



**ALLAMEH TABATABA'I UNIVERSITY
FACULTY OF LITERATURE AND FOREIGN LANGUAGES
EFL DEPARTMENT**

**Collocational Errors in the Written
English of Adult Iranian EFL Learners:
A Corpus-Based Analysis**

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of Master of Arts in TEFL

Supervisor: **Dr. Fahimeh Ma'refat**
Advisor: **Dr. Mehdi Nowruzi Khiabani**

Conducted By:
Mahdi Hajatpour

TEHRAN
February 2009

**IN THE NAME OF
GOD**

فرم گردآوری اطلاعات پایان نامه ها

کتابخانه مرکزی دانشگاه علامه طباطبائی

عنوان: تحلیل خطاهای نوشتاری زبان آموزان ایرانی در کاربرد همایندهای زبان انگلیسی

The Collocational Errors in the Written English of Adult Iranian EFL learners:
A Corpus-Based Analysis

نویسنده / محقق: مهدی حاجت پور بهبید

مترجم: ندارد

استاد راهنما: دکتر فهیمه معرفت استاد مشاور: دکتر مهدی نوروزی خیابانی استاد داور: دکتر غلامرضا تجویدی

واژه نامه:

کتابنامه:

کاربردی

توسعه ای

بنیادی

نوع پایان نامه:

سال تحصیلی: 1387-88

مقطع تحصیلی: کارشناسی ارشد

دانشکده: ادبیات و زبانهای خارجی

نام دانشگاه: علامه طباطبائی

محل تحصیل: تهران

گروه آموزشی: زبان و ادبیات انگلیسی

تعداد صفحات: 100

کلید واژه ها به زبان فارسی:

همایندها، واژه های همایندها، توانش همایندها، بیکره زبانی، زبانشناسی بیکره ای، زبانشناسی مقابله ای، تحلیل بیکره ای زبان

کلید واژه ها به زبان انگلیسی:

British National Corpus, Collocations, Collocational Competence, Contrastive Linguistics, Corpus Analysis, Corpus Linguistics, Linguistic Corpus

چکیده

الف. موضوع و طرح مسئله (اهمیت موضوع و هدف):

ورود رایانه به حوزه تحقیق و پژوهش را می توان زمینه ساز تحولی شگرف در علوم برشمرد. به موازات بهره مندی علوم مختلف از این فناوری، علم زبانشناسی نیز در شناخت و تجزیه عناصر مختلف زبانی و کمک به آموزش زبان دوم به زبان آموزان، تا جایی پیشرفته است که امروزه زبانشناسی رایانه ای جایگاه تثبیت شده ای را در کنار گرایشات مختلف این رشته بدست آورده است. یکی از موارد بهره مندی این رشته از رایانه، شناسایی واژگانی است که به دلیل تمایل به مجاورت با واسطه یا بی واسطه با یکدیگر در یک بافت زبانی مشخص، از آنان تحت عنوان «واژگان همایند» یاد می شود. اهمیت کاربرد صحیح واژگان همایند در نوشتار و گفتار به اندازه ای است که کلیه زبانشناسانی که در حوزه آموزش زبان دوم به تحقیق و مطالعه مشغولند، استفاده صحیح این واژگان را یکی از ملاکهای اصلی تشخیص گویشور بومی از گویشور غیربومی ارزیابی می کنند. تحقیق حاضر کاربرد همایندهای زبان انگلیسی را در انگلیسی نوشتاری زبان آموزان ایرانی مورد تحلیل قرار می دهد.

ب. مبانی نظری شامل مرور مختصری بر منابع، چارچوب نظری و پرسشها و فرضیه ها:

یکی از شاخصه های اصلی زبان گویشوران غیربومی زبان، استفاده از واژگانی است که هرچند به لحاظ زبانشناختی و معناشناختی انحرافی در همسویی آنها با استانداردهای زبان دوم مشاهده نمی شود اما از دید گویشوران بومی زبان به عنوان نمونه ای از «استفاده غیربومی از واژگان» قابل تشخیص می باشند. از این رو، با فرض مشکل آفرین بودن کاربرد همایندهای زبان انگلیسی برای زبان آموزان ایرانی، سؤالات زیر به عنوان سؤالات اصلی تحقیق مطرح گردید:

- 1- مشکل سازترین نوع همایندهای زبان انگلیسی برای زبان آموزان ایرانی کدامند؟ به بیان دیگر، آیا انواع مختلف همایندها به یک اندازه زبان آموزان ایرانی را با مشکل مواجه می کنند؟
- 2- آیا زبان آموزانی که در سطوح مختلفی از توانش زبانی قرار دارند، در کاربرد همایندهای زبان انگلیسی به یک اندازه با مشکل مواجه می شوند؟ به بیان دیگر، آیا مهارت زبانی زبان آموزان بر توانش همایند گزینی آنان تأثیرگذار است؟
- 3- میزان تأثیرگذاری زبان اول بر خطاهای نوشتاری زبان آموزان ایرانی در کاربرد همایندهای زبان انگلیسی چقدر است؟
- 4- آیا زبان آموزانی که در سطوح بالاتری از توانش زبانی قرار دارند، به همان نسبت از همایندهای بیشتری در انگلیسی نوشتاری خود استفاده می کنند؟

پ. روش تحقیق شامل تعریف مفاهیم، روش تحقیق، جامعه مورد تحقیق، نمونه گیری و روشهای نمونه گیری، ابزار

اندازه گیری، نحوه اجرای آن و شیوه گردآوری و تجزیه و تحلیل داده ها:

اصطلاح «همایند» به واژگانی اطلاق می شود که به گفته «فیرث» (1957)، استفاده همزمان و مکرر آنها در یک بافت زبانی مشخص به یکی از عادات زبانی گویشوران بومی زبان تبدیل شده و جایگزینی هر یک از آنها با واژگان مترادف، نوعی انحراف از گونه زبانی معیار به شمار می رود. بر همین اساس، هدف از تحقیق حاضر، ارزیابی توانش همایند گزینی زبان آموزان ایرانی از طریق جمع آوری، طبقه بندی و تحلیل خطاهای نوشتاری آنان در کاربرد همایندهای زبان انگلیسی است. به منظور تحقق این هدف، 87 زبان آموز ایرانی از دانشگاه علامه طباطبائی، دانشگاه تهران، دانشگاه آزاد اسلامشهر و مؤسسه آموزش زبان کیش در تحقیق حاضر شرکت کردند که پس از شرکت در آزمون زبان انگلیسی میسیگان، 72 تن از آنان برای شرکت در مرحله نهایی تحقیق انتخاب و در 3 سطح مبتدی، متوسط و پیشرفته طبقه بندی شدند.

پس از برگزاری آزمون تعیین سطح و تقسیم بندی شرکت کنندگان به 3 گروه مبتدی (Low)، متوسط (Mid)، و پیشرفته (High)، توانش زبانی زبان آموزان در زمینه انتخاب همایندهای صحیح در زبان انگلیسی از طریق ارائه موضوع و بررسی کنش نوشتاری آنان مورد ارزیابی قرار گرفت. در این راستا، کلیه ترکیبات 2 واژه ای تولید شده توسط زبان آموزان ایرانی از متون نوشتاری آنان استخراج و در 6 گروه مختلف شامل: فعل + حرف اضافه، اسم + حرف اضافه، صفت + اسم، قید + صفت، فعل + قید و اسم + فعل طبقه بندی شدند. پس از استخراج و طبقه بندی این ترکیبات، قابلیت همایندی آنان در زبان انگلیسی با استفاده از منابعی چون دیکشنری همایندهای زبان انگلیسی آکسفورد، دیکشنری همایندهای زبان انگلیسی Collins COBUILD English Dictionary و بانک اطلاعات زبانی بریتانیا British National Corpus مورد ارزیابی و خطاهای زبان آموزان ایرانی در کاربرد همایندهای زبان انگلیسی به لحاظ مشأ خطا (زبان اول یا زبان دوم) مورد تحلیل قرار گرفت.

ت. یافته های تحقیق:

تحقیق حاضر نشان داد که: 1) از میان 6 گروه مختلف هماینها، همایند «فعل + حرف اضافه» مشکل سازترین نوع همایند برای زبان آموزان ایرانی به شمار می رود؛ 2) زبان آموزانی که در سطح بالاتری از توانش زبانی قرار دارند، در انتخاب صحیح همایندهای زبان انگلیسی با دشواری کمتری مواجه می شوند؛ 3) زبان اول به میزان قابل توجهی (34/4 درصد) بر خطاهای نوشتاری زبان آموزان ایرانی در کاربرد همایندهای زبان انگلیسی تأثیر گذار است و 4) زبان آموزانی که در سطوح بالاتری از توانش زبانی قرار دارند، به همان نسبت از همایندهای بیشتری در انگلیسی نوشتاری خود استفاده می کنند.

