role of a refuser in each situation in English. After three months, the same students were asked to play the role of a refuser in the same situations, but in Persian. The time duration was given to reduce the probable effects of the first phase on the second.

The role play situations which were used in this study involved eight situations designed to elicit refusals as a response to four different initiating speech acts: suggestions, invitations, request, and offers. Each situation was based on two social variables: "relative power" and "social distance" between the interlocutors.

Data analysis

The refusals elicited were classified according to Beebe et al' (1990) classification of refusals (Appendix B). This is a widely used classification of refusals strategies in refusal research (Bardovi-Harlig and Hartfort 1990, Gass and Houck, 1999, Nelson et.al., 2002, Felix-Brasdefer 2006). To ensure the reliability of the coding, an inter-rater reliability technique was used; that is, each utterance was checked by another expert in Pragmatics.

In this study, an assessment of central tendency and dispersion for the research questions was conducted using descriptive statistics. SPSS 11.5 was used to do the calculations. To measure the difference in indirect strategies used by the two language groups, the mean of each strategy in each group with three different social statuses was calculated.

Results and discussion

The findings suggest similarity in terms of types of refusal strategies, but differences in terms of frequency and the content of strategies used in English and Persian. However, the number of indirect strategies was far greater than direct ones in both languages. Also, in the majority of cases in both languages direct strategies were commensurate with indirect ones (table 1). This supports the findings of previous research at least in that refusals operate by universal rules in terms of preference of indirect over direct strategies in communications between the people who have ongoing relationship (see Kasper & Rose, 2003; Blum-Kulka et al., 1989, for the universality of speech acts).

Ochs (1996, p. 425) maintains that "there are certain commonalities across the world's language communities and communities of practice" in people's social ends in communication. She cautions that this principle does not mean that all pragmatic principles

Table 1: *Means and standard deviations of total direct and indirect I strategy use by Iranian EFL learners' English and Persian refusals*

	LANGUAGE	N	Mean	Std. Deviation	Std. Error Mean
Total indirect strategy use	Persian	30	28.1167	6.99417	.90294
	English	30	21.4333	4.60790	.59488
Total direct Strategy use	Persian	30	4.0833	2.87179	.37075
	English	30	4.3000	1.95110	.25189

are shared across all cultures. Rather, this principle, as she says, makes "a common ground of socialization experiences" that interlocutors can employ as their common ground to realize "local ways of indexing and constituting social situations". This is what she has termed the Local Culture Principle which is formed through situationally specific values.

"Bald-on-record" refusals (in the terminology of Brown and Levinson, 1987) like 'no' 'I can't' or 'I won't' were observed very occasionally. This is in accordance to findings of many previous scholars (Beebe, et al., 1990; Kwon, 2003, Shokouhi and Khalili, 2008, Wannaruk, 2008, Izadi and Zuraidah, 2010) The majority of informants avoided a direct refusal (mere ***no***) and tended to provide reasons, explanations or excuses as a way to imply their lack of ability or unwillingness.

Another point concerning the first question of the research is the greater number of indirect refusals in the participants' Persian data in comparison to the English one (see Table1). This higher number of indirect refusals in participants' native language may be due to their far greater competence in Persian language in comparison to English as a foreign language. Another justification for this phenomena may be the cultural norms of Iranian society in which making a refusal directly even to someone of lower social status is considered as discourtesy. Apart from these points, one may interpret this lengthy combination of indirect strategies as these informants' awareness of the (though not seemingly) hierarchical nature of the society in which they live.

Concerning the type of strategies, expectedly, 'reasoning' was the most common strategy in both languages (Eslami, 2010). Since it was the most frequent strategy present in all the studies mentioned in the review section, this strategy type can be considered as a universal feature of refusals. Moreover, in social status distinction analysis, this strategy again gained the first place among other types of strategies. What differentiated the two tests, however, was the kind and length of reason in the two tests. Participants in English test tended to suffice to short structurally simple reasons. They employed a kind of "avoidance strategy" which has been frequently reported in SLA literature. While, in Persian, they elaborated on the reasons and provided lengthier justifications for their refusals.

Strategies 'consideration of addressee's feeling or opinion', 'statement of regret', 'letting interlocutor off the hook', and 'expression of gratitude' came to be the second to the fifth preferred strategies, respectively, in both languages. A strong motivation for using a number of 'mitigators' to soften a refusal, as has frequently been discussed in the literature, is 'face' considerations (Goffman, 1967; Brown and Levinson, 1978/87; Beebe, et al., 1990, Felix- Brasdefer, 2006, to name but a few). Respondents

of the present study attempted to soften their refusals with softening strategies like the statement of regret, providing explanations, gratitude, etc). As Beebe et al. (1990, p. 106) indicate such softeners are expected in face-threatening acts to save the face of the requester or inviter. Table (2) denotes the types and frequencies of the five most frequent strategies:

The second research question concerning the 'power' and 'distance' influence on the strategy types showed that in Persian, participants used more indirect strategies when making refusals to someone of higher social status. Each initiating act was considered to involve two situations in the role play scenarios: one with + power and distance, and the other with – power and distance. The first pair of each initiating act involved a situation where the power of the addressee was high and the social distance between the refuser and his/her addressee

Table 2: *Means and standard deviations of refusal strategy types in the two groups' refusals*

	LANGUAGE	N	Mean	Std. Deviation	Std. Error Mean
Reasoning	English	30	14.8348	3.50770	.45284
	Persian	30	13.5833	3.81030	.49191
Positive opinion	English	30	5.2833	2.29549	.29505
	Persian	30	4.1667	.49289	.06363
Statement of regret	English	30	5.0333	1.07304	.13853
	Persian	30	3.5833	1.01030	.49191
Gratitude	English	30	3.4167	1.00314	.09594
	Persian	30	3.0167	1.06551	.13756
Let interlocutor off the hook	English	30	2.7333	1.07304	.13853
	Persian	30	2.5000	2.82543	.36476

was great.

The two instances of powerful and distant addresses were found in employer-employee and teacher-student relationship. Unsurprisingly, the participants found these refusals challenging and employed a host of mitigating strategies in realizing them. The justification for this outcome may be the greater consciousness of the hierarchical nature of employer-employee and teacher-student relationship in Iran, where people tend to either defer to the individual with higher status and more power (Nelson et al., 2002, p. 183) or fear the probable negative consequences of their refusals (Izadi and Zuraidah, 2010), or a mixed sense of both. The following interaction shows an attempt by a female student (S) to refuse her male teacher and HOD's (T) suggestion to take a course with him (situation 1 of the role play scenarios; appendix A):

T: *xanome [name] mæn in term dærse xandæne se ro erae dadæm*
   Miss [name] this semester I have offered the course Reading 3
   mitunid in dærso ba mæn begirid
   you can take this course with me

S: *mmm jeddæn ostad? cheghædr heif shod=*
   Mmm seriously professor? What a pity
   Chunke vase terme bæ'd (.) man ye seri-
   Because for next semester I have a series of
   xeili æhvaz chun særæm sholughe ye seri kara daræm vase hamin vase terme bæ'd
   Very busy because I am in Ahvaz I've got a series of businesses to do for next sem
   bishtær vahedhaye omumi bayæd værdaræm=
   I have to take general courses more
   Fek nækonam betunam dærse xandæne se ro begiræm
   I don't think I can take Reading 3 course
   chun dærse sænginieh(.)
   Because it's a difficult course
T: *are are*
   Yeah yeah
S: *bayæd adæm amade bashe ye meghdar moshkel pish miad ba'd mmm væli*
   One must be ready   there might be a bit of a problem   but
   xob tælashæmo mikonæm væli kollæn fek nemikonæm
   ok I'll do my best   but generally I don't think so
T: *besyar xob enshalla termay ayænde dær khedmætetun hæstim*
   All right   ok hopefully next semester I'll be at your service
S: *xeili mæmnun ostad æz pishnæhadetun*
   Thank you very much professor for your suggestion
T: *xahesh mikonæm*
    It's all right

Being fully aware of the mechanism of interactions in teacher-student relationships, S tries her best to employ as many 'face saving' strategies as possible (Brown and Levinson, 1978/87; Felix-Brasdefer, 2006; Eslami, 2010).   The perceived social distance between the two participants makes the situation formal. What is more important than the formality of the situation, however, is the power relation between T and S. S, consequently, finds this refusal very challenging, and does not suffice to a prototypical reasoning. She tends to elaborate on her explanation/excuse/reason to persuade T why she is not able to take the course.

   S starts with a 'request for approval' strategy (*jeddæn ostad*: seriously professor*?*) following a 'pause filler' (mmm). The next strategy she uses is 'expression of regret' (*cheghædr heif shod*: What a pity). She shows (or pretends) to regret that she has missed a chance to take a course with T, which sets the scene for her refusal (pre-sequence). In refusal sequence, S is seemingly explaining why she feels pity, but actually she is refusing, using 'reasoning strategy'. Given that the context is very power-sensitive and S is fully aware of that, she presents an elaborate reasoning combined with 'statement of principle' strategy (*dærse sænginieh*: it's a difficult course, *bayæd adæm amade bashe*: one must be ready), negative ability (*fek nækonam (.) betunam dærse xandæne se ro begiræm*: I don't think I can take Reading 3 course*)* and 'self-defense' (*væli xob tælashamo mikonæm:* but I'll do my best). She then

finishes by 'expression of gratitude' (*xeili mæmnun ostad æz pishnæhadetun*: thank you very much for your suggestion professor).

The myriad of indirect strategies employed as well as the existence of other pragmatic and (pragmatic markers, address terms) paralinguistic features (tone of voice) (which is beyond the scope of the present study) all attest to impact of power of the addressee on the linguistic choice of refusal strategies on the part of the student.

Regarding the third research question, pragmatic failure would be almost improbable on the two role played languages, at least at the level of linguistic verbalization of refusals. The remarkable similarity in the type of refusal strategy application and employing reasons in combination with a variety of indirect strategies in making refusals in both tests is in accordance with Kasper (1997) and Kasper & Blum-Kulka (1993) who believe that the speech act strategies that are similar between two languages (in our case English and Persian) often result in pragmatic success. Nelson et al. (2002) maintain that the degree of this success depends on learners' ability to recognize the proper degree of sociopragmatic appropriateness. The sociopragmatic sphere of pragmatic competence is more likely to be prone to negative pragmatic transfer. This aspect of pragmatics was beyond the scope and methodological capacity of the present study. Another reason for lack of transfer in the data could be assigned to the proficiency level of the participants. As beginners and intermediate level students of English, they provided the minimum response to the initiating acts in the role plays, mainly prototypical refusals. It goes without saying that the less one speaks, the less likely the evidence of pragmatic transfer is. Pragmatic transfer might become more evident in English used by advanced learners of English, since they enjoy more fluency and accuracy in the foreign language (Takahashi & Beebe, 1987). Therefore, the results of this study cannot be generalized to all Iranian EFL learners.

Conclusion

The present research participants, being status aware, used more indirect strategies in the Persian test in comparison to the English one which could be due to their greater mastery over their native language. However, the other possible explanation is these learners understanding of power and distance relationship in a strictly (but not seemingly) hierarchical society.

Based on the findings, there seem to be more similarities than differences between the Persian and English in making refusals. As a result, we do not expect to and a lot of miscommunication between the two groups. However, situations of communication breakdown can be predicted. However, the other important facet of pragmatic competence is sociopragmatic competence which is more susceptible to negative pragmatic transfer (Kwon, 2003). This sociopragmatic interface of refusals requires more research especially those which use naturally occurring data.

The findings of this study have pertinent pedagogical implications in that they provide features of L2 refusals as compared with L1. Therefore, curriculum developers and practitioners in ELT may enjoy the results of the present study. The findings may help material developers to adjust teaching materials with Iranian EFL

students' need in developing communicative competence. It is also hoped that the results have contributed to a smoother cross-cultural communication.

References

Bardovi-Harlig, K. & Hartford, B. (1990). Congruence in native and nonnative conversations: status balance in the academic advising session. *Language Learning, 40*, 467-501.

Beebe, L. M., Takahashi, T. & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R.C. Scarcella, E. Anderson & S. Krashen (eds), *Developing Communicatve Competence in a Second Language,* 55-73. New York: Newbury House.

Beebe, L. M. & Cummings, C. (1996). Natural speech act data versus written questionnaire data: how data collection method affects speech act performance. In S.M. Gass & J. Neu (eds), *Speech Acts across Cultures*, 65-86. Berlin: Mouton de Gruyter.

Blum-Kulka, S. & Olshtain, E. (1986). Too many words: Length of utterance and pragmatic failure. *Studies in Second Language Acquisition,* 8, 47-61.

Blum-Kulka, S., J. House & G. kasper. (1989). Investigating cross-cultural pragmatics: an introductory overview. In S. Blum – Kulka, J. House & G. Kasper (eds), *Cross-Cultural pragmatics: requests and apologies*, 1-34. NJ: Ablex.

Brown, P. & Levinson, S. C. (1987). *Politeness: some universals in language usage.* Cambridge: Cambridge University Press.

Chen, X., Ye, L., & Zhang, Y. (1995). Refusing in Chinese. In G. Kasper (eds.), *Pragmatics of Chinese as a Native and Target Language,* 119-163. Honolulu, HI: University of Hawai'I Press.

Eslami, Z. R. (2010). Refusals: How to develop appropriate refusal strategies. In A. Martínez-Flor **& E.** Usó-Juan (eds.), *Speech Act Performance: Theoretical, empirical and methodological issues,* 217–236. John Benjamins Publishing Company.

Felix-Brasdefer, J. C. (2006). Linguistic Politeness in Mexico: Refusal Strategies among Male Speakers of Mexican Spanish', *Journal of Pragmatics 38*, 2158-87.

Goffman, E. (1967). *Interactional ritual*: *Essays in face-to-face behavior.* New York: Pantheon.

Izadi, A., & Zuraidah, M. D. (2010). Interlanguage pragmatics: strategy use in English and Persian refusals by Iranian EFL learners. *Iranian Journal of Teaching English as a Foreign Language and Literature, 2, 1, 3-23.*

Kasper, G. (1992). Pragmatic transfer. *Second Language Research*, 8, 203-231.

Kasper, G. (1997). Can pragmatic competence be taught? Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center. Retrieved at; http://www.nflrc.hawaii.edu/NetWorks/NW06/default.html.

Kasper, G. & Blum-Kulka, S. (1993). Interlanguage pragmatics: an introduction. In G. Kasper & S. Blum-Kulka (eds.), *Interlanguage Pragmatics*, 3-17. New York: Oxford University Press.

Kasper, G. & Rose, K. (2003). *Pragmatic Development in a Second Language*. Malden: Blackwell.

Kowan, J. (2003). Pragmatic Transfer and Proficiency in Refusals of Korean EFL Learners'. *Unpublished doctoral dissertation*: Boston University.

Nelson, G. L., Carson, J. Al-Batal, M. & El-Bakari W. (2002). Cross-cultural pragmatics: strategy use in Egyptian Arabic and American English refusals. *Applied Linguistics*, *23*, 163-189.

Ochs, E. (1996). Linguistic resources for socializing humanity. In J. Gumperz & S. Levinson (eds.), *Rethinking Linguistic Relativity*, 407-437. Cambridge: Cambridge University Press.

Sadler, R. W. & Eröz, B. (2001). I refuse you! an examination of English refusals by native speakers of English, Lao, and Turkish. *Arizona Working Papers in SLAT, 9,* 53-80.

Schmidt, R. 1993. Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (eds), *Interlanguage Pragmatics*, 21-42. Oxford: Oxford University Press.

Searle, J. (1969). *Speech Acts*. Cambridge: Cambridge University Press.

Shokouhi, H. & Khalili, M. (2008). Pragmatic transfer in learners' refusals: a case of gender distinction. *Journal of the Faulty of Literature & Humanities, Shahid Chamran University of Ahwaz*, 2, 215-252.

Takahashi, T., & Beebe, L. M. (1987). The development of pragmatic competence by Japanese learners of English. *Jalt Journal, 8*, 131-155.

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied linguistics, 4,* 91-112.

Wannaruk, A. (2008). Pragmatic transfer in Thai refusals. *Regional Language Center Journal, 39 (3),* 318-337.

Appendix A: Role play scenarios

1.Suggestion (+power, +distance)
You are a first semester senior at the university----------- and since pre-registration is next week; you are planning your schedule for your final semester. You have already put together a tentative schedule, but you need to get approval from your head of the department (HOD). You have taken one course with HOD during your first year, and have failed. You do not like to take any other course with him/her, as s/he is very strict. S/he says that s/he has offered a course you need to take and suggests that you take it.

2. Suggestion (-power, -distance)
You are sitting at a bench in the university campus waiting for your class to begin. At this time another student (same-sex) from the class that, you get along well with comes to join you. You have worked on an assignment in class and have gone out

together occasionally, but are not close friends since you have only known each other since the beginning of the semester. You begin to discuss different types of food and drink and you realize that you have similar taste. After about 20 minutes, you realize it is class time and are both getting your books together to walk over to the class, when s/he suggests skipping the class, and going to a coffee shop down the street to have a coffee, but you don't want to go.

3. Request (+power, +distance)

You have been working at a part-time job for extra spending money after school at the University bookstore since the beginning of the semester. The bookstore is open Saturday through Thursday from 9:00 a.m. until 7:00 p.m. You work from 3:00 p.m. to 7:00 p.m. Saturday through Thursday. You get along fine with your boss, but you are not friends and you do not socialize together outside work. It is Thursday evening at 6:45 p.m. and the boss approaches you and asks you to work extra hours (until 9:00 p.m.), but you can't stay.

4. Request (-power, -distance)

You are taking a course in literature this semester. You haven't missed this class once this semester and consider yourself a diligent student. So far you haven't a good average in the class, not because it is easy for you, but because you have worked very hard. Among your classmates, you have a reputation for taking very good notes. The professor has just announced that the midterm exam is next week. One of your close friends and classmates, who is taking the class, asks you for your notes. Although you love your friend but also believe that everybody should prepare his/her own notes. You are not really willing to lend your notes to anybody.

5. Invitation (+power, +distance)

You have been working for Dena Company as a sales representative for the last five years. You have a good working relationship with your boss although you do not socialize together outside the office. Your boss has always been supportive of your ideas and has been instrumental in your receiving a recent promotion. After working for him for three years, he has recently been promoted and will become the manager of the Dena Company's central office which will require his relocation to Tehran next month. He is having a party next Friday evening at a restaurant and is inviting you and other member of his sales group to celebrate his promotion and as a farewell, but you are unable to attend.

6. Invitation (-power, -distance)

You are walking across campus when you run into a good friend of yours whom you haven't seen for about a month. You and s/he have been studying in the same program at the University for three years, and have studied and written papers together in the past, but you don't have any classes together this semester since you have been doing an internship off-campus. He invites you to his 21st birthday party at his house next Friday night at 8:00 p.m... He tells you that a group of mutual friends that you both used to hang out with and whom you haven't seen since the semester started will also be there. You know that this would be a good opportunity to see everyone again and to celebrate this special occasion with him. Unfortunately, you cannot make it.

7. Offer (+ power, +distance)

You are a student at a university. You are working on a project and need your advisor's consultation. You go to your advisor's office for the meeting. When you entered the office, before starting a conversation, your advisor offers you some chocolate. But you are too shy to accept the offer.

8. Offer (-power, -distance)
You are at dinner table with one of your close friends. You have known each other for several years. Your friend is very happy to have dinner with you and is cooking a new dish that you don't like some of the ingredients. Your friend doesn't know that. You try some and prefer not to eat the special dish and at the table you just eat salad and some drink. Your friend insists on having more but you don't want to tell your friend that you don't like the food.

Appendix B: Classification of Refusals (Beebe, Takahashi & Uliss-Weltz, 1990)
A: direct:
1. Flat 'no'
2. Negative ability/willingness 'I can't/I won't"
B: Indirect
1. Statement of regret like "I'm sorry."
2. Wish like "I wish I could help you."
3. Excuse, reason, explanation like "I have an exam."
4. Statement of alternative which falls into two divisions:
    4-1. I can do X instead of Y like "I'd rather ….."
    4-2. Why don't you do X instead of Y like "Why don't you ask someone else?"
5. Set condition for future or past acceptance like "If I had enough money"
6. Promise of future acceptance like "I'll do it next time."
7. Statement of principle "I never drink right after dinner."
8. Statement of philosophy like "One can't be too careful."
9. Attempt to dissuade interlocutor:
9-1. Threat or statement of negative consequences to the requester like "If I knew you would judge me like this I never did that"
    9-2. Criticize the requester "It's a silly suggestion."
 9-3. Guilt trip (waiter to customers who want to sit for a while: "I can't make a living off people who just order tea"
9-4. Request for help, empathy, and assistance by dropping or holding the request like "I hope you understand my difficult situation"
    9-5. Let interlocutor off the hook "Don't worry about it."
    9-6. Self-defense like "I'm doing my best."
10. Acceptance functioning as a refusal:
    10-1. Unspecific or indefinite reply "I don't know when I can give them to you"
    10-2. Lack of enthusiasm "I'm not interested in diets"
11. Avoidance:
    11-1. Non-verbal (silence, hesitation, doing nothing and physical departure)
11-2. Verbal (topic switch, joke, repetition of past request, postponement and hedge); an example for postponement can be "I'll think about it."
12. Statement of positive opinion like "That's a good idea"
13. Statement of empathy "I know you are in a bad situation"
14. Pause fillers like" well" and "uhm"
15. gratitude/appreciation like "Thank you."

Title

Guiding a Persian EFL Instructor to Create Individualized Classroom Environments:

Novel Opportunities for ELT Practitioners

Authors

Ali Rahimi (Ph.D.)

University of Kashan, Iran.

Nabi. A Ebrahimi (M.A.)

Islamic Azad University, Arsanjan Branch, Iran.

Bio Data

Ali Rahimi, assistant professor in TEFL, received his Ph.D. from Shiraz University in 2004. He has taught at Isfahan, Kashan, Tehran Teacher Education and Shiraz universities. He is now supervising M.A. and Ph.D. students at Kashan, Tehran Teacher Education and Allameh Tabatabai universities. He has written and translated many books on critical discourse analysis, social psychology, globalization, translation theories, and teaching methodologies. He has published many articles and presented papers at national and international conferences.

Nabi.A Ebrahimi, M.A on TEFL from University of Kashan, Iran. He is now teaching in Islamic Azad University, Arsanjan Branch, Iran. He has presented some papers in national and international conferences. He is interested in educational reform studies, psycholinguistics, discourse analysis, and learning environment researches.

Abstract

Taking into account the sad fact that most of the researches carried out in the field of learning environment research are allocated to science and mathematics classrooms, we have endeavored to introduce the field of learning environment research to ELT practitioners. The idea presented here is that the insights gained through decades of learning environment research conducted basically in science and mathematics classrooms can be conducive to the improvement of English classroom environments. Furthermore, since this was the first in-depth study concerning learning environments

in Iran, it would elucidate the practical benefits of learning environment investigations by explaining how a learning environment instrument could be used to guide a Persian EFL instructor to foster a more individualized learner-centered classroom environment. 32 (F=26 and M=6) EFL students were given the Individualized Classroom Environment Questionnaire in its actual and preferred forms. The statistical analyses revealed that there was a significant difference (p<0.05) between the scores on actual and preferred forms, hence the students' dissatisfaction with their classroom environment. In addition, the dimensions from which this dissatisfaction arose were also explored and proposed to the instructor to make him able to meet his students' ideas and preferences. The ideas can be utilized by EFL teachers, materials developers and assessment specialists to create more innovative, creative, critical and democratic classroom environments as the sine qua non of modern successful learner based education.

Keywords: Learning environments, Individualized classroom environments, English classroom environments, Students' perceptions, Satisfaction, ICEQ; ELT.

Introduction

In his article "learning environment research: yesterday, today and tomorrow", Fraser (2002) drew attention to the lack of any published article that reports Asian teachers' attempts to use learning environment assessments to guide improvements in their classroom environments. Since 2002, much change has occurred and much research has been conducted in Asian countries such as Hong Kong (e.g. Kember, Ho & Hong, 2009), Korea (e.g. Fraser & Lee, 2009), Malaysia (e.g. Scott & Fisher, 2004), Singapore (e.g. Quek, Wong, Divaharan, Peer & Williams , 2007), Taiwan (e.g. Wanpen & Fisher, 2006) and Turkey (e.g. Ozkok, Walker & Buyukozturk, 2009), but it seems that learning environment research has remained stagnant in Iran. In Fraser's (2002, p.19) words, "the important practical benefit of such research has not yet been realized in [Iran] as no published article could be located that reported [Persian] teachers' attempts to use learning environment assessments to guide improvements in their classroom environments."

   In addition, much of the research conducted in the field of learning environments is related to science and mathematics classrooms. The contribution of EFL, ESL, and ELT researchers to this field of inquiry is not significant and few studies (e.g. den Brok, Fisher, Rickards & Bull, 2006; Wei, Brok & Zhou, 2009; Wilks, 2000, cited in Fraser, 2002, p.6) aiming to enhance English classroom environments can be found in the literature. It does not mean that studies in and attentions to learning environments

are a forgotten part of our profession as ELT researchers and teachers, but it is the unsystematicity of such research that demands more work and attention. In our view, the experiences gained through learning environment studies mostly carried out in science and mathematics classrooms can be of great assistance in improving English classroom environments. Learning environment studies in English classrooms can be directed in new and more systematic ways.

The aim of the present study is to respond to these deficiencies and to open new doors by introducing learning environment research to the vast field of ELT. Being one of the few ones conducting in English classrooms, this study is also the first one reporting the practical use of learning environment research in Iran.

Through this study, the field of learning environment research will be implicitly and explicitly introduced and the lines for future researches will be shown. The methodology and terminology of this study are also a reflection of common trends in the field. In addition, using Individualized Classroom Environment Questionnaire (ICEQ), this study also aims to assist a Persian EFL instructor to explore his students' level of satisfaction with their classroom environments. For this purpose, the following questions are going to be answered:

 (1) Are there any significant differences between the actual and preferred classroom environments as perceived by the participating Persian EFL students?
 (2) To what classroom environment dimensions (i.e. personalization, participation, independence, investigation or differentiation) is the probable dissatisfaction attributed?
 (3) How remote are the investigated English classroom environments as perceived by the participants from true individualized ones?

