This is to certify that the
thesis entitled

A CORPUS-BASED APPROACH TO ESL:
TEXTBOOK AND MATERIALS DEVELOPMENT AND
EVALUATION

presented by

Angela Komsic Super

has been accepted towards fulfillment
of the requirements for the

M.A. degree in TESOL

Debra M. Hardison
Major Professor's Signature

April 22, 2004
Date

MSU is an Affirmative Action/Equal Opportunity Institution

DATE DUE	DATE DUE	DATE DUE
MAR 1 2 2007		
0917 07		
JUL 2 7 2009		
0605 11		

A CORPUS-BASED APPROACH TO ESL: TEXTBOOK AND MATERIALS DEVELOPMENT AND EVALUATION

By

Angela Komsic Super

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Linguistics

2004

ABSTRACT

A CORPUS-BASED APPROACH TO ESL:
TEXTBOOK AND MATERIALS DEVELOPMENT AND EVALUATION

By

Angela Komsic Super

Taking both the advantages and limitations of using a corpus-based approach in language learning into account, I make the claim that corpora are most suitably utilized only if they are used for the purpose of curriculum and textbook development and evaluation. I focus this claim by providing the results of three studies that I have conducted.

The first study provides evidence for one way that corpora may most effectively be utilized (based on empirical data) for the purpose of textbook evaluation. The second study focuses on the ways in which ESL teachers and researchers may develop textbooks based on information (e.g. idioms) gathered from simple corpus concordance searching in relevant corpora, and how to most effectively transform basic search results into communicative teaching materials. The third study focuses on the ways in which a corpus-based approach may be used for the purpose of materials development in an ESP classroom.

Each study is introduced with justifications based on prior research done on communicative teaching methods in ESL. Each study is concluded with further implications and elaboration on how language teachers can easily and most effectively use corpora in their classrooms.

ACKNOWLEDGEMENTS

I would like to thank my primary advisor and professor, Dr. Debra Hardison, for her constant encouragement and guidance, without which this thesis would not have been possible. Dr. Hardison's dedication to research and to students' interests is truly admirable and greatly appreciated.

I would also like to thank my second committee member and professor, Dr. Charlene Polio, for encouraging me to explore my interests and giving me the initial idea to write a thesis in the first place.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (cont'd)

INTRODUCTION

The purpose of this thesis is to provide various models and guidelines to illustrate the many ways in which a corpus-based approach can most effectively be utilized in the TESOL (Teaching English to Speakers of Other Languages) field in general, and in ESL (English as a Second Language) textbook and materials development and evaluation in particular.

While much literature is available in the field of corpus linguistics, there is a paucity of materials that illustrate how corpora may most effectively and efficiently be used for communicative purposes in the language-learning classroom. Research that has been conducted most recently has focused on the use of corpora in the classroom by demonstrating the many ways in which corpus-searching can benefit second language learners independently (outside of the classroom) by guiding language students and teachers to conduct extensive searches which may provide information about collocation and phrase patterns, collocation and synonymy, connotation and semantic prosody, syntax, cohesion in texts, metaphor, and frequency. While such research has created a useful base for the incorporation of corpora in the classroom, it also has its limitations since such extensive data-driven searching is not only time-consuming and not always accessible, but also reserved for only the most motivated of language learners and teachers.

Taking both the advantages and limitations of using a corpus-based approach in language learning into account, I conclude by making the claim that corpora are most suitably utilized only if they are used for the purpose of curriculum and textbook

development and evaluation. I focus this claim by providing the results of three studies that I have conducted which demonstrate 1.) The ways in which corpora may be used for textbook evaluation, 2.) The ways in which corpora may be used for textbook development, and 3.) The ways in which corpora may be used for ESP (English for Specific Purposes) materials development.

The first study provides evidence for one way that corpora may most effectively be utilized (based on empirical data) for the purpose of textbook evaluation. This study compared an academic speech textbook, *Discussion and Interaction in the Academic Community* (Madden & Rohlck, 2000), with a corpus of academic speech, the Michigan Corpus of Academic Spoken English (MICASE). Frequencies of expressions and collocations were compared, and the results suggested that the textbook was *not* in fact representative of academic speech, as it had previously claimed. The MICASE corpus, as well as the COBUILD corpus, were then utilized in order to update the textbook and provide more realistic data and content for the textbook authors.

