



Western Washington University
Western CEDAR

The Edna L. Sterling Collection

Western Libraries and the Learning Commons

11-21-2005

The Lexical Approach: Collocation in High School English Language Learners

Claire Smith
clairsmith@aol.com

Follow this and additional works at: https://cedar.wwu.edu/sterling_collection



Part of the [Bilingual, Multilingual, and Multicultural Education Commons](#)

Recommended Citation

Smith, Claire, "The Lexical Approach: Collocation in High School English Language Learners" (2005). *The Edna L. Sterling Collection*.
1.
https://cedar.wwu.edu/sterling_collection/1

This Book is brought to you for free and open access by the Western Libraries and the Learning Commons at Western CEDAR. It has been accepted for inclusion in The Edna L. Sterling Collection by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.



GIFT OF
THE RACHEL ROYSTON
PERMANENT SCHOLARSHIP
FOUNDATION
OF THE
DELTA KAPPA GAMMA SOCIETY

This is an authorized facsimile, made from the microfilm master copy of the original dissertation or master thesis published by ProQuest Information and Learning Company's UMI® Dissertation Services.

The bibliographic information for this thesis is contained in the UMI® Dissertation Abstracts database, the only central source for accessing almost every doctoral dissertation accepted in North America since 1861.

UMI® Dissertation
Services

777 Eisenhower Parkway
P.O. Box 1346
Ann Arbor, Michigan 48106-1346 USA
800.521.0600 734.761.4700
web www.il.proquest.com



Printed in 2006 by digital xerographic process
on acid-free paper

THE LEXICAL APPROACH:

COLLOCATION IN HIGH SCHOOL ENGLISH LANGUAGE LEARNERS

by

CLAIRES SMITH

Presented to Educational Foundations and Leadership Department
and the Graduate School of George Fox University
in partial fulfillment of the requirements
for the degree of
Doctor of Education

11/21/2005

UMI Number: 3195526

Copyright 2006 by
Smith, Claire

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.



UMI Microform 3195526

Copyright 2006 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

George Fox University
School of Education
Newberg, Oregon

"The Lexical Approach: Collocation in high school English language learners,"
a doctoral dissertation prepared by Claire Smith in partial fulfillment of the
requirements for the Doctor of Education degree in the Educational Foundations
and Leadership department. This dissertation has been approved and accepted by:

<u>11-21-05</u>		Committee Chair
Date	Katsu Ozawa, Ph.D.	Associate Professor of Education
<u>11-22-05</u>		
Date	Gary L. Railsback, Ph.D.	Associate Professor of Education
<u>11/22/05</u>		
Date	Scot Headley, Ph.D.	Professor of Education
<u>11/30/05</u>		
Date	Robin Baker, Ph.D.	Provost

Abstract

Collocation involves knowing which words typically co-occur in fluent, native-like language. Studies across several settings have shown that collocation skills lag behind other aspects of second language learning. The present study looks at collocation skills in a linguistically diverse immigrant high school population in the USA, correlating scores on a collocation test (based on Bonk, 2000) with length of residence in the USA. The correlation was moderate for the group as a whole, but close to zero for the 19 – 36 month residence bracket. The second half of the study compares the collocation skills and the approaches to collocation of Russian- and Spanish-speaking students. Although they did not differ in overall level of performance, there was evidence of different approaches by the two language groups to collocation.

TABLE OF CONTENTS

CHAPTER 1.....	1
Introduction	1
Statement of the problem:	2
Part 1: the large group study.....	3
Research question.....	3
Hypothesis.....	3
Part 2: the small group quantitative study.....	3
Research question.....	3
Hypothesis 2a.....	4
Hypothesis 2b	4
Hypothesis 2c.....	4
Hypothesis 2d	5
Part 3: the small group qualitative study.....	5
Research question.....	5
Hypothesis 3a.....	5
Hypothesis 3b	5
Hypothesis 3c.....	5
Hypothesis 3d	6
Definition of terms.....	6
Limitations and delimitations	7
CHAPTER 2.....	9
Review of the literature.....	9
Early Approaches to Foreign Language Pedagogy	9
Grammar-and-Translation.....	9
Naturalistic.....	10
Communicative.....	11
Recent developments in linguistics	12
The Lexical Approach of Michael Lewis	14
Key aspects of the approach	14
Classroom applications.....	17
Evaluations of Lewis.....	19
Collocation.....	21
Definitions and categorization.....	21
Pedagogical aspects.....	25
Measuring collocation.....	28
L2 collocation performance	29
Development of L2 Collocation.....	33
Approaches to collocation.....	34
Conclusions from the literature review	36

CHAPTER 3.....	38
Methods	38
Research questions	38
Setting	39
Participants	40
The large group.....	40
The small group.....	41
Instrumentation	42
Procedure	42
Scoring and Analysis	44
Part 1: the large group study.....	44
Part 2: the small group quantitative study.....	45
Part 3: the small group qualitative study.....	45
CHAPTER 4.....	47
Results	47
Collocation Test Scores.....	47
Part 1: Correlation between Collocation Test scores and LOR.....	50
Part 2: Comparison of subgroup means.....	54
Part 3: Qualitative analysis of Section A.....	56
Breakdown of errors.....	56
Blank responses	58
L1 transfer.....	60
Use of basic verbs	60
Use of paraphrasing.....	62
Summary of findings.....	62
CHAPTER 5.....	64
DISCUSSION	64
The Large Group.....	64
The Small Group	67
Theoretical Implications.....	73
Contributions to the Field.....	73
Future Research	74
Classroom Applications	75
References	77
APPENDIX A COLLOCATION TEST.....	83
APPENDIX B ANSWER KEY	87
APPENDIX C LANGUAGE BACKGROUNDS.....	88

APPENDIX D LENGTH OF RESIDENCE.....	89
APPENDIX E DEMOGRAPHIC SHEET.....	90
APPENDIX F ADMINISTRATION INSTRUCTIONS	91
APPENDIX G COLLOCATIONS: RUSSIAN & SPANISH EQUIVALENTS	92
APPENDIX H PARAPHRASED ANSWERS BY LANGUAGE GROUP	94
FIGURE 4.1 <i>LOR vs Collocation Test Scores.....</i>	51
FIGURE 4.1 <i>LOR vs Collocation Test Scores for Low/Mid/High LOR.....</i>	53
TABLE 4.1 <i>Collocation Test scores.....</i>	49
TABLE 4.2 <i>Correlation Coefficients between LOR and Coll. Test scores.....</i>	52
TABLE 4.3 <i>Comparison of means for Russian- vs Spanish-speaking students....</i>	55
TABLE 4.4 <i>Section A errors by type and language group.....</i>	57
TABLE 4.5 <i>Number and location of blank errors.....</i>	59
TABLE 4.6 <i>Incorrect use of basic verbs by language group.....</i>	61

Chapter 1

Introduction

In the early 1990s, Lewis (1993) revitalized second- and foreign-language teaching with the publication of the seminal text: *The Lexical Approach*. Language was now to be seen as a collection, not of words, but of linguistic “chunks.” The fluent speaker was one who was aware of linguistic “chunks,” who had vast numbers of chunks stored in memory, and who could retrieve and combine them quickly and effectively. Teachers were to move away from the “Present – Practice – Produce” model of language instruction, “in favor of a paradigm based on the Observe – Hypothesize – Experiment cycle” (Lewis, 1993, p.vii).

Lewis presents several fundamentals of the lexical approach. One of them, collocation, is the subject of this investigation. Collocation involves knowing typical and frequent word combinations. Claims for non-linear development of collocation skill are examined for an immigrant, English Language Learner (ELL) high school population. In addition, the study compares the results of a subgroup of Russian-speakers versus Spanish-speakers in collocation skill and approach to collocation.

Lewis’ work is influential in the field of adult English as a Foreign Language (EFL). It is also gaining in popularity as an approach in the field of English as a Second Language (ESL) for both adults and school-aged populations. His ideas are still very new in the United States. I am suggesting here that the lexical approach, especially the teaching of collocation, has a great deal to offer American teachers teaching additional languages.

Lewis (2000) suggests that collocation promises a lot: a path from intermediate to

advanced levels of English language learning, native-like language production, a continual deepening of existing knowledge. Lewis shares much with constructivist learning approaches (as Bruner, 1990), where students build their own understanding of how the language works, rather than just passively receive it. He offers a sound and workable language pedagogy. At a time when there is pressure to mainstream ELL students quickly and effectively, the Lexical Approach, and collocation in particular, may be a vital component in the linguistic and academic assimilation of the nation's ELL children and youth.

Statement of the problem:

The current study examines whether collocation skill, a major component of the Lexical Approach, develops over time for immigrant ELL students of high school age. "Foreign-sounding" combination errors of vocabulary or lexis have been discussed extensively in the literature (Nattinger & DeCarrico, 1992; Fargal & Obiedat, 1995), and studied in post-secondary and EFL populations (Bonk, 2000; Gitsaki, 1999), but it is not known whether collocation is also problematical for school-age, immigrant, ELL learners.

This study also investigates whether students from different language backgrounds perform differently as regards collocation skill. Most of the experimental work in the field has focused on groups with the same first language, making comparisons across language groups difficult.

The present study is in three parts. The first part involves a large group of 143 high school, immigrant ELL students. Using the Collocation Test (based on Bonk, 2000) I measure collocation skill and correlate this with length of residence in months, to check for

continuity in development. A copy of the test is in Appendix A, and the answer key in Appendix B.

The second and third parts of the study comprise an in-depth analysis of a random, stratified subgroup of 30 students. The subgroup was selected from the pool of intermediates (ELL students who have attended American schools for at least 18 months or six complete quarters) whose stated language background is Spanish or Russian. For each of these language groups, 15 intermediate students were randomly selected.

In the second part of the investigation, I look for differences in Collocation Test scores between the two language groups. In the third part of the study, I look at the approach to collocation by the two language groups, by analyzing the type of collocation error (random, synonym or syntactic, or no response) made in section A of the test.

Part 1: the large group study

Research question.

The research question in Part 1 involves the large group of 143 students. Is there an increase in measured collocation skill over time, such that greater length of residence (LOR) in the target culture results in higher Collocation Test scores?

Hypothesis.

The researcher hypothesizes that collocation skill, as measured by the Collocation Test, will correlate highly and positively with length of residence for students in the large group.

Part 2: the small group quantitative study

Research question

The research question in Part 2 involves the small subgroup of 30 intermediate (LOR > 18months) students: 15 Spanish-speakers, and 15 Russian-speakers. Is there a difference in the total number of collocation errors made by intermediate students of the two language backgrounds? Is there a difference in the number of errors made by intermediate students of the two language backgrounds in section A (production of *noun + verb* collocations)? Is there a difference in the number of errors made by intermediate students of the two language backgrounds in section B (production of *verb + preposition* collocations)? Is there a difference in the number of errors made by intermediate students of the two language backgrounds in section C (recognition of *verbal* collocations)?

Hypothesis 2a

The researcher hypothesizes that there will be no significant difference in the total test scores of the Russian-speaking students vs the Spanish-speaking students.

Hypothesis 2b

The researcher hypothesizes that there will be no significant difference in the scores on section A (production of *noun + verb* collocations) of the Russian-speaking students vs the Spanish-speaking students.

Hypothesis 2c

The researcher hypothesizes that there will be no significant difference in the scores on section B (production of *verb + preposition* collocations) of the Russian-speaking students vs the Spanish-speaking students.

Hypothesis 2d

The researcher hypothesizes that there will be no significant difference in the scores on section C (recognition of *verbal* collocations) of the Russian-speaking students vs the Spanish-speaking students.

Part 3: the small group qualitative study

Research question

The research question in part 3 involves the same subgroup as in the previous question. Research question 3 only considers responses to section A of the Collocation Test.

This part of the investigation looks at how Russian-speaking and Spanish-speaking students approach an unknown collocation. Do they attempt all items, or leave some unanswered? Do they answer based on what they would have said in the native language? Do they resort incorrectly to easy, basic words, such as the delexicalized verbs *make*, *do*, *have*, *get*, *be*, *put*? Do they paraphrase (use a multi-word response or a non-verb response)?

Hypothesis 3a

The researcher hypothesizes that there will be no difference in the amount of blank responses of the Russian-speaking students vs the Spanish-speaking students.

Hypothesis 3b

The researcher hypothesizes that there will be no difference in the amount of synonym errors showing evidence of transfer from the native language (NL) for the Russian-speaking students vs the Spanish-speaking students.

Hypothesis 3c

The researcher hypothesizes that there will be no difference in the amount of

incorrect responses involving *make*, *do*, *have*, *get*, *be*, *put* by the Russian-speaking students vs the Spanish-speaking students.

Hypothesis 3d

The researcher hypothesizes that there will be no difference in the use of paraphrasing (a multi-word response or a non-verb response) by the Russian-speaking students vs the Spanish-speaking students.

Definition of terms

The researcher is responsible for the definitions of collocation errors. Other terms are from Lewis (1993) or are in very high usage in the foreign/second language field, or in education in general.

- ACTFL: American Council on the Teaching of Foreign Languages (professional body of Foreign Language teachers)
- Collocation: tendency of two words, or collocates, to co-occur (Lewis, 1993)
- Collocation Skill: degree of competence in manipulating collocation (Lewis, 1993)
- Collocation error: pairing of two words that do not typically collocate in natural speech
 - Blank collocation error: include items left blank by the test taker.
 - Grammar collocation error: a response involving the correct collocation, but with an error in grammar or spelling
 - Random collocation error: a response that bears no relation, semantically or syntactically, to the correct collocation; this category does not include items left blank by the test taker.
 - Synonym collocation error: a response that is a close synonym to the correct

collocation, but is not included in the Collocation Test Answer Guide as native-like

- Constructivist Learning: learning in which the pupil builds his own knowledge and understanding (as Bruner, 1990)
- English as a Foreign Language (EFL): English classes taught to non-native speakers in a setting outside of an English speaking country
- English Language Learners (ELL): students with a native language other than English, who are learning English
- English as a Second Language (ESL): a program of English instruction for ELL students
- English for Speakers of Other Languages (ESOL): synonymous with ESL, an updated term. Also used as an adjective before 'teacher' or 'lesson', etc.
- Intermediate: an ELL student who has been enrolled at an American school for at least six complete quarters (definition is for the current investigation only)
- K – 12: Kindergarten through grade 12 education (in the USA)
- Native language (NL): the first language, or mother tongue
- Teachers of English to Speakers of Other Languages (TESOL): professional body of English as a Second Language teachers
- Target language (TL): the language a student is endeavoring to acquire.

Limitations and delimitations

The major limitation of this study is that there is no guarantee that collocational skill

will transfer to authentic speech situations beyond the current testing. The test format in the current study combines fill-in-the-blank items and multiple choice items; there are no free response items.