The rest of the paper will be presented as follows. Background to the learning environment research, individualized instruction, ICEQ and some general information about some useful learning environment instruments for ELT practitioners are given in the literature review section. Next, in method section, methods and the procedures of the present study are described. In result section are shown the results of different t-tests each one answering the questions set by the study. An overall discussion of the results is provided in discussion section, where the investigated English classroom environments are assessed against individualized instruction ideas and where a detailed discussion of Persian EFL students' level of satisfaction with their classroom environments is presented. Finally, the conclusion and implications of the study are provided in the last sections.

The field of learning environment research & learning environment studies

A turning point in the history of learning environments studies occurred about 40 years ago when Herbert Walberg and Rudolf Moos started their research programs which consequently formed the cornerstone of modern learning environments researches. Walberg developed the Learning Environment Inventory as part of the research and evaluation activities of Harvard Project Physics, while Moos elaborated social climate scales for various human environments, including the Classroom Environment Scale (Fraser, 1998). In turn, Walberg and Moss's pioneering work built

upon earlier ideas of Lewin and Murray several decades before. In his field theory, Lewis acknowledged that both the environment and its interaction with personal characteristics of the individual are factors determining human behavior.

Another milestone in the development of the field of learning environments is the birth of *Learning Environments Research: An International Journal* (*LER*) giving shape to and opening new horizons for learning environments studies. The field is now an established one, able to assess learning environments in a variety of quantitative and qualitative methods, attracted by researchers, teachers, school administrators and administrators of school systems in different countries. It is equipped with a variety of economical, valid and widely-applicable questionnaires that have been developed and used for assessing students' perceptions of classroom environment. In the following, a brief account of learning environment is presented.

The concept of environment, as applied to educational settings, refers to the atmosphere, ambience, tone, or climate that dominates the particular setting (Dorman, Fraser & Aldridge, 2006, p.906). Fraser (1994, 1998), the prominent figure of the field of learning environments research, defines learning environments as both social and psychological in nature, and believes that they are 'determinants' of learning. Numerous research studies have shown that student perceptions of the classroom environment account for appreciable amounts of variance in learning outcomes, often beyond that attributable to background student characteristics (Dorman, 2001, p.244). As Barry Fraser (2002) explains, student perceptions of the classroom learning environment are important, should be of interest to classroom teachers, and can be fairly easily measured with classroom environment perception instruments existing in the well-established field of leaning environment research.

The above-mentioned definitions of learning environment embrace a vast array of hidden and unhidden aspects of a learning process and also the most important ones or 'determinants' of learning. It leads not only to the significance of learning environments researches but also to the comprehensiveness of such studies. In other words, the picture such studies present of any educational setting is hardly obtainable through other approaches with such a thoroughness and quickness. In our view, learning environment studies deserve more attention because they are simultaneously comprehensive and economical. These features have not been emphasized deservedly even by established learning environments researchers.

Individualized instruction and individualized classroom environments (ICEs)

No one can deny the positive effects of adopting individualized instructions and creating ICEs on learners' self-improvement, self-awareness, self-competition, motivation, autonomy, and satisfaction. In ideal ICEs, learners find everything in line with their preferences, interests, and favorites. By adopting such environments, most of the educational problems directly related to learners' satisfaction and motivation will be solved. Inevitable difficulties and serious obstacles will emerge if learners find their learning and its environment as irrelevant, demotivating, unsatisfactory, and meaningless.

Individualized instruction is a method of instruction in which content, instructional

materials, instructional media, and pace of learning are germane with the abilities and interests of each individual learner. It has been elaborated with the premises that each learner is unique and "is an individual who must be helped to find his or her way to become autonomous" (Williams & L.Burden, 1998, p.194) and learners have varied learning styles, learn at different rates, have varying socioeconomic backgrounds, and have diverse intellectual strengths (Dileo, 2007). Here the traits of the individual learner are given more consideration and learning is improved by varying the pace of instruction, the instructional method, and the content. In such settings, "learner achievements are independent of each other, everyone has an equal opportunity of gaining a reward of some kind", and "success or failure is more likely to be attributed to effort" (Williams & L.Burden, 1998, p.193).

"Both the anecdotal evidence of many teachers, as well as recent student feedback, conclude that students want relevant, hands-on, authentic learning experiences and they expect- indeed, demand- more responsibility for, involvement in and control over the construction of their learning journey" (Murdokh & Wilson, 2008, p.3).

Autonomy of the learners is also of great importance in all levels of education. Learner autonomy is at the heart of individualized classroom environment, and in Paulo Freire's view, a matter of ethics:

Respect for the autonomy and dignity of every person is an ethical imperative and not a favor that we mayor may not concede to each other. It is precisely because we are ethical beings that we can commit what can only be called a transgression by denying our essentially ethical condition. The teacher who does not respect the student's curiosity in its diverse aesthetic, linguistic, and syntactical expressions [......] transgresses fundamental ethical principles of the human condition. It is also in this sense that the possibility of true dialogue, in which subjects in dialogue learn and grow by confronting their differences, becomes a coherent demand required by an assumed unfinishedness that reveals itself as ethical.

In ICEs, students have opportunities to interact with teacher (personalization), they are encouraged to participate in the classroom (participation), they are treated individually by teacher (independence), they are encouraged to perform individual problem solving activities (investigation), and they have control over their learning (differentiation). All of these dimensions of ICEs differentiate them from conventional and traditional classroom environments. These are also the dimensions which form the five different scales of ICEQ, an instrument designed specifically to describe, classify and measure those learning dimensions which differentiate conventional classrooms from individualized or learner-centered ones. A different part has been allocated to ICEQ.

For more clarification, mentioning some of the benefits and ramifications of ICEs seems necessary:

(1) Individualized instruction allows a student who is above or below average to proceed at their own pace for optimal learning.

(2) Students do not have to repeat parts of a course that they have already mastered.

(3) Students learn the self-discipline and goal-orientation needed to motivate them and to

keep their progress on target.

(4) Students can check their own results on classwork and seek help when needed.

(5) Individualistic structures can be regarded as "providing a form of self-competition, but differ from competitive structures in that they are essentially goal oriented and involve the development of self-awareness" (Williams & L.Burden, 1998, p.194).

(6) Individualistic approach makes it possible to focus upon "the learning process and to identify personal strategies that are likely to lead to successful learning" (ibid, p.194).

Individualized English language teaching

The individualized teaching mode which is the compulsory alternative of the teacher-centered ELT has made vital contributions to the field. Although individualized teaching is of great adherence, but in actual teaching the learner centered mode has been to a great extent perceived as class centered or group centered teaching that does not deliberately concentrate on the individual learner (Sarigoz, 2008, p.51). Personal beliefs, Individuals' emotions and talents should be taken into account in order to personalize ELT for more efficient, individual friendly learning.

Individualized ELT is not a new concept. Prior to communicative era some practices were modified to pay more attention to individual learners. These activities included answering questions individually instead of answering them in chorus, or having limited and controlled dialogs about a given topic and eliciting the grammar rules (Sarigoz, 2008, p.54). Then the discussions about learner strategies and learner profile put more emphasis on individual differences and individual differences were taken into account in language teaching. Today, a more detailed understanding of each learner's personality, preferences and needs is recommended for effective language teaching.

There are a lot of studies emphasizing individualized ELT. McGroarty (1998) sees the multiple identities of language users as a challenge set by new schools of thought for applied linguists. Sarigoz (2008) states that "in order to cope with the new learning standards and the demands of today's complex societies, we need to develop teaching that goes far beyond dispensing information, giving tests, and grading" (p.57). He adds that teachers should understand how to teach in ways that respond to diverse learning approaches of the students. Breen & Littlejohn (2000) emphasize the need for creating classrooms which will provide learners to have a say in the management of their own language learning. Griffiths & Keohane (2000, p.1) also bring into foreground the importance of individualizing language learning by integrating feelings, opinions, and experiences of individual learners into lessons.

About ICEQ

The origin of the Individualized Classroom Environment Questionnaire (ICEQ) can be traced back to the study conducted by Rentoul & Fraser (1979) "which was concerned with the conceptualization of enquiry-based and open classroom learning environments" (Franz, 1990, p.8). Moss's dimensions of relationship, personal development and system maintenance were also used in the development of ICEQ. The initial development of ICEQ was guided by: the literature on individualized,

open and inquiry-based education; extensive interviewing of teachers and secondary school students; and reactions to draft versions sought from selected experts, teachers and junior high school students (Fraser, 1998, p.11). The outcome of Rentoul & Fraser's research and effort was ICEQ, the instrument we used in this study, which was designed specifically to describe, classify and measure those learning dimensions which differentiate conventional classrooms from individualized ones.

Studies conducted by Fraser (1984) and Fraser & Fisher (1983) confirmed the effectiveness of using actual and preferred forms of classroom environment instruments for comparing students' perceptions of actual and preferred classroom environment. With regard to this matter, ICEQ does not make an exception and it has also actual and preferred forms.

ICEQ consists of five scales consisting of personalization, participation, independence, investigation and differentiation which are the main dimensions of a learner-centered classroom environment. The short form of ICEQ (Fraser, 1990) consists of 50 items (25 on actual form and 25 on preferred form). There are 10 items per scale and each item is responded to on a Likert scale including almost never, seldom, sometimes, often and very often alternatives.

The names, definitions and some sample items of the five scales of ICEQ are provided below:

(1) Personalization- the extent to which students have opportunities to interact with teacher (e.g. "the teacher considers student's feelings")

(2) Participation- the extent to which students are encouraged to participate in the classroom (e.g. "students ask the teacher question")

(3) Independence- the extent to which students have control over their learning (e.g. "students choose their partners for group work")

(4) Investigation- the extent to which students are encouraged to perform individual problem solving activities (e.g. "students carry out investigations to answer questions which puzzle them")

(5) Differentiation- the extent to which students are treated individually by teacher (e.g. "different students do different works")

ICEQ has been used widely in studies conducted in different educational contexts in different countries such as Brunei Darussalam (Asghar & Fraser, 1995), Swede (Allodi, 2002) and especially in Australia (Franz, 1990 and Fraser & Fisher, 1982).

Although ICEQ has been elaborated and validated for use in science and mathematics classrooms, the theories and ideas underlying this instrument do not prevent other researchers and teachers in different educational levels and field of studies to benefit from it. Individualized instruction is a general idea applicable to all educational levels and fields. In addition, there are no special references to science or mathematics classrooms in the items and categories of ICEQ.

Other instruments for assessing learning environments

The field of learning environment research is among few fields of educational research that can boast such a rich array of validated and robust instruments. Most of these instruments have been elaborated and validated for use in science and

mathematics classrooms, but it does not mean that other researchers from other fields of study cannot use and benefit from such instruments.

With regard to the applications of these instruments in other fields of study, they can be divided into three groups. The first group involves those instruments concerning general ideas of education (e.g. ICEQ). They can also be used in classrooms other than science and mathematics. One thing should be noted here. We do not mean that validating and confirming the usefulness of such instruments is worthless. Validating them can increase the accuracy of the results and can help us to establish a new and independent field of classroom environment research specific to English classrooms. The second group involves instruments that should be reworded in one or two subcategories (e.g. Constructivist Learning Environment Survey (CLES)) and it is necessary to validate them for use in other educational levels and fields (like ELT). The third group contains those instruments elaborated for special classrooms (e.g. Science Laboratory Environment Inventory (SELI) and Geography Classroom Environment Inventory (GCEI)). They cannot be used in other classes but the experience provided by these instruments (especially SELI) is of worth. We can elaborate special instruments for special context of language teaching (e.g., language laboratory environment).

Table 1 presents some general information about five important instruments existing in the field of learning environments research. They are related to our first and second categories discussed above. For full details, the reader is directed to consult the related studies.

What is interesting about these instruments is their diversity. Each one has a specific underlying theory and is able to assess classroom environments with regard to that theory. In addition, having actual and preferred forms, these questionnaires will be of great assistance for researchers interested in exploring the levels of learners' satisfactions and investigating the environmental dimensions from which potential learners' dissatisfaction stem from.

If we want to use the instruments related to the second group, we should at first reword those items directly related to science and mathematics classrooms and then validate these instruments and confirm their usefulness for use in ELT or EFL classes. In this process, various analyses should be done with the aim of evaluating discriminant validity and factorial validity of each instrument and its ability to differentiate between the perceptions of students in different classes. We may have to change or omit some of the scales or items. For example, analysis of data from a Persian sample of 311 ELT students in 14 classes (Rahimi & Ebrahimi, under preparation) led to a final form of CLES-ELT (the name selected for the reworded and modified CLES for ELT classrooms) containing five five-item scales (i.e. the five-factor structure of CLES was replicated but the number of items reduced to 25).

Table 1 General information about some learning environment instruments

Instrument	Items per scale	Scales	Level	Related studies
Learning Environment Inventory (LEI)	7	Cohesiveness/Friction/ Favoritism/Cliqueness/ Satisfaction/Apathy/ Speed/Difficulty/ Competitiveness/Diversity/ Formality/Material Environment/Goal Direction/ Disorganization/Democracy	Secondary	Fraser et al., 1982; Walberg & Anderson, 1968
Classroom Environment Scale (CES)	10	Involvement/Affiliation/ Teacher support/ Task orientation/ Competition/ Order and organization/ Rule clarity / Teacher control / Innovation	Secondary	Fraser & Fisher , 1983; Moos, 1979; Moos & Trickett, 1987
Questionnaire on Teacher Interaction (QTI)	8-10	Helpful and friendly/Understanding/ Dissatisfied/Admonishing/ Leadership/Student responsibility and freedom/ Uncertain/Strict	Secondary/ Primary	Goh & Fraser, 1996; Riah et al., 1997; Waldrip & Fisher, 2003
Constructivist Learning Environment Survey (CLES)	6	Personal relevance/ Uncertainty/ Critical voice/ Shared control/ Student negotiation	Secondary	Taylor et al., 1997; Harwell et al, 2001
What Is Happening In This Classroom (WIHIC)	8	Student cohesiveness/Teacher support/Involvement/ Investigation/Task orientation/ Cooperation/Equity	Secondary	Huang et al., 1998, cited in Fraser, 1998

Method

Participants and data collection

ICEQ was administered among 32 (F=26 and M=6) second year EFL students studying at University of Kashan, Iran. They were from two classes taught by the same instructor. With regard to age, they were from 19 to 23.

ICEQ was given to the students at the beginning of the class. First, the instructor, being also one of the authors of this article, talked about the objectives of the study and got the idea across to the respondents that there was no right or wrong answers.

The instructor told them that they did not have to write their names on the given papers and the way they completed the questionnaires did not affect the instructor's attitude toward them and their class. Students were encouraged to think about how well each statement in the actual form of the questionnaire described what the class is like for them personally. In other words, they were encouraged to select from among the items which described their classroom environment as it was. They were also told that in the preferred form, they should select the items they preferred as the elements of an ideal classroom. Students were directed to read each statement carefully and they were given enough time to complete the questionnaire. All their questions were responded to attentively by the instructor. Then, they were encouraged to read the statements again, and at the end of data collection, the instructor thanked everyone for their participation and attention.

Data analysis

The students' responses to the Likert scale including almost never, seldom, sometimes, often and very often alternatives, were scored 1, 2, 3, 4 and 5 respectively. Each student was given a total score for each version of the questionnaire. These scores were sum of the scores given for each item. Five other scores each one being the sum of the scores given to items included in each scale were prepared for each form. In other words, each participant was given six scores for each form, a total one and five related to each scale.

The mean of all students' scores on each scale of the actual and preferred forms were also important and were included in the analysis.

Results

In this section, the questions set by the study will be explored one by one.

Question 1: Are there any significant differences between the actual and preferred classroom environments as perceived by the participating Persian EFL students?

With regard to the first question, the actual and the preferred forms were scored separately for each participant. In other words, each participant was given two total scores: one related to the preferred form and the other to the actual form. These total scores were the sum of the scores of each statement. Then two groups of scores were provided: total scores on actual form and total scores on preferred form for all participants. Then, SPSS was used to analyze the results and to conduct paired-sample t-test with the aim of determining whether there was a significant difference between the actual and preferred classroom environments as perceived by the participants. The results of the paired-sample t-test are provided in the Table 2.

As you can see from Table 2, the difference between total scores of all participants on actual form and total scores on preferred form is significant ($p < 0.05$). In other words, there is a significant difference between the actual and preferred classroom environments as perceived by the participating Persian EFL students.

Table 2 The results of the paired-sample t-test between total scores of all participants on the actual form and their total scores on the preferred form

	Paired Differences					
	Mean	Std. Deviation	Std. Error Mean	t	df	**Sig. P<0.05**
Pair 1 — total scores on actual form - total scores on preferred form	-9.81250	9.54932	1.68810	-5.813	31	**.000**

Question 2: To what classroom environment dimensions (i.e. personalization, participation, independence, investigation or differentiation) is the probable dissatisfaction attributed?

This question tries to determine where the difference stemmed from by taking into account the five dimensions of ICEQ. In the part we allocated to talking about ICEQ, it was mentioned that this questionnaire included five scales which were compatible with five dimensions of a learner-centered classroom environment: personalization, participation, independence, investigation and differentiation. To answer the second question, we gave each student five scores which were the sum of the scores of all statements in each scale.

Then we provided five groups of scores for each form of the questionnaire for all participants. In other words, we provided scores on personalization, participation, independence, investigation and differentiation dimensions for all students for each form.

These ten groups of scores (five derived from the actual form and five derived from the preferred form) were computed through SPSS for conducting different paired-sample t-tests between the scores of the same scales of the actual and preferred forms.

The results of these paired-sample t-tests are provided in Table 3. As it is clear, there are significant differences (p<0.05) between scores on personalization, participation, investigation and differentiation dimensions in the actual and preferred classroom environments. It can be asserted that overall dissatisfaction of Persian EFL students participating in this study with their learning environment originates from their dissatisfaction with each of these dimensions. Interestingly enough, there is no significant difference (p=0.83) between scores on independence dimension derived from actual form and those derived from preferred form. It can be inferred that the only sign of satisfaction in the English classroom environments perceived by EFL learners participating in this study is related to independence dimension.

Table 3 The results of different paired-sample t-tests between the scores of all participants on the five dimensions of actual and preferred form

		Paired Differences					
		Mean dif	Std. Deviation	Std. Error Mean	t	df	**Sig.** P<0.05
Pair 1	scores on *personalization* (Actual) - scores on personalization (*Preferred*)	-3.81250	3.41191	.60315	-6.321	31	**.000**
Pair 2	scores on *participation* (Actual) - scores on participation (*Preferred*)	-2.75000	2.88489	.50998	-5.392	31	**.000**
Pair 3	scores on *independence* (Actual) - scores on independence (*Preferred*)	-.09375	2.48037	.43847	-.214	31	**.832**
Pair 4	scores on *investigation* (Actual) - scores on investigation (*Preferred*)	-1.87500	2.37935	.42061	-4.458	31	**.000**
Pair 5	scores on *differentiation* (Actual) - scores on differentiation (*Preferred*)	-1.28125	3.02926	.53550	-2.393	31	**.023**

Question 3: How remote are the investigated English classroom environments as perceived by the participants from true individualized ones?

With regard to the last question, the answer is already clear. The mean differences in Table 2 represent the remoteness of each dimension of the English classroom environments under study from that of true individualized ones. These numbers are negative because the means of scores on different dimensions of actual classroom environment are less than the means of scores related to the dimensions of preferred classroom environments. In Fraser's (1998, p.21) words, "students preferred a more positive classroom environment than was actually present for all five environment dimensions". The pattern in which students prefer a more positive classroom learning environment than the one perceived as being currently present has been replicated in other studies (e.g. Margianti et al., 2001). The greater the mean difference, the remoter the related dimension will be from that of true individualized environments. As it is expected from the previous results of this study, the least amount of mean difference is related to the scores of independence dimension on actual and preferred forms. As it was mentioned this dimension is the only sign of participants' satisfaction

with their English classroom environments. For more clarification a linear graph (Figure 1) has been provided which demonstrates the remoteness of each dimension of actual classroom environments from that of true individualized ones. The coincidence point in Figure 1 represents the only sign of students' satisfaction and the only dimension which is in line with true ICEs.

In addition, estimates of the internal consistency of the five classroom environment scales (in the actual form) were calculated through SPSS by using Cronbach's alpha coefficient. Coefficients .84, .78, .86, .88, and .80 were obtained for Personalization, Participation, Independence, Investigation, and Differentiation dimensions respectively. All scales had good internal consistency for individual as unit of analysis.
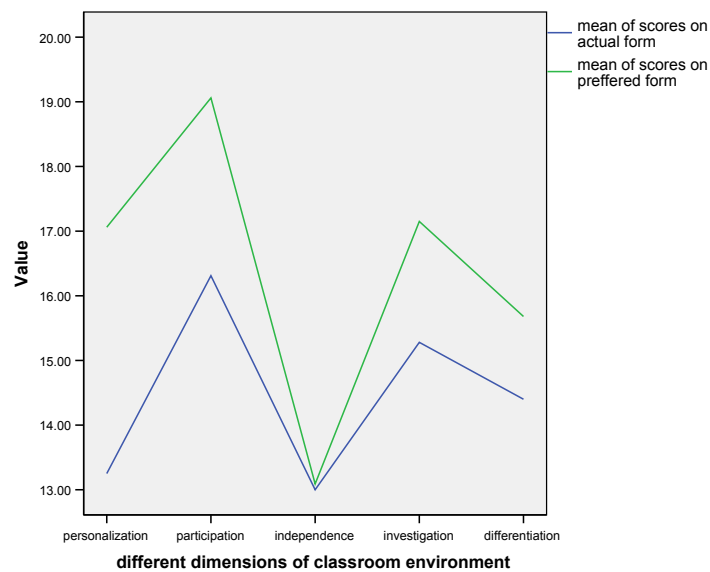
Figure 1 The remoteness of the investigated English classroom environments at University of Kashan, Iran, from true individualized learning environments. The remotest dimension is related to personalization and the closest is related to independence.

Discussion
Firstly, the results of this study indicated participants' dissatisfaction with their English classroom environments. Overall, the English classroom environments under study were not learner-centered and individualized and the interests and preferences of the students were ignored. In searching for the dimensions from which the students' dissatisfaction stemmed from, further investigation showed that lack of learner-centeredness in personalization, participation, investigation and differentiation dimensions of the classroom environments. If the instructor wants to create a learner-centered and individualized classroom environment which is in accord with his students' interests and preferences, he must take these dimensions as the main subjects of change and improvement. He should change his ideas and practices so that

the students have opportunities to interact with him (*Personalization*), be encouraged to participate in the classroom *(Participation)* and perform individual problem solving activities *(Investigation)*. He should also try to treat students individually (*Differentiation*). Much of his students' self-improvement, self-awareness, self-competition, motivation, satisfaction, and autonomy depends on these changes in practice and idea.

The results also brought about good news for the instructor. He was happy to see that he had created an environment in which the students perceive that they have control over their learning (*Independence*). Only with regard to the independence dimension, his classroom environment turned out to be in line with his students' interests and preferences.

The instructor was informed that the English classroom environments as perceived by his EFL students are too remote from true individualized ones. He should change and improve these classroom environments if he aims to train autonomous, self-directed and motivated learners all of which are the main goals of individualized instruction. It should be noted that changing classroom environments is not an easy task. Classroom environment can be conceived as a complete whole consisted of and affected by range of factors (e.g. how the role of instructor and students are defined, how learning is defined and what are the main aims of education). This leads to the difficulty of changing classroom environment and at the same time indicates comprehensiveness of classroom environment research. In other words, since in studying learning environments all these factors are taken into account, the pictures such studies present are general and comprehensive ones hardly obtainable by other means.

Conclusion

In this study we tried to promote and direct learning environment research in the field of ELT in new and systematic ways. We believe that experiences gained through decades of research mainly in science and mathematics classroom environments have a lot of potentialities for ELT practitioners. The instruments existing in the field of learning environment research were divided into three groups. The first group contained instruments that were related to general educational ideas and that could be used in all levels of education and different classes and in the present study we used one of these instruments (i.e. ICEQ). The second group included instruments originally developed and validated for science and mathematics classrooms. If we are going to benefit from such instruments (e.g. CLES), they should be reworded in some items and then they should be validated for use in English classrooms. In this process, various analyses should be done with the aim of evaluating discriminant validity and factorial validity of each instrument and its ability to differentiate between the perceptions of students in different classes. The third group involved those instruments elaborated exclusively for science or other classrooms. These instruments (e.g. Science Laboratory Environment Inventory (SLEI)) cannot be used in classes other than the ones they are originally elaborated for. But here we are directed to elaborate instruments evaluating psychosocial environments of special contexts

related to English teaching and learning. For example, we can elaborate and validate a questionnaire assessing language laboratory environments.

Diversity of these instruments is also of great assistance. Each one has a specific underlying theory and is able to assess classroom environments with regard to that theory. In addition, having actual and preferred forms, these questionnaires will be of great assistance for researchers interested in exploring the levels of learners' satisfactions and investigating the environmental dimensions from which potential learners' dissatisfaction stem from.

We also tried to respond to the deficiency proposed by Fraser (2002) about the lack of any published article that reports Persian researchers' attempts to use learning environment assessments to improve classroom environments.

We used ICEQ to determine whether there were significant differences between the actual and preferred classroom environments as perceived by some Persian EFL students, what dimension(s) of the investigated English classroom environments were (or are not) in accord with the participants' interests and preferences and finally, how remote the explored English classroom environments perceived by the participated Persian EFL students were from true individualized ones. The answers related to the first two questions were directly applicable by the instructor. Through this study he found that the English classroom environments in which he played the role of an instructor were not learner-centered and individualized and the interests and preferences of his students were not taken into consideration. More exploration revealed to him that his students' dissatisfaction of their learning environments stemmed from their dissatisfaction of personalization, participation, investigation and differentiation dimensions. He found that if he wanted to create a learner-centered and individualized classroom environment which was in accord with his students' interests and preferences, these dimensions were the main subjects of change and improvement. The result also showed him that the only sign of satisfaction in the English classroom environment as perceived by his students was related to independence dimension.