ث. نتیجه گیری و پیشنهادات:

در هر زبان، بسیاری از واژگان تابع مجموعه ای از محدودیتهای معنایی می باشند که تعامل آزادانه آنان با یکدیگر در یک بافت زبانی مشخص را با محدودیت مواجه می سازد. از این رو، با توجه به اهمیت کاربرد صحیح هماینها در موفقیت زبانی زبان آموزان، یکی از راهکارهای مهمی که می توان در عرصه آموزش زبان دوم از آن بهره جست، همزمان ساختن آموزش یا یادگیری یک واژه با همایندهای خاص آن واژه است. تحقیق حاضر نشان می دهد که یکی از مشکلات شایع زبان آموزان ایرانی، استفاده از واژگان مترادف در بافتهای زبانی متفاوت است؛ در صورتی که به عنوان مثال، افعال و صفتهای مترادف بعضاً با مجموعه مختلفی از اسمها همراه می شوند.

افزون بر این، با توجه به نقش غیرقابل انکار زبان اول در کاربرد ناصحیح همایندهای زبان انگلیسی، انتقال عادات زبانی از زبان اول به زبان دوم را نیز باید یکی از فاکتورهای مهم و اثرگذار در فرایند آموزش زبان دوم برشمرد. همچنین، ضعف مشهود زبان آموزان ایرانی در کاربرد صحیح همایندهای زبان انگلیسی و تأثیر بازدارنده آن بر مهارت های ارتباطی آنان با گویشوران بومی زبان ایجاب می کند که واژگان همایند به عنوان ضروری ترین قالب از «قالب های زبانی پیش ساخته» نظیر کلیشه ها و اصطلاحات زبانی در مواد درسی و کتب آموزشی مورد توجه قرار گیرند.

صحت اطلاعات مندرج در این فرم را بر اساس محتوای پایان نامه و ضوابط مندرج در فرم گواهی می نمایم.

نام استاد راهنما: دکتر فهیمه معرفت

سمت علمی: دانشیار

نام دانشکده: ادبیات و زبانهای خارجی

رئیس کتابخانه:

TO THE SOUL OF
MY
MOTHER

AND

THE CHALKY HANDS OF
MY
TEACHERS

ACKNOWLEDGMENTS

I am indebted to so many people that listing and thanking them all for the different ways they helped me is simply not possible in this limited space, but I thank each one of them in my heart. First and foremost, I am honored to thank Dr. Fahimeh Ma'refat, my supervisor, for her unflinching patience, constant encouragement, and invaluable comments during the preparation of this thesis. She has been extremely indefatigable, patient and supportive in all my years as a graduate student in Allameh Tabataba'i University and has enormously contributed to my understanding of EFL domain throughout her courses and in various classes with her effective teaching and constructive criticism. I shall not forget her name for the rest of my life.

I am deeply indebted to Dr. Mehdi Nowruzi Khiabani, my advisor, who kindly provided me with invaluable insight in his field of expertise and continuous encouragement. I also wish to thank him for his informative article on the sources of collocational clashes between Persian and English, and his helpful comments on an earlier draft of this paper.

My most sincere gratitude also goes to Dr. Gholamreza Tajvidi, a man of noble character with a heart of gold, who kindly accepted the responsibility for reading my thesis and made undeniably helpful revisions to the final draft of the paper.

I have learnt about English Language Teaching from many professors, all of whom are devoted to academic excellence and the promotion of the students' knowledge. I would like to appreciate Dr. Parviz Birjandi, Dr. Zia Tajuddin, Dr. Mansour Fahim and Dr. Mohammad Khatib all of whom, in one way or another, broadened my knowledge of teaching methodology and linguistics.

I can not thank my parents especially my late mother, Shahrbanoo, whose love for learning was always inspiring. Had not been for them, I could not have finished my master's degree. I am also very fortunate in having siblings like Katayoon, Homayoon and Muhammad, and nephews like Farhad who have continuously comforted me during our phone calls when I was away from home and have actively encouraged me to carry out my studies.

I owe special gratitude to Mrs. Arezoo Asgarifar, a wonderful spouse, who has provided me with abundant emotional support since the first moment I met her.