The investigation does not control for prior education in the home country. Chapter 5 includes discussion on this point.

The major delimitation of note in this study is that students were tested only once. The results therefore depict past learning, rather than growth in collocation skill over time. Only two language groups were represented in parts 2 and 3 of the investigation, limiting the amount of cross-cultural comparison possible. Also, the language groups were not necessarily culturally homogeneous: students with more than one prior language (e.g. Kazakh and Russian) were considered to belong to the more dominant language and culture.

Chapter 2

Review of the literature

I open this review with a look at the major approaches to teaching second and foreign languages. The most recent of these, the Lexical Approach (Lewis, 1993) I discuss in detail. Four key points are examined: Chunking, Collocation, Consciousness Raising, and Observe-Hypothesize-Experiment. I share classroom examples embodying a lexical approach and give evidence – both theoretical and experimental – that the approach is sound. Following this I turn to a discussion of collocation: definitions, classifications and importance. I conclude this section with a review of experimental work to date on collocation, which is shown to be sketchy and intuitive, and in need of more evidence.

Early Approaches to Foreign Language Pedagogy

Grammar-and-Translation.

An approach in education is a philosophical statement of WHY one teaches in a certain way. An approach is superordinate to syllabus (WHAT one teaches) and to methodology (HOW one teaches). Approaches in the field of second and foreign languages are based on findings in linguistics (the nature of language), psycholinguistics (the acquisition of language by the individual learner) or learning theory.

Until well into the 20th century, the predominant approach to the teaching of Foreign Languages was the classical Grammar-and-Translation approach (Richard-Amato, 2003). It is also known as the structural approach. Modern languages were treated in the same way as Classical Greek and Latin, where the major aim was to understand works of literature in the

original language. Languages were seen as composed of structures (the grammar), which needed to be mastered. Vocabulary was seen as slotting into the structures. Understanding, translating and using the syntactical structure led to supposed mastery of the language.

In the classroom, drill and translation took preference over teaching pupils to communicate in the target language. Delivery of instruction was usually in the form of Present-Practice-Perform. The curriculum was based around grammar, mostly around verb structures, and – in the case of English - progressed from supposedly “easier” present simple indicative tenses through the simple past and then onto more complex constructions with modals.

Current thinking sees two major problems with this approach. In the first place, grammar and vocabulary are not dichotomous, but far more interconnected than linguists originally believed. The second critique is that structural approaches deal with (grammatically) possible utterances, which are often a far cry from probable utterances by a native speaker.

Naturalistic.

A Naturalistic approach on the other hand aims to have the student learn the foreign language in the same way as he acquired his mother tongue. Second language acquisition theory looks at the developmental stages of learning to speak in the mother tongue, and adapts those to the second language situation. Methodologies range from a sink-or-swim situation, through various immersion programs, through to the child-centered natural approach (Krashen & Terrell, 1983).

The emphasis in the classroom is on acquisition of language, rather than learning.

This needs a language rich classroom environment, with optimal input, input that is just beyond his current level of English. It also needs a non-threatening environment, where mistakes are tolerated as part of the process. Noticeably absent is any stress on language study. The child dictates the pace. Listening and speaking are given priority in the early stages over the literacy skills.

This approach is the predominant approach within ESL for younger students. The natural approach blends well with the ethos of most primary grade classrooms. Although there are many admirable aspects to this approach, older students – especially those already literate in their own language - can benefit from specific language learning lessons. In ESL work with older children and adults, a lot is made of the potential transfer of knowledge and skills from the native language. Indeed, the demands of graduation, or of a job, may require a much faster pace than that afforded by the Natural Method.

Communicative.

In the 1970s, the Communicative Approach came into the forefront. (Richard-Amato, 2003). This approach saw language primarily as a means of transmitting and receiving meaning. Whereas the structural approach had stressed accuracy, the Communicative Approach stressed comprehensibility. The job of the student was to make himself understood in various situations.

The emphasis in the curriculum was on presenting the language that students were likely to need. The curriculum could be based on functions (e.g. greeting people, expressing regret, denying), notions (e.g. possession, frequency, intensity, and manner) - or academic content (e.g. the farm, light, houses). No longer did students have to wait for the second

year of the curriculum before they “were allowed” to use the past tense.

A communicative approach has now dominated the foreign language field for many years. Many new syllabi have been published, but some of these are communicative in name only; the thrust of the text is to cover a grammatical sequence. Content-based ESL and English for Special Purposes are both, in a special sense, communicative, inasmuch as they aim to give the students the language they need.

Recent developments in linguistics

In the 1980s and ‘90s there were several new developments in linguistics, which caused linguists to question the basis of language and language teachers to propose changes in what and how they taught. These are precursors to the new approach to language teaching popularized by Michael Lewis (1993).

The first of these was made possible by technological advances. Known as Corpus Linguistics (Stubbs, 2001), this branch of linguistics creates databases containing vast corpora (bodies) of authentic language. These databases generate concordance information, such as word frequency and word usage. One of the best known databases is the COBUILD Project at the University of Birmingham, UK (COBUILD, 1990). Most major languages now have their own corpora, and there are specialized corpora too. Corpus Linguistics has also provided the material for a new generation of dictionaries: Learner Dictionaries (e.g. Macmillan English Dictionary for Advanced Learners of American English, 2002), and Collocation Dictionaries (e.g. Oxford Collocations Dictionary for Students of English, 2002). Linguists have also used Corpora to explore literal and connotative meaning in texts, cultural keywords and loan words (Stubbs, 2001).

A second development was the emergence of pragmatics as a discipline within linguistics. The emphasis here is on the social and emotional context of the meaning. The learner may understand the words of a text or discourse, but does he understand the intent of the speaker, the evaluative implications, the relationship between the parties to the language?

A third development concerned the very basis of language. It had long been assumed that language, whether native or additional, was comprised of words held together by syntax, or rules of grammar. Gradually, linguists came to realize that words contain their own rules, especially rules of combination. Firth (1957) in his famous dictum referred to this as “the company [that] words keep”. Pawley & Snyder (1983) highlighted the difficulty of “native-like selection”: out of the many grammar-generated utterances that are possible in a given situation, how does the native speaker intuitively know which sounds right?. The foreign learner has to learn which is native-like, and while he is so learning, runs the risk of many foreign-sounding errors. Nattinger & de Carrico (1992) claimed that the mental lexicon comprised not just individual words but prefabricated, multi-word items. They claimed that the mature speaker had tens of thousands of these items stored in memory and ready for retrieval.

By the 1990s, the vocabulary-grammar dichotomy had become somewhat suspect, as linguistics began to realize that each word contained within itself its set of rules for use. Meaning of most individual words was limited, but it became sensible – and far richer – in the context of a short phrase. Word combinations had become the new focus (Nattinger & DeCarrico, 1992).

In 1993, Lewis published his seminal work: *The Lexical Approach*, in which he popularized this view of vocabulary as phrasal. The work also gave great prominence to authentic language and to meaning. The approach uses the best of the communicative and natural approaches, while meeting the major challenges of the structural approach head-on. In subsequent works, Lewis (1997; 2000) took the view from the theoretical and applied it to second and foreign language classrooms. Lewis has been instrumental in causing the profession to rethink its philosophy of language and language teaching.

The Lexical Approach of Michael Lewis

Key aspects of the approach.

In this section, I discuss four aspects of the Lexical Approach. The first of these is the notion of “chunking”. For Lewis, language is composed of “chunks” of words, and the fluent speaker is one who has a vast number of these chunks stored in memory, and can recall them as needed and combine them appropriately. The main classroom job is to get as many of these “ready-made” chunks into the learner’s long-term memory as possible. The aim of the language learning is to become so familiar with likely and probable combinations of and between chunks, that one can produce them effortlessly.

Lewis wants students to recognize and process chunks, and he stresses classroom activity that enhances this awareness on the student’s part. For Lewis, it is not important to know which category of lexical item one is dealing with. What does matter is that the student deal with it as a chunk, stored in the mental lexicon and ready for use. From an information processing perspective, this is a more efficient way to learn vocabulary (Thornbury, 2002).

A number of writers have suggested that stress and intonation patterns may be the basis of chunking ability (Marks, as cited in Williams, 2002). Lewis (1997) emphasizes that the Lexical Approach deals efficiently and economically with language, in that meaning, lexical phrases and intonation are dealt with together in the teaching and learning.

Chunking is particularly important at the upper levels of proficiency. Thornbury (2002) reinforces Lewis' views on this when he states: "The ability to deploy a wide range of lexical chunks both accurately and appropriately is probably what most distinguishes advanced learners from the intermediate ones" (p.116).

The second aspect of the Lexical Approach to be discussed is that of collocations. Collocations are one type of lexical item. In addition to single words (the traditional mainstay of vocabulary teaching), polywords (*more or less, and so on,*) and institutionalized expressions (*Good morning, No thank you, Please pass the _____*), the Lexical Approach emphasizes collocations. Collocations are conventional combinations of words, such as *abdicate + throne; suspicious + of; rancid + butter*. In addition to being the largest category of lexical items, collocations are the backbone of Lewis' approach. I will deal with this aspect of the Lexical Approach in greater detail in a later section.

The third key aspect in the Lexical Approach is Consciousness Raising. Lewis uses the term to mean the deliberate effort on the part of the teacher to make the student aware of how the target language works. The student should be led to notice both form and function of a lexical item. This may include pointing out language patterns and irregularities, drawing contrasts with the student's native language, letting the student deduce rules. For Lewis, awareness of chunks is paramount. Lewis also claims that becoming more aware of

the language increases the probability that the input of lexis will become intake, in other words, that learning will occur.

This aspect is in line with philosophy in K – 12 classrooms regarding metacognition and self-regulation of learning. However, it is not clear how or why heightened linguistic awareness works. Does it provide extra scaffolding for the acquisition of new vocabulary? Does heightened linguistic awareness forge a link between the stronger native language (NL) and the weaker target language (TL)? Or does it work simply because there is extra attention paid to those items in the classroom?

The fourth aspect of the Lexical Approach is Lewis' new paradigm for foreign language teaching. Traditionally there is the old "Present-Practice-Perform" model: the teacher presents an aspect of the target language (usually grammar rules), the students practice this, and after a set amount of learning time the student has to perform or produce the learning (usually on a test). Lewis proposes the "Observe-Hypothesize-Experiment" paradigm. The student first observes or notices an aspect of the language, then comes up with a 'hunch' as to how he might communicate in a given situation, which he then 'tries out' on a more advanced speaker of the language. The cycle repeats itself, as he then observes whether his communication was successful. The paradigm has the student build his own knowledge of the foreign language, rather than having it prepackaged for him (Lewis, 1993). It is a constructivist approach to learning, and is in line with much current thinking in US schools. Such an approach involves greater retention of the material, possibly because there is buy-in on the part of the student.

Classroom applications.

A lexical curriculum would be based on the most frequently occurring chunks, ideally based on evidence from corpora. However, Lewis states emphatically that there is no need to throw out old text books. It is rather a matter of teachers adjusting their thinking about language, and adding a few new ideas to one's teaching repertoire. Any textbook can be turned into a lexical one, if lexis becomes central, if the teacher concentrates on probable versus possible word combinations, and on getting a 'feel' for how the language works..

Lewis does not discard grammar. An extreme lexical view sees every word as having its own grammar. Lewis (2000) considers chunks the raw data by which learners perceive 'grammatical' patterns. Lewis (1993) has some novel ideas as to what should be taught as 'grammar': rather than the usual sequence of verbs (and he is particularly scathing about third conditionals and the ending -s in Third Person Singular present tense verbs!), he wishes to see the study of word formation, especially bound morphemes, of auxiliary verbs, of opposition and negation, the pragmatic nature of utterances and supra-sentential linking.

The literature is replete with suggestions on how to implement the Lexical Approach in the classroom (Lewis, 1997; Thornbury, 2002). The prime requirement is the centrality of lexis in the curriculum, - and this is as much a factor of teacher attitude as it is of classroom materials. A major classroom tool is the Lexical Notebook: a personalized learning record for the individual student. The notebook is far more than a list of new words: the words (or better, phrases) should be organized and annotated in a way that aids learning and retrieval by the student. Use of authentic language, especially language obtained from a concordance or other primary material, is encouraged. Activities to

enhance student skill with lexis are varied. Cloze passages feature prominently, and then there are collocation grids, matching exercises, training in dictionary work, collocation maps, and collocation dominoes. Thornbury (2002) also suggests activities such as analyzing the names of popular films, books and pop music groups, as many such names are in fact common collocates. Projects and task-based learning are encouraged by other workers (Willis & Willis, 1996) as they require the production of authentic language. Bliss (2004) includes an excellent suggestion whereby students have to listen and look out for collocates of a key word for a period of a week.

What makes an activity lexical? Lewis turns an ordinary vocabulary lesson into a lexical one by using authentic (not textbook) language, by having students focus on the language chunks, and by having them notice what does or does not co-occur with an item. His concern is with getting to know lexical items ‘more intimately’ and with working to deepen the mental lexicon. Lewis proposes a lot of language analysis-type exercises and activities, with predominance given to spoken language over written. Although many of the activities can be done with a variety of ages and levels, Lewis states that the beginning classes should be handled differently to intermediate and advanced classes. The principle goal in the early months is the intake by students of as much vocabulary as possible. At later levels, the emphasis is on deepening existing knowledge. One aspect of the deepening is to get to know a word’s collocates. Lewis claims that a lack of this is a major factor in the “intermediate plateau”, which students at an intermediate level have great difficulty passing.

Lewis’ work has been very influential, especially in the field of adult EFL. More recent editions of text books and dictionaries have taken his principles into consideration.

Lexical lessons, especially on collocation, appear markedly in on-line professional development literature for teachers. *onestopenglish* (Macmillan, 2005) and *Humanising Language Teaching* (Pilgrims, 2005) are two such websites containing frequent references to collocation.

Evaluations of Lewis.

Many writers in the field have provided support for Lewis and his approach. A lot of the support is in the form of expert opinion, rather than being based on experimental evidence. According to Kavaliauskienė & Janulevičienė (2001) and Davis & Kryszewska, (2003), the Lexical Approach allows the learner acquires accurate, *natural English*. It is also supposed to speed up the learning process (Davis & Kryszewska, 2003; Thornbury, 2002). Claims have also been made for higher EFL test scores under the Lexical Approach (Davis & Kryszewska, 2003; Thornbury, 2002). Williams (2002) cites improved stress and intonation when vocabulary is presented and learned in chunks.