In addition, the result showed the instructor that the English classroom environments mainly affected and formed by his ideas and approaches were too remote from true ICEs. He found that his classrooms had to be changed to become places that genuinely offer students a voice and a shared role in their learning process. Beyond gainsay, constructing such environments is complex and highly demanding. "The teacher must be well versed in pedagogy, learning theories, curriculum planning and design, time management and organization, quality assessment and team skills" (Murdokh & Wilson, 2008, p.1). The goal must be to help students grow into inquiring and resourceful individuals who can demonstrate what they know and, importantly, what to do when they don't know. In a world of rapidly escalating and changing knowledge, the students must be able to select, critique and apply their learning across a wide range of contexts as well as adaptable, flexible, resilient individuals with the ability to learn and relearn (Murdokh & Wilson, 2008, p.1).

The first one who was informed of the above-mentioned results was the instructor himself. He was interested in whether his classroom environment was learner-

centered and in accord with his students' ideas, whether his students perceived their classroom environments positively, what measure he should take to create a learner-centered classroom environment. He found the answer of his questions through this study. The classroom environments in which he played the most influential role were remote from true ICEs. The finding showed him the remoteness of the English classroom environments from learner-centered ones and the dimensions that should be focused upon for change and improvement. Given the importance of individualized language teaching, he inferred that the failure of his students in English classrooms is mainly due to non-individualized instruction. It is far from reality to expect students to be proficient when they find their learning irrelevant, demotivating, meaningless and unsatisfactory. He was suggested that it would be better to add more individualized and personalized dimensions into his planning and practice and he was directed to put students at the center of teaching. The syllabus used in his classrooms should be more eclectic and more flexible to cover different learners' interests, beliefs and needs. The lesson plans he prepares for the classroom should reflect multiple intelligences and cognitive differences. His students should be encouraged to take part in class decisions. He was suggested to gain students' feedbacks through dialogues in and out of the classroom. In doing so, much of his learners' dissatisfaction could decrease. He could be hopeful that such changes in theory and practice ended up in more individualized environments.

The study also tried to provide new ways for looking at learning environment research in the field of ELT. Experiences gained through decades of learning environment research mainly in science and mathematics classrooms were presented and some of the useful instruments existing in the field were introduced for use in the field of ELT. In addition, directions for further research were provided implicitly or explicitly. This line of research can open up fruitful areas in ELT.

References

Allodi, M.W. (2002). A two-level analysis of classroom climate in relation to social context, group composition, and organization of special support. *Learning Environments Research, 5*, 253–274.

Asghar, M. & Fraser, B.J. (1995). Classroom environment and attitudes to science in Brunei Darussalam. *Journal of Science and Mathematics Education in Southeast Asia, XVIII*(2), 41– 47.

Breen, M. P. & A. Littlejohn. (2000). The significance of negotiation. In M. P. Breen & Littlejohn (Eds.) *Classroom decision-making: Negotiation and process syllabuses in practice* (pp.5-38). Cambridge: Cambridge University Press.

Den Brok, P., Fisher, D., Rickards, T., & Bull, E. (2006). Californian science students' perceptions of their classroom learning environments. *Educational Research and Evaluation, 12*, 3–25.

Dileo, John. (2007). *Individualized instruction*. Retrieved August 22, 2008 from http://www.dropoutprevention.org/effstrat/individualized_instruction/overview.htm

Dorman, J.P. (2001). Association between classroom environment and academic efficacy. *Learning Environments Research, 4*, 243–257.

Dorman, J.P., Fraser, B. J., & Aldridge, J. M. (2006). Using students' assessment of classroom environment to develop a typology of secondary school classrooms. *International Education Journal, 7*(7), 906-915.

Franz, J. M. (1990, July). *Design education in architecture: a psychosocial analysis.* Paper presented at Architectural Science and Design in Harmony, Joint ANZAScA / ADTRA Conference, Sydney.

Fraser, B.J. (1984). Differences between preferred and actual classroom environment as perceived by primary students and teachers. *British Journal of Educational Psychology, 54*, 336–339.

Fraser, B.J. (1990). Individualized Classroom Environment Questionnaire. *The Australian Council for Educational Research Ltd., Sidney, AUS.* Retrieved August 14, 2008 from http://www.emc.cmich.edu/charactered/iceq.htm

Fraser, B.J. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 493–541). New York*:* Macmillan.

Fraser, B.J. (1998). Classroom environment instruments: Development, validity and applications. *Learning Environments Research, 1*, 7– 33.

Fraser, B.J. (2002). Learning environment research: yesterday, today and tomorrow. In Goh, S.C, & Khine, M.C (Eds.), *Studies in educational learning environments: an international perspective* (pp.1-25). World Scientific Publishing Co. Pte. Ltd.

Fraser, B.J., Anderson, G.J. & Walberg, H.J. (1982). *Assessment of learning environments: manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI)* (3rd version). Perth, Australia: Western Australian Institute of Technology.

Fraser, B. J., & Fisher, D.L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal, 19*(4), 498-518.

Fraser, B.J., & Fisher, D.L. (1983). Use of actual and preferred classroom environment scales in person-environment fit research. *Journal of Educational Psychology, 75*(2), 303–313.

Fraser, B.J. & Lee, S.U. (2009). Science laboratory classroom environments in Korean high schools. *Learning Environment Research, 12*, 67–84.

Goh, S.C. & Fraser, B.J. (1996). Validation of an elementary school version of the Questionnaire on Teacher Interaction. *Psychological Reports, 79*, 512– 522.

Griffiths, G. & Keohane, K. (2000). *Personalizing language learning.* Cambridge: Cambridge University Press.

Harwell, S.H., Gunter, S., Montgomery, S., Shelton, C. & West, D. (2001). Technology integration and the classroom learning environment: Research for action. *Learning Environments Research, 4*, 259–286.

Kember, D., Ho, A. & Hong, C. (2009). Characterising a teaching and learning environment capable of motivating student learning. *Learning Environments Research*, *13*(1), 43-57.

Margianti, E.S., Fraser, B.J. & Aldridge, J. (2001). *Investigating the learning environment and students' outcomes in university level computing courses in Indonesia*. Paper presented at the annual conference of the Australian Association for Research in Education, Fremantle, Australia.

McGroarty, M. (1998). Constructive and constructivist challenges for applied linguistics. *Language Learning, 48*, 591-622.

Moos, R.H. (1979). *Evaluating educational environments: procedures, measures, findings and policy implications*. San Francisco, CA: Jossey-Bass.

Moos, R.H. & Trickett, E.J. (1987). *Classroom Environment Scale manual (2nd ed.)*. Palo Alto, CA: Consulting Psychologists Press.

Murdokh, K. & Wilson, J. (2008). *Creating a learner-centered primary classroom*. Routledge, Taylor & Francis e-Library.

Ozkok, A., Walker, S.L. & Buyukozturk, S. (2009). Reliability and validity of a Turkish version of the DELES. *Learning Environment Research, 12*, 175–190.

Quek, C., Wong, A., Divaharan, S., Liu, W., Peer, J. & Williams, M.D. (2007). Secondary school students' perceptions of teacher–student interaction and students' attitudes towards project work. *Learning Environment Research 10*, 177–187.

Rentoul, A., & Fraser, B.J. (1979). Conceptualization of enquiry-based or open classroom learning environments. *Journal of Curriculum Studies, 11*(3), 233–245.

Riah, H., Fraser, B.J. & Rickards, T. (1997). *Interpersonal teacher behaviour in chemistry classes in Brunei Darussalem's secondary schools*. Paper presented at the International Seminar on Innovations in Science and Mathematics Curricula, Bandar Seri Begawan, Brunei Darussalam.

Sarigoz, I.H. (2008). Towards individual centered foreign language teaching. *Journal of Language and Linguistic Studies, 4*(1), 51-64.

Scott, R., & Fisher, D.L. (2004). Development, validation and application of a Malay translation of an elementary version of the Questionnaire on Teacher Interaction. *Research in Science Education, 34*, 173–194.

Taylor, P.C., Fraser, B.J. & Fisher, D.L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293– 302.

Walberg, H.J. & Anderson, G.J. (1968). Classroom climate and individual learning. *Journal of Educational Psychology, 59*, 414– 419.

Waldrip, B.G & Fisher, D.L. (2003). Identifying exemplary science teachers through their classroom interactions with students. *Learning Environments Research,6*, 157–174.

Wanpen, S., & Fisher, D.L. (2006). Creating a constructivist learning environment in a computer classroom in Thailand using the CLES. In D. L. Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments: World views* (pp. 297–312). Singapore: World Scientific.

Wei, M., Brok, P. & Zhou, Y. (2009). Teacher interpersonal behavior and student achievement in English as a Foreign Language classrooms in China. *Learning Environment Research, 12*, 157–174.

Wheldall, K., Beaman, R., & Mok, M. (1999). Does the Individualized Classroom Environment Questionnaire (ICEQ) measure classroom climate?. *Educational and Psychological Measurement, 59*(5), 847-854.

Williams, M., & Burden, R.L. (1998). *Psychology for language teachers: a social constructivist approach*. Cambridge: Cambridge University Press.

Title

Assessing Iranian Students' English Ability: Semantic Competence vs. Pragmatic
Competence

Authors

Reza Keshavarz (M.A.)

Islamic Azad University, Shiraz Branch, Iran

Rasool Roozegar (Ph.D. candidate)

Islamic Azad University, Marvdasht Branch, Iran

Bio Data

Reza Keshavarz, M.A. of TEFL at the Islamic Azad University, Shiraz Branch,
Shiraz, Iran. At present, he studies at department of English language. His main
research interests are language teaching and the application of Research modeling in
social science. He has endeavored to work as a teacher and a tutor during his studying
at University.

Rasool Roozegar, Ph.D. candidate of Statistics at Shiraz University, Shiraz, Iran. He
is a lecturer at the Islamic Azad University, Marvdasht Branch, Marvdasht, Iran. His
main research interests are teaching, testing and the application of Statistical modeling
in all sciences, especially social science research. He has published and presented
papers in national and international conferences and journals.

Abstract

Visiting the problem of current English assessment in Iranian classroom settings a
recommendation was made to assess English language ability. Two kinds of test that
is semantic and pragmatic competence tests were made and administered to students
of Shiraz Iran Language Institute. Among them, 70 are advanced level students (aged
from 17 to 19) and 173 are high level students (aged from 13 to 15) which are
composed of 55 high1, 59 high2 and 59 high3 students. And the ratio of male and
female is roughly equal (126 male and 117 female). They were both males and
females sharing the same mother tongue and none of them had ever been in an
English speaking country. The data were analyzed through statistical procedures and
the results revealed that Iranian EFL students are poor in pragmatic issues and they

have a lot of problems. As hypothesized in terms of item percent correct, items measuring both semantic and pragmatic competence will become easier as students progress from high level to advanced levels. Whereas in terms of item discrimination, semantic items are more discriminating than pragmatic items in the current context.

Keywords: Pragmatic Competence, Semantic Competence, item percent correct, item discrimination, EFL learners.

Introduction
Preliminaries
The development of pragmatic and sociolinguistic rules of language use is important for language learners. It is necessary to understand and create language that is appropriate to the situations in which one is functioning, because failure to do so may cause users to miss key points that are being communicated or to have their messages misunderstood. Worse yet is the possibility of a total communication breakdown and the stereotypical labeling of second language users as insensitive, rude, or inept people (Thomas, 1983).

Semantic competence comprises the ability to organize the structure of language into grammatically correct sentences and to understand the meaning of words. This competence can be further broken into two types: grammatical and textual. The grammatical competence consists of knowledge of vocabulary and syntax, whereas textual competence consists of knowledge of the conventions for joining words together to form a meaningful text according to the rules of cohesion and rhetorical organization. Items designed in traditional testing formats are mainly aimed to measure this type of competence.

Pragmatic competence pertains to understanding linguistic signals used in communication and knowing how they are used to refer to persons, objects, ideas, and feelings. The concept of pragmatic competence includes illocutionary competence and sociolinguistic competence. The illocutionary competence refers to the ability to perform a function by means of speaking (e.g., asking someone to leave so that you may get rest)." Sociolinguistic competence is the sensitivity to, or control of the conventions of language use that are determined by the features of the specific language use context; it enables us to perform language functions in ways that are appropriate to that content." (Bachman, 1991) In other words, sociolinguistic competence refers to sensitivities to differences in naturalness and cultural registers in a communication context. This type of competence is often overlooked in the assessment of communicative ability.

Objectives of the study
This study seeks to gain some understanding about the general level of pragmatic competence of Iranian students in comparison with their semantic competence in the use of the English language. That is to say, the study aims to find out if those learners

who are semantically competent, have the same ability to deal with pragmatic issues. In other words, does being semantically competent guarantee coping with pragmatic problems? To get a better coverage of the whole age spectrum, we sampled students from different levels of English language Iran Language Institute (aged from 17 to 19) and an instrument designed to measure both semantic competence and pragmatic competence will be constructed and then administered to these students. The statistics (percent correct and discrimination) of all the items making up the instrument will form the basis for testing our hypotheses of interest.

Research hypotheses

H1. With respect to percent correct, we hypothesize that the semantic items will show higher percent correct than the pragmatic items because students are more familiar with this type of items than with the other kinds. Of course, if the English education has done its job right, items measuring both semantic and pragmatic competence will become easier as students progress from year 7 of the middle school to college.

H2. With respect to item discrimination, semantic competence will be more discriminating than pragmatic competence in the current context of the Iranian English education.

Null hypotheses

H1. The semantic items will not show higher percent correct than the pragmatic items.

H2. Semantic competence will not be more discriminating than pragmatic competence.

Literature Review

Introduction

The issue of "pragmatic English ability" and how it may be assessed in the learning of English as a foreign language in the Iranian context has for a long time been a strong interest of us. We strongly suspect that the English education throughout all levels of formal schooling in Iran is failing to develop students' functional abilities in using English as a foreign language. Though neither of the authors actually came from the field of teaching English as a foreign language, we decided to embark on this research, both as informed observers of current teaching practices and as concerned parents of current school-age children, in the hope that some light may be shed in regard to uncovering why the system fails to produce functional students. We chose the area of formal assessment as our stepping stone into understanding how the English education operates in Iran nowadays.

Our interest was initially triggered by the contents of some major English examinations my then middle-school age sister brought home periodically. From these tests we were often struck that the abilities required of students to score high do not necessarily equate with the ability to function in an ordinary daily living English language setting.

Our interest was reinforced when we were solicited by the parents of the students in her form class to serve as the teacher's assistants to help her maintain classroom order

on a one-hour per week basis. While serving in this role we chatted with mainstream English teachers who had students approximately the same age as ours (then 14-15). We mentioned our concern in regard to both the methods of teaching and the contents of assessment, but to our surprise our suggestion to bring the contents of English education closer in line with developing students' pragmatic English ability was met largely with strong resistance. To their credit indeed, most of these teachers do teach in the belief that they teach in ways that are most beneficial for their students in the Iranian environment. However, we did find solace in some parents in the course of working with the school who shared our views. Most of these parents are not grass-root kind of people. Rather, they are well educated and mostly have had overseas experiences.

Bardovi-Harlig (1998) found that the following factors have a direct influence on the acquisition of pragmatic competence: input, instruction, level of proficiency and length of stay living in the L2 culture, and the L1 culture. Shortcomings pertaining to input may be found in academic materials such as textbooks or even the instructor. The present author's cursory analysis of several textbooks reveals that speech acts are not presented at all; therefore, primarily using textbooks to teach students pragmatic information about a language may be ineffective. Also significant to the type of input available is the input provided by the instructor. For instance, the instructor may consider it appropriate to use one register when speaking to the learner but inappropriate for the learner to produce an utterance using the same register. This case may be illustrated in the case of imperatives, where the instructor uses an informal register to address the student but expects to be addressed in a formal register. This choice, then, is dependent upon what the instructor considers appropriate according to his/her understanding of the cultural norms of the target language.

Instruction may also be influenced by the instructor's emphasis on similarities and differences between the L1 and L2. The amount of exposure to specific pragmatic features may have an effect on the learner's pragmatic awareness. As testing assessments have typically evaluated the learner's linguistic competence through grammar-oriented tasks, the learner may not feel it is necessary to prepare for tasks based on pragmatic understanding of these forms. As such, the pragmatic component of language learning is neglected.

Types of Items Incompatible with Developing Students Pragmatic Ability

With regard to pronunciation, a lot of effort is devoted to linking English sounds with similar sounds in Iranian but the real serious problem is the lack of knowledge that certain words can be pronounced differently in different parts of the world. Take the following item for instance:

Which of the following has the same pronunciation as in data?

(1)Table (2)Pat (3)Mat (4)Gather (5)Cat

The school provided correct answer is (1), according to the North American way of pronunciation. But in fact, all the remaining four options are also correct, if you come from England or the British Commonwealth countries such as Australia. To say the

other options are wrong is a blatant rejection of the diversity our world is showing us today.

What bothers us most is the type of items that require inflexible choices of words for objects or expressions. It also happens to be the most common type of ridiculous items in our observation. We were struck by this item that appeared in my sister's first major English exam:

How are you?
 (1) Good.
 (2) Fine.
 (3) Fine, thank you.
 (4) Fine, thank you. And you?

The correct answer was given as (4), because apparently it was copied from a conversation from the standard textbook. Other options are deemed incorrect simply because they are not identical to what the textbook says. Another common form of inflexibility is the choice of prepositions. Instead of teaching students the meanings of prepositions, many teachers' present prepositions in certain combinations of words and only these combinations are deemed as correct answers in a test. Consider this item:

I am responsible _____ you.
 (1) for
 (2) of
 (3) with
 (4) to

The answer that the teacher is looking for is (1), but actually (4) makes perfect sense too albeit it takes on the opposite meaning. It is difficult to convince Iranian teachers that items of this sort with only one sentence can take on more than one correct answer. Another example is, while "It is very kind of you" is being wholeheartedly accepted as a legal sentence, "I kind of like you" is a no-no according to most English teachers because the preposition "of" is taught to be followed by a noun but not by a verb in its original form.

The type of items that I found most amusing, but certainly not conducive to developing students' pragmatic English ability is probably the kind of items that are considered to be grammatically correct, whereas in fact they would most likely confuse even the native English speakers. This can be best illustrated by a multiple choice item found in one of my daughter's major exams during her last year:

"Both John and Mary are not happy." means:
 (1) John is not happy.
 (2) Mary is not happy.
 (3) John and Mary are not happy.
 (4) One of them is happy.

The correct answer is (4) according to the strict grammatical rule, but to most people who are familiar with the English language, (3) appears to be the most appropriate answer, albeit it is an awkward form of expression. Items of this sort abound in school English tests across Iran. It is certainly no surprise that students who

are taught to focus on grammatical structure of sentences cannot produce satisfactory results when real-life English language competency is required. This phenomenon seems to go contrary to the worldwide trend in teaching English as a foreign language: The integration between teaching and assessment should focus on the "language as a tool for communication rather than on language knowledge as an end in itself" (Brindley, 1995, p.158).

Assessing Communicative Ability

It is important to point out that the designs for assessment of communicative ability differ from the traditional test designs in two major features: it is situation specific and it is production oriented. The former refers to the fact that by the definition of communication the contents of assessment are specific to a certain audience (people of certain occupations, areas, etc). The latter refers to the need that the communication requires oral or written performance. The production orientation feature of communicative language testing cannot be measured by traditional dichotomously scored item formats, and thus will not be a concern in the test design for this study. Instead, our current study will suffice by investigating the situation specific feature.

By excluding the production orientation feature from our test design, we by no means downplay its importance in the assessment of communicative language ability, of which production is an integral part. On the contrary, we uphold its importance, and the paucity of test designs aiming to measure this part of communicative ability calls for more research into coming up with valid assessment instruments for measuring students' abilities in this regard. Nevertheless, this endeavor is beyond the scope of our current study. In this study, we chose to use two types of competence to represent communicative ability: semantic competence and pragmatic competence. Our conceptualization was taken in part from Bachman's (1991) framework for communicative ability.

Semantic Competence versus Pragmatic Competence

Semantic competence comprises those abilities involved in organizing the structure of language for recognizing grammatically correct sentences and understanding the meaning of words. This competence can be further broken into two types: grammatical and textual. The grammatical competence consists of knowledge such as vocabulary and syntax, whereas textual competence consists of knowledge of the conventions for joining words together to form a meaningful text according to rules of cohesion and rhetorical organization. Items designed in traditional testing formats are mainly aimed to measure this type of competence.

Pragmatic competence pertains to understanding linguistic signals used in communication and how they are used to refer to persons, objects, ideas, and feelings. The concept of pragmatic competence includes illocutionary competence and sociolinguistic competence. The illocutionary competence refers to the ability to perform a function by means of speaking (e.g., asking someone to leave so that you may get rest). The sociolinguistic competence is "the sensitivity to, or control of the conventions of language use that are determined by the features of the specific

language use context; it enables us to perform language functions in ways that are appropriate." (Bachman, 1991, p.94) In other words, sociolinguistic competence refers to sensitivities to differences in naturalness and cultural registers in a communication context. This type of competence is often overlooked in the assessment of communicative ability.

The Current Study and Research Hypotheses

Pragmatics is the branch of linguistics which studies those aspects of language meaning which cannot be predicted from linguistic knowledge alone. An American philosopher Paul Grice, referred to by linguist Jean Aitchison (1999) as the "father of pragmatics", provided a philosophical basis for assessing pragmatic competence in a second language (L2) discourse setting. Grice emphasized four rules of a conversation that help human beings communicate efficiently to one another: quantity, quality, relevance and manner. For the purposes of this study, the description offered by Grice has been adopted for assessing pragmatic competence in the setting of learning English as a foreign language. We do recognize that although pragmatic competence as a whole can not be effectively tested using a written testing format, there are aspects of language that can be tested for that give a good indication of the progress language users have made towards achieving pragmatic competence. Grice supplied a good starting point from which we can test students' pragmatic competence by testing on the additional information given in a discourse beyond what can be explained by linguistic knowledge alone, namely, by applying the four rules of conversation.

The unique dimensionality of pragmatic competence in the assessment of communicative ability for students who learn English as a foreign language and its associated psychometric properties have been confirmed in a most recent study (Chou & Chen, 2005). They showed that items can be designed to measure the pragmatic aspect of the English communicative ability with good reliability and discrimination. In this study we seek to gain some understanding about the general level of pragmatic competence of the Iranian students in comparison with their semantic competence in the use of English language. To get a better coverage of the whole age spectrum, we sampled students from two levels of English language education: high levels (aged from 13 to 15) and advanced levels of Iran Language Institute (aged from 17 to 19). An instrument designed to measure both semantic competence and pragmatic competence will be constructed and then administered to these students. The statistics (percent correct and discrimination) of all the items making up the instrument will form the basis for testing our hypotheses of interest. The following two major research hypotheses are then postulated:

1. With respect to percent correct, we hypothesize that the semantic items will show higher percent correct than the pragmatic items because students are more familiar with this type of items than with the other kind of items. Of course, if the English education has done its job right, items measuring both semantic and pragmatic competence will become easier as students progress from Year 7 of the middle school to college.

2. With respect to item discrimination, semantic competence will be more discriminating than pragmatic competence in the current context of the Iranian English education.

Methodology

Introduction

This chapter deals with issues such as the participants and instruments as well as data collection procedures and statistical procedures for data analysis.

Participants

Participants in the present study were students of Shiraz Iran Language Institute. Among them, 70 are advanced level students and 173 are high level students which are composed of 55 high1, 59 high2 and 59 high3 level students. They were both males and females sharing the same mother tongue and none of them had ever been in an English speaking country. And the ratio of male and female is roughly equal (126 male and 117 female).

Instrumentation

Three settings were given as three sections in the form of three mutually independent work samples from which students must find answers to a number of questions asked according to information given. The first two sections are related to pragmatic competence in which the correct answers are not given in the text explicitly, but instead, one must apply the four rules in reaching a correct answer. The third section measures semantic competence, whose answers can be found directly in the text. All items take on the form of multiple choice formats, in which there is only one most appropriate answer out of four or five given options. The test was officially administered in two stages. In the first stage, a test of 19 items in the form of three work samples was given to high level students of Iran Language Institute. In the second stage, an additional item was added to the third section measuring semantic competence, and the test was then given to some advanced level students. The additional item was added to make the best use of space, as well as to bring the total number of items to a more familiar number, 20 instead of 19, following the advice of a colleague who reviewed the results from the first stage data analysis.

Test design

Our test design possesses the two following features. Firstly, it conforms to what McNamara (1996) described as a "weak performance test" in that getting the knowledge-based accurate answer is not as important as making the appropriate language choice in a discourse. A weak performance test therefore, tests the ability of an individual to communicate in a general sense without referring to any specific knowledge area. A "strong performance test", in contrast, would be a test that success would be judged in terms of both the subject area knowledge and the English language proficiency.

Secondly, our design catches the essence of what Lado (1961) proposed in the design of language tests. Lado thought language tests should concentrate on testing "control of the problems", which refer to the units and patterns of the target language that do not have a counterpart in the student's native language, ….….." (Lado, 1961, p.24). Our test items were designed in a way that there are no direct counterparts for the words and phrases in Iranian as they appear in the items, but the expressions are so commonly used in English language that if one fails to understand what they mean in a discourse, one can not reasonably be perceived as functional in the English language.

The instrument is designed in a way that there will be no controversy in scoring for all native English speakers as well as for people with a functional level of knowledge in using English as a foreign language. Given the authors' familiarity with the English-speaking conditions in Southeast Asia, had such a test been given to students of middle school level in Hongkong, Malaysia or Singapore, it would most likely yield percent-corrects greater than .7 for all items.