The second study focuses on the ways in which ESL teachers and researchers may develop textbooks based on information gathered from simple corpus concordance searching in relevant corpora, and how to most effectively transform basic search results into communicative teaching materials. This study specifically looked at which idioms are most common in academic speech. Two other researchers and I read through the entire MICASE corpus and noted all idiomatic expressions based on our criteria for the term "idiom". What were found were 600 occurrences of idioms in the MICASE, providing empirical evidence of idioms in academic speech, should any researcher choose to develop a textbook focusing on this topic. This study provides a substantial

amount of evidence for language teachers in academic environments who would like to teach idioms, without relying solely on idiosyncratic intuition. Furthermore, this study sets a methodological framework as to the most efficient and productive searching techniques (as well as limitations) for language researchers choosing to create textbooks using an empirical corpus-based approach.

The third study focuses on the ways in which a corpus-based approach may be used for the purpose of materials development in an ESP classroom. In this study, I created a written corpus of writing taken from 14 native speaker graduate students in the M.A. TESOL program at Michigan State University for the purpose of providing a tool with which non-native speakers in the M.A. TESOL department may improve their writing in English in their graduate classes. Some of the sub-corpora that were created from the master corpus include various corpora for each core TESOL course. This study serves solely as a model, and introduces the many ways in which corpora may best be utilized to create a specific curriculum that best fits the needs of the students, as well as considerations and limitations that need to be taken into account if creating corpora for materials development (e.g. what data to collect, how to collect it, what to do with it, etc.).

Each study will be introduced with justifications based on prior research done on communicative teaching methods in ESL. Each study will be concluded with further implications on how language teachers can easily and most effectively use the methods described in their classrooms.

Overall, it is assumed that, if found effective, the methods studied could be adopted by many teachers without special training.

CHAPTER 1
BACKGROUND: CORPUS LINGUISTICS

Taking a 'Corpus-Based Approach'

Corpus linguistics is currently a growing discipline that is slowly integrating itself into various spheres of language analyses and applied linguistics. Corpus linguistics technology requires a computer that can store a collection of natural text files (the corpus), and then apply software to those files to produce frequency lists, lists of key words, and, most importantly, strings of words showing which words co-occur (or collocate) with others in natural language. Therefore, to take a 'corpus-based approach' means to utilize a corpus or various corpora as a means of conducting language research based on empirical data. A corpus allows one to analyze actual usage patterns in natural texts, thereby making extensive use of computers for analysis, saving much time on behalf of the researcher. Often it is possible for a computer to analyze language through simple programming methods, rather than spending extensive periods of time reading through texts and trying to find language patterns or real-world examples of language in various contexts. The text files in a corpus may consist entirely of written texts (as in the Helinski Corpus of English Texts), entirely of transcriptions of speech (as in MICASE – the Michigan Corpus of Academic Spoken English), or of both (as in the Bank of English / COBUILD).[1] These corpora are typically constructed on certain principles that lead to appropriate sampling, and they can vary greatly in size. The Bank of English corpus roughly contains over 400 million words from various sources (i.e. newspapers, magazines, books, etc.) taken from both American and British media, whereas small specialized corpora, especially those devoted to single genres, such as telephone calls or

British teen slang, can be significantly smaller. The pros and cons of large diffuse corpora and small narrow ones is a matter of current debate (Simpson & Swales, 2001). However, from a more global and historical perspective:

> Many of the developments in corpus linguistics over the last 15 years are due to the work of European scholars, with particularly active groups in the United Kingdom and Scandinavia. Despite the work of Douglas Biber at Northern Arizona, Michael Barlow at Rice University, and important corporist groups at the University of California, Santa Barbara, and the University of Pennsylvania, North America has generally lagged behind. For many years now, for example, the Europeans have had several annual or biennial conferences on corpus linguistics, under such acronyms as ICAME (International Computer Archive of Modern English) and TALC (Teaching and Language Corpora). In North America, the first national symposium devoted to this kind of linguistics was held at the University of Michigan in May 1999. (Simpson & Swales, 2001, p. 2)

The various areas of applied linguistics that are currently implementing a corpus-based approach are: lexicography, grammar analysis (e.g. syntax, semantics, pragmatics), sociolinguistics (e.g. dialects, registers, etc.), language acquisition, style (e.g. writing), and various pedagogical applications (e.g. language teaching).

Corpus Linguistics and Language Learning

Given the recent emergence of corpus linguistics in North America, one particular area that has just recently been gaining attention is the area of corpus linguistics and language teaching. There has been a paucity of materials that illustrate how corpora may most effectively and efficiently be used in the language learning classroom, and many language teachers are not even aware of the infinite possibilities that the corpus-approach may offer to language learners. Nevertheless, as technology in the classroom increases at a steady rate, the use of corpora may prove to be extremely useful in the development of language teaching materials in the future, given the appropriate guidance.