Last but not least, I greatly appreciate the help of all English language learners, among many others, who provided me with needed data.

ABSTRACT

This thesis reports on an exploratory study that analyses the use of different types of collocations by Persian-speaking learners of English in order to provide answer to four research questions regarding the degree of difficulty the different types of collocations cause for Iranian EFL learners in relation to their language proficiency levels, and the negative transfer of L₁ collocational patterns to L₂ setting. For this purpose, an attempt was first made to define collocations and to dwell on their significance in language learning. Then, the types of errors that the learners made when producing collocations were identified and analyzed as to their source and frequency of occurrence among Iranian EFL learners.

The data for this study came from the written productions of 72 adult Iranian EFL learners with the female-male ratio being slightly in favour of males (37 vis-à-vis 35). At the time of the study, all the participants had been learning English under formal instruction for at least four years at university or two year in language learning institutes. The age of the participants ranged from 21 to 29.

The analysis of the corpus revealed that prepositional collocations in general and verb-preposition collocations in particular constitute the most challenging type of collocation for Iranian learners of English. The results also indicated that collocations do not present as much difficulty for high proficient learners as they do for less proficient ones. Moreover, each of the 6 collocational categories (i.e. verb-preposition, noun-preposition, adjective-noun, adverb-adjective, verb-adverb and verb-noun) presented different degrees of difficulty for EFL learners of different proficiency levels. Verb-preposition collocations, for example, were the most problematic category for low proficient learners, while they were the least difficult one for high proficient learners. Instead, what posed the greatest difficulty for Iranian advanced learners of English was the choice of appropriate adjective to modify the nouns.

As far as the sources of errors were concerned, 34.4 % of errors were due to first language interference that goes far beyond what earlier (small-scale) studies have predicted. It also turned out that the less proficient learners tend more strongly towards isolated words rather than multi-word units, whereas the more proficient learners prefer to use more word combinations and ready-made constructions as a way to achieve greater fluency in the second language.

The study ends with the discussion of the results and their pedagogical implications for foreign language learning in Iran, most importantly the role of L₁-L₂ differences and that of language teachers in highlighting the collocational restrictions of the words. The research reported in this study reveals how corpus-based analysis can shed light on the nature of collocations and the extent to which the foreign language learners are aware of the collocational properties of the words.

The Iranian EFL learners' rather unsatisfactory performance in different types of collocations implies a general unawareness of the semantic range and selectional restrictions imposed on English lexicon. This problem may stem from their habit of learning English vocabulary as isolated words rather than in combination with other words. Thus, the teaching of collocations inevitably needs to be integrated with the teaching of vocabulary, and the foreign language learners need to know that not every word combination that is easily understood can be as easily produced in English, nor in any other language. In addition, given the empirical evidence that first language transfer is a real phenomenon that also affects production of collocations in the second language, the selection of collocations for teaching as well as the way a language teacher teaches them should be with reference to L₁.

TABLE OF CONTENTS

Page

Dedication	I
Acknowledgment	II
Abstract	III
List of tables	VII
List of figures	VII

Chapter 1:

Introduction

1.1 Background	1
1.2 Statement of the Problem	2
1.3 Significance of the Study	3
1.4 Research Questions	5
1.5 Limitations of the Study	5
1.6 Definition of Key Terms	5

Chapter 2:

Review of the Related Literature

2.1 Importance of Collocations	7
2.2 Definition and Types of Collocations	12
2.2.1 Definition	12
2.2.2 Types	15
2.3 The Place of Vocabulary in Language Teaching	18
2.4 Collocations in Generative Linguistics	22
2.5 Empirical Studies	27
2.5.1 Collocations and Language Proficiency	27
2.5.2 Collocations and Language Transfer	29
2.5.3 Iranian EFL Learners' Problems with Collocations	30

Chapter 3:

Methodology

3.1 The Research Questions	34
3.2 Participants	34
3.3 Materials and Instrumentation	
3.3.1 Learners' Corpus	35
3.3.2 Dictionaries of Collocations and BNC	35
3.3.3 Michigan Test of Language Proficiency	36

3.4 Data Collection and Analysis Procedures	
3.4.1 Measuring Language Proficiency	36
3.4.2 Data Collection	37
3.4.3 Data Analysis	38

Chapter 4:

Results and Discussion

4.1 Analysis of EFL Learners' Collocational Competence	40
4.1.1 Collocational Errors Made by Iranian EFL Learners	41
4.1.2 Collocational Errors across Different Language Proficiency	44
4.2 The sources of Errors in Learner-made Collocations	48
4.2.1 Incorrect Collocations due to Interlingual Transfer	48
4.2.2 Incorrect Collocations due to Overgeneralization	51
4.2.3 Incorrect Collocations due to Collocation Structure	52

Chapter 5:

Conclusion and Educational Implications

5.1 Summary of Research Findings	54
5.2 Pedagogical Implications	57
Suggestions for Further Research	65

References

Appendixes

Appendix I: List of Collocations Produced by Adult Iranian EFL Learners

Appendix II: List of Collocational Errors by Adult Iranian EFL Learners

LIST OF TABLES

Table 4-1: Syntactic Patterns of Collocations	41
Table 4-2: Frequency and Percentage of Collocational Errors	42
Table 4-3: Types of Errors in Collocations	43
Table 4-4: Collocational Errors by Learners of Different Proficiency Levels	44
Table 4-5: Adverbial Collocations across Different Proficiency Levels	45
Table 4-6: Interlingual Collocational Errors across Different Proficiency Levels	48
Table 4-7: Intralingual Collocational Errors across Different Proficiency Levels	52
Table 4-8: Collocational Errors due to Unfamiliarity with Collocation Structure	53

LIST OF FIGURES

Figure 4-1: Distribution of Collocations among Different Proficiency Levels	46
Figure 4-2: Errors of Different Collocation Types in Different Proficiency Levels	47

CHAPTER 1

INTRODUCTION

1.1 Background:

For decades teaching professionals were constantly preoccupied by how to segment the language into pieces and specify what learners actually learn when they learn a language. Pure and applied linguists often used to divide the language into such traditional units as sounds, words, and rules of grammar and discourse. But since the 1930s, a number of prominent figures have come to advise those involved in the profession against this classification and urged English teachers to recognize that there are particular words which tend to occur in the company of other words and that achieving language fluency depends on learning to use these word groups. In 1933, Palmer (as cited in Kennedy, 2003) was the first to adopt the term *collocation* for these recurring groups of words.

However, despite the fact that collocations (i.e. word combinations such as *fragile peace*, *strong economy*, *bitter disappointment*, *fast colour*, *harsh criticism* or *mass destruction*) are an important part of a native speaker competence, and the need for their inclusion in foreign and second language teaching is widely acknowledged today (see for example Hussein, 1990; Bahns & Eldaw, 1993; Farghal & Obiedat, 1995; Hills, 1999; Nesselhauf, 2003), a considerable amount of efforts in applied linguistics research are still being concentrated on the grammatical and phonological levels, while the lexical level does not arouse the same degree of interest and very little attention is paid to an amazing property of lexis, i.e. a tendency on the part of some lexical items to co-occur, otherwise known as collocation.

In fact, until relatively recently, the true complexity of collocations was largely hidden. Over the past two decades, however, huge developments in computerized corpus linguistics and the availability of wide collections of text in electronic form has offered new insights into how words or groups of words are distributed in a given language. Sophisticated software for the analysis of first or second language corpora has allowed researchers and L₂ professionals to explore more deeply the nature of collocations and in doing so, according to Kennedy, “challenge syntax-based approaches to language description and pedagogy.” (2003, p. 469).

Wray (2002) rightly states that collocations are of particular importance for learners who wish to achieve a high degree of competence in the second language, but they are also of some importance for learners with less ambitious aspirations, as they not only enhance accuracy but also fluency. However, although some suggestions on the teaching of collocations have been made in recent years, it is largely unclear how and specially which of the great number of collocations in a language should be taught.

To answer these questions satisfactorily, it is doubtless essential to identify the problems that the EFL learners have in dealing with collocations. The present study aims at scrutinizing the issue of collocations as an important, yet largely unresearched area of linguistic competence, in Iranian EFL settings.

1.2 Statement of the Problem:

The EFL learners' language is characterized by linguistically incorrect and/or contextually inappropriate forms and expressions. Both types of deviations are labelled 'errors' when they result from a lack of competence in the language. But in addition to linguistic and pragmatic deviations, the learners' interlanguage also exhibit certain forms that are linguistically and pragmatically correct but still sound 'unnatural' or 'strange' to a native speaker. (Emery, 1987; as cited in Mahmoud, 2005). It is usually the case that even advanced EFL learners have various problems with collocations in their oral or written productions. (Bahns & Eldaw, 1993; Taiwo, 2004; William, 2000).