Subjects

We sought and acquired cooperation from six English teachers (two from high1 level, high2 level, and high3 level, respectively) in Iran language institute for the purposes of the current study. The instrument was given by the first author following all the procedures of a standardized test administration. A total of 243 test papers were collected, of which 55 are from high1 level, 59 from high2 level, and 59 from high3 level. The ages of advanced level students ranged from 17 to 19 and a total of 70 test papers were collected.

There is roughly equal representation of both genders across all levels in high level students. It is expected that the advanced level students would outperform the high level students due to the nature of the admission process of selecting students into the higher education institution. All students were given 30 minutes time to complete the test, but for all high level students, most students (over 80%) completed in less than 20 minutes, whereas for advanced level students, most students completed in less than 15 minutes.

Statistical procedure for data analysis

Two types of statistics are analyzed concerning the psychometric properties of an item: percent-correct and item discrimination. The former is measured by the percentage of examinees answering the item correctly, also commonly referred to as the difficulty level of the item. The higher the percent correct is, the easier the item is. The latter is measured by point-biserial correlation coefficient which gives indication of the strength of association between the item score and the total scale score. Discrimination in the context of psychometric properties means the ability of an item to separate high and low ability students and is thus a desirable trait. The higher the point-biserial correlation, the more discriminating the item is. In our current study, the percent correct and discrimination statistics for all twenty items are given in the Table

below (next page), which provides the empirical basis for all subsequent statistical analyses.

A split-plot ANOVA design is used to analyze the effect of work sample (Type) and the effect of level of schooling (Level) with Type serving as the whole-plot treatment variable and Level serving as the sub-plot variable. The two dependent variables here are percent correct scores and discrimination indices. It is noteworthy to mention that the units of analyses in our design are the items of the test, not the students taking the test. Type is a whole-plot variable because the items in three work samples are different items. Each item is used four times, which correspond to measurements taken at high1, high2, high3 levels and advanced level time points sequentially. This design is commonly referred to as the repeated measures design in the context of social science research methodology (Shaughnessy, Zechmeister and Zechmeister, 2006). The advantage of using such a design lies in its ability to test both the main effects (Type and Level) and their perceivable interaction effect simultaneously.

Table 1 Item analysis for three work samples by four levels of schooling

Item No	Percent Correct				Discrimination			
	High1	High2	High3	Advanced	High1	High2	High3	Advanced
1-1	.213	.170	.120	.409	.186	.372	.246	.594
1-2	.398	.318	.481	.743	.580	.666	.712	.498
1-3	.351	.511	.480	.771	.524	.706	.610	.625
1-4	.480	.569	.491	.928	.624	.661	.739	.483
1-5	.196	.321	.372	.591	.620	.581	.624	.666
2-1	.271	.349	.510	.641	.441	.501	.445	.596
2-2	.232	.427	.406	.692	.641	.546	.618	.475
2-3	.309	.418	.364	.718	.590	.641	.481	.290
2-4	.359	.444	.439	.462	.350	.248	.209	.080
2-5	.158	.191	.157	.101	.224	.228	.534	.541
2-6	.423	.542	.438	.984	.631	.286	.634	.292
2-7	.439	.434	.539	.671	.350	.472	.301	.229
2-8	.231	.310	.275	.678	.320	.468	.551	.564
3-1	.464	.474	.384	.918	.718	.871	.577	.702
3-2	.390	.388	.281	.931	.752	.791	.792	.518
3-3	N/A	N/A	N/A	.588	N/A	N/A	N/A	.584
3-4	.194	.198	.179	.512	.562	.643	.712	.659
3-5	.545	.468	.377	.799	.780	.886	.831	.669
3-6	.321	.318	.368	.879	.694	.734	.865	.531
3-7	.174	.171	.102	.743	.684	.634	.532	.501

Data Analysis and Interpretation

Item Percent Correct Analysis

The results of ANOVA for item percent correct are reported in the following table. It is not surprising that Level was found to be such a highly significant effect because students are expected to improve as they receive more education. Despite the non-significant indication for the main effect of Type, its interpretation is complicated by the presence of the significant interaction effect between Type and Level (F=10.070, P<.000). We shall take a more detailed look at the nature of this interaction effect by examining the plot of means for three work samples across four levels of schooling as shown in the following Figure.

As seen, the interaction effect may mainly be attributed to the differential role semantic competence had played across four levels of schooling in comparison to the pragmatic competence. The percent correct associated with semantic items is roughly the same or lower than the pragmatic items for high level students (high1 to high3). This situation turned around in which semantic items become easier as compared to their pragmatic counterparts.

Table 2 ANOVA results for percent correct analysis

Source	df	SS	MS	F	P>F
Type	2	.007	.004	.524	.595
Item(Type)	3	1.665	.555	78.339	.000
Level	6	.151	.025	3.542	.006
Type* Level	16	1.213	.076	10.070	.000
Error	48	.340	.007		

Note. Item (Type) is used as the error term for testing the effect of Type, Whereas Error is used to test the effects of Level and Type*Level.
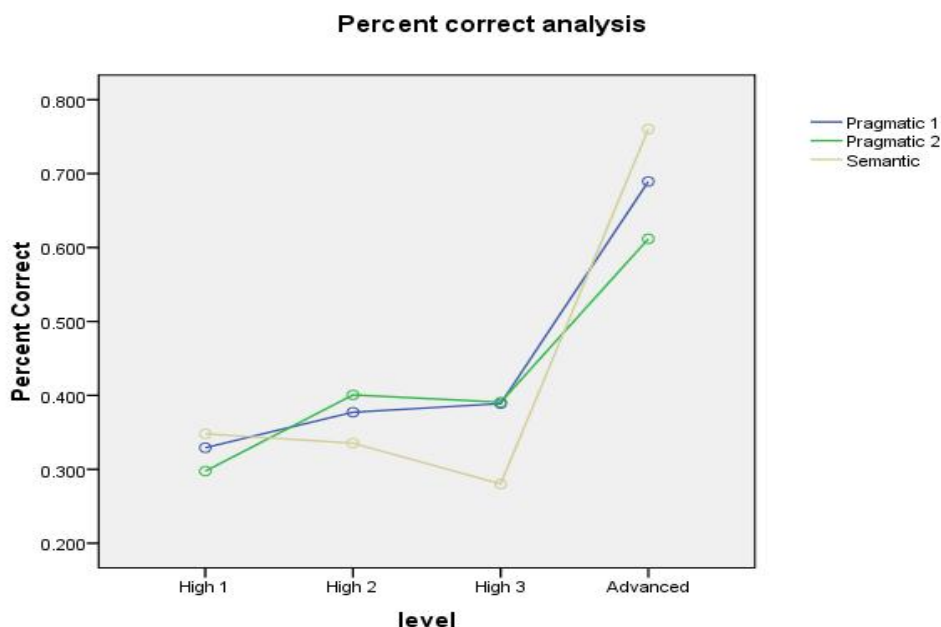
Figure 1 Percent correct analysis

Item Discrimination Analysis

The results of ANOVA for item discrimination are reported in the following table. No interaction effect was found. Both Type and Level main effects are significant, with Type as a highly significant main effect (F=33.642, P<.000) and Level as a moderately significant effect (F=1.827, P<.055). The plot of means for three work samples across four levels of schooling is shown in the following Figure.

As seen, Semantic items are more discriminating for the high level students. This result is certainly not surprising because high level students are more used to items measuring semantic competence than they are to the items measuring pragmatic competence. After all, the so called high English ability students at the high level are commonly perceived to mean those who can memorize answers to fixed patterns of tests in which the contents have all been previously exposed to them (all words must have been taught in class). For the advanced level students, items measuring semantic competence seem to be equally discriminating as the pragmatic items. This is also expected because students tend to increase their semantic competence as they progress from high levels to advanced level. At this stage, students with higher semantic competence will also tend to have higher pragmatic competence.

175

Table 3 ANOVA results for Discrimination analysis

Source	df	SS	MS	F	P>F
Type	2	.944	.472	33.642	.000
Item(Type)	16	.688	.043	2.883	.002
Level	3	.077	.026	1.827	.055
Type* Level	6	.067	.011	.799	.576
Error	48	.673	.014		

Note. Item (Type) is used as the error term for testing the effect of Type, Whereas Error is used to test the effects of Level and Type*Level.
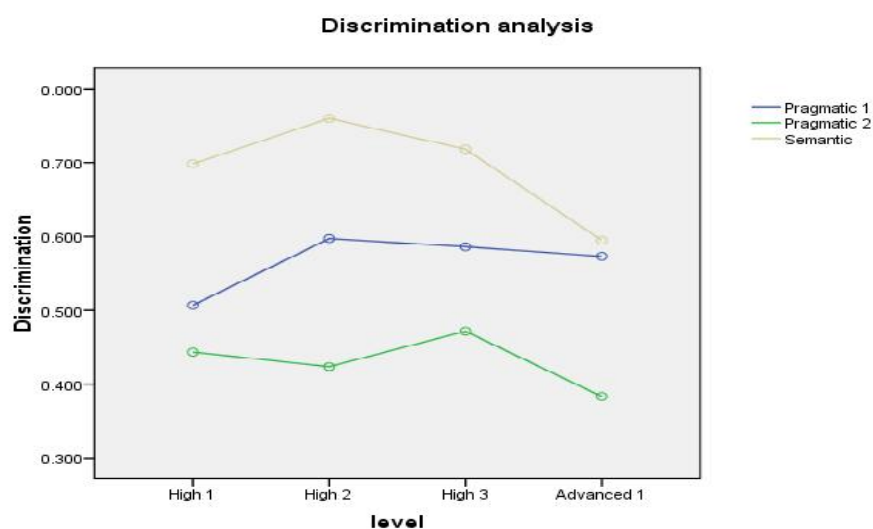

Figure 2 Discrimination analysis

Results

General English Ability

Despite the lack of representativeness of our sample, it is disappointing to see that Iranian students' English communicative ability is poor, as evidenced by an average of 6.651 on a 0-19 scale for the high level students, barely higher than a third of the total possible score. Given that these results were based upon students chosen from the two highest social economic status areas in the country, the general condition of English communicative ability may be much worse for the rest of the country. The results from the advanced level sample are no more comforting either, with an average of 13.661 on a 0-20 scale. Despite the fact that their mean score is substantially higher than that of the middle schools' students, we must also take into account their age difference (they are on average 4 years older than the high level students) and they are considered as top of the cream in their age group when compared to the rest of the population their age. If their performance is seen as shown, the rest of the population would certainly perform much worse.

For both high level and advanced level samples, the total raw scores are skewed slightly in opposite directions (skewness indices of 1.066 and -.898 for high levels and advanced level, respectively). The opposite directions of skewness were expected because the test is considered difficult for the high level students, and is relatively easy for the advanced level students. The internal consistency based reliability indices as measured by Cronbach's alpha coefficient are .854 and .632, respectively. This pattern of opposite skewness is observed across all three work samples. In terms of item percent correct, the average of High level samples was .35, whereas it was .69 for the advanced level sample. With respect to item discrimination, the average of high level samples was .57, whereas it was .50 for the advanced level sample. It is noteworthy that while our items are considered difficult for the high level students, they nevertheless do possess reasonable discriminatory power when it comes to assessing English communicative ability.

Conclusion

Our research hypotheses are shown to be supported. Three conclusions can be reached based on the results of the current study:

1. Items measuring both semantic and pragmatic competence will become easier as students progress from high level to advance level. This is a positive interpretation of our empirical findings.

2. Judging from percent correct statistics while taking into account the age factor, Iranian students English ability in terms of both semantic and pragmatic competence is far from being satisfactory.

3. With respect to item discrimination, semantic items are more discriminating than pragmatic items. This phenomenon is undesirable because it implicates the current English assessment fails to take into account students pragmatic competence in the current notion of English assessment.

References

Ackerman, T. A., Gierl, M. J., & Walker, C. M. (2003). *Using multidimensional item response theory to evaluate educational and psychological tests*. Educational Measurement: Issues and Practice. Fall, 37-51.

Aitchison, J. (1999). Linguistics. Chicago: NTC/Contemporary Publishing. Bachman, L. F. (1991). *Fundamental considerations in language testing*. Hong Kong: Oxford University Press.

Brindley, G. (1995). *Language Assessment in Action Research Series 8.* Sydney: NCELTR Macquarie University.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics. 1*, 1-47.

Chou, T. & Park, J. (2005). An application of multidimensional IRT to the assessment of Taiwanese students' English pragmatic competence. *Psychological Testing (Taiwan). 52*(2), 95-118.

Fraser, C. (1988). *NOHARM: An IBM PC computer program for fitting both unidimensional and multidimensional normal ogive models of latent trait theory*. Armidale, Australia: The University of New England.

Fraser, C., & McDonald, R. P. (2003). *NOHARM: A Windows program for fitting both unidimensional and multidimensional normal ogive models of latent trait theory.* Armidale, Australia: The University of New England.

Kasper, G., & Rose, K. (2001). *Pragmatics in Language Teaching*. Cambridge: Cambridge University Press.

Lado, R. (1961). *Language testing*. London: Longman.

McDonald, R. P. (2000). A basis for multidimensional item response theory. *Applied Psychological Measurement*, *24*, 99-114.

McNamara, T. F. (1996). *Measuring second language performance*. London: Longman.

Reckcase, M. D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement, 9*, 401-412.

Reckcase, M. D., & McKinley, R. L. (1991). The discrimination power of items that measure more than one dimension. *Applied Psychological Measurement, 14*, 361-373.

Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2006). *Research methods in psychology.* Taipei: McGraw-Hill.

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics, 4*, 91-112.

Title

The Effect of Image Generation on Remembering Story Details in Iranian EFL learners

Author

Mahboobeh Abhaji Ezabadi (M.A.)

Iran Language Institute, Yazd, Iran.

Abstract

The purpose of the present study is to determine the effect of the construction of mental images on FL learners' ability to recall narrative passages. To carry out the investigation, 60 female students majoring in English Translation at Taft Islamic Azad University, Iran, were chosen and assigned to two groups based on the placement test results. Students in the experimental group were directed to construct images for the story, while control subjects were instructed to "do whatever you can or have to" in order to remember the story. To analyze the obtained data, the mean and standard deviation of scores were calculated. The unpaired t-test was also done to compare the means of different groups. Results indicated that visualizers remembered the story better, both shortly and long after reading the story.

Keywords: Imagery, Memory, Dual coding theory, Visualization, Mental imaging, L2 reading strategies, Cognitive reading strategies, Comprehension and recall of literary texts.

Introduction

The human brain is divided into two hemispheres; left and right. Each hemisphere has specialized capabilities and each processes different types of information. The left hemisphere operates in a rational, analytical manner. It processes verbal information. It is the seat of language and logical thinking. It organizes and categorizes

information. The right hemisphere operates in an intuitive, holistic manner. It processes visual and spatial information. It is the seat of creativity and imagination. It combines separate elements to form coherent wholes.

Note taking, lecturing, reading, and analytical thought are left-brain skills. Recognizing patterns, configurations, shapes, and forms; intuitive thought; and visualization are right-brain skills. Students spend most of their time developing left-brain (verbal) skills. Those students who are also thinking pictorially, focusing on the whole instead of on separate parts, and trusting intuition, are taking advantage of both of the ways in which the brain makes learning possible.

According to Allan Paivio of the University of Western Ontario, who has done a great deal of research on memory, if you commit facts and ideas to memory through words only, you are using only half of your brainpower.1 When a fact or idea that you memorized through words (and stored in the left hemisphere of your brain) is also memorized through a picture or sketch (and is stored in the right hemisphere of your brain), you set up a powerful combination in your memory. You can draw on this combination later when you need to recall the fact or idea.

The dual coding theory proposes that meaning can be represented by two separate coding systems: one system (verbal) specializing in language and the other (non-verbal or imagery system) dealing with non-linguistic events (Paivio, 1971; 1983; 1991; Sadoski and Paivio, 2001; 2004). These two systems can operate independently, for example activity in one but not the other: reading without mental images. They can also operate in parallel, for example separate activity in both at the same time: reading with unrelated images, and they can operate in a connected integrated way: reading with related images (Sadoski and Paivio, 2001; 2004). Sadoski and Paivio (2001; 2001) hypothesize that the verbal system is organized in a way that favors abstract, sequential and logical thought whereas the non-verbal system holds concrete sets of information (such as images), which are free form logical restraints and better at parallel processing of spatial information. Central to the dual-coding theory is that although the two systems perform independent functions, they can also perform in an integrated way. The key component to this interaction is the hypothesis that language can evoke imagery, and imagery can evoke language (Sadoski, Paivio and Goetz, 1991).

For holistic learning to take place, the left side of your brain must cooperate and harmonize with the right side. This cooperation is easy to gain if you make it a habit to convert words into actual pictures or diagrams in your notes, or to convert words into mental pictures or images on the blackboard of your mind. Verbal descriptions often lend themselves quite easily to visual representation. Unfortunately, in many textbooks, pictures and diagrams are not as numerous as they should be. When a key concept lacks a picture, you can act as your textbook's illustrator by drawing what you read.

The role of visual imagery as a comprehension strategy can be explained by the "conceptual peg hypothesis" whereby mental images serve a key role in organization and retrieval from meaning by acting as "mental pegs" to which associated information can be "hooked" (Sadoski et al., 1991; Sadoski and Paivio, 2001).

Deficits in either the verbal or non-verbal systems could cause difficulties with processing and representing meaning. Representing ideas in a text with visual images might aid the integration of story events and ideas and facilitate the construction of a meaning-based representation of a text (Linden and Wittrock, 1981). For that reason, the effects might be specific to tasks that are dependent on integration, for example answering questions about story content (Oakhill and Patel, 1991), structuring of story events (Center et al., 1999) and comprehension monitoring (Gambrell and Bales, 1986).

Training people to build visually based representations of the content of text enhances comprehension. It is a technique that is used spontaneously by good comprehenders (Sadoski, 1985) and can be taught to young children in a relatively short space of time (Pressely, 1976; Johnson-Glenberg, 2000; Sadoski and Willson, 2006).

The mental imagery that we experience while reading, either spontaneously or induced by instruction has powerful effects on comprehension, memory, and appreciation for text.

Paivio's Dual Coding theory is one of the major theories in teaching. This theory maintains that cognition consists of the operations of two separate but interconnected mental coding systems: a verbal system for language and a non-verbal system that deals primarily with imagery. According to this theory, there are two codes and memory systems (imaginal and verbal). Paivio believes that information maybe coded and stored in one or both (Paivio, 1990, p. 53). Images will help the recall of events and details.

Schema theory is another theory that has been popular in offering an explanation of reading comprehension and memory. This theory maintains that a schema is an abstract knowledge structure that guides the way text information is assimilated, how inferences are made, and how text is remembered.

Context-availability theory is the other theory in text comprehension and memory. According to this theory concrete language has more prior knowledge connections than abstract language, and when abstract language is sufficiently familiar or presented in context, its comprehension and recall should be equal to concrete language.

Using images is helpful in learning both, in monolingual and bilingual situations. According to Pressely and Miller (1987) "No more experiments are required to substantiate the positive effect of pictures on children's learning". In addition to their effect on learning, images have a significant role in remembering too (Anglin, 1987). Braden (2001, p. 3) claims that picture arouse interest and curiosity, make reading more enjoyable and create positive attitudes toward reading.

According to two studies by Sadoski (1983, 1985), imagery of a key event in the story (its climax) was related to total recall and to deeper levels of comprehension, such as recognition of the story's theme.

Long, winograd, and Bridge (1989) used think aloud methodology and found that mental imagery occurs as a spontaneous and consistent process in reading and that imagery is related to interest in reading.

Gambrell and Jawitz (1993) investigated the relative effectiveness of inducing mental imagery, attention to story illustrations, or both together. The group instructed to form mental images read an unillustrated version of the story. The group instructed to attend to illustrations read the standard, illustrated version. The group instructed to do both read the illustrated version. A control group was instructed to read and remember the unillustrated version. Children in the group instructed to form mental images of their own as well as attend to illustrations significantly outperformed all the other groups on several measures of comprehension and recall. The imagery-only group outperformed the illustrations-only group on recall of story structure elements and complete recall of the story. The control group had the lowest performance on all recall tasks despite the fact that it was the only group explicitly instructed to read to remember.

Significance of the Study

The role of imagery in cognition and learning has been a weighty topic for Greek philosophers, the British empiricists and modern researchers of cognition and neuro-cognition (Solso, 1995). Teachers and other educational practitioners are interested in this field too, and have wanted to know if pictures and imagery have any positive effect on learning. Research generally indicated that imagery facilitates students' learning on basic memory tasks.

This study tries to explore the relationship between image generation and remembering story details. The results will be useful for EFL teachers and institutes who are involved in the education of people. According to the investigation done, in EFL learning situations, if materials are accompanied with pictures, they will be recalled better immediately and long after reading.

Purpose of the study

The present study investigates the effect of image generation on remembering story details. It tries to find out whether readers have different degrees of recall when they are dealing with pictures and mental images. In this study, the researcher tries to find out whether generating pictures while reading stories has any superiority over merely reading stories in terms of recall of details. Thus, the following questions are raised:

1. Does practice of visualization bring about any variation in EFL learners' recall of narrative texts shortly after reading?
2. Does practice of visualization bring about any variation in EFL learners' recall of narrative texts long after reading?

On the basis of the above research questions, the two following null hypotheses are formulated. Hence, the statistical results of the study would either support or reject the null hypotheses.

1. Encouraging EFL learners to visualize in the process of reading narrative texts has no effect on their recall of such texts shortly after being read.
2. Encouraging EFL learners to visualize in the process of reading narrative texts has no effect on their recall of such texts long after being read.

Method

Participants

To carry out the investigation, 92 female students majoring in English Translation at Taft Islamic Azad University were given a placement test. Based on the results, 60 students whose scores fell within one standard deviation below and above the mean were chosen and assigned to two groups in such a way that there was no statistically significant difference between them.

Instrumentations

The material used in this study included a 4-paragraph short story, "Portrait of Peace" by Linda Spitzer. To prevent any comprehension problem, some words were explained in a box at the top of the story.

Research procedure

The control group (non-illustrators) was asked to read the story carefully and do whatever they can or have to, to remember the story.

The students in the experimental group (illustrators) were also instructed to read the story and illustrate the details of the story with pictures and drawings. It was emphasized that the artistic aspect of the pictures was not the point.

The time allocated for both groups to do the task was 20 minutes. After an interval of about 15 minutes, the participants in both groups were asked to write the story in their own words and they were said not to be worried about grammatical correctness and spelling. Furthermore, they were asked to answer 13 short-answer questions related to the story.

To explore long term retention of the ideas, students were asked to write whatever they remember from the story and answer the questions one month later. The only information given to remind them was the title of the story.

Review of Related Studies

The finding of better memory for pictures compared to words was reported as early as the 19th century (Kirkpatrick, 1894). Kirkpatrick demonstrated that real objects were better remembered than either written or spoken words both tested immediately, and at a 3-day delay. This picture superiority effect (PSE), as it has come to be called, is a robust phenomenon with numerous demonstrations of the basic finding that pictures are better recognized and recalled than their labels (e.g., Brady et al., 2008; Madigan, 1974, 1983; Nelson, Reed, & Walling, 1976; Nickerson 1965, 1968; Paivio, 1991; Paivio & Csapo, 1973; Paivio, Rogers, & Smythe, 1968; Shepard, 1967). In addition, the number of pictures that can be remembered is striking. Standing (1973) showed that people can remember thousands of unique pictures with great accuracy.

Pictures and visualization (Visual Mnemonics) is a way students can use to promote memory (Thompson, 2002). Learners can pair pictures with words they need to learn. Flashcards with pictures or symbols are a good way of memorizing words. As soon as individuals see a particular picture, they remember the word that goes with

it. Sometimes instead of using real pictures, learners can visualize the word they need to remember.

Pictures may engage greater elaborative processing because they may be associated with more symbolic codes in conceptual memory (verbal and imaginal; Paivio, 1986), or with a more distinctive imaginal code (Nelson, 1979), than words. Thus, the mnemonic advantage of pictures over words is believed to stem from encoding differences between the two symbolic formats.

A number of researchers have suggested that pictures receive more extensive semantic processing than do words (Intraub & Nicklos, 1985; Nelson et al., 1977; Smith & Magee, 1980; Weldon & Roediger, 1987; Weldon, Roediger, & Challis, 1989); therefore, pictures benefit from deeper or more elaborate levels of processing (see Craik & Lockhart, 1972). Consistent with this explanation, research has shown that pictures can be semantically categorized faster than words (Potter & Faulconer, 1975; Smith & Magee, 1980) and that the picture superiority effect can be eliminated or reversed when orienting tasks increase the semantic or elaborative processing of words (see, e.g., Durso & Johnson, 1980). All of the different explanations of the picture superiority effect share the fundamental assumption that the memorial representation of pictures is in some way more elaborate, distinctive, or meaningful than the representation of words.

Several studies revealed that adults who used imagery remembered more of the content of a prose passage than subjects who did not (Anderson & Hidde, 1971; Anderson & Kulhavy, 1972).

In a study (1975), Paivio found that people recalled more successively repeated pictures than successively repeated words. Paivio believes that this is because people recall pictures better than words. This result is known as the picture superiority effect (Nelson, Reed, & Walling, 1976; Paivio, Rogers, & Smythe, 1968) and may be because pictures access meaning more quickly and completely than words (Smith & Magee, 1980; Nelson, 1979).

Following Levin (1976), maximizing what children learn can generally be accomplished through the use of techniques which "concretize" what is to be learned. Levin claims that pictures are more concrete than words in that they provide learners with a closer approximation to their environments. For children's and adults' learning of unconnected materials, pictures have been shown to be superior to words, in tasks involving both recognition and recall memory.

According to Oxford and Crookall (1990), "learners have better recall and appropriate use of the words when they learn them coded dually [verbally and visually] than when the words are coded in a single manner" (page 16). They concluded, "Additional pictorial cues are effective and efficient in helping learners make associations between pictures and words" (Oxford and Crookall, 1990, p. 17). Similarly, Chun and Plass (1996, p. 17) maintained that words illustrated both verbally and pictorially were learned better than words explained verbally. They explained that "organizing information in working memory seems to be aided by learners making connections between the verbal and visual systems, and this helps in linking information to components of the mental model in long term memory" (p. 17).