Further theoretical support for Lewis comes from the area of Language Acquisition, specifically L1 acquisition. The young child first acquires and then produces language as undifferentiated wholes ("chunks") and only later analyzes the language, such that new, creative utterances become possible. Krashen & Terrell (1983) believe second languages are also learned in this manner. Although he and Lewis part company over the issue of learning versus acquisition, there is sufficient in common to provide credibility for Lewis.

The advent of corpus linguistics has also lent support to Lewis. Corpora provide language researchers with a wealth of language as it is actually used. Maley, in the introduction to Morgan & Rinvolucri (2004), states:

Current thinking, based in large part on the analysis of computer corpora, has emphasized the importance of collocation and therefore the fact that vocabulary is largely phrasal. Words hang together in typical clusters rather than exist in splendid isolation. (p.3)

Additional support for Lewis comes from work in the field of linguistic awareness, which has much in common with Lewis' Consciousness Raising. Helping students become aware of how language functions, especially the comparison across two or more languages, has become part of many syllabi. The American Council of Teachers of Foreign Languages has included linguistic awareness (4.2 Comparing (TL and NL) languages) as part of their proficiency guidelines (ACTFL, 1986). Experimental work has shown that linguistic awareness is an additional factor in second language acquisition. Hawkins (1999) points out differences in levels of Language Awareness in pupils, not related to intelligence, but related to the degree of sophistication of the mother tongue. Ridley (1991) likewise noticed a variation in linguistic awareness not related to level of proficiency.

To date, however, there has been very little work on defining exactly what counts as linguistic awareness, or how it interplays with general proficiency in learning an additional language. Classroom practice in this area is generally ad-hoc, with the teacher pointing things out as they occur. However, linguistic awareness is a large and solid field, and as such it lends support to Lewis in his insistence on noticing and consciousness raising.

Lewis has his critics. His initial book (Lewis, 1993) contained a lot of theory and not much practical application. With the publication of the second and third books in the series (Lewis, 1997; Lewis, 2000), this criticism no longer holds.

Another criticism of the Lexical Approach is that it is too vague (Thornbury, 2002). Lindstromberg (2003) decries it as not useful. A further difficulty is that Lewis has not really explained to the reader how to determine when a chunk begins and ends in connected language. Making lexis the center of language teaching has not been universally accepted. Many teachers (and students) do not wish to ‘let go’ of grammar, and the place it had traditionally held in language courses. Lindstromberg (2003) prefers a more generative view, including grammar and word-level meaning alongside lexis. He considers the rate of learning that would be needed under a lexical approach to be overwhelming.

A criticism of a different nature was made by Waring (1997). He claims that Lewis focuses on input of language and the transfer to intake, but does not deal with output by the student. He also critiques Lewis for not including references to memory studies, memorization, mnemonics, and so on, and the role they have been shown to play in vocabulary learning. Lewis, however, does understand the demands of the classroom, especially in countries where exit exams dictate a curriculum founded on a grammatical sequence. All in all, Lewis has presented his approach, and left it to the individual teacher whether to implement strongly, weakly, or not at all.

Collocation

Definitions and categorization.

‘Collocation’ is a major idea within the Lexical Approach. Word choice is not free and open, but is severely limited by what goes before and after. The term *a fortune* typically follows a verb such as *amass, inherit or squander*. With the verb *to shrug*, virtually the only possible noun combination is *shoulders*. *Heavy* collocates with *rain* and *burden*, but not

with achievement. We may be *bitterly disappointed* or it may be a *bitterly cold* day, but we do not typically say *bitterly sad* or *bitterly shocked*. Thornbury (2002) uses the expression "hunting in packs" to describe this behavior of words. Knowing these typical combinations is the hallmark of a native speaker. The task of the foreign language learner is to get a 'feel' for collocations, for which words go together.

Lewis defines collocation as strength of association, or the probability of two words occurring together (Lewis, 1997). Williams (2002) refines this definition to: "two or three word clusters which occur with more than chance regularity" (p.1). Bahns & Eldaw (1993) propose a linguistic definition: collocations [are word combinations that] lie between free expressions and idioms.

There have been many classifications of collocations. Many linguists classify collocations by syntactical class. The simplest formulation is into lexical and grammatical categories. Lexical collocations involve one open word (noun, verb, adverb or adjective) combined with another open word. Grammatical collocations are comprised of an open word combined with a preposition, clause, infinitive or gerund. Of the two, lexical collocations bear more meaning and feature more prominently in both lessons and research.

Lewis (2000) mentions many kinds of collocation. He highlights 8, which are of special use to the learner: adjective + noun; noun + noun; verb + adjective + noun; verb + adverb; adverb + verb; adverb + adjective; verb + preposition + noun. His Collocations Dictionary (Hill & Lewis, 1997) is organized into two sections and includes some of these categories mentioned above: starting from a noun (verb + noun; noun + verb; adjective + noun; phrases) and starting from a verb or adjective.

Gitsaki (1999) gives and uses over 30 types of collocations. However, they are all forms of the lexical/grammatical classification, and she reduces her analysis to this dichotomy. Gitsaki based her categories on the system devised by Benson, Benson and Ilson (1997) in their Collocations Dictionary. Interestingly, in spite of this rigid categorization, their dictionary is organized alphabetically.

Other workers categorize collocations along one or more continua. One such continuum is degree of fixedness, that is, the extent to which the collocates can be replaced by another word. At one end the items are unchangeable (e.g. *chip off the old block*). At the other end are free combinations, words which have seemingly unlimited combination possibilities (e.g. *red* + virtually any concrete object). In between are semi-fixed items (e.g. *ham and eggs*), where the items could be reversed or *bacon* substituted for *ham*, and colligations (syntactical patterns such as *I am ____ing*). Some workers consider collocations to encompass this whole range. Others restrict use of the term collocation to items that fall in the middle of this category.

Nation (2001) has 10 attributes, along which he defines collocations. Some of the more important are: frequency of co-occurrence, adjacency, collocational specialization, fossilization, opaqueness and uniqueness of meaning. Renouf & Sinclair (1991) have measured the degree to which either part of a collocation is bound.

Some workers classify idioms and phrasal verbs as collocations. A major distinction between these two categories and collocations is the degree of opaqueness: the meaning of a collocation generally derives from the meaning of its constituent parts, whereas idioms and phrasal verbs are overwhelmingly metaphorical. As long as this distinction is made clear to

students, there is no reason why all three should not be learned as lexical chunks.

An important classification is by strength of collocation. Hill (in Lewis, 2000) distinguishes between unique, strong and weak collocations. Unique collocations are set combinations, such as *clove + garlic*, or *duckbilled + platypus*. A strong association is found between *world* and *war*: given one of these collocates, the chances of finding the second in close proximity are pretty high. A weaker collocation exists between *war* and *long*. There are plenty of other adjectives that could combine with *war*, and very many more nouns that could be *long*. Collocations such as *rancid + butter* are very strong in one direction but less so in the other: given the word *rancid*, *butter* is about the only possible noun combination. However, starting with *butter*, *rancid* is only one of several possibilities.

Strength of collocation is not the same as frequency of collocation. Frequency refers to the degree of usage, whether it is something that native speakers typically and often say. A collocation can be strong and frequent, or strong and infrequent, or weak and frequent. Although a fourth category (weak and infrequent) does exist, Lewis dismisses this as having little value to the teacher (Lewis, 1997, p.30). Frequent collocations should be part of the classroom lesson, but strong collocations may be the ones that need most attention. Another aspect of collocation is the collocational range of a word: which words, and how many different words, can combine with it. A related idea is that of which combinations are impossible or not native-like. *Pay* collocates with *bill*, *debt*, *costs* and so on, but does NOT collocate with *discount*. We speak of *serious* in relation to *illness*, *relationship*, or *matter*,

but we do NOT combine it with *advantage*. It is especially important for students to know non-collocates that might be acceptable combinations in their L1s.

Pedagogical aspects.

Within the field of vocabulary learning, the learning of collocations is being given increasing prominence. Both Nation (2001) and Thornbury (2002) devote time and attention to the area. Teaching vocabulary in (lexical) context is more effective than the teaching of individual words (Thornbury, 2002). At intermediate and advanced levels, it is increasingly important to have students review words they already know, with the view to enlarging the collocation range. Collocation is slowly making its way into ESL reading materials in the USA (Bliss, 2004), although there is still a very long way to go. Proficiency in collocation use is “native-like”: being native-like is the ultimate goal of language instruction, especially in the ESL field, where the aim is to mainstream students.

Collocations are presumed by Lewis to be acquired from rich input, either from listening or from reading. Stockdale (2000) considers that the key to learning collocations is the constant recycling of vocabulary through more and broader collocates. Editors of collocation dictionaries imply that collocations are learned through need: the student has to express an idea in speech or writing, and consults a dictionary for possible or likely combinations of words.

In recent years, several dictionaries of collocations have been published. (Benson, Benson & Ilson, 1997; Hill & Lewis, 1997; *Oxford collocations dictionary*, 2002; Stockdale, 2000) Whereas regular dictionaries focus on definitions, collocation dictionaries focus on words that frequently accompany the headword. A typical entry for “sun” might

include: *sunstroke*, *sunshine*, *midday sun*, *sunbathing*, *sun lotion*, *in the sun*, *sunrise*, and so on. Collocation dictionaries are also available in other languages, such as Russian (Benson & Benson, 1993), Spanish (Bosque, 2004), and French (Beauchesne, 2001).

Stockdale (2004) recommends using both a standard bilingual dictionary and a collocations dictionary, as one then gets both meaning and use of the headword. However, the two are not equal: he states that classroom time should be allocated 90/10 in favor of collocation.

Woolard (2004) has recently produced a series of text books for training in collocation. He claims that most current textbooks are not equipped to handle collocation. His *Key Words for Fluency* series is primarily for individual study rather than as a classroom text. It is based on his Key Word approach, where a noun is taught in conjunction with its left and right co-texts.

Opposition to teaching collocation states that the number of possible collocations in a language is so vast, that a student would in fact have more learning, and more memorizing, to do than in a traditional vocabulary class. Various workers have countered these complaints: Nation (2001) suggests concentrating on collocates of high frequency words. Bahns (1993) suggests that a contrastive approach (comparison of L1 and L2) would help with the sheer volume of possible collocates. Bahns and Eldaw (1993) recommend spending time on collocates that cannot be easily paraphrased. Howarth (1998) suggests focusing on collocates that are semi-fixed phrases, as these are the ones that cause the most difficulties. Porto (1998) suggests concentrating on the more general, less-fixed collocations, as these have greater potential for analysis in the classroom. Hill (in Lewis,

2000) suggests focusing on medium strength collocations; words such as *mistake* or *meeting*, that combine with basic words they already know: *make* and *hold*.

Including collocation in the curriculum is important for several reasons. The first reason is the widespread difficulty faced by non-native speakers in the accurate combination of words. Even in cases where the learner knows the individual words, collocations are still likely to be problematic.

Secondly, there is a need for learners to get beyond the “intermediate plateau” that Lewis (1993) describes. Such students can cope in most situations, but they tend to “avoid” or “talk around” the more challenging tasks of advanced language learning. Collocation instruction is especially motivating for upper level students (Williams, 2002). Hill, in Lewis (2000), reminds us that many complex ideas, such as advanced academic concepts, are expressed lexically.

A third reason is that having a knowledge of frequently occurring collocates deepens vocabulary knowledge, and increases fluency and aids stress and intonation (Williams, 2002). It is essential for accurate, natural sounding English (Kavaliauskienė & Janulevičienė, 2001.)

A final reason is that collocation errors are more disruptive to the communication than most grammatical errors. The result is unnatural sounding expressions or quaint or archaic phrasing. Williams (2002) warns that a non-native speaker with many collocation errors will start to “grate” on the listener or reader. Again, collocation determines whether or not the speaker is accepted as native-like.

Measuring collocation.

Productive knowledge of collocation is often measured by a fill-in-the-blanks format, whereby the student has to provide a suitable collocate for a given word. These tend to be de-contextualized sentence-based tests. Productive knowledge of collocation is also measured using translations into the target language or by assessing free writing, although in the latter there is no guarantee that the subject will actually produce any collocations that he knows. Gitsaki's (1999) investigation included both fill-in-the-blanks and paragraph writing.

Receptive knowledge of collocation is typically measured by a matched items test. Another aspect of receptive collocational knowledge is identification of collocations that do NOT work. Typical tests for this include odd-man-out exercises, where the student selects from a group of possible responses the item that does NOT collocate with the key item. Bonk's (2000) test, the basis for the test in the current investigation, includes one section of receptive collocation knowledge, (where for each question the student has to pick the one collocation out of four which is not natural English) as well as two sections of productive fill-in-the-blanks.

Many tests of collocation are teacher-created, used to assess classroom learning. (Thornbury, 2002). Two workers in the field, Bonk (2000) and Gitsaki (1999), have designed and published collocation tests.

Bonk's (2000) test is designed as a normative test. It is in three parts: Cloze style noun + verb production, Cloze style verb + preposition (phrasal verbs), and multiple choice selection of inappropriate collocation. The test was based on native speaker validation. In

Bonk (2000) he provides technical detail of validity and reliability: he claims “preliminary evidence for the construct validity of this test” (p.23) such that the test measured a construct “not covered by TOEFL-like measures of [linguistic] proficiency” (p.34). Overall reliability score was given as .83, although this was lower for Section B. Bonk also covers the results of a pilot study, and gives suggestions for future use of his test.

Gitsaki (1997) designed a test for her investigation based on the lexical and grammatical collocations occurring in the students’ English (EFL) textbooks. Beyond noting that the textbooks were in wide circulation, she does not present any technical details.

L2 collocation performance.

There is widespread agreement in the literature that collocation is a difficult skill for non-native speakers. A lot of the evidence is anecdotal or even intuitive: teachers have known for a long time about mis-selection of words that make the student ‘sound foreign’.

Several studies that measured collocation skill in second language learners used native speaker performance as the standard against which the non-native speakers were compared. Bonk (2000) and Gitsaki (1999) both used educated native speaker response in the development of their tests. Hasselgren (1994) compared Norwegian students in advanced EFL classes and British students. Granger (1998) and Howarth (1998) compared written work by native and non-native speakers of French.

It is hardly surprising that the non-native speakers fell woefully short of the native levels of performance, in both quantity and quality of collocation performance. Researchers such as Krashen and Terrell (1983) and Selinker (1972), who subscribe to a naturalistic

view of language acquisition, would not see this as a problem. They would claim that collocational errors should be hailed as evidence of progress in the target language. Such errors would be examples of creative interlanguage, not of a deficit in L2 performance. The point remains though, that such utterances – however quaint, unusual or endearing - are not what a native speaker would use. We must remember where the students are headed, especially students in an ESL situation, where the goal is integration into mainstream academia.