In terms of having difficulty with recurring groups of words, Iranian learners of English are no exception. Even after years of foreign language study at universities or

language learning institutes, the majority of Iranian EFL learners still make collocational errors and this is in spite of the fact that they seem to have sufficient lexical or grammatical knowledge. Such erroneous expressions like *hard question*, *heavy tea*, *fixed colour*, *to take fish* and *to interview with*, just to name a few, are not due to poor lexical or grammatical knowledge. These problems, as Koosha & Jafarpour (2006) have rightly referred to, arise largely from lack of collocational knowledge among Iranian EFL learners, the inadequate emphasis given to collocational patterns in their textbooks, and the type of instructions they receive. Moreover, the compounds that words keep have not usually been a focus of teaching and research in our country and language teachers are more accustomed to providing definitions rather than collocations in their classroom presentations.

According to Thomas (1984, p.187), such production difficulties among EFL learners are “hardly surprising, given the vast scope and very idiosyncratic use of lexical items and collocations.” One level of difficulty with collocations is the apparent lack of rules. In the case of grammar, once we know the rules we can make new, correct sentences all the time, simply by slotting in new vocabulary in the right places. But with collocations, there are limits to the changes we can make to the combination of words and more importantly, there are no hard and fast rules students can turn to. Actually, according to Prodromou (2003, p. 24), “the only rule is that there are -virtually- no rules.”

1.3 Significance of the Study:

It has recently been suggested that for the purpose of English as an international lingua franca, "idiomaticity" (i.e. use of such prefabricated forms as idioms, collocations, proverbs, catchphrases and clichés) places obstacles in the way of communication between native and non-native speakers. (Seidlhofer, 2001; as cited in Prodromou, 2003) Sinclair (1991) – the first to use corpus-based evidence - also considers idiomaticity, in all its varieties, the most important factor in reaching 'native-like fluency'. However, although it is generally accepted that collocations are both indispensable and at the same time problematic for foreign language learners and they therefore should be guaranteed a proper place in second language acquisition, especially for adult learners, learners' difficulties with collocations have not been investigated in detail by EFL practitioners so far.

In fact, while the field of language teaching and learning is relatively rich in studies of foreign language learners' linguistic and pragmatic errors, research on 'strangeness' of linguistic forms and expressions seems to be lagging behind and systematic and in-depth analyses of Iranian EFL learners' lexical errors to determine if they produce 'unnatural' word combinations are relatively rare. The need for such studies stems from the fact that although post-intermediate and advanced learners of EFL may not face structural and pragmatic problems, their language sounds 'odd' in the eyes of native speakers of English. Therefore, studies are needed such as that reported by Mahmoud (2005) where Arab university students' collocation errors were analysed.

The purpose of the present study was to collect, classify and analyse the collocational errors Iranian EFL learners made when producing collocations, to find out if knowledge of collocations could differentiate between different levels of EFL learners' proficiency (i.e. low, mid and high) and to determine the extent to which Iranian EFL learners' collocational competence is affected by their L₁, thus adding one more ring to the still short chain of studies in the area of lexis in general and the area of collocations in particular. An analysis of collocational errors can reveal the problems that EFL learners make and the causes of these problems in that area and help teachers, syllabus designers and EFL specialists find appropriate ways of dealing with them in the EFL courses.

The important role that collocations play in the successful and native-like performance of EFL learners on the one hand and the problems that Iranian EFL students have with collocations of various types on the other, highlight the significance of the present study. Furthermore, there is an abundant stock of word combinations in English that represent innumerable collocations, and the mastery over them can strongly affect Iranian EFL learners' fluency as well as accuracy in both speaking and writing.

Based on the findings of this study, some suggestions can also be made on how to teach collocations. First, if the use of collocations comes to be highly correlated with EFL learners' language proficiency, collocations should be considered as an important factor in determining the overall proficiency of English students. And second, if EFL learners' L₁ comes to be highly influential in the production of collocations, then not only should the selection of collocations but also their teaching should be with reference to L₁.