In the past 15 years of corpus linguistics and language learning research, there have been two general approaches to using corpus-based materials: Teachers can either analyze the corpora themselves for materials design or assessment, or they can decide to introduce them into the classroom and train students in their use. In the first case:

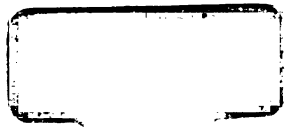
> ...teachers might use corpus-based investigations to (i) determine the most frequent patterns in a particular domain; (ii) enrich their knowledge of the language, perhaps in response to questions raised in the classroom' (iii) provide "authentic data" examples; and (iv) generate teaching materials. (Barlow, 1996, p. 30)

In the second case:

> ...teachers may also wish to have their students explore corpus materials, either in following a path of investigation determined by the teacher (so that the students come to understand a particular pattern of usage such as *say* versus *tell* or the collocations of *bright*) or in exploring an issue in a more open-ended way. (Barlow, 1996, p. 30)

Furthermore, one early study conducted by Johns (1991a, 1991b) has described a number of ways teachers can create materials and exercises for use in the classroom, and he has also developed a theory of what he calls "data driven learning" (DDL). The basis of this kind of learning is inductive acquisition on the part of students "through the process of analyzing the patterns of language use of specifically selected items as revealed through corpora" (Tribble & Jones, 1990; Johns, 1991a). Advocates of this method argue that this provides a more meaningful context in which to learn grammar (as opposed to simply being taught grammar rules) and appropriate word usage. Johns outlines three general effects of adopting the DDL approach:

> [It] can have a considerable influence on the process of language learning, stimulating enquiry and speculation on the part of the learner, and helping the learner also to develop the ability to see patterning in the target language and to form generalizations to account for that patterning. The second main effect of DDL is on the role of the teacher, who has to learn to become a director and



> coordinator of student-initiated research... The third... is a revelation of the place of grammar in language learning and language teaching... The DDL approach makes possible a new style of "grammatical consciousness raising" (Rutherford, 1987) by placing the learner's own discovery of grammar at the center of language learning. (Johns, 1991a, p. 2-3)

Taking the possibilities of data-driven language learning into account, Aston (1997) follows up by suggesting that approaches to using corpus resources in the classroom can be divided into two general areas, those for *reference* and those for *browsing*:

> On the one hand, they might be treated as a reference tool, which could be looked up to provide examples and therefore clarify doubts on particular problems which had arisen in other language activities. From this perspective the corpus could be seen as complementing the grammar, the dictionary, and the encyclopedia. On the other hand, a corpus might be treated as a course of activity in itself, a hypertext to be browsed in, where the user passes from one text or concordance to another, and where, rather then being determined by a preselected goal, progressive discoveries occur on a negotiated step-by-step basis. (as cited in Partington, 1997, p. 7)

Nevertheless, while it is true that such research has created a useful base for the incorporation of corpora in the classroom, it also has its limitations since such extensive data-driven searching is not only time-consuming and not always accessible, but also reserved for only the most motivated of language learners and teachers.

Some core researchers who have contributed to developing the area of corpus linguistics and language learning beyond DDL include Partington (1998), Biber, Conrad, Reppen (1998), and Kennedy (1998), who have focused their implementation of the use of corpora in the classroom by demonstrating the many ways in which corpus searching and concordancing can benefit second language learners independently (outside of the classroom) by guiding language students and teachers to conduct extensive searches which may provide information about collocation and phrase patterns, collocation and synonymy, connotation and semantic prosody, syntax, cohesion in texts, metaphor, and

frequency, all of which may be used to create language learning materials and textbooks. Collectively, their work has greatly contributed to corpus linguistics and language learning research being conducted today.

Partington's book, *Patterns and Meanings: Using Corpora for English Language Research and Teaching* (1998), was considered a breakthrough when it was first published since it bridged the gap between previous corpus-based studies and the possibility of taking a corpus-based approach for more communicative language learning. Partington focused his early research by going beyond data-driven language learning, and offering various demonstrations of how it is possible to interrogate a medium-sized corpus for purposes such as those described above (DDL), but also for purposes of enriching the teacher's awareness of language (e.g. with concordance searching), refining and developing syllabi and materials (e.g. developing authentic materials based on empirical data), as well as creating activities as a means of developing students' study skills and understanding of the target language (in this case, English). While these techniques had been speculated upon for years, this text was the first to publicly bring attention to instruction for language teachers when looking at various concordance searches and the illustration of how search results may easily be developed into general communicative language learning materials.