Other studies have shown collocation skill to be low relative to a student's other measures of language proficiency. In other words, the student is being compared to himself. Vocabulary level is a case in point. Collocation ability has been shown to lag behind measures of general vocabulary (Bahns & Eldaw, 1993). Even when subjects knew the items as individual words, there were still difficulties with the collocations. Zughoul & Abdul-Fattah (2003) and Fargal & Obiedat, (1995) both found collocation deficiencies, even in cases of students with large English vocabularies. Bonci (2002) had similar comments about her subjects studying Italian.

Students with excellent syntactic skills likewise could be poor in collocation skill (Zughoul & Abdul-Fattah, 2003; Fargal, & Obiedat, 1995). Likewise, students with advanced writing skills may be deficient in collocation. Granger (1998) and Howarth (1998) both conducted research using corpora data. Compared to native speakers, the non-native speakers (NNS) used fewer collocations, and the ones they did use were less successful as combinations: we are back again at the unnaturalness of NNS language.

There is debate in the field as to whether collocation skill is independent from

overall level of proficiency. At the University of Hawaii, Bonk (2000) found collocation ability to correlate moderately highly ($r=0.73$) with ESL proficiency, in that those with higher TOEFL-like proficiency scores outperformed those in lower. Gitsaki (1999) has demonstrated the development of collocation competence in Greek-speaking EFL students. She showed that this skill grew alongside general level of proficiency in English. Other workers have demonstrated a lack of relationship: Bonci (2002) showed that collocation skill level varied widely amongst students of a given proficiency level, such that one cannot talk of “beginning collocation skills”. Howarth (1998) also found a wide range of collocation ability within a given proficiency level, as did Bahns & Eldaw (1993), Zughoul & Abdul-Fattah (2003) and Fargal & Obiedat (1995).

There is the danger of circularity when an investigator compares collocation skill and general L2 proficiency level. Part of the definition of “advanced” speaker involves making native-like collocation choices. Consider in this respect the scoring guidelines for the Language Assessment Scales (DeAvila & Duncan, 1990), a widely used ESL proficiency test. For the oral Story Retelling and the Let’s Write! Sections, there is a maximum score of “3” (“5” being the highest) for any response including non-native-like expressions. These two sections are weighted heavily in the overall scores, such that a student making collocation error would be at best a *Limited English Proficient*.

To avoid this, an investigation should measure proficiency in ways that do not include collocation, such as single-word vocabulary tests. Whether this would be a good measure of proficiency is of course another matter.

Learners from many different language backgrounds have been shown to be

deficient in collocation knowledge. Most studies examined a monolingual group, where all the students had the same L1. Collocation studies have covered a variety of native languages, of varying linguistic distance from English. Although this seems to suggest that collocation is a widespread problem, each of the experimenters used a different instrument to measure collocation, making comparisons impossible.

A few studies dealt simultaneously with more than one language group. Birkup (1992) looked at collocation in two language groups, Polish speakers and German speakers learning English. She attributed the lower Polish scores to cultural differences, especially in school expectations, rather than the greater linguistic distance of the Polish language to English. Bonk (2000) also had a multilingual group, the students being predominantly from South East Asian groups. However, he did not specifically compare across language groups.

The majority of studies have used students in a University or college setting, and mostly at advanced levels of English. However, some work with younger students has also shown the same difficulties as experienced by adults. Hasselgren (1994) included final year high school Norwegian students, and Gitsaki (1999) studied younger teenagers.

It is not clear whether collocation develops according to the amount or length of exposure to the target language. Gitsaki (1999) believes it does. Length of time (in years) in English class was related to the growth in collocation skill. Bonk (2000) in his pilot study found a correlation between Collocation Test score and LOR of 0.39. He did not include this measure in his larger study. Bonk does not give figures for length of residence, only that everyone arrived after the age of 13. He does suggest, though, that quality of exposure to the target language is more important than quantity. This area of inquiry is in need of

further research. The current investigation seeks to obtain further data on this matter.

Development of L2 Collocation.

To date the only study to look at collocations developmentally is Gitsaki (1997).

This study looked at performance at one point in time of three different proficiency levels. It did not follow individuals through collocational development longitudinally.

Gitsaki's (1997) study compared collocation skill at three age levels: students in their first year of studying English (11 – 12 year olds), second-years (12 – 13 year olds) and third-years (13 – 14 year olds). She looked at the development of collocation skill, by comparing typical first-year performance with that of the typical second-year, which in turn was compared to the performance of the typical third-year student. Her results showed there is a developmental aspect to collocation, that it is non-linear, with a wide range within a given proficiency level. Most importantly, Gitsaki claimed that the nature of the collocations mastered at any stage was different. Within collocations, grammatical collocations were shown to be more easily acquired than lexical collocations.

Gitsaki was principally interested in type of collocation mastered by her various levels. This is good, because her claims that she demonstrated development of collocation skill are hampered by the instrument used. Gitsaki gave a different test to each of the three levels, making comparisons between the levels invalid.

Gitsaki claims that the three levels of students represent different proficiency levels. Although she collected samples of written work to check for overall differences in proficiency, she did not place students in groups based on this assessed proficiency level. It is more truly a convenience grouping by age/grade in school.

There is one other problem with Gitsaki's design. The instrument (writing test) used to determine proficiency is also used as part of the measure of collocation. This is circular.

Gitsaki's study provides a rich amount of data on collocation. It also suggests a way of looking at the development of this aspect of language development. Her use of this population – Greek middle school students learning English – was another unique contribution to the research literature.

Approaches to collocation.

We have seen how non-native speakers struggle with fluent and natural word combinations. The question arises as to the reason for this. What is going on in the students' heads as they attempt a collocation? Why are there so many errors?

Many studies and much theoretical work point to interference from the native language as being responsible for collocation errors. The interference may or may not be consciously realized. The most obvious form of interference is direct translation, word for word, into the target language of what the student would have said in the native language. Hasselgren (1994) refers to this as 'transliteration'. This may be unwitting: the student simply did not realize that the target language handles this instance in a different manner, or it may be the only thing the student can come up with.

Interference is also exhibited in "gridding" (term used by Hasselgren 1994): when the collocational ranges of a word in L1 is narrower or wider than the collocational range of the equivalent word in L2, a mis-selection may be made. Hasselgren (1994) also mentions cognates, especially false cognates, and misleadingly close words or expressions as likely to result in interference.

Another explanation for collocation errors is that students tend to ‘play it safe’ by resorting to what they do know. Taiwo (2004) recognizes the tendency of students to use the most common or general form of a semantic item, even when another form would be more precise (*walk* vs *hike* or *stroll*). Hasselgren (1994) also found high usage of what she terms “lexical Teddy Bears”. These are core words, usually words of high frequency and with little meaning on their own, words that students typically learn in the first year of study (*get, do, make, have, be, put*). She interpreted this to mean that students tend to stick with the familiar (and easy) when formulating collocations. Other indications that students select easier but less effective combinations are found in both Granger (1998) and Howarth (1998).

A third reason behind collocation error is that sometimes the student simply avoids the collocation altogether. This is seen particularly in productive collocations tests, where a student leaves the item blank. Biskup (1992) found the Polish students had more omissions than the German students. She interpreted this as meaning that the Poles were less willing to take risks than the Germans, and she related this to factors of cultural background.

A fourth reason is related to the risk-taking in the previous paragraph. Some students, when faced with a collocation they do not know, may talk their way around the collocation. Biskup’s (1992) German students, for example, produced some creative paraphrasing. In such situations, the students are using what they do know of the target language to create something approximating the desired collocation. It may be a long or a short description of the desired collocation, or use of a different part of speech, or it may be

acted out for the examiner. In any case, the student here is attempting to get at the spirit of the collocation.

Conclusions from the literature review

The literature has shown that collocation skills are generally poor in second language students and that they lag behind other aspects of language such as vocabulary size, syntactical accuracy and comprehension. Previous studies also suggest a non-linear development curve, with a wide range of collocation skill within a given proficiency level. Although exposure to the target language is a necessary condition for acquisition of collocation, it is not clear whether exposure, *per se*, is a sufficient condition.

The primary aim of the current study is to investigate whether collocation skill increases over time in high school immigrant ESL students. The study correlates scores on the Collocation Test with Length of Residence in the USA, measured in months.

The study improves on previous work in that it uses a larger group of students and considers a wider range of the ESL spectrum. This study uses the same test for all students. This study uses LOR as the independent variable. Exposure to English for two students with the same LOR can be considered equal: the amount of time is the same in both cases, as is the (academic) setting over that time spread. By using LOR as the independent variable, it avoids the pitfall of circularity seen in the literature where proficiency is the independent variable. The study extends previous work, which is mostly with advanced EFL University students, to the immigrant, high school ESL situation.

The secondary aim of the current study is to compare collocation skill in intermediate subjects (LOR = 19+) of two different language backgrounds. Most previous

studies have involved a single L1. The present study looks at overall performance on the Collocation Test and the approach to collocation in Spanish- versus Russian-speakers.

Chapter 3

Methods

The primary aim of the current study is to investigate the development of collocation skill over time in high school immigrant ESL students, by correlating scores on the Collocations Test with length of residence. The secondary aim is to compare collocation skill (overall performance and the approach to collocation) in intermediate subjects of two different language backgrounds.

Research questions

The research question in Part 1 involves the large group of 143 students. Is there an increase in measured collocation skill over time, such that greater length of residence (LOR) in the target culture results in higher Collocation Test scores?

The research question in Part 2 involves the small subgroup of 30 intermediate (LOR > 18months) students: 15 Spanish-speakers, and 15 Russian-speakers. Is there a difference in the total number of collocation errors made by intermediate students of the two language backgrounds? Is there a difference in the number of errors made by intermediate students of the two language backgrounds in section A (production of *noun + verb collocations*)? Is there a difference in the number of errors made by intermediate students of the two language backgrounds in section B (production of *verb + preposition collocations*)? Is there a difference in the number of errors made by intermediate students of the two language backgrounds in section C (recognition of *verbal collocations*)?

The research question in part 3 involves the same subgroup as in part 2. Is there a difference in the way Russian-speaking and Spanish-speaking students approach an unknown collocation? Do they attempt all items, or leave some unanswered? Do they answer based on what they would have said in the native language? Do they resort incorrectly to easy, basic words, such as the delexicalized verbs *make*, *do*, *have*, *get*, *be*, *put*? Do they paraphrase (use a multi-word response or a non-verb response)?

Setting

Two suburban comprehensive high schools provided students for this study. The two schools were located in the Pacific Northwest, one in Oregon and one in Washington. They were both large public schools, grades 9 – 12, and were very similar in terms of socio-economic status. The percentages of students on free and reduced lunch were 68.5% at school A and 56% at school B. Each school housed the district High School ESL program. The most recent school profiles show the ESL population at 20% of the student body at school A, and 18% at school B. Each school had a diversity of language backgrounds in the ESL population: in both schools Russian-speakers comprised the largest subgroup, followed by Spanish-speakers. The ESL program was similar at the two schools, with an emphasis on English immersion and content based ESL. Neither school was the researcher's current place of employment. The two school districts were selected based on appropriateness and convenience of access (Berg, 2004), with their ELL populations representative of this metropolitan area as a whole.

Participants

The large group.

Students selected for the study were to be immigrants to the USA, currently enrolled in a high school ESL program. Upon arrival in the USA, the level of English had warranted placement in a newcomer or beginning level ESL class. Length of residence had to be at least two full quarters (six months) for the individual to be included in the study. A total of 156 students were tested. Eight students were excluded from the analysis because their LOR scores were less than 6 months. Two other students were excluded because they were not immigrants. A further three students were excluded for reasons connected with the testing situation. The remaining 143 students formed the large group for Part 1 of the study.

School A provided 50 students for the study, and School B provided 93. A total of eight ESL teachers provided access to students in 11 classes. Three classes were designated beginning ESL, five were intermediate and three were advanced. No advanced classes were available to the researcher at School A, due to scheduling conflicts. 20% of students in the large group came from the beginning classes, 55% from the intermediate classes and 25% from the advanced classes.

In the large group, there were 59 males and 84 females. There were 43 freshmen, 54 sophomores, 36 juniors and 10 seniors. All but eight students had been placed in the age-appropriate grade, or one academic grade below their age-appropriate placement. On the day of testing, all but three students were between the ages of 14 and 19 inclusive.

The break-down of students into language groups was as follows. Russian-speakers were the largest language group, at 71 students. This total included 34 Russian-Ukrainian

bilinguals and one Russian-Moldovan bilingual. There were 38 Spanish speakers, predominantly from Mexico. Of the remaining 34 students, 11 identified themselves as Romanian speaking, and eight as Chinese speaking. There were three Arabic speaking students, two each of Khmer, Swahili and Vietnamese, and a singleton of each of Chamorro, Chuukese, French/Creole, Korean, Portuguese, and Serbo-Croatian. The few students indicating multiple languages were categorized with the dominant language. Appendix C shows these in tabular form.

Length of Residence scores ranged from 6 months to 98 months. No newcomers ($LOR < 6$ months) were included in the design. Forty students had LOR between 6 months and 18 months. This is typically the beginning ESL range. 50 students had LOR between 19 and 36 months, or the typical intermediate range. 53 students had LOR in the typical advanced range at 37 + months, having resided in the USA for three years or more. The average LOR was 33.67 months, with a SD of 20.51. This information is presented in Appendix D in tabular form.

The small group.

The students for the second and third parts of the study were selected as follows: all intermediate students (LOR of 19+) with a designated L1 of Spanish or Russian were identified. There were 30 Spanish-speakers and 49 Russian-speakers so identified. From the pool of Spanish-speakers, 15 were randomly selected, with each candidate having an equal chance of being chosen. From the pool of Russian-speakers, 15 were likewise selected. This gave a small group of 30 intermediate students. Students in the small group were identified by a two-letter code: RA, RB, RC etc. for the Russian-speakers and SA, SB, SC etc. for the

Spanish-speakers, according to the order they were selected.

Instrumentation

A sheet requesting necessary demographic data was designed by the researcher. The sheet was on white colored paper. A copy is in Appendix E.

The Collocation Test was based on the test designed by Bonk (2000). The current version involved a few changes in vocabulary to better suit the climate and culture of the two high schools ("spouse" instead of "wife"; "steak" instead of "wine"). None of the changes affected the collocation being measured.

The Collocation Test was rated independently by four ESL professionals as "not unduly favoring one language group culturally or linguistically over any other language group". The items in Section A were also analyzed by speakers of Russian and Spanish, to determine the equivalents in those languages of the verb + noun collocates. None of the 17 items had a one-to-one correspondence between the collocation in English and the equivalent in either Spanish or Russian.