When learning a large amount of new information, such as a foreign language, mnemonics are said to be a useful tool, for they function as "memory aids" (Higbee, 1977) that relieve the burden on learners' short-term memory by associating the new information with something familiar (Ericsson, Chase, and Faloon, 1980).

According to two studies by Sadoski (1983, 1985), imagery of a key event in the story (its climax) was related to total recall and to deeper levels of comprehension, such as recognition of the story's theme.

In a classroom setting, Pressley (1976) taught third-grade children a mental imagery strategy to help them remember stories. The children were given practice constructing images for progressively longer prose passages (sentences, paragraphs, stories) and were shown slides depicting good examples of images for the passages. Controls were told to do whatever they could to remember and did not see the slides. Both groups then read a 950-word story with alternating printed and blank pages. The imagery group was reminded regularly to form images on the blank pages and the control group was reminded regularly to do whatever they could to remember when they saw the blank pages. On a 24-item short-answer test, the imagery group outperformed the control group. There were no differences in reading times for the passages.

In a somewhat similar study, Gambrell (1982) gave first and third graders short stories to read in segments. Before each segment, children in the experimental group were told to make pictures in their heads to help remember, while the controls were told to think about what they read in order to remember it. After reading each segment, the participants were asked a prediction question ("What do you think is going to happen next?"). Responses were scored for factual accuracy and number of accurate predictions. Third graders in the imagery group reported twice as many facts and made twice as many accurate predictions as controls. Although first-grade imagers also outperformed controls on both measures, the differences were not statistically significant. Gambrell and other researchers have speculated that with beginning readers, the burden of verbal processing may inhibit simultaneous formation of images. Possibly, very beginning readers may do better reading and forming images successively, as in the structure of the Pressley study.

Sadoski (1985) had third and fourth graders read an unillustrated basal reader story aloud and then answer a series of comprehension questions, retell the story, and report any images recalled from the story either before or after the retelling. The story included a particularly dramatic climax. Children who were questioned prior to story recall and reported a climax image recalled more of the story than those who didn't report a climax image. There was no such effect for children who recalled their imagery after recalling the story. Sadoski suggested that the climax image functioned as a conceptual peg for subsequent story recall.

McDermott and Roediger (1994) have reported that imagery can promote priming on implicit memory tests. When subjects were given words during a study phase and asked to form mental images of corresponding pictures, more priming was obtained on a picture fragment identification test compared to a study phase in which subjects performed semantic analysis of the words. The authors concluded that imagery is

perceptual in nature; that is, imagery engages some of the same mechanisms used in perception and thereby produces priming (McDermott & Roediger, 1994).

Sadoski (1983, 1985) asked the subjects to report the mental images that occurred while reading a text. Subjects as young as ten years old reported a variety of spontaneous images, mostly consistent with the meaning of the text (Sadoski; 1983, 1985; Sadoski, Goetz, & Kangiser, 1988). Images are frequently unrelated to story illustrations (in case of illustrated stories), but have the same vividness as images related to these illustrations. Sadoski et al. (1988) found that subjects reported similar degrees of mental imagery and affect at certain paragraphs of the story, with considerable agreement on the kind of pictures and feelings the story evoked. This result suggests that –while allowing for individual and group differences (e.g., in education and profession)—it is possible to predict the occurrence of images in a text, to measure the imagery evokingness of texts, and even to design texts with imagery cues specific to an audience.

In his early studies, Paivio (1986) found that words with higher imagery value were remembered better than ones with lower imagery value. Sadoski showed that this is true for entire texts (Sadoski & Quast, 1990). According to Sadoski and Quast (1990), the best-recalled parts of a text two weeks later were not connected to the importance of the parts, as many researchers had hypothesized, but rather to imagery and affect ratings and paragraph length. They note that the "importance ratings may tend to reflect the reader's reconstruction of the author's idea hierarchy," while imagery and affect ratings "may tend to reflect the construction of personal meanings" (p.271).

More limited research has addressed the effects of visual complexity on recall memory. Sampson (1970) found pictures better than words in both immediate and delayed free recall situations. Ritchey (1982) reported an advantage in recall for outline drawings over detailed drawings. Conversely, Alfahad (1990) found realistic color visuals to be superior to black and white or line drawing visuals in a recall memory task. A study conducted by Jesky (1984) suggested the superiority of color over black and white and both color and black and white visuals respectively over line drawing images in a recall task.

In a review of pictorial research related to science education, Holliday (1973) concluded that pictures in conjunction with related verbal material can facilitate recall of a combination of verbal and pictorial information. It is suspected that pictures can increase comprehension in some cases; however more empirical evidence is needed (p. 210).

There is now substantial support for the claim that prose-relevant pictures do contribute to increased recall of prose materials, particularly if the subjects are young children (Holliday, 1975; Holliday & Harvey, 1976; Holliday & Thursby, 1977; Levie &Lentz, 1982; Levin, 1981; Levin, Anglin & Carney, in press; Levin & Lesgold, 1978; Willows, Borwick, & Hayvren, 1981). Levie and Lentz (1982) conclude that "illustrations facilitate learning and recall of the information in written text that is depicted in illustration" (p. 231). Levie and Lentz found that the average improvement for groups reading with pictures was 36%.

According to Park and Gabrieli pictures are inevitably remembered better than words on tasks of recall and recognition (Park & Gabrieli, 1995). This "picture superiority effect" is an "established memory phenomenon," in that experiments have repeatedly shown that "memory for pictorial stimuli is extremely accurate, durable, and extensive compared to that for verbal stimuli" (Noldy, Stelmack, & Campbell, 1990). In Shepard's (1967) experiment, subjects recognized old word stimuli 90% of the time, sentences 88% of the time, and pictures 98% of the time. Pictures become even easier to remember when the objects are not just side-by-side but are shown interacting, e.g., a car crashing into a tree (Wollen & Lowry 1971).

Rogers (1967) found improved recall for abstract nouns using imagery instructions but not for concrete nouns. Later Gupton and Frinche (1986) found imagery instructions improved recall for high imagery words but not for abstract.

A theory to explain why pictures are memorable says that the processing of pictures in the brain needs "additional allocation of attentional resources or effort" (Noldy, Stelmack, & Campbell 1990). Noldy, Stelmack, & Campbell's (1990) EEG recordings of brain ERP (Event-Related Potential) waves showed that it took longer to name a picture than to read the verbal label of the picture. Park & Gabrieli's (1995) participants also named pictures more slowly than they read words. Investigations of elementary learning processes, such as free-association reactions to words, drawings, and objects, have since the 1940s found a longer reaction time to pictures than to words (Otto, 1962).

Pictures are more complex than the words that label the pictures, so more time and attention is needed to identify, or "name," a picture. We spend more time looking at pictures (or real-life objects) before we can name them, so we remember pictures better. We spend less time looking at words in sentences, so we don't remember the sentences exactly— though we remember the gist. Pictures are also more distinctive and more unique than the words that label them, which further make pictures more memorable.

In their study of a German language-learning class, Plass, Chun & Leutner (1998) showed how vocabulary learning was affected by individual preference of learning style, classified as "visualizers" or "verbalizers." Using a multimedia computer story, college students could look up marked words by hearing the pronunciation plus seeing either a written translation of the word, or a picture (half the time a video clip) depicting the word.

Visualizers recalled illustrated words much better than unillustrated words. Verbalizers recalled illustrated propositions only slightly better than unillustrated ones. The best results occurred when both illustrations and translations were looked up, with only a small difference in results for visualizers and verbalizers.

Pictures may also "enhance the long-term retention of the words" since the "dualcoding effect" and the "greater effort" made "to process information and establish the relationship between the various sources of information" may make learned information "more resistant to memory loss" (Solman & Wu 1995).

Pictures are more perceptually rich than words, and this visual distinctiveness lends them an advantage in memory. To the extent that subjects also encode the stimulus as

a verbal label, subjects have two codes for pictures: in addition to the perceptual features of the stimulus such as color, shape, and texture, subjects also store a verbal label (similar to the representation for a studied word), that enriches the memory trace and provides redundancy. Picture illustrations are included in textbooks because they corroborate text and are often more effective than text alone for problem-solving transfer (Mayer, 1984; 1989; 1993; Mayer, Steinhoff, Bower, & Mars, 1995).

Stevick (1986) believes that "words that have come into our heads from reading or listening commonly leave us with pictures, sounds and feelings in our minds." Similarly, Denis (1982) suggests that the reading of descriptive or narrative texts by many readers "is accompanied by a sequence of visual images that express the semantic content of the text" (p. 540). In a similar comment, Brewer (1988) maintains that descriptive and narrative texts, in particular, tend to produce imagery in the mind during reading.

In one of the few studies done in the 1970s, Anderson and Kulhavy (1972) reported that high school seniors who received imagery instructions before studying a text recalled no more 'factual' content than did a non-instructed control group. However, on further probing, through giving the subjects a questionnaire after the study, the researchers discovered that not all of the students in the imaging group actually created images (only 50% did), and many in the control group did create images (about one-third)! Comparing those subjects from both groups who actually used imaging with those who did not showed significant differences in favor of the imaging strategy.

Scoring and Data Analysis
The original story was carefully read by the researcher to identify and list the significant propositions therein (See Appendix B). The ideas in the students' reproductions which matched the listed propositions were tallied, no matter where in students' writings they turned up. Each proposition correctly remembered, scored one point. To assure the reliability of the scoring, one other scorer scored the free recall and short-answer questions.

To analyze the obtained data, the mean and standard deviation were calculated. The unpaired t-test was also done to compare the means of different groups.

Results
The original story was carefully read to identify and list the main ideas of the story. The ideas in students' reproductions which matched the main ideas were considered, no matter where in students' writings turned up. The scores were obtained on the basis of one point for each proposition correctly remembered. The obtained scores (the sum of ideas) of the recounted details were entered in the statistical software for statistical descriptive and inferential information.

As we can see in Tables 1 & 2, the differences are most for paragraph two and especially for paragraph three. It should be noted that the most visualization task took place in reference to these two paragraphs.

Table 1 Means and standard deviations of the two groups' scores for immediate free-recall (N=30)

	Paragraph 1 (7 ideas) Mean SD		Paragraph 2 (11 ideas) Mean SD		Paragraph 3 (19 ideas) Mean SD		Paragraph 4 (4 ideas) Mean SD		Whole story (41 ideas) Mean SD
Experimental	3.33	1.73	5.63	2.31	9.1	3.5	2	0.98	20.06 6.90
Control	3.56	1.38	5.16	2.49	7.2	3.41	2.06	1.01	18 7.19

Figure 1 Means and standard deviations of the two groups' scores for immediate free-recall (N=30)

a) Bar Graph   b) Line Graph

Table 2 Means and standard deviations of the two groups' scores for delayed free-recall (N=30)

	Paragraph 1 (7 ideas) Mean	SD	Paragraph 2 (11 ideas) Mean	SD	Paragraph 3 (19 ideas) Mean	SD	Paragraph 4 (4 ideas) Mean	SD	Whole story (41 ideas) Mean SD
Experimental	3.26	1.14	3.7	2.18	6.23	3.11	1.63	1	14.73 5.81
Control	3.06	1.17	2.6	1.85	4.2	2.75	1.43	0.9	11.1 5.18

Figure 2 Means and standard deviations of the two groups' scores for delayed free-recall (N=30)

a) Bar Graph   b) Line Graph
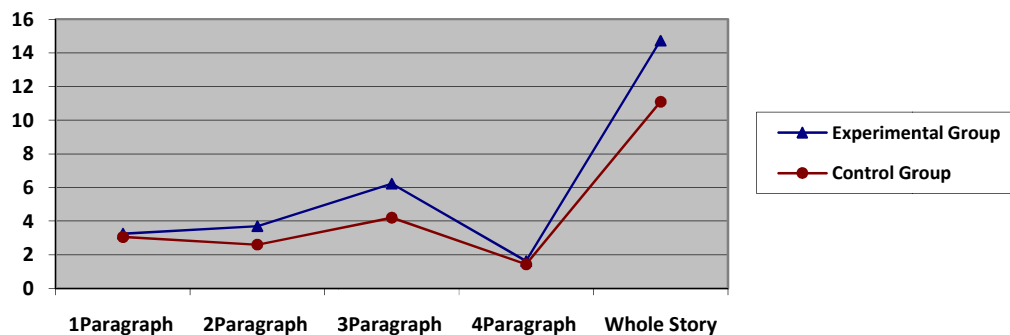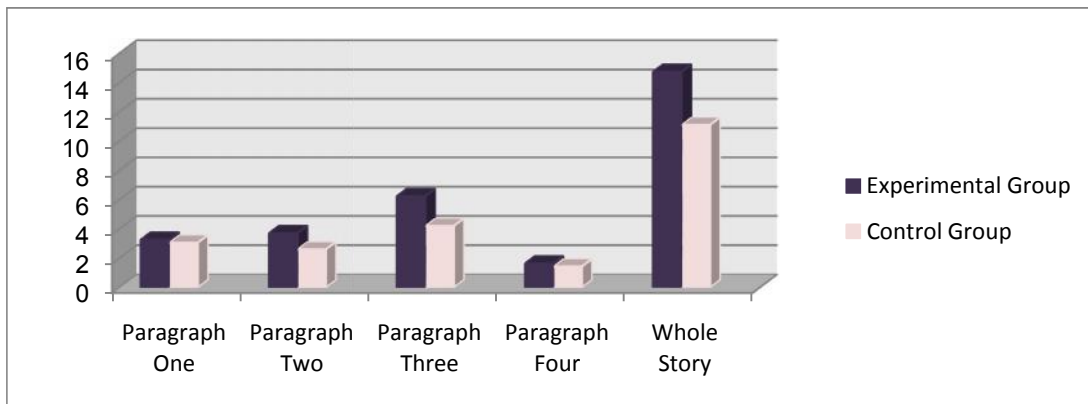
On both the immediate short-answer question test (Table 3) and the delayed short-answer question test (Table 4), the experimental group seems to have done better than the control group on comprehension and recall.

Table 3 Unpaired t-test on immediate short-answer questions

	Mean	SD	SEM	t-value	d-f	P-value
Experimental	8.067	2.087	0.381	2.5162	58	0.0147
Control	6.767	1.911	0.349			

Table 4 Unpaired t-test on Delayed short-answer questions

	Mean	SD	SEM	t-value	d-f	P-value
Experimental	8.25	2.521	0.46	3.1127	58	0.0029
Control	6.1	2.82	0.515			

For immediate free-recall, the difference between the experimental group and the control group favors the experimental group, but the difference is not significant (Table 5). For delayed free-recall, the difference between the experimental group and the control group favors the experimental group and the difference is significant (Table 6).

Table 5 Unpaired t-test on immediate free-recall test

	Mean	SD	SEM	t-value	d-f	P-value
Experimental	20.07	6.90	1.26	1.1355	58	0.2608
Control	18	7.19	1.31			

Table 6 Unpaired t-test on delayed free-recall test

	Mean	SD	SEM	t-value	d-f	P-value
Experimental	14.73	5.81	1.06	2.5076	58	0.015
Control	11.17	5.19	0.95			

To see whether the test type brings about any variation in the amount of recalled items, the percentage of the mean scores was obtained. As it's shown in Table 7 and Figure 3, we came to the conclusion that short-answer questions test type improves the recall of the information. One of the reasons of this may be the hint that is provided in the stem of the "short-answer questions" test type.

Table 7 Percentage of mean scores

	Experimental Group		Control Group	
	Short-Answer	Free-Recall	Short-Answer	Free-Recall
Immediate	62	48.93	52	43.9
Delayed	63.46	35.93	46.92	27.07

Figure 3 Percentage of mean scores

a) Bar Graph   b) Line Graph
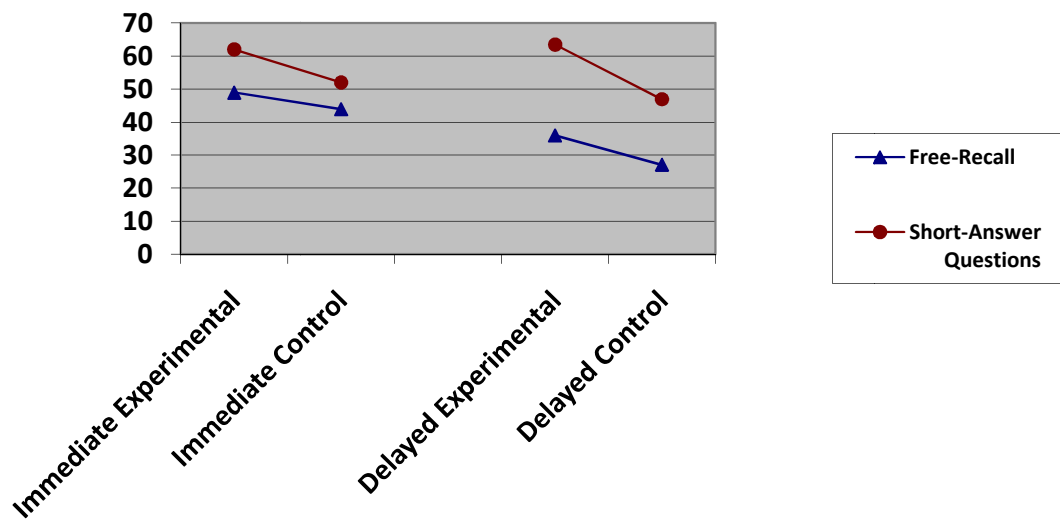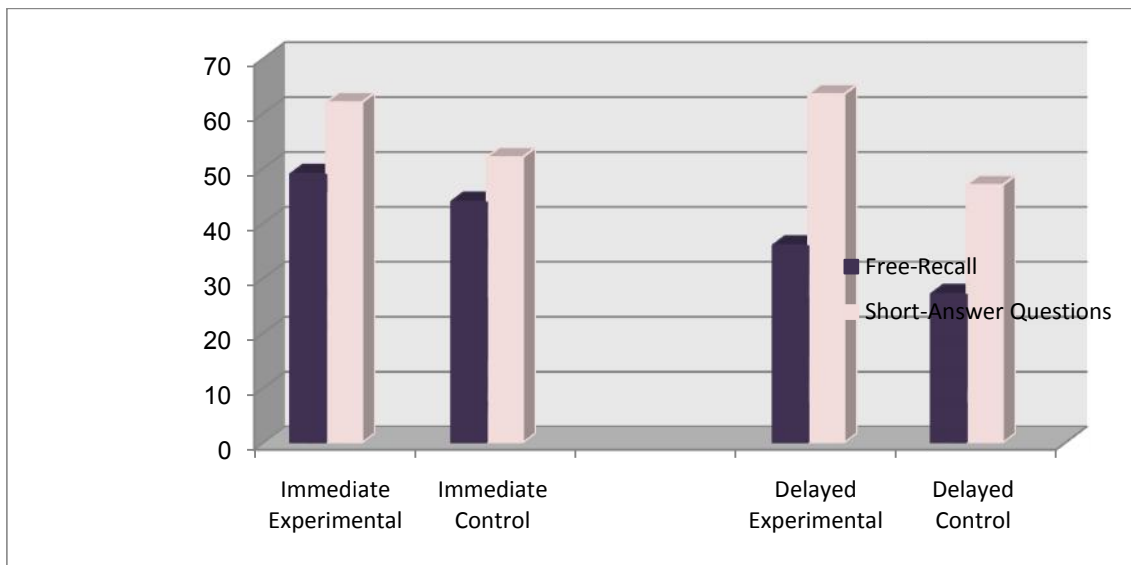
Conclusion

Through the findings of this study, both of the null hypotheses were rejected. The results obtained in this research supported the findings of other studies concerning the effect of visualization and image generation on remembering narrative texts.

The obtained mean scores by the experimental and control groups on the recall tests, as displayed in above tables, indicate that the experimental group by attempting to visualize in the process of reading the short story obtained a significantly greater mean than the control group.

Concerning the findings of this study, it is concluded that applying visualization in the process of reading narrative texts does improve EFL learners' recall of such texts. To put it another way, the visualizers significantly outperformed the non-visualizers on the recall tests.

Theoretical Implications

Why should ESL and EFL teachers use visuals in the classroom? There are specific reasons why teachers should use advance organizers in the classroom. First, the enhancement, expansion, and promotion of higher order thinking skills. These thinking skills are analyzing, synthesizing, evaluating, categorizing, sequencing, relating, conceptualizing, comparing and contrasting. Secondly, better comprehension and understanding of complex learning materials and concepts. When using a graphic organizer, students are able to make connections with prior knowledge and new knowledge they are acquiring to increase learning. Thirdly, long-term retention of information is facilitated by the use of organizers. They support the long-term memory process as the information enters into it through the working memory. While the information must be repeated or rehearsed to stay in working memory, storing the information after being elaborated, classified, organized, connected, is a process owned by the long-term memory. When using graphic organizers, old information is retrieved and linked to new information in order to facilitate the comprehension of new knowledge. Organizing the data graphically aids students in their attempts to establish relevant connections regarding the acquisition of knowledge. This process is a primary teaching and learning goal (Gil-Garcia & Villegas, 2003).

It has long been documented that humans have several distinct intelligences and that each intelligence is relatively independent of the others. Research has shown that any significant achievement involves a blending of the intelligences; however, research indicates that these intelligences are valued by cultures around the world, though not always to the same degree. Three of main selections of intelligence suggest a relationship between language learning and the arts. It can be hypothesized that linguistic, musical and spatial intelligence has the potential to work together to help enhance the learning of other languages through the use of picture prompts and visual cueing. This is because linguistic intelligence suggests that learners show a level of sensitivity to language and the relations among words. Musical intelligence reflects the learners' ability to relate to the arts. Human beings also demonstrate spatial intelligence in which they are able to observe; from mental images; as well as to make relationships to metaphors and gestalts (Canning-wilson, 1999).

Research has shown that imagery facilitates learning. Moreover, imagery combined with texts make subjects more likely to think about the process of the language more fully. Overall pictures help us as individuals make sense of output and input surrounding us in our daily lives.

Visual images allow us to predict, infer, and deduce information from a variety of sources. Moreover, pictures can bring the outside world into the classroom; thus, making situations more real and in turn helping the learner to use appropriate associated language. Furthermore, the uses of a visual can be used to create a social setting or to immerse a learner into a new or familiar world that cannot otherwise be created in the classroom environment.

Improved retention of details, which according to this study can be enhanced as a result of image generation while reading, has always been a big concern of the language teachers. So, it will be useful for both EFL learners and teachers to use pictures and visualization strategies in the process of learning. The results of this study may also have implications for other issues related to second language learning.

Pedagogical Implications
In using pictures in EFL classes we should consider several points:

First, check for ambiguity. Just as we choose realistic, conversational forms of language for our students to speak, we should choose realistic, unambiguous pictures for our students to see.

Second, don't clutter pictures with nonessential details. A picture with too many details looks like a mass of color or lines to a student, and he will have trouble deciding just what it is that the teacher wants him to see. If the idea behind the picture is to provide the context for language use, extraneous, nonrelevant material must be omitted and only the bare necessities included.

Third, be sure the picture is culturally recognizable. This point is one which is probably more pertinent for those working in specific areas overseas. There is no doubt that the students must, at some point, learn culturally different items, but too many cannot be introduced at one time or at the beginning. When too many are introduced too soon, none will be mastered.

Lastly, be aware of differences in the interpretation of color, shape, and the direction of eye-movement. We assume that the color of mourning is black; in Viet Nam it is white. In China the color of mourning is white, and red is worn for marriage. So if a picture is used by a practitioner in the classroom should:

- Be able to be interpreted
- Be to the point
- Show reasonable judgment
- Enhance learning and sensory acuteness
- Not indicate violent acts, appear overcrowded, stereotype in any form or offer too many distracters
- Be able to force full or partial student production
- Help to clarify the gist of a message

- Help offer familiarity with certain cultures
- Aid in the role of recall
- Be authentic
- Be sequenced
- Be used to test a knowledge of meanings of different words in isolation, as well as vocabulary in a sentence context
A visual may be ineffective in a learning environment when:
- The visual is too small
- Stereotype visuals are used to represent people
- Poor reproduction are created
- The picture is too far away from the text illustration
- The image offers too much information related or unrelated to the picture
- The purpose of the picture is unclear and doesn't compliment the text

Therefore before using a visual with a text or a lesson, a practitioner must ask himself or herself the following set of questions:

- Why are you using this particular visual? How will it enhance the lesson?
- What are different methods for using the pictures as part of a language lesson?
- How can this picture be best taught?
- How could this picture be used in a teaching situation in a future lesson as reinforcement?
- How could the picture be interpreted?
- What is the relationship between the item being taught and the visual prompt?

Therefore, it is important that EFL, ESL and F/SL practitioners realize that how one uses a picture is as important as the visual chosen. Furthermore, it is important that the visual supplements the text instead of detracting from the text (Canning-Wilson, 1999). It is the belief of this paper that visuals can aid in the learning and eventually the acquiring of a second language if used properly and that visuals can serve to enhance the learning process.

References

Alfahad, F. N., (1990). *The interactive effects of cognitive functioning and visual realism on visual memory recall tasks*. (Unpublished doctoral dissertation, University of Pittsburgh).

Allan, D. (1985). *Oxford Placement test*, Oxford: Oxford University Press.

Anderson, R.C. & Hiddo, J.L. (1971). *Imagery and sentence learning*. Journal of Educational Psychology, 62, 526-530.

Anderson, R.C., & Kulhavy, R.W. (1972). *Imagery and prose learning*. Journal of Educational Psychology, 63, 242-243.

Brady, T.F., Konkle, T., Alvarez, G.A. and Oliva, A. (2008). *Visual long-term memory has a massive storage capacity for object details*. Proceedings of the National Academy of Sciences, 105(38), 14325-14329.

Brewer, W. F. (1988). *Postscript: Imagery and text genre*, Text 8, pp. 431-438.

Clark, J.M., & Paivio, A. (1991). *Dual coding theory and education*. Educational Psychology Review, 3(3), 149-210.