1.4 Research Questions:

This study sought to find answers to the following questions:

1. What are the most problematic types of collocations for Iranian EFL learners? In other words, are word combinations of all types (i.e. verb-preposition, noun-preposition, adjective-noun, adverb-adjective, verb-adverb and verb-noun) equally difficult for Persian-speaking learners of English?
2. Do collocations exert the same degree of difficulty for different levels of language proficiency among Iranian EFL students?
3. To what extent is the use of collocations affected by Iranian EFL learners' L₁?
4. Is there more idiomaticity (particularly the use of collocations) in the language of more proficient EFL learners?

1.5 Limitations of the Study:

First, this study was exploratory to a large degree and, due to the lack of previous studies in the field, the methodology could not be free of faults. Second, in this study the analysis of learners' collocational errors was only based on their written productions. Thus, although judging learners' collocational competence by their written productions has the advantages of ascertaining their true mastery over the linguistic collocations and decreasing the possibility of errors caused by factors other than their lack of competence (e.g. distraction), it has the disadvantages of failing to weigh their “comprehension” of collocations in context and also not reflecting their mastery over the collocations in verbal communication.

1.6 Definition of the Key Terms:

Collocations:

Collocations are not easily defined. In the linguistic literature, they are often discussed in contrast with free word combinations at one extreme and idiomatic expressions at the other, collocations occurring somewhere in the middle of this spectrum. Nonetheless, an agreed-

upon definition, according to Williams (2002), is that “collocation is a lexical unit consisting of a cluster of two or three words from the same or different parts of speech.”

Longman Dictionary of Language Teaching and Applied Linguistics (1992, p.62) defines collocation as “the way in which words are used together regularly.” Based on this definition, collocation refers to the restrictions on how words can be used together, for example which verbs and nouns are used together, or which adverbs are used with particular verbs or adjectives. For example, in English the verb *do* collocates with *damage*, *duty*, and *wrong*, but not with *trouble*, *noise*, and *excuse*. Similarly, *high* collocates with *probability* but not with *chance*. We say a *high probability* but a *good chance*. Further elaboration on the definition of collocation will appear in chapter 2.

Corpus linguistics:

Corpus linguistics is simply the study of language through corpus-based research, but it differs from traditional linguistics in its insistence on the systematic study of authentic examples of language in use. (Sinclair, 1997)

Error analysis:

Error analysis is a branch of applied linguistics whose primary focus is on the evidence that Language learners' errors provide with an emphasis on the underlying process of foreign language acquisition. Keshavarz (1997) suggests that the field of error analysis can be divided into two branches: theoretical and applied.

Theoretical analysis of errors concerns the process and strategies of language learning and its similarities with first language acquisition. It also tries to decode the strategies adopted by the learners such as overgeneralization and simplification and reach a more general conclusion that concerns universals of language learning process.

Applied error analysis, on the other hand, concerns organizing remedial courses and developing appropriate materials and teaching strategies based on the findings of the theoretical error analysis.

Corpus analysis:

Longman Dictionary of Language Teaching and Applied Linguistics (1992, p.88) defines corpus analysis as “the analysis of a collection of materials that has been made for a particular purpose, such as a set of textbooks or a sample of sentences or utterances, for their linguistic features.”

CHAPTER 2

REVIEW OF RELATED LITERATURE

2.1 Importance of Collocations:

In the past, vocabulary in general and prefabricated forms such as collocations, idioms and clichés in particular were often given little priority in second language teaching profession, but recently there has been a renewed interest in the nature of vocabulary and its role in learning and teaching. Traditionally, vocabulary learning was often left to look after itself and received only incidental attention in many language learning textbooks and language teaching programs. Thus, although the course curriculum was often quite specific about different aspects of teaching such as grammar, pronunciation, speaking, reading or writing, very little specification was given to the role of vocabulary in promoting communication among non-native or native and non-native speakers. However, the status of vocabulary now seems to be changing.

For one thing, the notion of a word has been broadened to include lexical phrases and routines, and it has been suggested that in the initial stages of learning these play a primary role in communication and acquisition. In addition, access to lexical corpora has made it possible for applied linguists to access huge samples of language in order to find out how words are used both by native speakers and by foreign language learners. Such studies has enabled instructors and applied linguists to identify common patterns of collocations and other lexical phrases that are inseparable from a native speaker's lexical competence.

Richards and Renandya (2002) consider vocabulary a core component of language proficiency which provides much of the basis for how well learners speak, listen, read and