For convenience and ease of administration, the three sections of the test were colored coded. Section A was blue, section B yellow and section C pink. A plain copy of the test is in Appendix A. The answer key is included as Appendix B.

Procedure

The researcher contacted the two school districts well in advanced of the proposed test date. All school contacts were made through a designated teacher. A testing date was set for each of the schools and the designated contact set up a schedule for class visits. About a

week prior to the testing date, students were informed about the study and that participation was voluntary. One school district required parental notification, the other did not.

The researcher administered all the testing. Students were assessed in 12 sessions, each session dealing with between 6 students and 24 students. The testing took place either in the ESL classroom, or in an empty hallway nearby. Testing was done during the student's ESL class time, and each session took between 40 and 50 minutes.

The test administration was as follows. (Verbatim instructions are included as Appendix F.) The researcher introduced herself and explained the study as "an investigation of how people learn vocabulary in ESL classes". They were thanked for their participation. The test was then administered according to the instructions in Appendix F. As each student finished, the four sets of sheets (demographic sheet, section A, section B and the multi-page section C) were stapled together, with the demographic sheet on top.

The researcher identified all test packs as follows: all were labeled (on the demographic sheet) with the code for the school, class period, and the name of ESL teacher. Individual test packs were then numbered in the order in which papers were collected in that session. The researcher completed this identification at the end of each administration, before proceeding to the next test session.

At the end of each testing day, the researcher left with her thanks and made arrangements to follow up with any staff member or student who desired feedback on the results.

Scoring and Analysis

Part 1: the large group study.

The demographic sheets were checked for date of arrival, to exclude students who had been born in the USA and those who had been in the country less than 6 months. Other responses (home language, date of birth, grade level, gender) were checked for completeness. The few discrepancies involved a follow-up contact to the teacher concerned.

The tests were scored according to the scoring guide in Appendix B. In sections A and B the response had to be in English and totally correct in spelling and grammar (e.g. verb tense) to be counted as correct. Inconsistencies in upper/lower case were not penalized. If more than one response had been written in the blank, the one on the line itself was accepted, unless this had clearly been crossed out. In section C, if more than one response had been circled for a particular item, the answer to that item was considered as incorrect, unless one circle had clearly been scratched out or erased. Particular care was given to the scoring of section C, as this involved selection of a "wrong" answer by the student.

Scores for each section were entered at the top of that section. In addition scores for sections A, B and C were transferred to the demographic sheet. The total (out of 50) was then calculated, and entered on the demographic sheet.

The test scores were checked by an undergraduate assistant, who randomly selected around 75 tests to rescore.

LORs were calculated next. A table was drawn up for this to facilitate this procedure. Only whole months were counted: the month of arrival and the month during which testing took place were excluded from the calculation.

The following data were then entered into an Excel spreadsheet: school code; name of teacher; period; student pack #; date of testing; date of arrival in USA; gender; NL; score on section A; score on section B; score on section C; total score; LOR.

Pearson's r correlations were run on the Excel program for the whole group, between total Collocation Test score and Length of Residence.

In addition, correlation scores were also run on subgroups of interest: low/mid/high LOR; school A/school B; Russian/Spanish/other language background; males/females.

Part 2: the small group quantitative study

Part two of the study did not involve any additional scoring. The researcher transferred the necessary data from the part one spreadsheet to a new 30-line spreadsheet. An extra heading denoted the student's small group identification (SA-SO, or RA-RO).

For each of Total Test score, Section A score, Section B score, and Section C score a t-test was run on the Excel program, to compare the mean scores of the Russian-speakers vs the Spanish-speakers. A two-tailed test was used, with $\alpha = 0.05$ and 28 df.

Part 3: the small group qualitative study.

Part three did involve additional scoring. For each of the 30 students in the small group, all incorrect responses from section A were identified. I then analyzed each response, and coded it according to whether it was a grammar error (G), a synonym error (S), a random error (R) or whether the response had been left blank (B). Grammar errors involve a spelling or syntactic variation of the correct item. Synonym errors involve the substitution of another word, such that the sense of the collocation is intact. This includes direct (but faulty) translations from L1. Random errors are responses that do not reflect the sense of the

intended collocation. Synonym and random errors may also involve a spelling or syntactical error.

A new 30-line spreadsheet was created, with two columns for each question in section A. The first of these contained the actual error (e.g. "*taked*"), the second contained the error code (G, S, R or B). Correct responses on section A were not entered into this spreadsheet: that data point was left empty.

Part three of the study involved qualitative analysis of the actual responses in section A. The number of errors by each of the language groups was recorded and broken down into type of error (Grammar, Synonym, Random or Blank). I next conducted an analysis of the blank errors: the number of blank errors and the location (were the items at the beginning, the middle or the end of section A?), comparing these across the two language groups. Following this, I looked for evidence of (faulty) L1 transfer by each of the language groups, based on their synonym errors. The Spanish analysis of L1 transfer was done by the researcher. The Russian analysis of L1 transfer was done by the researcher with assistance from a Russian colleague. The next analysis involved the extent to which the synonym and random errors by the two language groups involved incorrect use of *make, do, have, get, be, put*. Finally I looked at the extent to which the two language groups used paraphrasing (non-verb responses or multi-word responses) in their synonym and random errors.

Chapter 4

Results

The first set of data in this chapter is scores on the Collocation Test for the whole group and also for strategic subgroups. Following this, I will present the results for Part 1, Part 2 and Part 3 of the study.

For Part 1, I show the correlation statistic between Collocation Test score and Length of Residence for the whole group ($N = 143$). Although the score for the whole group is of principal interest, I will also share subgroup correlations that merit discussion in chapter 5. For Part 2, I show the results of the comparison of means (Russian-speakers versus Spanish-speakers) for the intermediate subgroup ($n = 30$). For Part 3, I share qualitative information on the type and quality of errors made by the subgroup ($n = 30$) on section A.

Collocation Test Scores

The mean Collocation Test score for the whole group was 14.78, with a standard deviation of 8.01. The lowest score was 0, the highest score was 36 out of a possible maximum score of 50.

There were no significant differences in mean Collocation Test score for the males versus females. Nor was there a significant difference in mean scores for the different language groups, although the Spanish speakers' performance was the lowest by several

points. The Russian speakers' mean performance was very close to that of all other speakers' mean performance. School B had a higher mean than did School A, but this difference is not significant.

There were differences in mean scores in the expected direction for low/middle/high levels of LOR. An ANOVA test (between groups df =2) gave an F statistic of 11.8453 and a p-value of 0.0000, indicating a significant difference between groups.

Table 4.1 presents Collocation Test scores for the whole group and for subgroups of interest.

Table 4.1

Collocation Test scores

	Mean	Standard deviation	Range of scores
Whole group	14.78	8.01	0 - 36
by Gender			
Males	14.93	7.46	0 - 32
Females	14.67	8.42	0 - 36
by School			
School A	11.54	5.71	0 - 31
School B	16.51	8.54	0 - 36
by Language			
Russian	15.63	8.62	0 - 36
Spanish	12.95	7.44	0 - 32
All others	15.02	7.14	5 - 32
by LOR			
Low (6-18)	10.38	5.94	0 - 26
Mid (19-36)	14.90	8.34	0 - 32
High (37+)	17.98	7.60	4 - 36

Part 1: Correlation between Collocation Test scores and LOR

Hypothesis 1 stated that collocation skill, as measured by the Collocation Test, will correlate highly and positively with LOR for students in the large group.

For the group overall, the Pearson's r statistic was 0.4577. This indicates a moderate correlation between Collocation test score and length of residence. Hypothesis 1 predicted a high correlation. We have some support, but not total support, for Hypothesis 1. Results indicate that students with higher LOR in the USA tend to have somewhat higher test scores. Figure 4.1 shows this information in graphic form.

Correlations for the subgroups of gender, school and language were similar. Most of these fell in the Moderate range of correlation, with r scores between 0.40 and 0.57. Two of the subgroups of interest fell in the Low range: all males (0.3366) and all School A (0.2261). When the LOR scores were divided into three (6 – 18 months, 19-36 months and 37+ months) the correlations varied dramatically. Of particular interest is the Mid LOR scores (0.0199) which is in the Negligible range. Possible reasons for this very low correlation for intermediate students (LOR 19 – 36 months) will be discussed in chapter 5.

Table 4.1 and Figure 4.2 present this information in tabular and graphic form respectively.

Figure 4.1

LOR vs Collocation Test Scores

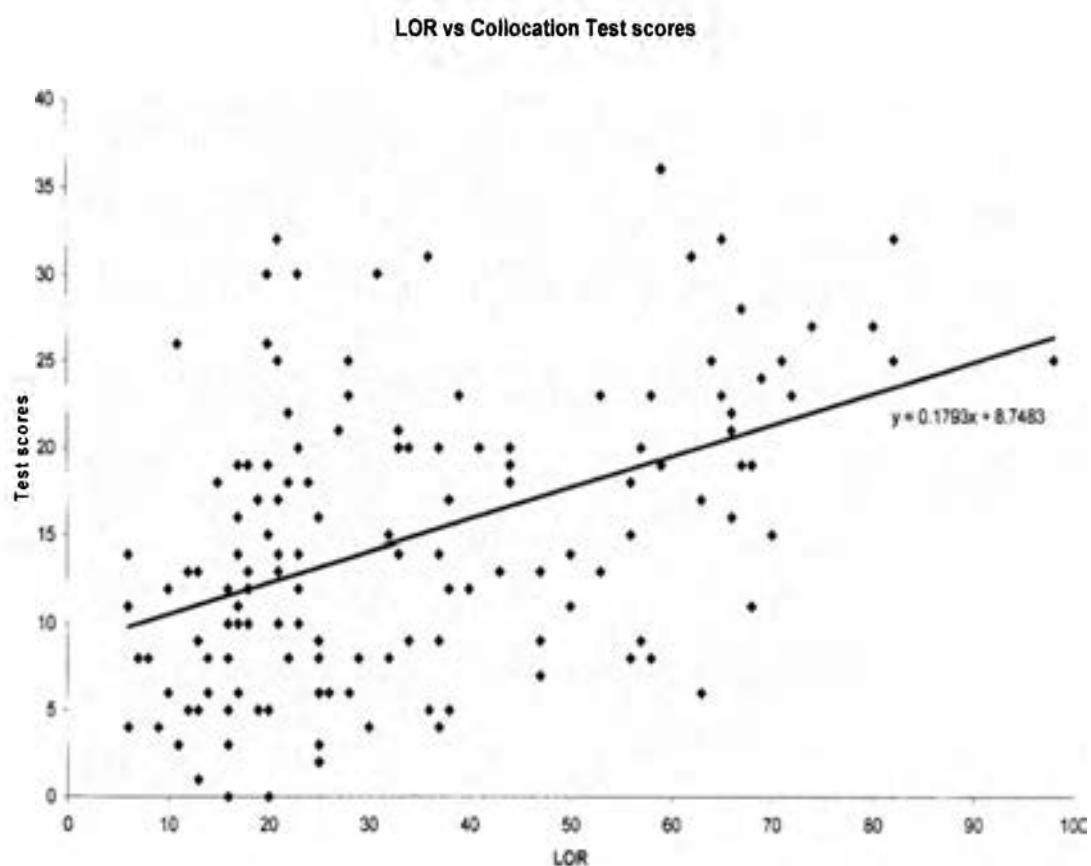


Table 4.2

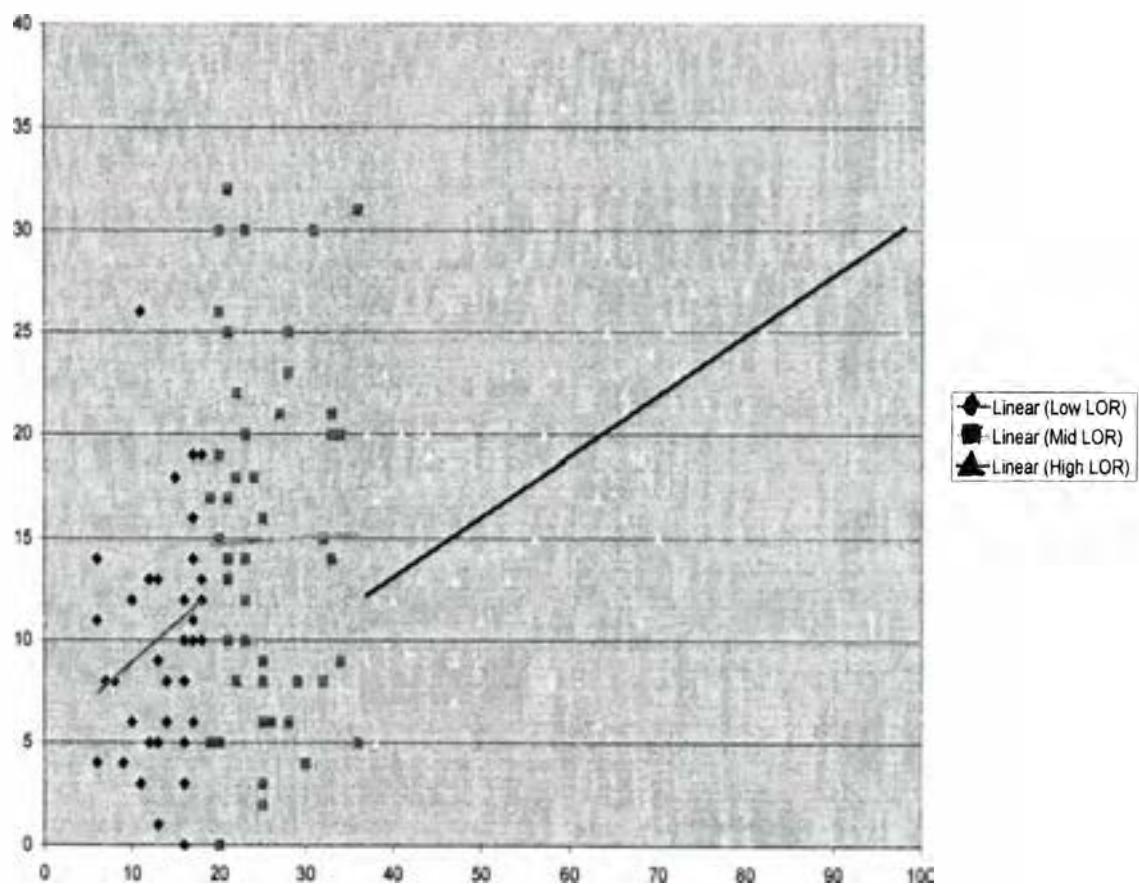
Correlation Coefficients between LOR and Collocation Test scores

Grouping	#	Pearson's r	Interpretation **
Whole group	143	0.4577	Moderate correlation
Males	59	0.3366	Low correlation
Females	84	0.5324	Moderate correlation
School A	50	0.2261	Low correlation
School B	93	0.4615	Moderate correlation
Russian	71	0.4912	Moderate correlation
Spanish	38	0.5407	Moderate correlation
All other langs.	34	0.4167	Moderate correlation
Low LOR: 6-18 months	40	0.2350	Low correlation
Mid LOR: 19-36 months	50	0.0199	Negligible correlation
High LOR: 37 + months	53	0.5495	Moderate correlation

** based on Best & Kahn (1986), p. 240

Figure 4.2

LOR vs Collocation Test Scores for Low/Mid/High LOR



Part 2: Comparison of subgroup means

Hypothesis 2a states that there will be no significant difference in the total test scores of intermediate students of Russian versus Spanish language backgrounds.