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

Durso, F. T., & Johnson, M. K. (1980). *The effects of orienting tasks on recognition, recall, and modality confusion of pictures and words*. Journal of Verbal Learning & Verbal Behavior, 19, 416-429.

Ericsson, K. A., Chase, W. G., and Faloon, S. (1980). *Acquisition of memory skill*. Science 208: 1181–82.

Gambrell, L. B. (1982). *Induced mental imagery and the text prediction performance of first and third graders*. In J. A. Niles & L. A. Harris (Eds.), New inquiries in reading research and instruction (31st yearbook of the National Reading Conference, pp. 131-135). Rochester, NY: National Reading Conference.

Gupton, T., & Frinche, G. (1986). *Imagery, mediational instructions, and noun position in free recall of noun-verb pairs*. Journal of Experimental Psychology, 461-462.

Higbee, K.L. 1977. *Your memory, how it works and how to improve it*. Englewood Cliffs, NJ: Prentice-Hall.

Holliday, W. G. (1973). *Critical analysis of pictorial research related to science education*. Science Education, 57, 201-214.

Holliday, W. G. (1975). *The effects of verbal and adjunct pictorial-verbal in science instruction*. Journal of Research in Science Teaching, 12. 77-83.

Holliday, W. G., & Harvey, D. A. (1976). *Adjunct labeled drawings in teaching physics to junior high school students*. Journal of Research in Science Teaching, 13, 37-43.

Holliday, W. G., & Thursby, P. M. (1977). *Teaching visual concepts in the classroom*. Paper presented at the meeting of the association for educational communications and technology, Miami.

Intraub, H., & Nicklos, S. (1985). *Levels of processing and picture memory: The physical superiority effect*. Journal of Experimental Psychology: Learning, Memory, & Cognition, 11, 284-298.

Jesky, R. (1984). *The interactive effects of pictorial presentation and cognitive style on a visual recall memory task* (Doctoral dissertation, University of Pittsburgh, 1985). Dissertation Abstracts International, 44, 11077.

Kirkpatrick, E.A. (1894). *An experimental study of memory*. Psychological Review, 1, 602-609.

Levie, W. H., & Lentz, R. (1982). *Effects of text illustrations: a review of research*. Educational Communication and Technology Journal, 30, 195-232.

Levin, J. R. (1976). *What have we learned about maximizing what children learn?* In J. R. Levin & V. L. Allen (Eds.), Cognitive learning in children: Theories and strategies. New York: Academic Press.

Levin, J. R. (1981). *On Functions of Pictures in Prose*. In Pirozzolo Francis J., & Wittrock, Merlin C. Neuropsychological and cognitive processes in reading (pp 203-228). New York: Academic Press.

Levin, J. R., & Lesgold, A. M. (1978). *On pictures in prose*. Educational Communication and Technology Journal, 26, 233-243.

Levin, J. R., Anglin, G. J., & Carney, R. N. (in press). *On empirically validating functions of pictures in prose*. In D. M. Willows & H. A. Houghton (Eds.), illustrations, graphs and diagrams: Psychological theory and educational practice. New York: Springer-Verlag.

Madigan, S. (1974). *Representational storage in memory*. Bulletin of the Psychonomic Society, 4, 567-568.

Marcus, N., Cooper, M., & Sweller, J. (1996). *Understanding Instructions*. Journal of educational psychology, 88 (1), 49-63.

Mayer, R.E. (1984). *Aids to text comprehension*. Educational Psychology, 19, 30- 42.

Mayer, R.E. (1989). *Models for understanding*. Review of Educational Research, 59, 43-64.

Mayer, R.E. (1993). *Illustrations that instruct*. In R. Glaser (Ed.), Advances in Instructional Psychology, Volume 5. Hillsdale, JH: Erlbaum.

Mayer, R.E., Steinhoff, K., Bower, G., & Mars, R. (1995). *A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text*. Educational Technology Research and Development, 43(1), 31-41.

McDermott, K., & Roediger, H. (1994). *Effects of imagery on perceptual implicit memory tests*. Journal of Experimental Psychology: Learning, Memory and Cognition, 20, 1379-1390.

Nelson, D. L. (1979). *Remembering pictures and words: Appearance, significance, and name*. In L. S. Cermak & F. I. M. Craik (Eds.), Levels of processing in human memory (pp. 45-76). Hillsdale, NJ: Erlbaum.

Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). *Learning to order pictures and words: A model of sensory and semantic encoding*. Journal of Experimental Psychology: Human Learning & Memory, **3**, 485-497.

Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). *Pictorial superiority effect*. Journal of Experimental Psychology: Human Learning & Memory, **2**, 523-528.

Nickerson, R.S. (1965). *Short-term memory for complex meaningful visual configurations, a demonstration of capacity*. Canadian Journal of Psychology, 19, 155-60.

Nickerson, R.S. (1968). *A note on long-term recognition memory for pictorial material*. Psychonomic Science, 11, 58.

Noldy, Nancy E., Stelmack, Robert M., Campbell, Kenneth B. (1990). *Event-Related Potentials and Recognition Memory for Pictures and Words: The Effects of Intentional and Incidental Learning*. Psychophysiology, 27 (4), 417-428.

Otto, W. (1962). *The Differential Effects of Verbal and Pictorial Representaions of Stimuli upon Responses Evoked*. Journal of Verbal Learning and Verbal Behavior, 1, 192-196.

Paivio, A. (1975). *Coding distinctions and repetition effects in memory*. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 179-214) (Vol. 9). New York: Academic Press.

Paivio, A. (1979). *Imagery and verbal processes*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A. (1991). *Dual coding theory: Retrospect and current status*. Canadian Journal of Psychology, 45, 255-287.

Paivio, A. & Csapo, K. (1973). *Picture superiority in free recall: Imagery or dual coding?* Cognitive Psychology, 5(2) 176-206.

Paivio, A., Rogers, T.B., & Smythe, P.C. (1968). *Why are pictures easier to recall than words?* Psychonomic Science, 11, 137-138.

Park, Sunghi M. & Gabrieli, John D.E. (1995). *Perceptual and Nonperceptual Components of Implicit Memory for Pictures*. Journal of experimental psychology: learning, memory & cognition, 21 (6), 1583-1594.

Plass, Jan L., Chun, Dorothy M., & Leutner, Detlev. (1998). *Supporting Visual and Verbal Learning Preferences in a Second-Language Multimedia Learning Environment*. Journal of educational psychology, 90 (1), 25-35.

Potter, M. C., & Faulconer, B. A. (1975). *Time to understand pictures and words*. Nature, 253, 437-438.

Pressley, G.M. (1976). *Mental imagery helps eight-year-olds remember what they read*. Journal of Educational Psychology, 68, 355-359.

Ritchey, G.H. (1982). *Notes, comments and findings: Pictorial detail and recall in adults and children*. Journal of Experimental Psychology: Learning, Memory and Cognition, 2, 139-141.

Rogers, T. B. (1967). *Coding instructions and item concreteness in free recall*. Unpublished master's thesis, University of Western Ontario.

Sadoski, M. (1983). *An exploratory study of the relationships between reported imagery and the comprehension and recall of a story*. Reading Research Quarterly, 19(1), 110-123.

Sadoski, M. (1985). *The natural use of imagery in story comprehension and recall: Replication and extension*. Reading Research Quarterly, 20, 658-667.

Sadoski, M., Goetz, E.T., & Kangiser, S. (1988). *Imagination in story response: Relationships between imagery, affect, and structural importance*. Reading Research Quarterly, 23(3), 320-336.

Sadoski, M., & Quast, Z. (1990). *Reader response and long-term recall for journalistic text: the roles of imagery, affect, and importance*. Reading Research Quarterly, 25(4), 256-272.

Sampson, J.R., (1970). *Free recall of verbal and non-verbal stimuli*. Quarterly Journal of Experimental Psychology, 22, 215-221.

Shepard, R.N. (1967). *Recogniton Memory for Words, Sentences, and Pictures*. Journal of Verbal Learning and Verbal Behavior, 6, 156-163.

Smith, M. C., & Magee, L. E. (1980). *Tracing the time course of picture–word processing*. Journal of Experimental Psychology: General, 109, 373-392.

Solman, Robert T. & Wu, Huei-Min. (1995). *Pictures as Feedback in Single-Word Learning*. Educational psychology, 15 (3), 227-244.

Standing, L. (1973). *Learning 10,000 pictures*. Quarterly Journal of Experimental Psychology, 25, 207-222.

Stevick, E. (1986). *Images and options in the language classroom*. Cambridge: Cambridge University Press.

Thompson, Irene. (2002) *Vocabulary Learning Strategies*.

Weldon, M. S., & Roediger, H. L., III (1987). *Altering retrieval demands reverses the picture superiority effect*. Memory & Cognition, 15, 269-280.

Weldon, M. S., Roediger, H. L., III, & Challis, B. H. (1989). *The properties of retrieval cues constrain the picture superiority effect*. Memory & Cognition, 17, 95-105.

Willows. D. M., Borwick, D., & Hayvren, M. (1981). In T. G. Waller & G. E. Mackinnon (Eds.), *Reading research: advances in theory and practice* (vol. 2) New York: Academic Press.

Wollen, K.A., & Lowry, D.H. (1971). *Effects of Imagery on Paired-Associate Learning*. Journal of Verbal Learning and Verbal Behavior, 10, 276-284.

Yates, F.A. (1966). *The art of memory*. London: Routledge and Kegan Paul.

Title

The Effect of the Method on the Trait: Investigating the Function of No-Error Options in Grammar Error-Identification Items in Admissions Tests

Authors

Roya Khoii (Ph.D.)

Islamic Azad University, North Tehran Branch, Iran.

Nazli Shamsi (M.A.)

Islamic Azad University, North Tehran Branch, Iran.

Bio Data

Roya Khoii is assistant professor of TEFL at Islamic Azad University, North Tehran Branch**.** She received her Ph.D. in TEFL in 1997 from Islamic Azad University, Science and Research Complex. She has written several articles and books and translated more than 25 books from Persian into English. Her main research interests are teaching and testing foreign language skills and CALL.

Nazli Shamsi holds an M.A. in TEFL from Islamic Azad University, North Tehran Branch, Iran. She has been teaching general English in Kish Language School in Tehran for seven years. Her research interests are testing and teaching foreign language skills.

Abstract

Every year many institutes, universities and organizations in many countries, including Iran, use the scores of the TOEFL, SAT, or their subtests for admission purposes. Given the fact that these tests play a too great role in determining the future life of the testees, they are expected to enjoy a high level of construct validity. This study was dedicated to investigating the effect of the test method on trait by comparing the construct validities of two different formats of error-identification grammar tests as used in the TOEFL (with four options for each item) and SAT (with a no-error option as the fifth option) while using a multiple choice grammar test as the criterion. After administering all the three tests to 131 Iranian EFL learners, the results were statistically analyzed, and it was found that while the TOEFL error-identification test had the highest level of construct validity, the no-error option in the SAT error-identification test reduced its construct validity to a considerable degree. Thus it was concluded that including no error options in admissions test does not allow an accurate evaluation of examinees' grammatical knowledge.

Introduction

One of the fundamental uses of testing in an educational program is to provide information for making decisions with regard to individual admissions. Generally speaking, educational admissions are based on measurement contributing data to institutional decisions about whether and on what basis to admit students for study in an institution, college, university or program. In this regard, admissions-test scores serve as a kind of "common metric" for expressing students' ability or preparedness on a common scale. When the number of applicants exceeds the openings available, the admissions-tests results are used to identify, select and admit the ones with the potential for success (Whitney, cited in Linn, 1993).

Every year many students stake their futures on university admissions tests; however, only a few of them are admitted. In such situations the measurement device which is a test should be constructed with utmost care to represent data of high credibility. Every test which is administered deserves some kind of evaluation because not all tests are well-developed, nor are all testing procedures wise and beneficial. Test developers are responsible for providing defensible and clear interpretations of test scores and encouraging their appropriate use. In order for a decision to be fair, our tests must be accurate in the sense that they must indicate what they are designed to indicate; that is, they must enjoy a high level of construct validity.

Scholars in the field of language testing have continuously tried one means or another to find a reliable, valid and practical measure of different aspects of second or foreign language. Many scholars such as Gergely (2007), McNamara and Roever (2006), Bachman and Palmer (1996), propose that although there are some arguments against using multiple choice (MC) techniques, the demands of an educational context with large numbers of test-takers, and the need for fast marking might make the use of these tests inevitable. The TOEFL (The Test of English as a Foreign Language) and the SAT (Scholastic Aptitude Test) are two standardized multiple-choice tests which are used as admissions tests. Many universities, institutions and organizations in different countries including Iran use the scores on these tests or on their modified versions for admission into MA programs. Different parts of these tests are also used to determine students' knowledge of different language skills and areas such as listening comprehension, grammatical competence, and reading comprehension.

TOEFL

Test of English as a Foreign Language (TOEFL) is designed to measure English proficiency to determine whether or not a student will be able to succeed at the college level. The first TOEFL was conceived and administered to nonnative speakers of English who were applying to North American universities in the late 1950s and 1960s (Chapelle et al. 2009). Nowadays, many educational institutions in

English speaking nations require TOEFL scores from foreign students who have not satisfied other requirements, such as studying at the college level in an English speaking country for at least two years (Smith, 2009).

The main purpose of the TOEFL is to assure that all students have reached a baseline knowledge level of the English language. According to Chapelle et al. (2009), members of the TOEFL revision project follow Messick's belief that language proficiency theory is the basis for score interpretation for large-scale tests with high-stake outcomes; however, according to most language assessment specialists, there is no best way of defining language proficiency that the TOEFL project could adopt. According to Carroll (1980), there are two approaches in defining language proficiency. The integrative approach, which suggests there is such a factor as overall proficiency, and the discrete-point approach, which attempts to break up knowing a language into a number of separate skills and further into a number of distinct items making up each skill.

Expanding Carroll's idea of integrative approach in 1980s made the TOEFL developers attempt to determine how communicative competence could be measured by the TOEFL and yielded to the introduction of TWE (Test of Written English), which was a sign of a move toward testing of real language abilities (Chapelle 2009). Although TWE was offered as a separate test from TOEFL, it was soon included as a part of TOEFL test administration.

In 1961 the issues of construct and content validation of the TOEFL were seen not only as technical terms, but also as issues to be communicated to test users. However, the increased use of the TOEFL in intensive English programs and as part of university admissions tests outside the United States, suggests more accurate considerations of its validity. Chapelle (2009) demonstrates the validity of TOEFL score interpretation as an indicator of academic English language proficiency which is used for admissions decisions.

SAT

According to Whitney (cited in Linn, 1993), one of the tests that is most commonly used in making general admissions decisions at the undergraduate level is the Scholastic Aptitude Test (SAT). Although the SAT people sought to clear up the confusion in a press release that declared "SAT is not an initialism; it does not stand for anything," many people still believe that the letters S-A-T stand for Scholastic Aptitude Test (Pacenza, 2010). The design of SAT was based on the IQ test which was created by the French psychologist Alfred Binet in 1905. Since the aim of SAT was to identify talented students from underprivileged backgrounds, it was thought of as a test that could measure an innate ability (aptitude), rather than abilities that these students might have developed through school (Pacenza, 2010). Wayne Camara (cited in Collage Board.com) suggests that the SAT does not measure any innate ability; rather, it measures "developed reasoning," which is described as the skills that students develop not only in school but also outside of school. Quoting from a recent publication of the College Board, Geiser and Studley (2001) indicate that the SAT has proven to be an important predictor of success in college. Its validity as a predictor of

success in college has been demonstrated through hundreds of validity studies. These validity studies consistently find that "high school grades and SAT scores together are substantial and significant predictors of achievement in college" (Camara and Echternacht, 2000, p. 9). However, as ETS (2007) revealed, the psychologist Claude Steele has found that the SAT measures only about 18 percent of the things that it takes to do well in school, and thus it is not a very good predictor of how a student will do in college.

Testing grammar

In recent years, grammar has regained its importance in language teaching. Nowadays, it is believed that grammar is too important to be ignored, and without the knowledge of grammar, it is not possible to create many new sentences (Swan, 2005). Thornbury (1999) defines this terminology as a partial study of what forms or structures are possible in a language. Richards, Platt and Weber (1985) define grammar as a description of the structure of a language and the way in which linguistic units such as words and phrases are combined to produce sentences in the language, and, according to Gee (1999), grammar is a set of devices that speakers and writers use to design, shape or craft their sentences and texts for effective communication. Therefore, the lack of grammatical ability sets limits to language proficiency. Despite all the assortments of defining grammar, it is clear that grammatical knowledge is a necessary condition for making appropriate written expressions and structures.

There are different ways for testing grammar. One of the most widely used types of items in objective tests is the multiple-choice (MC) test. According to Gergely (2007, p. 66), "The declining fortunes of grammar appear to have affected the technique most closely associated with it: multiple-choice (MC)." He also argues that there has been little research in recent literature on MC format test, especially in connection with grammar testing. However, the usefulness of MC items is limited. Gergely (2007) proposes that MC items leave a part of the test untested and are, therefore, unsatisfactory for many testing purposes. More specifically, in relation to grammar testing, the MC format is seen unsuitable because the concept of grammar has broadened over the years. Another shortcoming of MC tests is the difficulty of writing good MC items. Wolf, et al. (1991) also criticized these items since they are chosen to distinguish between students rather than representing the construct being assessed. Despite all the arguments against multiple –choice grammar tests, these items are used in commercial tests such as the TOEFL and SAT. Beside the fact that the scores of these tests are objective, and they can be scored easily, these tests can help teachers and students to locate the areas of difficulty since they test the knowledge of each component separately (Heaton, 1990).

Gergely (2007) clarifies some different types of MC items as:

- *Standard type*: A four-choice sentence-based item which presents an incomplete sentence stem followed by four multiple-choice options for completing the sentence.
- *Multitrak items*: Items which have the unacceptable choice as the correct answer, which is the issue of "testing negatively". The error-identification grammar tests which are used in the TOEFL and SAT admissions tests are included in this group. Unlike the standard type, such items do not require the students to complete the sentence; instead, they have to find the part containing an error.

Madson (1984) believes that error-identification items are useful for testing the grammar points which have few logical options. There is also another format of error-identification items in which students are given correct sentences together with the incorrect ones and are required to choose the fifth option, which will be the No-Error option if the sentence does not contain any error. Heaton (1990) believes that in practice this method does not work too well, since many students tend to regard every sentence as having an error. Another argument against this method is its emphasis on more negative aspects of language learning. It is believed that simply recognizing an error is not sufficient, and the students ought to be encouraged to concentrate on recognizing and producing the correct form. This argument is supported by many scholars who suggest that exposing the students to incorrect forms of language is undesirable. Nevertheless, since such items are related to the skills required for checking, editing, and proof-reading any report, article, paper or essay, they are used in many tests such as the TOEFL and SAT.

The effects of testing methods

The two major factors interacting in the process of testing are trait and method. Trait is the knowledge that is being measured, and method refers to the specific procedures or techniques for assessing the trait. Different testing methods produce different degrees of difficulties for test takers. Therefore, each has a significant effect on students' scores in that construct.

Bachman (1990) notes, "Characteristics of the test method can be seen as analogues to the features that characterize the context of situation" (p.111). Since a given trait can be assessed through different methods, these methods can have different effects on that trait and the test-takers' scores. Stansfield (1986) also believes that different testing methods which aim at measuring the same trait can be significantly different from one another and, thus, lead to different scores for the test taker. Testers' familiarity with one method can be one reason of this difference.

The item format may limit or prevent certain construct elements from being included in the test, or otherwise interfere with it, causing distortions in the scores with the possible result that they no longer reflect the construct very well (Gergely, 2007). In other words, the format may make candidates think in certain, possibly undesirable ways.

According to Shohamy (1984), a characteristic of a good test is one in which the method has little effect on the trait. Therefore, Bachman (1990) emphasizes the need

of developing a framework for delineating the specific features or facets of test method in order to understand variation in language test performance. Cohen (1984) gives credit to an attempt for a closer fit between how test makers intend for their tests to be taken, and how test-takers actually take them. It is also mentioned that changing the format of the tests or training the testees to deal with different formats can be involved in this attempt.

Although the usefulness of Multiple-Choice items has been proved to be limited, they are still common in many educational contexts because they meet the criteria of practicality and can provide a wide range of scores to produce a fewer number of candidates with the same score. In such contexts, it is desirable to investigate the effects of these tests on students' performance and to find out which format is a better indicator of the students' abilities on the construct the test is going to test. Therefore, MC tests require close investigation of their validity.

There are two formats of error-identification tests used in Iranian university admissions tests. The error identification section of the State Universities admissions test is similar to the SAT error-identification tests, in which the students are exposed to both correct and incorrect sentences, and each item involves a no-error option as the fifth option. However, this part in the Islamic Azad University Admissions Test is similar in format to TOEFL grammar tests, and lacks the no-error option.

The following represent typical items appearing on the sample TOEFL and SAT error-identification tests as used in Iranian university admissions tests:

Sample error-identification item on the TOEFL:
-   Halifax is the <u>largest</u> city and <u>chief port</u> of Nova Scotia and <u>is</u> the eastern terminus of
    Canada's two great <u>railway system</u>.   B   C
                                        A
                                    D
Sample error-identification item on the SAT:
-   The crowd, <u>which </u>clamored for the play to begin, <u>were</u> surprisingly <u>rowdy</u> for a
    Broadway <u>audience</u>. <u>No error</u>   B   C
                                        A
                        D   E
Reviewing the multiple-choice items used in internationally established ESL/EFL tests (e.g. TOEFL, TOEIC, SAT, Cambridge ESOL exams), one almost always finds four or five options per item. However, Shizuka, et al. (2006) proves that different number of options performed nearly the same in university admissions tests.
Validity
Messick (cited in Linn, 1993) defines validity as "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness and actions based on test scores or other modes of assessment."(p.13). His view of validity refers to the degree to which we are justified in making an inference to a construct from a test score, rather than a property of a test. That is, the behavioral inferences that one can extrapolate from test scores is of immediate focus (Swaim, 2009). For validating an inference, not only the validation of score meaning is required but also the validation of value implications and action outcomes for particular applied purposes and of the social consequences of their use

are important (Messick cited in Linn, 1993). In order to be valid, the inferences made from scores need to be "appropriate, meaningful, and useful" (Gregory, 1992). According to Mahmoodian (2000), construct validity is the milestone in current developments of test validation. It is the ongoing process of demonstrating that a particular interpretation of test scores is justified. Messick (cited in Linn, 1993) states that construct validity embraces all forms of validity evidence.

In 1995 Messick argued that since content and criterion–related evidence contribute to score meaning, they can be recognized as aspects of construct validity. A construct in scientific research is a type of concept used to describe events that share similar characteristics (Borg & Gall, 1989). In other words, construct validity refers to the extent to which the psychological reality of a trait or construct can be established. In establishing construct validity, no cut-off measure exists as to what are acceptable correlations between the construct and related entities in statistical order. Once again, this becomes a qualitative judgment that the researcher must make (Crocker & Algina, 1986; DeVelice, 1991; Gregory, 1992). One important issue with regard to test validity is whether the context can affect test validity. Sometimes a test which is valid for a purpose in one setting might be invalid or have a different validity in a different setting, or it might be invalid for another purpose in the same setting (Messick, cited in Linn 1993). Cronbach (1982) notes that although certain kinds of tests may work well in some situations, they cannot be generalized in other situations unless their validity in the new settings is checked. Since validity is the only essential justification for test interpretation and use, professional judgments are required about the tests' validity in each measurement enterprise (Messick, cited in Linn 1993). Therefore, although the TOEFL and SAT subtests have proved to be valid means for testing different aspects of language competence, their validity should be investigated before using them for admissions purposes with different nationality EFL learners.

Research questions
This study aimed at answering the following questions:
1. Does the use of the "No- error" option in the error identification sections of the TOEFL and SAT have a significant effect on test validity?
2. Do the students' scores on the TOEFL error identification part correlate with their scores on the SAT error identification part?
3. Is there a significant difference between students' scores on the TOEFL error identification part and their scores on the SAT error identification part?

Method
Participants
131 students at two different universities in Iran, Islamic Azad University, North Tehran branch, and Allame Tabataba'i University, participated in this study. All the participants were undergraduate male and female Iranian students majoring in English translation, literature, and teaching who had passed at least 100 credits. The rational behind selecting senior students was to have more proficient individuals. All the participants were between 22 and 30 years old.

Instrumentation

In order to examine the performance of the test-takers on the different formats of grammar tests and also estimate the tests' construct validity, three tests were utilized in this study. The first was a 30 item multiple-choice grammar test (modeling the TOEFL grammar section) in which the students had to choose the best option in order to complete each sentence. This test, which was a modified form of the TOEFL test and had been standardized by the researchers, functioned as the criterion measure. All the items had been selected from the TOFEL preparation books since the researchers did not have access to the original test. Moreover, what was intended here was a study of the effect of the test method and form on the measurement of the trait.

The second was an error-identification test including 30 items. Here, the students had to identify the structural error, and each item consisted of one and only one error. This test was a modified version of the TOEFL error-identification test, and its items had been selected from the TOFEL preparation books published by well-known publishers.

The third was an error-identification test including 30 items. Here, the students had to identify the structural errors. In this test each item included a No-error option, since the items might or might not have contained an error. This test was a modified version of the error- identification part of the SAT, and its items had all been selected from Barron's Verbal Workbook for the New SAT.

The last two instruments were two different forms of multitrak grammar tests. All the instruments in this study tested the same structural points since they all shared the same table of specifications.

The criterion-measure, a multiple-choice test consisting of 45 items, constructed by the researchers and revised by a testing expert, was administered to 27 students who had the same characteristics of the target group. The time for answering the items was 30 minutes. After scoring the papers, the item facility (IF) and item discrimination indexes (ID) of the items were calculated and their choice distributions (CD) were examined. Items with IFs between 0.3 and 0.7 and IDs beyond 0.4 were considered to be acceptable. After discarding the poor items, 30 items remained on the test. The test was rescored on the basis of the remaining items and its reliability was calculated using the KR-20 formula. Since the reliability of the test was 0.84, it was decided that the test was suitable to be used as the criterion measure.