Hypothesis 2b states that there will be no significant difference in the scores on section A (production of *noun + verb* collocations) of intermediate students of Russian versus Spanish language backgrounds.

Hypothesis 2c states that there will be no significant difference in the scores on section B (production of *verb + preposition* collocations) of intermediate students of Russian versus Spanish language backgrounds.

Hypothesis 2d states that there will be no significant difference in the scores on section C (recognition of *verbal* collocations) of intermediate students of Russian versus Spanish language backgrounds.

The data in this section refers to the mean scores on the Collocations Test made by the small subgroup of 30 intermediate students. Means for Russian-speakers and for Spanish-speakers are presented for the whole test, and for sections A, B, and C. Table 4.3 presents this data, together with the results of t-test analyses.

Table 4.3

Comparison of means for Russian- vs Spanish-speaking students

Test section	Subgroup	Mean	SD	t stat	df	Crit. 2-tail	Significant
Total	Russian	17.8	6.99	1.75	28	2.048	no
	Spanish	13.07	7.77				
Section A	Russian	4.8	2.31	1.53	28	2.048	no
	Spanish	3.4	2.69				
Section B	Russian	6.47	3.42	1.38	28	2.048	no
	Spanish	4.8	3.17				
Section C	Russian	6.53	2.33	1.77	28	2.048	no
	Spanish	4.87	2.80				

There were no significant differences in the performance of the Spanish small group versus the Russian small group on any of the sections, or on the total Collocation Test scores. Each comparison used a two-tailed t-test, with 28 degrees of freedom and an alpha level of 0.05.

We thus have support for Hypotheses 2a, 2b, 2c and 2d, in that the data do not indicate a significant difference between the two language groups on the overall scores or in any of the Sections.

Part 3: Qualitative analysis of Section A

This section of the results refers to Section A answers made by the subgroup of 30 intermediate students. I will present first a breakdown of errors by category and by language group. Next I will extract from this breakdown four further sets of data: number of items not attempted on the test, items showing evidence of L1 transfer, the use of basic delexicalized verbs and the use of paraphrasing. In each case, I will state the relevance of findings to the respective Hypothesis.

Breakdown of errors.

There were 17 items in section A and 30 students, giving a total of 510 responses. Of those, 123 were correct responses. There were 25 grammatical errors (i.e. the correct answer but with inaccurate spelling or grammar) and 73 responses were blanks (i.e. student did not attempt an answer). There were furthermore 131 random errors (i.e. response had no semantic relation to the collocation) and 158 synonym errors (i.e. responses had a semantic bearing on the collocation). Table 4.4 gives this detail in tabular form.

Table 4.4

Section A errors by type and language group

Category	# Russian	# Spanish	# total
Correct response	72 (28%)	51 (20%)	123 (24%)
Item left blank	27 (11%)	46 (18%)	73 (14%)
Grammar error	11 (4%)	14 (5%)	25 (5%)
Synonym error	85 (33%)	73 (29%)	158 (31%)
Random error	60 (24%)	71 (28%)	131 (26%)
TOTAL	255 (100%)	255 (100%)	510 (100%)

Blank responses.

Hypothesis 3a states that there will be no difference in the amount of blank responses of the Russian-speaking students vs the Spanish-speaking students.

Hypothesis 3a concerns the number of items not attempted by the two language groups. Many more items were left blank by the Spanish-speakers (46) than by the Russian-speakers (27). Six individual Spanish-speakers each had five or more blanks, compared to only two Russian-speakers with five or more blanks. Seven Russian-speakers and six Spanish-speakers attempted every item (that is, they had zero blanks). Four Spanish-speakers had multiple blanks at the end of the test, compared to only one Russian-speaker.

Table 4.5 gives this information for students with one or more blank responses.

The data here do not support Hypothesis 3a. We can conclude that the Russian-speakers attempted all questions to a greater degree than did the Spanish-speakers. Also, the Spanish-speakers produced more blanks at the end of the test, compared to the scattered distribution of blanks by the Russian-speakers.

Table 4.5

Number and location of blank errors

Russian student #	# blanks	Location of blanks	Spanish student #	# blanks	Location of blanks
RA	3	mixed	SA	9	end
RC	4	end	SC	7	end
RE	8	mixed	SE	1	end
RF	5	mixed	SH	1	beginning
RH	1	end	SJ	5	end
RK	1	end	SK	10	mixed
RL	2	mixed	SM	5	mixed
RM	3	mixed	SN	1	end
			SO	7	end

L1 transfer.

Hypothesis 3b states that there will be no difference in the amount of synonym errors showing evidence of L1 transfer for Russian-speakers vs Spanish-speakers.

Hypothesis 3b has to do with L1 transfer. Of the 158 synonym errors, there were 38 synonym errors that showed evidence of (unsuccessful) L1 transfer. Russian-speakers were responsible for only 4 of these: 2 in Question 6 (*see a dream*), and 1 in each of Question 7 (*turn attention*) and Question 12 (*take a decision*). Spanish-speakers showed evidence of L1 transfer 34 times. Problematical questions were: Question 4 (*take your appetite*), Question 7 (*put attention*), Question 10 (*writing a diary*), Question 12 (*take a decision*), Question 14 (*do for and make a favor*) and Question 16 (*clean/wash your feet*).

The data do not support Hypothesis 3b, in that Spanish-speakers showed far more evidence of L1 transfer in their (incorrect) responses. Russian-speakers showed evidence of a broader search for the correct response, in that they relied far less than did the Spanish-speakers on what they would have said in their native language.

Use of basic verbs.

Hypothesis 3c states that there will be no difference in the amount of incorrect responses involving *make, do, have, get, be, put* by the Russian-speakers vs the Spanish-speakers.

Hypothesis 3c deals with the extent to which the two groups resort to easy, basic words, such as *make, do, have, get, be, put*. There were 111 incorrect responses using *make, do, have, get, be, put*. 59 of these were made by the Spanish-speakers and 52 by the Russian-speakers. Table 4.6 gives this information by verb and by language group.

Table 4.6

Incorrect use of basic verbs by language group

Verb	Russian-speakers	Spanish-speakers	<u>totals</u>
Make	11	12	23
Do	38	8	11
Have	21	14	35
Get	4	2	6
Be	11	16	27
Put	2	7	9
<u>totals</u>	52	59	111

The data here provide support for Hypothesis 3c: the two groups show a similar tendency to resort to easy, delexicalized verbs, particularly *make*, *have* and *be*. The Spanish-speakers, however, show this tendency slightly more than the Russian-speakers, except for *have*.

Use of paraphrasing.

Hypothesis 3d states that there will be no difference in the use of paraphrasing by the Russian-speakers vs the Spanish-speakers.

Hypothesis 3d deals with the extent to which the two language groups paraphrase (use a multi-word response or a non-verb response). In section A of the Collocation Test, all questions needed a single-word verb as answer. There were 43 answers that were not verbs, and a further 7 answers were multi-word verb answers. Spanish-speakers were responsible for 27 of these errors, and Russian-speakers for 23. Appendix H gives full details of these answers, together with an indication of whether the response was a semantic fit (i.e. caused the sentence to make sense).

The data thus support Hypothesis 3d, in that the two language groups show similar extent of use of paraphrase in their attempts to provide a sensible answer. The Spanish-speakers were ahead of the Russian-speakers (52% to 35%) in making the non-verbs or multi-word items fit the context. In Chapter 5, I will discuss the quality of these responses.

Summary of findings

The current study has shown a moderate correlation overall between the Collocation Test scores and Length of Residence, with widely differing correlation coefficients for low, mid and high levels of LOR when analyzed separately. The subgroup analysis shows no

difference in overall level of performance on the Collocation Test for intermediate students of Russian- and Spanish-speaking backgrounds. The qualitative analysis showed that the Russian-speakers were more likely to attempt an item, and the Spanish-speakers were more likely to fall back on L1. The two groups showed similar tendencies to use very basic words and to use paraphrasing.

Discussion

This chapter begins with a discussion of the whole group results. I draw comparisons to the work of other researchers and discuss interesting points of the present data. I then turn to the results of the small group analysis, with an in-depth treatment of the approach to collocation by the two language groups. Following this there is a set of theoretical conclusions from the study, and statements as to the study's contribution to the field. I end with suggestions for future research and applications for the classroom.

The Large Group

For the group overall, the current data support the original question, in that they show collocation skill does increase over time. Greater LOR is associated with increased CT scores. For a school-based population, this is hardly surprising, as English proficiency in general would be expected to increase over time, and collocation is part of that proficiency. What is remarkable though is that the correlation is only to a moderate degree (Pearson's $r = 0.4577$).

The reason that the correlation is not higher is because of the very wide range of CT scores. We have more recent arrivals outscoring the long-term residents; there are very low scores over most of the LOR range; the highest scores fall in both mid-term and long-term brackets. This very wide range of CT scores supports and reinforces results from other

studies, where we saw wide diversity of collocation skill, even within the same – supposedly leveled – ESL class (Bonci, 2002; Howarth, 1998).

What the research suggests is that school-based exposure to English does not in itself guarantee acquisition of collocation. Some other factor must be responsible for the range of collocation skill. Maybe it is a characteristic such as motivation, attentiveness, or language awareness that gives one student a keen sense of collocates, while another student - who arrived the same month – makes do with getting his point across.

The breakdown of data into low, mid, and high LOR reveals an interesting pattern as to how collocation skill develops over time. Students with 3 or more years residence in the USA showed a far higher correlation between LOR and CT (0.5495), suggesting that gains in collocation skill at this point in time are steadier and more even. Students with less than 18 months residence have a very low correlation (0.2350), suggesting a wide range of collocation skill at this level. However, it is the scores of the mid-range LOR students that are of greatest interest: these students showed a virtual absence of relationship between CT score and LOR (0.0199).

This absence of relationship in those crucial mid-range months suggests performance 'all over the board'. We need to ask why some students are still at the same levels as the recent arrivals, while others have zoomed ahead to the levels of the long-term students. Is the school system simply not identifying students low in this aspect of English? Is the curriculum at fault? Does the system not 'care' about collocation, either because professionals know too little about it, or because its importance is still a mystery?

The 18 – 36 month time interval is crucial in an ESL program. It is typically the

time frame in which an ESL student is no longer exempt from standardized or benchmark testing along with mainstream peers (Oregon Department of Education, 2005). The upper limit of this time interval is frequently the point at which major mainstreaming is expected of an ESL student. The [Washington State] Office of Superintendent of Public Instruction (2005) states in its yearly report:

The state program is intended to provide temporary support services for up to three years,ELL students may be served in the TBIP program for longer than three years if school districts can document that the students remain limited English proficient (p.33)

We need to pay particular attention to ESL students in this crucial time frame. We need to ensure that all our ESL students in that mid-range LOR have access to fluent, natural sounding English that some of them seem to acquire readily and early on. We need to be on the alert for students that are not acquiring such fluency. Otherwise we may be dooming some of them to the 'intermediate plateau' that Lewis (2000) describes. They need and deserve more in their second language learning.

This non-linear development of collocation skill reflects what was seen in the literature review: the intermediate stage of language acquisition is particularly problematical. It would appear to be that collocation does not increase step by step with other measures of language proficiency, and that some students are definitely being 'left behind' in this area. Students may have adequate vocabulary, syntax and can function in a survival mode, and can perform well enough to be promoted to the next level in an ESL sequence. But for many students collocation skill lags behind.

At first glance, the collocation scores themselves may also seem depressingly low. The highest score by any student was 36 out of 50 and there were two scores of 0. However, the current data, Bonk's (2000) data and my pilot studies suggest that there is a developmental trend in collocation skill.

The present study shows increasing means (10.38, 14.90 and 17.98) for low, mid-level and high LOR respectively. Bonk's (2000) data gave an average score for ESL undergraduate students of 25. In my pilot studies, former ESL students (exited from the program) scored in the low 30's, and native English speakers in secondary schools scored on average in the low 40's. Adult native English speakers (college students or graduates) scored 45 or higher. If the test were repeated over time, it would be expected that scores for ESL students would improve: whether and how they ever reach 'native' levels is not known, but would be of great interest.

The Small Group

The small group analysis showed that Russian- and Spanish-speakers did not differ statistically in their overall performance. Although this result is not surprising, it is gratifying. We do not have to treat different language groups differently, or hold different expectations. Collocation skill is not related to linguistic closeness, it is not 'harder' for speakers with a linguistically remote L1.

The Collocation Test scores obtained appear to be fair measures for the two language groups. Analysis of the test items by the researcher and four other ESL teachers did not find any question that unduly favored one language group over the other, in terms of vocabulary or cultural familiarity.

The qualitative analysis of responses to section A shows that a variety of collocations errors are made by both the Russian-speaking and the Spanish-speaking students when confronted with an unknown collocation. Russian-speaking students tended to make more synonym errors, and Spanish-speakers tended to make more random errors. Many more items were left blank by the Spanish-speakers. Grammar errors were approximately equal for the two small groups. This suggests that the two language groups have different approaches to dealing with collocation. The next section discusses four aspects of approach to collocation.

The first aspect of approach to collocation is whether a student attempts an item. To what extent do the two language groups have a stab at an answer, especially when the answer is not known?

There were 73 blank responses that were distributed across items and students. We have seen that Spanish-speaking students had almost twice as many blanks as Russian-speaking students (46 to 27). Many of the blanks by the Spanish-speakers occurred toward the end of the test. This suggests that the student had insufficient time to complete the section. This may have been because the student worked more slowly, and did not get round to later items before time was called. However, as most of the 143 students finished working each section within the time allotted, the slowness probably reflects less knowledge of collocation itself.

There were also many blanks in mid-test. These items were passed over presumably because the collocation was not known. In general, the Russian-speaking students passed over fewer items, that is, even if a collocation was unknown, they were more likely to fill in

the blank.