After standardizing the criterion test, it was administered to 131 students along with the other two tests (two forms of error-identification tests). These tests included 90 items altogether and 70 minutes was allocated to administering them. It is worth mentioning that the participants had not been informed about the test beforehand, so there was no preparation of any kind for the exam.

Scoring

Each part of the final test was corrected and scored separately, which provided 3 sets of scores for each individual. One point was given to each correct answer and no point to incorrect ones. There was no penalty for incorrect answers.

Results

As mentioned previously, the present study was conducted in order to investigate the function of no-error options in grammar error-identification items in admissions tests, particularly with reference to test validity. This part entails the results of the statistical analyses of the collected data.

Descriptive statistics

After administering the final test, which consisted of 30 MC items, 30 error-identification items each with a No-error option, and 30 error-identification items lacking no-error options, to 131 participants, the descriptive statistics were calculated for each part of the final test as follows:

Table 1 Descriptive statistics of the tests

	N	Range	Min	Max	Mean	Std. Deviation	Variance
multiple-choice	131	25.00	3.00	28.00	15.3817	4.90441	24.053
TOEFL	131	21.00	3.00	24.00	13.1832	5.08512	25.858
SAT	131	23.00	0	23.00	11.2366	5.24453	27.505

Reliability

Cronbach's alpha was used in order to estimate the reliability quotients of the three tests. Cronbach's alpha is a coefficient used to rate the internal consistency (homogeneity) or the correlation of the items in a test together. If a test enjoys strong internal consistency, most measurement experts agree that it should show only moderate correlation among items. For exploratory purposes 0.60 is accepted; for confirmatory purposes 0.70 is accepted; and 0.80 is considered good (Garson, 2010). The Cronbach's alpha for all the tests was calculated using the SPSS program as follows:

Table 2 : Cronbach' Alpha reliability of the three tests

Test	Cronbach's Alpha	N of Items
MC	.70	30
TOEFL	.701	30
SAT	.768	30

As the results show, all the tests enjoyed acceptable reliability coefficients. However, coefficient alpha, the customary index of reliability in Marketing, is said to underestimate the reliability of a multidimensional measure (UCLA Academic Technology Service, 2010).

Construct Validity

One of the most extensively used approaches in the construct validation of language tests is factor analysis (Bachman, 1990). Factor analysis attempts to identify underlying variables or factors that explain the pattern of correlations within a set of observed variables (Farhady, 1983). Therefore, in order to further investigate the construct validity of the three measures which were used in this study (MC grammar test, the TOEFL and SAT error-identification structure tests), the scores of the participants on these measures were subjected to a factor analysis. This analysis was done to determine whether all these measures shared common variance and thus could be said to tap the same underlying construct. It is worth mentioning that the purpose of all the three tests was to measure the grammatical knowledge of the students. To ensure higher precision, a principal axis factoring (PAF), as opposed to a principal components factoring (PCF), was employed to extract the initial factors. There are many ways to determine how many factors to extract (Thumpson & Daniel, 1996). However, according to many scholars such as Sharma (1996), Zwick and Velicer (1986), the eigenvalue-greater-than-one was selected as the extraction rule. The eigenvalue-greater-than-one rule suggests that those factors whose eigenvalues (sum of squared loadings) are less than unity be excluded from the analysis. Table 3 shows that only one factor with eigenvalue more than one (2.13) was extracted, and all the tests loaded on the same underlying factor, that is, factor 1. Factor 1 also explained 60.18% of the total variance; in other words, more than half of the variance produced by the measures entered into the analysis was due to Factor 1, which can be best interpreted as accounting for students' grammatical knowledge.

Table 3. Total variance explained by factor analysis

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.132	71.081	71.081	1.806	60.187	60.187
2	.594	19.790	90.871			
3	.274	.9.129	100.000			

Note: The extraction method is Principal Axis Factoring.

Table 4 indicates that almost all the measures enjoyed high loadings on Factor 1 (i.e., they had high correlations with it). The highest belonged to the TOEFL error-identification test (.97) and the lowest to the SAT error-identification test (.59).

Table 4. Results of factor analysis

Tests	Factor 1
MC test	.703
TOEFL	.978
SAT	.592

Although the results can be regarded as evidence that all the tests measured the same construct to a large extent, it seems that the No-error option of the SAT error-identification reduced the construct validity to a considerable amount; therefore, it was concluded that the No-error option had a considerable effect on the construct validity of the test.

Criterion-related validity
In order to provide more information as to the construct validity of the tests including multitrak items, their correlations with the criterion test (i.e., multiple-choice) were calculated using the SPSS program:

Table 5. Correlation coefficients among the scores of the three tests

		TOEFL	SAT	MC test
TOEFL	Pearson Correlation	1	.581[**]	.690[**]
	Sig. (2-tailed)		.000	.000
	N	131	131	131
SAT	Pearson Correlation	.581[**]	1	.417[**]
	Sig. (2-tailed)	.000		.000
	N	131	131	131
MC test	Pearson Correlation	.690[**]	.417[**]	1
	Sig. (2-tailed)	.000	.000	
	N	131	131	131

Note: Correlation is significant at the 0.01 level (2-tailed).

The table reveals that the correlation between the TOEFL and MC tests of grammar was 0.69, while the correlation between the SAT and the MC tests of grammar was 0.417, which, although significant, was not promising. The results indicated that the TOEFL enjoyed a higher level of criterion-related validity than the SAT. Therefore, it was concluded that the no-error option could have affected the criterion-related validity of the SAT error-identification test.

Correlation
The second question of this study targeted the pattern of correlation between the two types of error identification tests. Table 5 indicates that the correlation coefficients between the TOEFL and SAT error identification tests equaled +.58, which was significant at the 0.01 level (2-tailed test, at 129 degrees of freedom). Therefore, there was a statistically significant correlation between the students' scores on the TOEFL and SAT error-identification structure tests, but it was not very high. The coefficient of determination which was 0.34 showed that there was only 34 percent of shared variance between the two sets of scores. Thus, it was concluded that the students' scores on the TOEFL error-identification test had a moderate correlation with their scores on the SAT error-identification test.

Mean comparison
Finally, in order to provide an answer to the third research question, a one-way analysis of variance (ANOVA) and post-hoc Tukey's test were carried out to compare the students' mean scores on the MC grammar test and the TOEFL and SAT error-identification tests (Table 6).

Table 6. ANOVA results for mean differences among the three grammar tests

Source of variance	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	977.420	2	488.710	19.526	.000
Within Groups	9761.42	390	25.029		

According to Table 6, the obtained F ratio equaled 19.52, which was significant at p<.000 level (the degree of freedom was 2.390 for all the tests) suggesting that the differences among the means were significant. However, the significance of the F ratio in the analysis of variance merely indicates that there is a significant difference among the means of the compared groups as a whole; it indicates that there is at least one significant difference between the means of at least one pair of the groups compared (Brown, 1988). In order to find out which two means are significantly different from each other, post hoc or follow-up tests are required. The highlighted numbers in table 7 delineate the differences between the students' means on different tests.

Table 7. Multiple comparisons

(I) test	(J) test	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
MC	TOEFL	1.97710*	.61816	.004	.5227	3.4315
	SAT	3.86260*	.61816	.000	2.4082	5.3169
TOEFL	MC	-1.97710*	.61816	.004	-3.4315	-.5227
	SAT	1.88550*	.61816	.007	.4311	3.3398
SAT	MC	-3.86260*	.61816	.000	-5.3169	-2.4082
	TOEFL	-1.88550*	.61816	.007	-3.3398	-.4311

Note: The mean difference is significant at the 0.05 level.

Table 8 summarizes the results of the Tukey's procedure to show which differences were significant.

Table 8. Summary of Tukey's results

Test	N	Subset for alpha = 0.01		
		1	2	3
SAT	131	11.2366		
TOEFL	131		13.1221	
MC	131			15.0992
Sig.		1.000	1.000	1.000

The results reveal that the differences between the means for all the tests were significant when alpha equaled 0.01. Therefore, it was concluded that there was a significant difference between the students' scores on the different formats of grammar tests. In other words, the MC grammar test proved to be the easiest while the SAT error identification test was found to be the most difficult one (Figure 1).
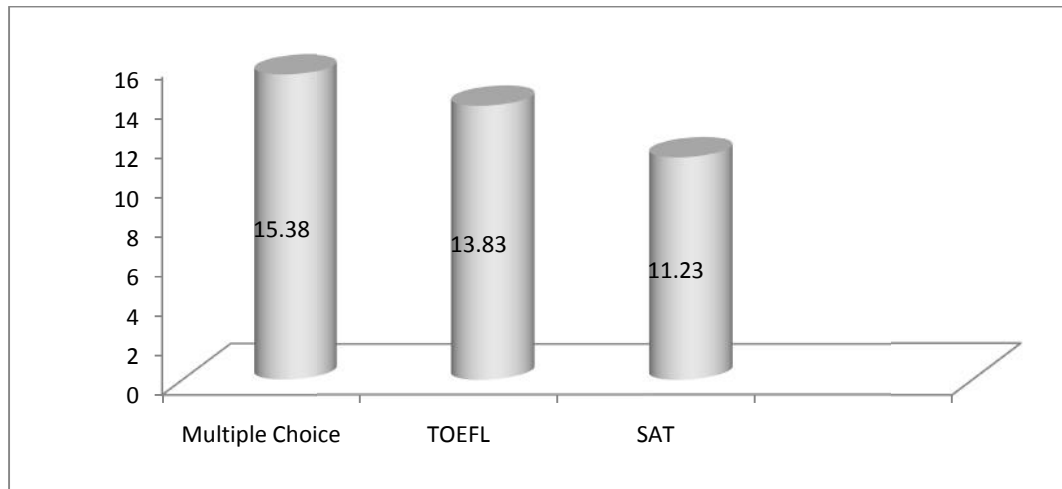
Figure 1. Comparison of the means on three tests (N=131)

Conclusion

In sum, the data analysis and consequent results indicated that different test formats produce different results, confirming the findings from numerous research studies which have demonstrated that the methods which are used to measure language ability influence performance on language tests (Bachman, 1990). The significant difference between the students' mean scores on multitrak items and standard multiple-choice tests also confirmed Gergely's findings that multitrak items "require a different kind of thinking", and students had to read and consider each response option carefully and draw on various kinds of grammatical knowledge to respond correctly. Geregly (2007) also indicated that though double-blank and standard MC items may be used efficiently at level B1 (knowing enough language to get by (*Common European Framework of Reference for Languages,* CEFR, 2001)), above this level, they tend to supply the language tester with less information about the candidates. His study revealed that multitrak items provide more information about candidates in the ability range represented by levels B2, which specifies a "sufficient range of language to be able to give clear descriptions (CEFR, 2001) and C1, which includes an appropriate formulation from a broad range of language (CEFR. 2001). Since the participants of this study were senior students majoring in English translation, literature, and TEFL, their level could be defined as B2 and above, and the highest construct validity which belonged to the TOEFL error-identification test confirmed Gergely's finding. However, the existence of the No-error option in this kind of test could violate its construct validity.

In Iran access to higher education is based on a supply and demand model, and social stratification corresponds to the number of years of formal education, which is directly associated with the prestige of the highest-reached educational institution. Therefore, it is not surprising that the number of candidates who take part in MA Admissions tests of Iranian universities increases every year. Considering the effects of these tests on individuals' lives, future, and careers, the large amount of money and

energy spent in the construction and administration of the tests, and their possible undesirable backwash effects, this study highlighted the need for compensating for the effect of different methods and options which are used in testing a given construct, especially in high stake tests such as admissions tests. The results also point out the need for a closer examination of the error-identification section of the state universities' admissions tests in Iran. Since the no-error option of such tests reduces their validity to a considerable amount, the rationale behind employing this option should be re-examined.

The multiple comparison of the means of the three tests revealed that MC items are easier for the students to answer while multitrak items with no-error options are much more difficult. However, the error identification test conforming to the TOEFL in format proved to be not only highly construct valid but midway between the other two tests in terms of difficulty. A test which is too difficult or too easy to answer does not provide us with reliable information about competency levels. Language test developers are always advised to construct tests which are at the right level of difficulty. The statistical characteristics of the error-identification test used in this study revealed that this format could be the most appropriate for testing the knowledge of grammar, especially when high stake decisions are to be made.

On the basis of the findings, the higher construct validity of the TOEFL error-identification tests revealed that it is a valid means of assessing grammatical knowledge. Since finding plausible distracters for standard multiple-choice items is not an easy job, especially concerning certain grammar points which yield few logical options, error-identification grammar tests can be a good replacement for multiple-choice grammar tests. Given the results of this study, the researchers recommend using the TOEFL error-identification tests for assessing students' grammar knowledge in high stake admissions tests.

References

Bachman, L. & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bachman, L. F. (1990). *Fundamental consideration in language teaching.* Oxford: Oxford University press.

Borg, W. R., & Gall, M. D. (1989). *Educational research: an introduction* (fifth edition ed.). White Plains, NY: Longman.

Brown. H. D, (2004). *Language assessment principles and classroom practices.* USA: Longman.

Camara, W. & Echternacht, G. (2000). The SAT I and high school grades: utility in predicting success in college. Research Report RN-10, *College Entrance Examination Board*. New York.

Carroll, J.B. (1980). *Testing communicative performance: an interim study*. Oxford: Pergamon Press.

Chapelle, C.A , Enright, M.K, & Jamieson, J.M. (2009). *Building a validity argument for the test of English as a foreign language.* New York: Routledge, Taylor & Francis group.

Cohen, A. D. (1984). On taking language tests: what the students report. *Language Testing*.1(1), 70-81.

College Board .com, INC, (2010): *SAT identifying sentence errors.*

Council of Europe. (2001). *Common European framework of reference for languages: learning, teaching, assessment.* Cambridge: Cambridge University Press.

Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.

DeVelice, R. F. (1991). *Scale development*. New Bury Park, NJ: Sage Publications.

Educational Testing Service (ETS). (2007). *Test and score data summary for TOEFL® internet-based and paper-based tests, the TOEFL® Test — quality beyond measure*. Retrieved 16th February, 2010 from http://www.ets.org/toefl

Farhady, H. (1983). On the plausibility of the unitary language proficiency factor. In J.W. Oller, Jr., (Ed.). *Issues in language testing research* (pp. 11-28). Rowley, MA: Newbury House.

Garson , David .(2010). *Factor Analysis*. From Statnotes: Topics in multivariate analysis. Retrieved June, 12 , 2010 from http://www2.chass.ncsu.edu/garson/PA765/factor.htm

Gee. J.P. (1999). *An introduction to discourse analysis theory and method*. London: Routledge.

Geiser. S & Studley. R. (2001). *SAT I vs. SAT II: predictive validity and social stratification* (University of California, Office of the President). Abstract retrieved May 20, 2009.

Gergely. D. (2007): Investigating the performance of alternative types of grammar items [Electronical version]. *Language Testing*, Vol. 24, No. 1, 65-97. Retrieved May, 2, 2009.

Gregory, R.J. (1992). *Psychological testing: history, principles and applications.* Boston: Allyn and Bacon.

Heaton, J.B. (1990). *Writing English language tests*. New York: Longman.

Linn, Robert. L. (1993). *Educational measurement.* American Council on Education. New York: Oryx Press.

Madson, Harold. S. (1983). *Techniques in testing*. Oxford: Oxford University Press.

Mahmoodian, B. (2000).*On the consequential validity of TOEFL, ethical concerns and considerations*. Unpublished master's thesis, University of Tehran. Tehran. Iran.

McNamara, T. & Roever, C. (2006). *Language testing: the social dimension*. Oxford: Blackwell.

Messick, S. (1995). Validity of psychological assessment. *American Psychologist* (50) 741–49.

Pacenza, Matt. (2010): *Whiter and wealthier, how the SAT and the end of affirmative action threaten campus diversity.* Retrieved 12th April, 2010 from http://www.journalism.nyu.edu

Richard, J., Platt, J,. & Weber, H., (1985). *Longman dictionary of applied linguistics*. Longman: London.

Schizuka. T , Takeuchi. o & Takeuchi. T. (2006). A comparison of three- and four-option English tests for university entrance selection purposes in Japan. *Language Testing* 23(1)35-57.

Sharma, S. (1996). *Applied multivariate techniques*. USA: John Wiley & Sons Inc.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing*, 1(2), 147-170.

Smith, S.E. (2009). *What is TOEFL?* Retrieved 23 May, 2009 from http://www.wisegeek.com/what-is-the-toefl.htm

Stansfield (1986). A history of the test of written English: the developmental year. *Language Testing*, 3(2), 224-234.

Swaim, V.S. (2009). *Determining the number of factors in data containing a single outlier: A Study of factor analysis of simulated data.* Unpublished dissertation. Louisiana State University and agricultural and mechanical college.

Swan, M. (2005). *Grammar*. Oxford : Oxford University Press.

Thornbury, S (1999). *How to teach grammar*. Edinburgh, Pearson UCLA: Academic Technology Services, Statistical consulting group. *Annotated SPSS Output Factor Analysis*. Retrieved June, 2010 from http://www.ats.ucla.edu/stat

Thumpson & Daniel (1996). Factor analysis evidence for the construct validity of the scores: A historical overview and some guidelines. *Educational and Psychological Measurement*, 56 (2). 197-208.

Wolf .D, Bixby, J, Glenn, J., & Gardener, H (1991). To use their minds well: investigation of new forms of student assessment. *Review of Research I.*

Zwick, W. R. & Velicer, W. F, (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99 (3), 432-442.

Title
The Relationship between Critical Thinking and Deductive/Inductive
Teaching of Grammar to Iranian EFL Learners

Authors

Mansoor Fahim (Ph.D.)

Islamic Azad University, Science and Research Branch, Tehran, Iran

Shahla Azarnioushi (M.A.)

Islamic Azad University**,** Science and Research Branch, Tehran Iran

Bio Data

Mansoor Fahim, associate professor of TEFL at Allameh Tabataba'I University, Tehran, Iran from 1981 to 2008. At present, he runs Research methods, psycholinguistics, Applied Linguistics, Second Language Acquisition, and seminar classes at M.A. level. Also, First Language Acquisition, Psycholinguistics, and discourse Analysis at Ph.D. level. He has published several articles and books mostly in the field of TEFL and has translated some books into Persian.

Shahla Azarnioushi, M.A. in TEFL at Islamic Azad University, Science and Research Branch, Tehran, Iran. She is a lecturer at Allameh Tabataba'i University and School of International Relations Affiliated to Ministry of Foreign Affairs. She is interested in action research and reflective teaching practices.

Abstract

Recently, the interest of both teachers and researchers in the field of foreign language learning and teaching has increasingly focused on the learner, including the strategies which an individual uses in learning and communicating. The problem under investigation is to see whether there is any relationship between the critical thinking ability of language learners and their performances using rule driven/ discovery learning approaches to teaching grammar. After the homogenizing process, 73 learners were taught during two periods of eight sessions. During the first period, the researcher taught the group deductively and during the second period, inductively. At the end of each period, a grammar test was administered to measure the grammar knowledge of the learners. The results of the analyses for the collected data showed

that there was a positive correlation between the critical thinking ability of the learners and their grammar test scores in the inductive period. However, as for the deductive teaching method, no special relationship could be found between the critical thinking ability of the learners and their grammar test scores. In other words, the results of the study indicated that learners with a higher critical thinking ability prefer inductive methods of teaching grammar while in deductive methods of teaching grammar, there seems to be no difference between learners with high or low critical thinking abilities.

Keywords: Critical Thinking, Cognitive Style, Deductive Reasoning, Inductive Reasoning

Introduction

Over the last decade the number of studies concerned with the effects of learner traits such as critical thinking on learning a foreign/second language has increased considerably. This is in part due to the fact that non-linguistic factors can strongly influence language learning. In fact the goal of critical thinking is to establish a disciplined "executive" level of thinking to our thinking, a powerful inner voice of reason, to monitor, assess, and reconstitute–in a more rational direction–our thinking, feeling, and action. Benesch (1999) believes that when we become critical thinkers we develop an awareness of the assumptions under which we, and others, think and act. We learn to pay attention to the context in which our actions and ideas are generated. We become skeptical of quick-fix solutions, of single answers to problems, and of claims to universal truth. According to Cheung (2002) critical thinking suggests an integration of being critical and the thinking process. An adequate conceptualization of critical thinking should combine cognitive thinking skills, motivational dispositions, behavioral habits, and ideological beliefs.

As Beach (2004) says, "The art of Socratic questioning is important for the critical thinker because the art of questioning is important to excellence of thought". What the word 'Socratic' adds is "systematicity," "depth," and a keen interest in assessing the truth or plausibility of things.

Beach (2004) believes that there is a special relationship between critical thinking and Socratic questioning because both share a common end. Critical thinking gives one a comprehensive view of how the mind functions (in its pursuit of meaning and truth), and Socratic questioning takes advantage of that overview to frame questions essential to the quality of that pursuit.

Inductive and Deductive approaches in teaching grammar

There are many theoretical approaches that have been developed to promote the students' success in learning new information. In TESOL (Teaching English to Students of Other Languages), there are two main theoretical approaches for the presentation of new English grammar structures or functions to ESL/EFL students: inductive approach and deductive approach. The more traditional of the two theories, is the deductive approach, while the emerging and more modern theory, is the inductive approach.

The deductive approach represents a more traditional style of teaching in which the grammatical structures or rules are dictated to the students first. Thus, the students learn the rule and apply it only after they have been introduced to the rule. For example, if the structure to be presented is present perfect, the teacher would begin the lesson by saying, "Today we are going to learn how to use the present perfect structure." Then, the rules of the present perfect structure would be outlined and the students would complete exercises, in a number of ways, to practice using the structure. In this approach, the teacher is the center of the class and is responsible for all of the presentation and explanation of the new material.

The inductive approach represents a more modern style of teaching where the new grammatical structures or rules are presented to the students in a real language context. The students learn the use of the structure through practice of the language in context, and later realize the rules from the practical examples. For example, if the structure to be presented is the comparative form, the teacher would begin the lesson by drawing a figure on the board and saying, "This is Jim. He is tall." Then, the teacher would draw another taller figure next to the first saying, "This is Bill. He is taller than Jim." The teacher would then provide many examples using students and items from the classroom, famous people, or anything within the normal daily life of the students, to create an understanding of the use of the structure. The students repeat after the teacher, after each of the different examples, and eventually practice the structures meaningfully in groups or pairs. With this approach, the teacher's role is to provide meaningful contexts to encourage demonstration of the rule, while the students evolve the rules from the examples of its use and continued practice.

However, in both approaches, the students practice and apply the use of the grammatical structure, yet, there are advantages and disadvantages to each in the EFL/ESL classroom. The deductive approach can be effective with students of a higher level, who already know the basic structures of the language, or with students who are accustomed to a very traditional style of learning and expect grammatical presentations. The deductive approach however, is less suitable for lower level language students, for presenting grammatical structures that are complex in both form and meaning, and for classrooms that contain younger learners. The advantages of the inductive approach are that students can focus on the use of the language without being held back by grammatical terminology and rules that can inhibit fluency. The inductive approach also promotes increased student participation and practice of the target language in the classroom, in meaningful contexts. The use of

the inductive approach has been noted for its success in EFL/ESL classrooms world-wide, but its disadvantage is that it is sometimes difficult for students who expect a more traditional style of teaching to induce the language rules from context. Understanding the disadvantages and advantages of both approaches, may help the teacher to vary and organize the EFL/ESL lesson, in order to keep classes interesting and motivating for the students.

The Deductive Approach towards teaching grammar:

The deductive approach to grammar teaching provides grammar rules before anything else. Many students and teachers are more comfortable with a rule-based approach to grammar and, though some teachers would argue against this approach, it is likely to continue to be used extensively classrooms.

The Inductive Approach towards teaching grammar:

The Inductive thinking proceeds from a specific case, or from cases, to the general. This is the opposite of deductive thinking. In inductive thinking, the individual makes a number of observations which are then sorted into a concept or generalization; the individual does not have prior knowledge of the abstraction but only arrives at it after observing and analyzing the observations.

Like deduction, the process of induction is a very common and often unconscious process in humans. From the day babies are born, they take in sensory impressions and try to make sense of them. Because of the lack of language in young children, this learning must be inductive. The burden, then, for parents who wish to provide learning environments for young children is to insure that the child has ample amounts of data to inductively process into meaningful information. The same is true with older children when dealing with ideas where they lack the terminology.

In fact, the inductive approach is a teaching strategy which uses data to teach pupils concepts and generalizations. In this approach, the teacher presents pupils with data; pupils are asked to make observations of the data and, on the basis of these observations, to form the abstraction being taught.

Implementing the approaches in teaching grammar

The teacher often begins by presenting one of the selected examples or illustrations of the abstraction to be taught. The teacher asks the pupils to make as many observations from the example as possible. Many of the observations may not be related to the generalization the teacher has in mind. However, they are accurate observations, and as they are made, the teacher might beneficially write them on the board for transparency.

After the pupils have exhausted the number of observations they can make, or the teacher in the interest of time decides to move on, a second example is presented. Again the students make as many observations as possible. They may already begin to notice similarities between the first and second example which will tend to narrow the range of observations for succeeding examples.

The effectiveness of deductive and inductive approaches in teaching grammar
The effectiveness of the **deductive** and **inductive** approaches, aiming at maximizing the students' opportunity to practice thinking skills, has been investigated in empirical studies. *Deductive* learning is an approach to language teaching in which learners are taught rules and given specific information about a language. Then, they apply these rules when they use the language. This may be contrasted with *inductive* learning in which learners are not taught rules directly, but are left to discover or induce rules from their experience of using the language. It should be mentioned that these two techniques encourage learners to compensate for the gap in their second language knowledge by using a variety of communication strategies. A number of research studies, likewise, have reported that successful learners often adopt certain learning strategies such as seeking out practice opportunities or mouthing the questions put to other learners. Inductive and deductive models offer this chance to learners because these two models foster a cooperative atmosphere among students.