A second aspect of approach to collocation is the quality of answer. How broadly does each language group search? Do they put down a very general term ('do', 'go') or look for something more precise? Do they translate based on mother tongue usage?

The Russian-speakers seemed to make a broader search of their mental lexicons, when looking for an answer that would fit. This is seen in their willingness to complete most of the questions, their lesser reliance on L1 transfer, and their higher number of synonym responses. There were several individual answers by Russian-speakers that pointed toward thoughtfulness of response: *draw lots of pictures* (*c.f.* the Russian proclivity in Art), *working a diary, untie your shoes*, and – a truly delightful response - *the meeting will be short!* By contrast, the Spanish-speakers seemed to make a narrower search, tending to prefer a basic item, or relying strongly on L1 transfer, or leaving an answer blank.

A third aspect is the degree to which students relate their response to something in the question. In this study, this is overwhelmingly the case. The number of (correct + grammatical + synonym responses) outweigh the number of (random + blanks) 2:1. In addition, many of the random responses were understandable and context-related. The students appear to have cooperated with the testing, to the point of trying to find an appropriate answer from their store of English.

Between them, the Russian- and the Spanish-speakers only produced 18 responses which did not have any (apparent) connection to the item. The Russian-speakers produced 7 of these: discuse one's appetite (Q4), like no difference (Q8), wright the meeting (Q9), eating under and very a diary (Q10), and watch a favor (Q14). The Spanish-speakers gave:

I'm and *am* pictures (Q3), *late the risk* (Q5), *paited attention* (Q7), *select* and *in my bed a diary* (Q10), *cat* and *cafeteria a nap* (Q11), *eat a decision* (Q12) and *xo* and *before a spectacle of oneself* (Q15). These bizarre responses account for only 3.5% of the total responses to Section A by the subgroup. It concords well with Bonk's (2000) study, where students entered "semantically appropriate responses the great majority of the time" (p.25).

Individual follow-up might have generated some explanations of the bizarre responses: for example, was '*cat a nap*' an incomplete rendering of '*catch a nap*'? Was '*paited attention*' an inaccurate past tense rendering of the (also incorrect) '*put attention*'? It would also be interesting to know why the students involved did not leave these 18 items blank.

There were also responses which indicated broader, almost lateral, thinking on the part of certain students. Even though the given collocate was not correct, the responses showed a surprising depth of vocabulary. *Untie your feet* must have come from the connection of feet/shoes/laces. *Ashamed a spectacle of himself* showed the student had picked up on the negative connotations of the first half of the sentence. The many responses of *at the risk of losing....* suggest the students had previous acquaintance with the adjectival phrase *at risk*. This breadth of English knowledge bodes well for the teaching and acquisition of the actual collocations.

A fourth aspect of approach to collocation is technical quality of response. Many of the errors were foreign sounding, the type of error not found in a native speaker. Particularly problematical was the control over verb form, especially the use of the present or basic (dictionary) form of the verb when another form was needed. This section also considers

control over spelling and control over part of speech.

The 25 grammatical errors in the small group were of the following types: there were 15 uses of the present tense in place of a past, 2 uses of a past tense in place of a present, 1 incorrectly formed past tense ("*taked*"), 1 use of a gerund instead of an infinitive and 1 incorrect negative ("*it doesn't makes*"). In addition, there were 92 responses from Russian-speakers and 99 from Spanish-speakers using the wrong form of verb (whether or not they had the correct collocate). In questions requiring a past tense answer, students differed widely in whether they even attempted a past tense. Question 9, which required a passive construction, was especially difficult. No student was able to come up with a passive. Question 15, which required a gerund, had only 3 students give this form.

This analysis of verb tense usage is not surprising in an ESL context. Even if they have been formally taught the structures, students tend to stick to the familiar (the basic form of the verb) when producing language, until they are developmentally ready for something more complex (Thornbury, 1999). It also seems that students in the present study were focusing their time and energy on vocabulary rather than syntax. We are reminded that professionals would do well to work fully on collocations, so that vocabulary and syntax are both intact.

Control over spelling however was much sounder. In Section A the subgroup had only 16 misspelled answers out of 510. This figure indicates that the students knew the word they wrote, whether or not it was the correct collocate.

Finally there is the matter of control over part of speech. There were 50 answers that were not single-word verbs, and of these only 22 gave an answer which fit the context. This

is the ability to manipulate language, to get by with what one does know, to give a reasonable alternative. However, in the current study, many were not successful substitutions: students used a preposition or a noun that could not possibly fit. Possibly these items were not understood by the students. On the other hand, if the items were understood by the student, we have an absence of self-correction on the part of the student. This self-correction is something that could reasonably be expected after 18 months of English study.

Looking at the overall performance on Section A, the subgroup had only 24% correct. A superficial glance at this statistic does not look exciting. However, a more encouraging observation is the fact that there were so many grammar and synonym errors. Grammar errors indicate that the student knew the collocation, but could not quite come up with the right form. Synonym errors indicate an understanding of the item. The student was substituting appropriately for a specific word he did not have access to at the time.

When we consider the numbers of correct and 'nearly correct' (grammar and synonym errors) we have 66% ($72 + 11 + 85 = 168/255$) for the Russians and 54% ($51 + 14 + 73 = 138/255$) for the Spanish-speakers. These percentages reflect two facts: the students understood over half of the items, and the students were successful in communication. In itself, the level of reading comprehension is gratifying. Section A did not lend itself to correct guessing of answers. The Russians in particular understood enough to respond correctly or 'nearly correctly' in 2/3 of cases.

They also expressed themselves well. It may not have been idiomatic or the most natural English, but the students found the words to express what needed to be said on

average 60% of the time. Communication, the major goal of language learning, occurred.

Theoretical Implications

Several theoretical implications can be drawn from the current study. Firstly, collocation in general is problematical for immigrant ESL students, as it has been shown to be in foreign language situations. This is so, even though there is greater exposure to the target language, and more pressure from the educational community to become fluent. Collocation seems to be particularly difficult for many intermediate students.

Secondly, many of the students are functioning in a survival mode: they appeared to understand the context and could express themselves, often appropriately, but could not get the collocation quite right. Their language endeavors concentrated on successful communication rather than on native-like fluency.

Thirdly, collocation skills seem to be equally difficult for speakers of different L1s. L1 transfer plays a surprisingly minor role in collocation for students at this age level and in this immigrant language learning situation.

Contributions to the Field

The current study has contributed a large amount of data on collocation skill. The number of students studied was far larger than in most comparable studies. It also included a wider range of ESL levels. The present study was also unique in its comparison of Russian and Spanish language groups.

This study used LOR rather than an independent measure of English ability, against which to compare collocation skill. This preempts any charge of circularity in design, and

is a strong point of the current study. Any future study that attempts to measure collocation skill at the various ESL proficiency levels must take great care to have proficiency measured in a way that does not include collocation.

The present study did not control for educational background of the students or parents. It is possible that some students had significant gaps in their educational history, prior to coming to the USA. This was not considered a major factor in the present study, as the academic demands of the Collocation Test were small and all students had received continuous schooling as per the LOR score. However, a future study might wish to investigate this factor.

Future Research

Future research could profitably concentrate on longitudinal work on the development of collocation skill within the individual. Ideally this should be combined with an analysis of coping strategies. Of greatest use would be a study over 5 years (or more), which would follow the subject through several proficiency levels and see if and how that individual reached native levels.

Another area of interest for future investigation is to track collocation skill along with other aspects of English proficiency, such as vocabulary size and quality.

A group that has not been well studied is former ESL students, those exited from ESL programs because they have met criterion. It would be useful to know if part of that 'criterion' included native-like handling of collocations.

A final area of research possibilities is investigating the effectiveness or otherwise of deliberate teaching of collocation skill. What does a teacher need to do on the classroom

for students to gain in collocation skill? Does the teaching of collocation skill have results in other areas, such as writing?

Classroom Applications

A major classroom application is for collocation to become part of the curriculum. This does not so much involve a new syllabus or updated course books, but rather the change in the teacher's mind regarding vocabulary. Instead of single words, there should be a focus on lexis: multi-word utterances, showing typical and successful word combinations.

This study points to the importance of including collocation work daily in the ESL classroom. Collocation analysis, collocation games, collocation practice are all appropriate activities. Every classroom should contain and use a Collocations Dictionary. There should also be contrastive work: students being taught to compare collocations in their own language with what they now need to say in the target language. Another learning area is noting which words do NOT collocate with a given word: Lewis (2000) gives a diacritical system for denoting these pairs. Another aspect is review: returning to words the students already know, and studying their other uses and their other collocations.

This study lends support to Lewis' (2000) proposal that we teach beginners and intermediates in fundamentally different ways. Beginners should be primarily exposed to the language, particularly via listening comprehension, whereas intermediates should be encouraged to deepen their existing knowledge. Collocation fits very well in this deepening.

Perhaps the major implication of the study is the plea to teachers not to be content with a survival level of English as a Second Language. Communication is of course important, but it is not enough for the student to just get his point across. Especially in an

immigrant ESL situation, we are aiming to mainstream the student, to have him speak fluently and be as literate as his native-born classmate. Collocation is a skill that will help him get to this level – let's use it!

References

- American Council of Teachers of Foreign Languages. (1986). *Defining and developing proficiency: Implementation and concepts*. New York: ACTFL.
- Bahns, J. (1993). Lexical collocations: a contrastive view. *English Language Teaching Journal*. 47(1), 56 – 63.
- Bahns, J. & Eldaw, M. (1993). Should we teach ESL students collocations? *System*, 21(1), 101 - 114.
- Beauchesne, J. (2001). *Dictionnaire des cooccurrences* [Collocations dictionary]. Montreal, Quebec: Guérin.
- Benson, M. & Benson, E. (1993). *Russian-English Dictionary of Verbal Collocations*. Philadelphia: John Benjamins Publishing Company.
- Benson, M., Benson, E. & Ilson, R. (1997). *The BBI Dictionary of English Word Combinations*. (revised) Philadelphia: John Benjamins Publishing Company.
- Berg, B. L. (2004). *Qualitative research methods for the social sciences* (5th ed.). Boston: Allyn and Bacon.
- Best, J. & Kahn, J. (1986). *Research in Education* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Biskup, D. (1992). L1 influence on learners' renderings of English collocations: A Polish/German empirical study. In P.Arnaud & H. Bejoint, (Eds.), *Vocabulary and Applied Linguistics* (pp.85-93). London, UK: Macmillan.
- Bliss, A. (2004). *Lexical Collocations in Academic ESL Reading Textbooks*. Unpublished master's thesis, Portland State University, Portland OR.

- Bonci, A. (2002). Collocational restrictions in Italian as a Second Language: a case control study. *Tuttitalia*, 26, 3 – 14.
- Bonk, W. (2000). Testing ESL learners' knowledge of collocations. ERIC Document ED442309. Retrieved July 7, 2004 from ERIC database.
- Bosque, I (2004). *REDES: Diccionario combinatorio del español contemporáneo* [REDES Dictionary of Combinations in Contemporary Spanish]. Madrid, Spain: Ediciones SM.
- Bruner, J. (1990). *Acts of Meaning*. Cambridge, MA: Harvard University Press.
- COBUILD (1990). COBUILD Corpus. University of Birmingham, UK. Available from <http://titania.cobuild.collins.co.uk/form.html>
- Davis, P. & Kryszewska, H. (2003). Chunking for beginners: Some recipes. *Humanising Language Teaching*, March 2003. Available from <http://www.hltmag.co.uk>
- DeAvila, E. & Duncan, S. (1990). *Language Assessment Scales*. Monterey, CA: CTB McGraw-Hill.
- Fargal, M & Obiedat, H. (1995). Collocations: a neglected variable in EFL. *IRAL*, 33(4), 315-331.
- Firth, J. (1957). *Papers in Linguistics*. London, UK: Oxford University Press.
- Gitsaki, C. (1999). *Second language lexical acquisition: A study of the development of collocational knowledge*. Bethesda, MD: International Scholars Publications.
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A.P. Cowie (ed). *Phraseology – Theory, analysis and applications*. Oxford: Clarendon Press.

- Hasselgren, A. (1994). Lexical teddy bears and advanced learners: a study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, 4 (2), 237-260.
- Hawkins, E. (1999). Foreign language study and language awareness. *Language Awareness*, 8 (3 & 4), 124-142.
- Hill, J. & M. Lewis, eds. (1997). *Dictionary of selected collocations*. Hove, UK: Language Teaching Publications.
- Howarth, P. (1998). The phraseology of learners' academic writing. In A.P. Cowie (ed). *Phraseology – Theory, analysis and applications*. Oxford, UK: Clarendon Press.
- Kavaliauskienė, G. & Janulevičienė, V. (2001). Using the lexical approach for the acquisition of ESP vocabulary. *The Internet TESL Journal*, 7 (3). Retrieved July 7, 2004 from <http://iteslj.org/>
- Krashen, S. & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Oxford University Press.
- Lewis, M. (1993). *The lexical approach*. Hove, UK: Language Teaching Publications.
- Lewis, M. (1997). *Implementing the lexical approach*. Hove, UK: Language Teaching Publications.
- Lewis, M. (2000). *Teaching collocation*. Hove, UK: Language Teaching Publications.
- Lindstromberg, S. (2003). My goodbye to the Lexical Approach. *Humanising Language Teaching*, March 2003. Available from <http://www.hltmag.co.uk>

- Macmillan English dictionary for advanced learners of American English.* (2002). Oxford, UK: Macmillan.
- Macmillan. (2005). *One Stop English Professional Support*. Available from
<http://www.onestopenglish.com>
- Morgan, J. & Rinvolucri, M. (2000). *Vocabulary* (2nd ed). Oxford, UK: Oxford University Press.
- Nation, I. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nattinger, J. & DeCarico, J. (1992). *Lexical phrases and language learning*. Oxford, UK: Oxford University Press.
- Oregon Department of Education. (2005). Assessing New Limited English Proficient Students: executive memorandum. Retrieved October 30, 2005 from
http://www.ode.state.or.us/pubs/memos/2003_04/029_04.htm
- Office of Superintendent of Public Instruction [Washington State]. (2005). Educating English Language Learners in Washington State, School Year 2003-2004: Annual report of the State Transitional Bilingual Instruction Program. Available from
<http://www.k12.wa.us/>
- Oxford collocations dictionary for students of English.* (2002). Oxford, UK: Oxford University Press.
- Pawley, A. & Snyder, F. (1983). Two puzzles for linguistics theory: nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (eds.) *Language and Communication* (pp. 191-226). London, UK: Longman.

- Pilgrims Language Schools. (2005). *Humanising Language Teaching*. Available at
<http://www.hltmag.co.uk>
- Porto, M. (1998). Lexical phrases and language teaching. *Forum*, 36 (3) 22-26. Available at: <http://www.eli.ubc.ca/teachers/journals.html>
- Renouf, A. & Sinclair, J. (1991). Collocational frameworks in English. In K. Aijmer & B. Altenberg (eds) (1991) *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. (pp. 95-110) London, UK: Longman.
- Richard-Amato, P. (2003). *Making it Happen: from Interactive to Participatory Language Teaching* (3rd. ed). White Plains, NY: Pearson Education, Inc.
- Ridley, J. (1991). Strategic competence in second language performance: A study of four advanced learners. *Center for Language and Communication Studies: Occasional Paper* #28. Trinity College Dublin (Ireland): CLCS.
- Stockdale, J. (2000). *Mr. Stockdale's dictionary of collocations*. Lingvo DSL. ERIC Abstract ED447693. Retrieved July 7, 2004 from ERIC database.
- Stockdale, J. (2004). Definition plus collocation in vocabulary teaching and learning. *The Internet TESOL Journal*, 10(5). Available from <http://iteslj.org/>
- Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford, UK: Blackwell.
- Taiwo, R. (2004). Helping ESL learners to minimize collocational errors. *The Internet TESOL Journal*, 10(4), April 2004. Available from <http://iteslj.org/>
- Thornbury, S. (1999). *How to teach grammar*. Harlow, UK: Longman.

- Thornbury, S. (2002). *How to teach vocabulary*. Harlow, UK: Longman.
- Waring, R. (1997). Review: Michael Lewis 'Implementing the Lexical Approach'. Retrieved October 30, 2005 from University of Wales Swansea, Vocabulary Acquisition Research Group, Virtual Library Web site:
<http://www.swan.ac.uk/cals/calsres/vlibrary/rw97a.htm>
- Williams, B. (2002). Collocation with advanced levels. In *English*. Autumn 2002. Available from <http://www.britishcouncil.org>
- Willis, J. & Willis, D. (1996). *Challenge and Change in Language Teaching*. Oxford, UK: Heinemann.
- Woolard, G. (2004) *Key words for fluency*. London, UK: Heinle.
- Zughoul, M.R. & Abdul-Fattah, H.S. (2003). Collocational strategies of Arab learners of English: a study in lexical semantics. ERIC document ED479650. Retrieved July 7, 2004 from ERIC database.

Appendix A Collocation Test

Section A. Put the **verb**, which best completes the sentence, in the blank.

EXAMPLE: He was accused of spreading rumors about the other employees.

1. I'm sorry, but I can't give it to you at that price: I'm trying to _____ a business here, not a charity organization.
2. She _____ the subject every time I asked her about her boyfriend; she obviously didn't want to talk about him.
3. I _____ lots of pictures on my vacation, but none of them came out nice.
4. Don't eat chocolate before dinner, it'll _____ your appetite.
5. If you don't make up back copies of all the files on your computer, you'll be _____ the risk of losing all your data during a power failure.
6. I _____ a bad dream last night.
7. If I want a better grade, I will have to _____ attention to what the teacher is saying.
8. It doesn't _____ any difference if you go left or right here, the two streets will meet again after only 1 mile.
9. The meeting will be _____ on Friday in room 1302.
10. I've been _____ a diary for over ten years now; I love to look back at it now and see what I was thinking about.
11. I like to _____ a nap every afternoon after lunch.
12. My spouse would never _____ such a serious decision without first discussing it with me.
13. We all _____ the feeling that she didn't really want to come to our party.
14. She _____ me a big favor and took care of the kids for the afternoon when I had to go to the doctor's office.
15. Whenever he has to sing a solo, Max ends up _____ a spectacle of himself.
16. Please _____ your feet before you come into the house, your shoes look dirty.
17. Orimco, the company that made these chemicals, _____ out of business a year ago.

Section B. Fill in the blanks with **prepositions** to complete the sentences. The meaning of the expression is in parentheses at the end of the line.

EXAMPLE: She had a lot of free time and decided to take up piano lessons.

(to start)

1. We ordered steak and two appetizers, and the bill came up \$78. (to total)
2. Don't worry, you can depend on me! (to trust)
3. I'm sorry, I can't drive you all the way to school, but I can drop you off at the bus stop if you like. (to drive someone to a place)
4. It's taken me more than a month to get over this cold. (to recover)
5. On Independence Day, they set off so many fireworks the sky almost disappeared! (to make explode)
6. The criminals who held up the Bank of Hawaii last week still have not been caught. (to rob)
7. Sylvester Stallone's new movie is scheduled to come out on April 19. (to appear)
8. My doctor told me I had to give up alcohol and cigarettes if I wanted to stay healthy. (to stop using)
9. Patty was Ron's girlfriend for a long time, but they broke up two weeks ago. (to end a relationship)
10. Many of the birds in the area were killed off by local hunters. (to exterminate)
11. I've finished writing up the minutes of the last meeting, would you please look them over and see if I've made any mistakes? (to check)
12. We've been talking about this for over an hour now; let's move on to another topic. (to go)
13. My father never studied English formally, he just picked it up when he came here ten years ago. (to learn)
14. Other kids always pick on her because she's so overweight. (to tease)
15. He takes after his father, he has the same hair and eyes. (to look like)
16. She told me that she wanted to go to Alaska and work on a fishing boat for the summer, but I talked her out of it. (to convince not to do)
17. Grandfather has been feeling very sad, let's go to his house and cheer him up. (to make happier)

Section C. Each of the four sentences is using the underlined verb in a different way. One of them is not really a correct usage of that word. **Circle the letter corresponding to the least acceptable sentence.**

EXAMPLE: a. I am writing my annual letter.

- b. Peter wrote the university entrance exam yesterday.
- c. The toddlers got black markers and wrote on the walls.
- d. You must write on the lines.

1. a. You are covered by your doctor's diagnosis.
 b. Another soldier covered Peter with a machine gun as he started forward.
 c. With this insurance policy, we will cover you in case of any accident.
 d. Wow, we've been driving fast! We've covered almost three hundred miles in one day!
2. a. After his wife divorced him and his son died last year, Mark took a heart attack.
 b. Our plane took off two hours late due to heavy snow on the runways.
 c. The robbery took place at about 3:30 a.m.
 d. Children have to be taught how to take turns, or they will continue to be selfish.
3. a. Don't put the blame on me!
 b. Can you put your signature at the bottom of this letter, please?
 c. It's OK, I think I can put the risk.
 d. 25 convicted murderers were put to death in Texas last year.
4. a. I can see up your plans, you want to take my house away from me!
 b. Her whole family was at the train station to see her off.
 c. She saw to it that the animals were fed at the same time every day.
 d. Ever since I got hit on the head with that rock, I've been seeing double.
5. a. I didn't leave my apartment until after two o'clock.
 b. My Aunt Janice left me this beautiful tea set.
 c. Roberta left her husband for another man.
 d. Hitler left his mind in the days before his suicide.
6. a. Are the Johnsons throwing another party?
 b. She threw him the advertising concept to see if he liked it.
 c. The team from New Jersey was accused of throwing the game.
 d. The new information from the Singapore office threw the meeting into confusion.
7. a. Let's drop the subject. I don't want to talk about it any more.
 b. He kept dropping hints that he wanted a job, too.
 c. Bob said he would bring the computer over to my house, but I dropped his offer.
 d. She didn't like the professor, so she decided to drop the class.

Appendix A

8. a. I don't know if I can stand another winter in this place.
b. She stood a good chance of winning the race.
c. I've already explained where I stand on that issue.
d. He stood three roses in the vase.

9. a. They play the radio really loud, even late at night.
b. Money plays a large part in politics, especially when it comes to rich countries.
c. What movies are playing at the Cinemat tonight?
d. This book played great impact on the intellectuals of Eastern Europe.

10. a. I don't think you can pick a conclusion from this data.
b. Tom is always trying to pick fights with Ronnie.
c. His pockets were picked while standing in line at the theater.
d. Picking berries in the woods is a great way to spend a Sunday afternoon.

11. a. They got big success when they finally discovered how to market their products in Japan.
b. If you get good results from this medication, let me know.
c. He got cancer as a result of his thirty years of smoking cigarettes.
d. We got to New York on the same day that the big snowstorm began.

12. a. He didn't feel that a videocassette would make a very good present, so we bought something else.
b. I made the table while Mom finished cooking dinner.
c. Seven and nine make sixteen.
d. The train crossing signals are starting to flash - do you think we can make it?

13. a. I'm not sure if they do repairs in that shop or not.
b. We did our best, but it wasn't good enough to win.
c. What profession does he do?
d. She didn't do very well on the test.

14. a. We're going to run some tests on your mother to see if the accident affected her brain.
b. All the newspapers in the country ran the story on their front pages.
c. A car ran over my cat yesterday.
d. You need to run more vitamins and minerals in your diet.

15. a. Could you please move your car? It's blocking my driveway.
b. I was very moved by her latest film.
c. The situation seems very bad now for you, but it may move any time.
d. The neighborhood was getting too dangerous, so they moved.

16. a. Bob pulled a terrible trick with my computer, and now I can't get it to work again.
b. A TV crew pulled up right outside my apartment last night!
c. I want to make sure that everyone in this company is pulling their own weight.
d. She's never pulled anything like this before.

Appendix B Answer key

Section A

1. run

2. avoided, changed

3. took

4. spoil, ruin

5. running

6. had

7. pay

8. make

9. held

10. keeping

11. take

12. make

13. got

14. did

15. making

16. wipe

17. went

Section B

1. to

2. on

3. off

4. over

5. off

6. up

7. out

8. up

9. up

10. off

11. over

12. on

13. up

14. on

15. after

16. out

17. up

Section C

1. a

2. a

3. c

4. a

5. d

6. b

7. c

8. d

9. d

10. a

11. a

12. b

13. c

14. d

15. c

16. a

Appendix C

Language backgrounds

Language	Number of speakers	Percent of total
Russian/Ukrainian	71	50%
Spanish	38	27%
Romanian	11	8%
Chinese	8	6%
Arabic	3	2%
Khmer	2	1%
Swahili	2	1%
Vietnamese	2	1%
Chamorro	1	<1%
Chuukese	1	<1%
French/Creole	1	<1%
Korean	1	<1%
Portuguese	1	<1%
Serbo-Croatian	1	<1%

Appendix D

Length of Residence

Length of Residence	Number of students	Percent of total
0 – 5 months	0	0%
6 – 18 months	40	28%
19 – 36 months	50	35%
37 months +	53	37%

Appendix E

Demographic sheet

Collocation Investigation, Mrs. Claire Smith: George Fox University/Vancouver School District

Name: _____ male/female _____

School: _____ grade: _____

Student ID #: _____

Date of birth: (month) _____ / (day) _____ / (year) 19_____

Today's date: (month) _____ / (day) _____ / (year) 20_____

What is your native language(s): _____

What language(s) do you speak at home? _____

Were you born in the United States of America? Yes _____ / No _____

If No, date of arrival in USA: (month) _____ / (year) _____

Appendix F Administration instructions

Good morning! I am (*name*) and I work at (*school or district*). I am conducting a research experiment at George Fox University on how people learn vocabulary in a second language. I am interested in how people who come to America from other countries learn English. Thank you for being part of my experiment. My contact information is on the board, if you decide you want more details of what I find out.

The first sheet is white. I need some details about you; these will be kept confidential. (*hand out the white demographic sheets*) Please fill in the blanks. Raise your hand if you have a question. (*pause*) Please leave the white sheet turned over on your desk.

Next we have three different subtests, which we will do one after the other. All the questions are in English, and you will answer in English. We should be finished by (*30 minutes time*). (*hand out blue-colored Section A*)

On the blue sheet, please fill in today's date at the top. (*pause*) There are 17 items in this section. For each item, please fill in the blank with the VERB that best completes the sentence, for example: (*read example at top of Section A*). Any questions? (*pause*) Please start. (*allow 8 minutes; if everyone is through earlier, proceed*) Please leave the blue sheet turned over on your desk. (*hand out yellow-colored Section B*)

The yellow sheet is similar, but this time you fill in the blanks with a preposition, such as: ON, IN, TO, UP. Please put the date at the top. There are 17 items in this section. Any questions? (*pause*) Please start. (*allow 8 minutes; if everyone is through earlier, proceed*) Please leave the yellow sheet turned over on your desk with the blue sheet. (*hand out pink-colored Section C*)

The pink section has sixteen items. Please put the date at the top. This section is different, so please follow as I read the instructions. (*read instructions at the top of Section C*) So, for each question chose the sentence which is NOT idiomatic, or good, English. Any questions? (*pause*) Please start. (*allow 10 minutes*) Please leave the pink section with the blue and the yellow sheets. I will come round and collect all four sections from you. Again, thank you for participating.

Appendix G

Collocations: Russian & Spanish equivalents

	English collocation	Exact translation of Russian equivalent	Exact translation of Spanish equivalent
1	To run a business	To manage/direct a business	To carry/direct a business
2	To change/avoid the subject	To avoid the question	To turn the sheet/page
3	To take a photograph	To photograph	To take out a photograph
4	To ruin/spoil one's appetite	To damage/ruin appetite	To take away from one the appetite
5	To run the risk	To risk	To run a risk/run with the risk
6	To have a dream	To see a dream	To dream with
7	To pay attention	To turn attention	To put/lend attention
8	To not matter	To be no difference	To be same
9	To hold a meeting	To lead a meeting	To happen/ 'realize' (meeting as subject)

Appendix G

10	To keep a diary	To lead/keep a diary	To carry/write a diary
11	To take a nap	To nap	To throw oneself a siesta
12	To make a decision	To take/accept a decision	To make a decision
13	To get the feeling	To feel	To feel
14	To do someone a favor	To show someone a favor	To make/do someone a favor
15	To make a spectacle of oneself	To play a spectacle	To put oneself a spectacle
16	To wipe one's feet	To dry/wipe one's feet	To clean one's shoes
17	To go out of business	To go out of/close one's business	To close the business

Appendix H

Paraphrased answers by language group

	Russian response	semantic fit ?	Spanish response	semantic fit ?
Q2			in	No
Q4			make lost	Yes (1)
Q5	at (x3), in (x2), in danger of slow	Yes (6) No	In, at, on late	Yes (3) No
Q7	turn my	Yes (1)	school	No
Q9	short wright, at, day	Yes (1) No	8:00 p.m., cool, first, in (name of school), interesting, in school take place	Yes (7)
Q10	under, very	No	in my bed	No
Q11			cat, cafeteria	No
Q14			likes to have, did for	No
Q15	with (x6), to, in	No	as, ashamed before, xo, with (x2)	Yes (2) No
Q16			don't put	No
Q17	with	No (1)	have been	Yes (1)
total	23	8	27	14