Purpose of the study
The purpose of the present study is to explore the relationship between critical thinking as an independent variable and inductive/deductive teaching of grammar as dependent variables. That is, the researcher wanted to see whether teaching grammar operates identically on different subjects who show pronounced differences in the above dichotomy. It's worth mentioning here that critical thinking is a cognitive skill which exists and influences the way we think. It does have some impact on nearly anything in our lives. Critical thinking is a technique for evaluating information and ideas. It is also a technique for deciding what to accept and believe. Concerning the objectives of the study, the following research questions were propounded:

1. Do learners with more critical thinking ability show any
    significant preference towards deductive teaching of grammar?
2. Do learners with less critical thinking ability show any significant
    preference towards inductive teaching of grammar?
3. Do learners with more critical thinking ability show any significant
    preference towards inductive teaching of grammar?
4. Do learners with less critical thinking ability show any significant
    preference towards deductive teaching of grammar?

In keeping with the above research questions, the following null hypotheses were proposed:

H0(1): Learners with more critical thinking ability do not show any
     significant preference towards deductive teaching of grammar.
H0(2): Learners with less critical thinking ability do not show any
     significant preference towards inductive teaching of grammar.
H0(3): Learners with more critical thinking ability do not show any
     significant preference towards inductive teaching of grammar.

H0(4): Learners with less critical thinking ability do not show any significant preference towards deductive teaching of grammar.

Method

participants

To investigate the hypotheses, the researcher selected an experimental random sample for the collection of the data. The selected sample for this research included 120 students from Allame Tabatabaee University. It is worth mentioning that all 120 subjects were translation students from both sexes. At first, the TOEFL test was administered in order to determine the level of proficiency of the learners and make a group as homogeneous as possible. Through considering the normal distribution of the subjects' scores on the proficiency test, those subjects whose scores were one standard deviation above and below the mean were decided to be in the final group of subjects. Consequently, 73 of them were found to be homogenous and were selected for the purpose of this research. The next step was to determine the level of critical thinking ability among the subjects. Administering the critical thinking questionnaire, the researcher found that out of the 73 subjects, 34 students were critical thinkers while 39 of them had less critical thinking abilities.

Instrumentation

The TOEFL test:

The TOEFL test was employed to determine the subjects' level of   English language proficiency. The proficiency test included 70 items. It took the subjects 45 minutes to answer. As it was specified by the test book, 25 minutes was needed for the grammar section and 20 minutes for the vocabulary section. All the items were in the multiple choice format.

The critical thinking questionnaire:

A critical thinking questionnaire including 30 items was administered to the subjects to evaluate the skills of analysis, inference, evaluation, inductive reasoning, and deductive reasoning. The critical thinking questionnaire was adopted from the thesis of Naieni (2005). She states that this questionnaire was provided from Peter Honey. Honey (2004, cited in Naieni, 2005) states that their mission is to help organizations to be successful through learning. As Naieni (2005) has stated, the English version of critical thinking questionnaire was translated by her to guarantee the full comprehension of the questions by the subjects. As Naieni (2005) claims, the Persian version of the questionnaire was scrutinized by some experts at Azad University, and they made necessary modifications. The corrected version went through a pilot study and was administered to about 20 students from Kish Institute. As she states, after getting the feedback from these students, the reliability of the questionnaire was found to be 0.86. And, the reliability of this questionnaire was acceptable.

A grammar test

In order to measure the grammar knowledge of the learners, a grammar test was administered.

Procedure
In order to accomplish this research, different steps were followed. First, the proficiency test and the Persian version of the critical thinking questionnaire were given to one hundred and twenty students from Allame Tabatabaee University. Through considering the normal distribution of the subjects' scores on the proficiency test, those scores which were one standard deviation above and below the mean were decided to be in the group of subjects. Later, in order to determine the level of critical thinking among the subjects the critical thinking questionnaire was administered and the researcher found that out of 73 subjects, 34 were critical thinkers and 39 had less critical thinking abilities.

The researcher could not have all the 73 subjects in one class. Therefore, the subjects were randomly put into 3 different classes. The teaching time for the whole semester was 16 sessions. The researcher divided the semester into two subdivisions of 8 sessions. During the first 8 sessions, the researcher chose the first 6 units from the English teaching series known as "passages" and taught them deductively to all the subjects in the three classes. The time allocated to each session was 90 minutes. Because of the purpose of this study, in each session, 45 minutes was allocated to teaching grammar. The other 45 minutes was used to teach the other parts of the unit. During the second 8 sessions, the researcher chose the second 6 units from "passages" and taught them inductively to all the subjects in the three classes. The time allocated to each session was 90 minutes. Because of the purpose of this study, in each session, 45 minutes was allocated to teaching grammar. The other 45 minutes was used to teach the other parts of the unit.

Validity Indices
The correlation coefficients between the critical thinking test (CTT) and deductive and inductive tests with the proficiency test are used as the validity indices of the former tests. As displayed in Table 1, all of the correlation coefficients are labeled as significant by receiving one or two asterisks.

Table 1: Validity Indices

		CRITICALTHINKING	DEDUCTIVE	INDUCTIVE
THE PEOFICIENCY TEST	Pearson Correlation	.244(*)	.560(**)	.672(**)
	Sig. (2-tailed)	.037	.000	.000
	N	73	73	73

The critical value of Pearson r at 71 degrees of freedom is .22. All of the observed values of r are higher than the critical value of .22, thus it can be claimed that the instrument employed in this study enjoys statistically significant validity indices although the correlation between the TOFFL and CTT is not as high as the other tests.

Reliability Indices
The KR-21 reliability indices are calculated for the instruments employed in this study. As displayed in table 2, all of the tests administered in this study enjoy high reliability indices.

Table 2: Reliability Indices

	N of Items	Mean	Variance	KR-21
THE PEOFICIENCY TEST	75	50.3562	109.316	.86
DEDUCTIVE	40	32.8630	15.620	.64
INDUCTIVE	40	34.4658	15.058	.70

Figure 1. **Graph Showing the Score Distribution of the Proficiency Test**
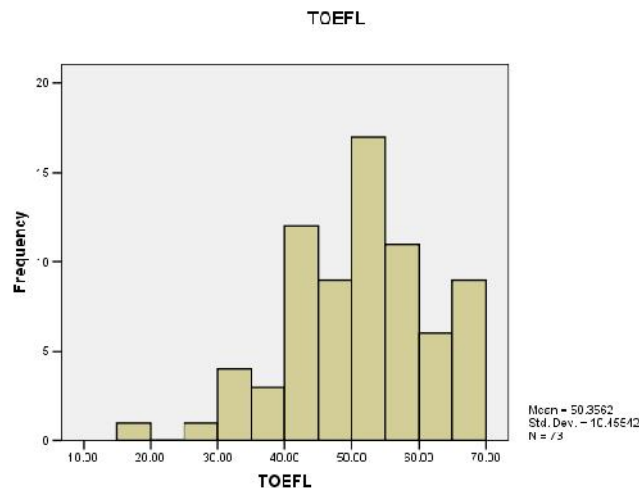
The above diagram illustrates the grades of the proficiency test for the homogenized group. As it can be seen on the diagram, for the 73 subjects, the mean score for this group is 50.35. The majority of the marks have gathered between 40 and 60. Also, there are no marks above 70. The details have been summarized in the following table.

**Table 3: Statistics for the grades of the proficiency test**

Parameter	N	Mean	Minimum	Maximum	variance
Value	73	50.3562	19	70	109.316

Table 4. **Descriptive Statistics of the Critical Thinking Questionnaire**

| Parameter	N	Mean	Minimum	Maximum
Value	73	110	78	138

Based on the students' total scores on the Critical Thinking Questionnaire, they were divided into two groups of high and low critical thinking learners. The mean of the Critical Thinking Questionnaire was 110. Those who scored 110 and below form the low critical thinking group (39 subjects) and the rest of the subjects form the high group with 34 subjects. As it is shown in the above table, the maximum and minimum scores obtained by the participants were 138 and 78 respectively. The mean was 110. Figure 2 shows the frequency and the distribution of the scores of the participants on the critical thinking questionnaire.

Figure 2. **Graph Showing the Score Distribution of the Critical Thinking**
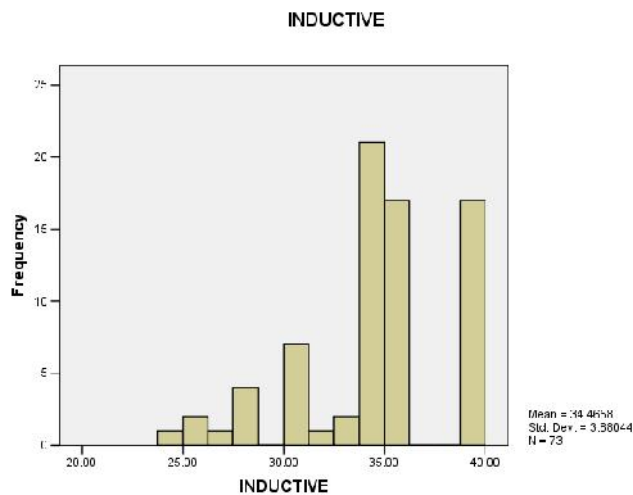   **Questionnaire**
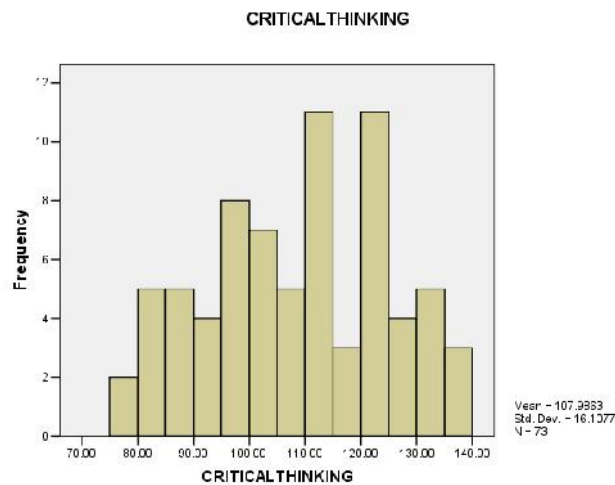
Figure 3. **Descriptive Data for the Deductive Test**
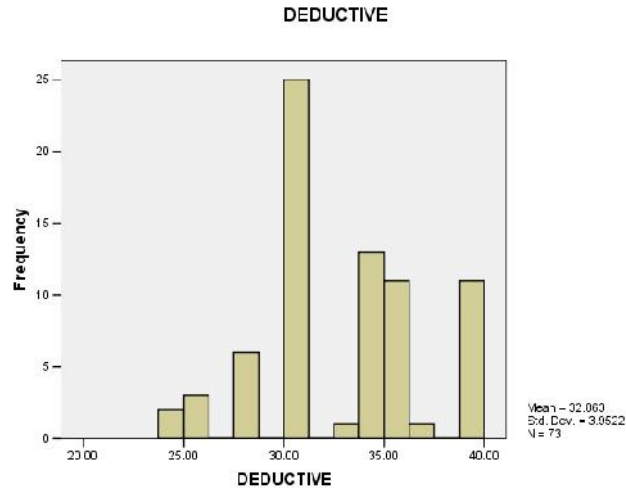
DEDUCTIVE

The above diagram illustrates the grades of the students in the grammar test after being taught deductively. As it can be seen on the diagram, for the 73 subjects, the mean score for this group is 32.863. The majority of the marks have gathered between 30 and 32. Also, there are no marks above 40.

Figure 4. **Descriptive Data for the Inductive Test**

INDUCTIVE

The above diagram illustrates the grades of the students in the grammar test after being taught inductively. As it can be seen on the diagram, for the 73 subjects, the

mean score for this group is 34.46. The majority of the marks have gathered between 34 and 36. Also, there are no marks above 40.

Testing the Hypotheses
The Proficiency Test and the Critical Thinking Questionnaire
To examine the relationship between the critical thinking ability of the learners and their language proficiency level, the researcher used the Pearson Product Moment Correlation Coefficient for the scores of the two tests. In fact correlation shows the degree of togetherness of the two variables. Table 5 illustrates the results.

*Table5.* **Correlation of the Critical Thinking Questionnaire and Language Proficiency Level**

		CRITICALTHINKING
THE PEOFICIENCY TEST	**Pearson Correlation**	.244(*)
	Sig. (2-tailed)	.037
	N	73

* Correlation is significant at the 0.05 level (2-tailed).

As Table 5 shows, there is a positive relationship between the critical thinking ability and the language proficiency level of the participants. The value of correlation (r=0.244) indicates how closely critical thinking ability and language proficiency level of the learners are related.

The following chart displays the scatter plot for the proficiency test and Critical Thinking Questionnaire. The plot is neither curve not funnel shaped. This graph indicates that the plot is not curve linear; hence, the Pearson correlation between the TOEFL and Critical Thinking Questionnaire can be calculated.

Figure 5. **The Pearson correlation between the TOEFL and Critical Thinking Questionnaire**
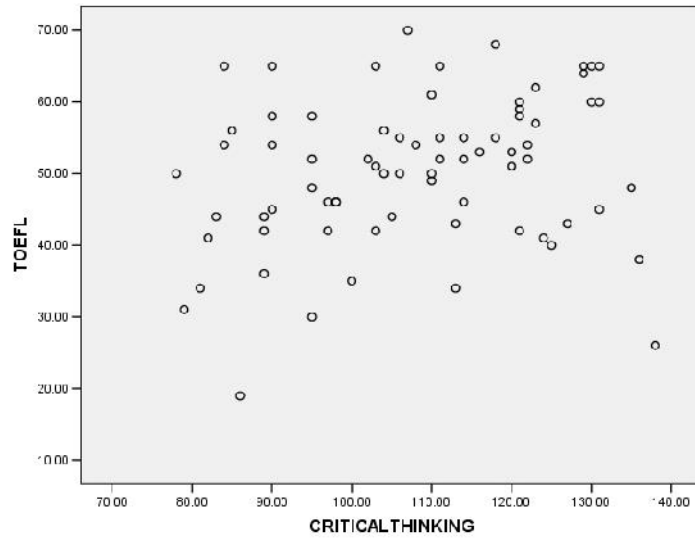
The Proficiency Test and the Inductive Test

To examine the relationship between the language proficiency level of the learners and their performance in the inductive test, the Pearson Product Moment Correlation was used in order to describe the relationship between these two variables. Table 6 shows the results.

Table 6. **Correlation of the language proficiency level and the Inductive Test**

		INDUCTIVE
	Pearson Correlation	.672(**)
THE PEOFICIENCY TEST	Sig. (2-tailed)	.000
	N	73

As it is obvious in table 6, the value of correlation is 0.672. It means that there is a positive relationship between the language proficiency level of the learners and their performance on the grammar test.

The following chart displays the scatter plot for the TOEFL and the Inductive Test. The plot is neither curve not funnel shaped. This graph indicates that the plot is not curve linear; hence, the Pearson correlation between the TOEFL and the inductive test can be calculated.

Figure 6. **The Pearson correlation between the TOEFL and the inductive test**

228
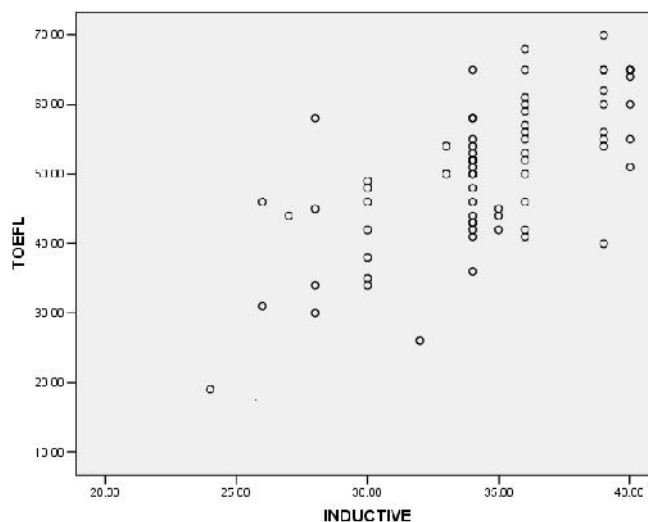
The Proficiency Test and the Deductive Test

In order to examine the relationship between the critical thinking ability of EFL learners and their performance on the deductive test, the researcher used the correlation coefficient in order to reveal the degree of go togetherness of these two variables. Table 7 reveals the results.

Table 7. **Correlation of the language proficiency level and the deductive test**

		DEDUCTIVE
	Pearson Correlation	**.560(**)**
THE PEOFICIENCY TEST	**Sig. (2-tailed)**	**.000**
	N	**73**

   ** Correlation is significant at the 0.01 level (2-tailed).

   According to Table 7, there is a positive relationship between the language proficiency level of the subjects and their performance on the deductive test. The value of correlation (r=0.506) means that the subjects below the mean on the TOEFL test are below the mean on the deductive test; subjects above the mean on the proficiency test are above the mean on the deductive test. In other words, there is a positive relationship between the language proficiency level of EFL learners and their performance on the deductive test.

   The following chart displays the scatter plot for the TOEFL and Deductive Test. The plot is neither curve not funnel shaped. This Graph indicates that the plot is not curve linear hence the Pearson correlation between the TOEFL test and the Deductive test can be calculated.

Figure 7. **The Pearson correlation between the TOEFL test and the Deductive test**
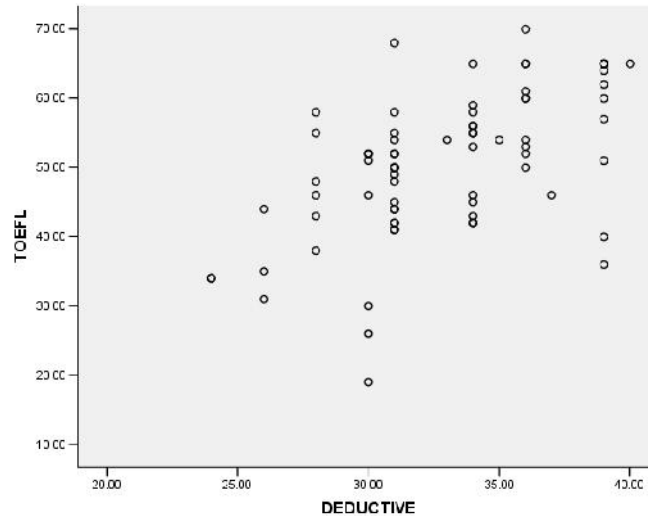
Discussion

Critical thinking is one of the cognitive abilities which "increases the probability of a desirable outcome" (Halpern, 1996 cited in Fowler, 2004, 3). According to Bachman (1990), cognitive characteristic is one of the sources which affect the performance of the test takers in language tests. Also Bachman states that, "There is a general agreement among language acquisition researchers and theorists that cognitive factors influence language acquisition" (p. 275). Furthermore, Bachman (1990) claims that, "The effects of cognitive characteristics on performance in language tests raise questions about the relationships between these constructs and our definitions of language abilities" (p. 279). A rather detailed discussion on the conclusions and implications of this study appears in the next and final chapter. For the analyses of data, based on the design of the study, the statistical packages for social sciences were employed to perform t-test. That is, this statistical technique was used to figure the effects of test takers' critical thinking ability. An independent t-test was used to compare the mean scores of the high and low Critical Thinking Questionnaire scores on the Deductive test. The t-observed value 1.97 does not exceed the critical value of t at 71 degrees of freedom, i.e. 1.99.

Table 8: Independent T-Test Deductive Test by the Critical Thinking Questionnaire:

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
DEDUCTIVE	Equal variances assumed	.576	.450	1.978	71	.052	1.79789	.90912	-.01485	3.61063
	Equal variances not assumed			1.960	66.371	.054	1.79789	.91713	-.03302	3.62880

Based on these results it can be concluded that there is not any significant difference between the high and low Critical Thinking Questionnaire mean scores on the deductive Test. Thus the null-hypothesis as no significant difference between the high and low Critical Thinking Questionnaire groups on the Deductive Test is supported. In other words, irrespective of one's high CTT or low CTT score, they do not perform differently on the Deductive Test. The descriptive statistics for the high and low CTT groups on the Deductive Test are displayed in table 9.

Table 9: Descriptive Statistics Deductive Test by CTT

	THINKING LEVEL	N	Mean	Std. Deviation	Std. Error Mean
DEDUCTIVE	HUIGH	34	33.8235	4.13023	.70833
	LOW	39	32.0256	3.63815	.58257

It should be noted that the underlying assumption of the independent t-test, homogeneity of the variances of the two groups, is met. As the results of the Levene's Test of Equality of Variances indicate, the F-value of .75 has a probability of .45. Since the probability is greater than the .05 level proposed by the researcher, it can be concluded that the two groups of high and low CTT enjoy homogeneity of variances on the Deductive Test. That is why the first row of Table 4 - Equal variances assumed – is reported.

An independent t-test was used to compare the mean scores of the high and low CTT on the Inductive test. The t-observed value 3.83 exceeds the critical value of t at 71 degrees of freedom, i.e. 1.99.

Table 10: Independent T-Test Inductive Test by CTT

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
INDUCTIVE	Equal variances assumed	4.052	.048	3.838	71	.000	3.20211	.83441	1.53835	4.86587
	Equal variances not assumed			3.931	67.915	.000	3.20211	.81466	1.57644	4.82778

Based on these results it can be concluded that there is a significant difference between the high and low CTT groups' mean scores on the Inductive Test. Thus the null-hypothesis as no significant difference between the high and low CTT groups on

the Inductive Test is rejected. In other words, the high CTT group outperformed the low CTT group on the Inductive Test.

The descriptive statistics for the high and low CTT groups on the Inductive Test are displayed in table 11.

Table 11: Descriptive Statistics Inductive Test by CTT

	THINKING LEVEL	N	Mean	Std. Deviation	Std. Error Mean
INDUCTIVE	HUIGH	34	36.1765	2.84414	.48777
	LOW	39	32.9744	4.07487	.65250

It should be noted that the underlying assumption of the independent t-test, homogeneity of the variances of the two groups, is not met. As the results of the Levene's Test of Equality of Variances indicate, the F-value of 4.05 has a probability of .048. Since the probability is lower than the .05 level proposed by the researcher, it can be concluded that the two groups of high and low CTT do not enjoy homogeneity of variances on the Inductive Test. That is why the second row of Table 6 - Equal variances not assumed – is reported.

As the results of the analyses indicate, there is a cognitive style bias operating in conjunction with the grammar teaching method. That is, the critical thinking cognitive style as an individual cognitive characteristic can affect performance on learning grammar. The results in the first statistical analysis show that learners with a high level of critical thinking ability prefer the teaching methods in which the teacher directly goes to the examples and then later on tries to focus on the rules.

However, no special difference could be seen between the scores of learners with high and low critical thinking abilities on the deductive test.

The fact that learners with a high level of critical thinking ability show preference towards inductive methods of teaching was very much fascinating to the researcher. So in other words, the teachers for language learners with high critical thinking abilities should emphasize the fact that the method in the classroom should be inductive. In other words, learners with high critical thinking abilities seem to have lack of interest in rules, regulations and fixed standards. It is noteworthy here that the present research findings can seriously cast doubts on our interpretations of the scores in grammar classes.

The results of the analyses for the collected data showed that there was a significant difference at the level of $p<0.05$ between the grades of the students in the inductive final test but not in the deductive final test. In other words, the results of the study indicated that there was a cognitive style bias operating in conjunction with learning different grammatical structures regarding the critical thinking ability of the learners.

Learners with high critical thinking abilities prefer inductive methods of teaching grammar whereas in the deductive test, no special difference could be seen among the grammar scores of learners with high or low critical thinking abilities. In other words, the first and fourth null hypotheses could not be rejected while the second and third null hypotheses were rejected.

The findings of this study have both theoretical and pedagogical implications. From the standpoint of practice, the present research has got direct relevance to the improvement of language teaching in practice.

Theoretically speaking, as the findings of the study illustrate, the results of the exams in grammar teaching classes are to some extent distorted on the grounds that they are practically at the mercy of the cognitive characteristics of the language learners. This insight into the nature of teaching can encourage language teachers to reconsider teaching grammar. Moreover, the present research results warn language teachers and test developers against soft-pedaling on non-linguistic factors such as critical thinking. That is, instead of ignoring such factors, they must identify them to minimize their potential effects.

Since the old days, language learners have been supposed to be equal in the way they learn the grammar. That is, individual factors have always been neglected in this regard. Hence, the findings of the present study not only call the above method of testing into question but also call for a judicious coordination between the above variables. This can be best achieved by dividing the learners according to their cognitive styles. Furthermore, the present research results can be beneficial to syllabus designers, curriculum and test developers on the ground that by dint of such studies, they will be able to make slight modifications on their approaches to both language teaching and testing. That is, they can adapt their teaching and testing styles to students' cognitive styles.

References

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford : Oxford University Press.

Beach, D. (2004). The role of Socratic questioning in thinking, teaching,& learning. Retrieved September 26, 2005, from http://www.criticalthinking.org?resources/articles/the-role-socraatic-questioning-ttl.shtml

Benesch, S. (1999). Thinking critically, thinking dialogically. *TESOL Quarterly. 33(3),* 573-580.

Brown, H. D. (2000). *Principles of language learning and teaching.* (4th ed.). Longman: Longman.

Cheung, C. K., (2002). Assessing university students' general and specific critical thinking. Retrieved September 26, 2005, from http://www.findarticles.com /p/articles/mi_m0FCR/is_4_36/ai_96619957/print

Fowler, B. (2004). Critical thinking across the curriculum project. Retrieved November 24, 2004, from http://www.kcmetro.cc.mo.us/longview/ ctac/definitions.htm

Freeley, A. J., & Steinberg, D. L. (2000). *Argumentation and debate: Critical thinking for reasoned decision making.* Stamford: Wadsworth.

Naeini, J. (2005). *The effects of collaborative learning on critical thinking of Iranian EFL learners.* Unpublished M.A. Thesis, Islamic Azad University, Central branch, Tehran, Iran.

Richards, J. C., & Renandya, W.A. (2002). *Methodologies in language teaching. an anthology of current practice* Cambridge University Press: CUP.