Another researcher who has contributed to the shift from DDL to using corpora to develop communicative language learning materials is Kennedy (1998). According to Kennedy (1998), there are a number of ways the language in a corpus can influence language-teaching pedagogies:

> However valuable may have been the focus on learning language as communication (particularly in increasing language fluency), there has been

> growing recognition that systemic accuracy is also relevant, and consequently there has been a more recent revival of interest among language teachers in what is being learned: the content of language teaching... Curriculum designers and classroom teachers need to have access to this information through better reference materials and syllabuses... First it can influence the content of language teaching by affecting the selection of what to teach, the sequencing of pedagogy, and the weight given to items or parts of the language being taught, thus contributing directly to the content of instruction. Secondly, through the consciousness-raising of teachers about language and language use, it can show that likelihood of occurrence, or frequency of use, is an important measure of usefulness. (Kennedy, 1998, p. 281)

Most notably, Kennedy was one of the first language researchers to use a corpus for English for Specific Purposes (ESP) (see Chapter 4 *Background*), as well as one of the first to integrate corpora for the purpose of developing teaching pedagogy, noting not only the importance of teaching the collocations of highest frequency for ESP, but also noting the importance of teaching those collocations which teachers have known to cause the greatest difficulties among their students. Kennedy's contribution to corpus linguistics and language learning, apart from numerous specialized studies, has been in the area of teaching pedagogy and teaching objectives, illustrating how easily (and abundantly) corpora provide natural texts that may be used to create language-learning materials for every specialized classroom.

Lastly, another group of pioneer researchers who have contributed to the development of corpus linguistics in communicative language learning are Biber, Conrad, and Reppen (1998). Their contribution to the area of corpus linguistics and communicative language learning over the years focuses primarily on register variation and ESP, and language acquisition and development. More specifically, various studies that they have conducted have illustrated the creations of various corpora made up of student writing. Until the time of their research, it was most common for teachers to

analyze specialized published corpora, rather than create corpora on their own from students' work for analysis. Some particular studies conducted by Biber, Conrad, and Reppen (1998) include the creation of corpora to compare the writing of third grade students with the writing of sixth grade students in order to examine development during a three-year time span, as well as the creation of corpora to compare the writing of native speakers with that of non-native speakers. The idea that teachers could create their own corpora (whether for ESP or in order to analyze their own students' writing) was not a new idea, but Biber, Conrad, and Reppen (1998) focused on illustrating how easily this task could be accomplished given the right software, as well as some justifications for why language researchers and teachers should consider this option.

In the last five years, there has been yet another breakthrough in the field of corpus linguistics and language learning, simply in that taking a 'corpus-based approach' has become more accepted in the field of language learning, and more accessible (with most university libraries now subscribing to various core corpora, and still other corpora becoming available online). Due to the increase in distribution and accessibility, teachers are increasing their use of corpora in the classroom (See Appendix A for samples of communicative language learning materials using corpora). However, while taking a 'corpus-based approach' has been proven to be empirically sound and data reliable, many researchers and language teachers question whether or not taking a 'corpus-based approach' is more time-consuming than it is worth. It is true that teachers can create quick exercises for students to do in class, focusing on data-driven language learning, although it is difficult to say whether the use of corpora in this manner will ever become the prevailing method of implementation. However, it may be possible to summarize the

background research in corpus linguistics by stating that one common advantage of using corpora lies in textbook and materials development and evaluation. Many currently-used textbooks contain only invented examples and their descriptions are based apparently upon intuition or second-hand accounts, whereas other books, such as the books produced by the Collins COBUILD project, are explicitly empirical and rely for their examples and descriptions upon corpora and other sources of real life language data (McEnery & Wilson, 1996). Therefore, while the majority of language teachers may not choose to incorporate corpora directly into their own classrooms, how would they respond to textbooks based on corpora?

Motivated teachers and students who choose to use corpora for DDL can surely benefit in their language studies if they successfully take on the "learners as researchers" persona, but realistically this may include only a small percentage of people. Therefore, taking both the advantages and limitations of taking a 'corpus-based approach' in language learning into account, and in an attempt to increase corpora use in the classroom, I conclude by making the claim that corpora are most suitably utilized only if they are used for the purpose of curriculum and textbook development and evaluation, rather than for the purpose of data-driven language learning. In this case, corpora are utilized for the purpose of including natural texts into language textbooks based on empirical data, thereby improving the textbooks empirically. Following are three studies that demonstrate this possibility.

Notes:

[1] For more information, see http://www.ling.upenn.edu/mideng, http://www.hti.umich.edu/micase, http://titania.cobuild.collins.co.uk/boe_info.html

CHAPTER 2
TEXTBOOK EVALUATION

Background

Teachers of English as a Second Language are often encouraged to increase their use of authentic materials in the classroom and to deviate from using their course textbook as a sole means of providing input to their students. However, there are several definitions of authenticity in materials. Rogers and Medley (1988) use the term 'authentic' to refer to all language samples that "reflect a naturalness of form, and an appropriateness of cultural and situational contexts that would be found in the language as used by native speakers" (p. 468). Porter and Roberts (1987) state that authentic texts are those "whose instances of spoken language ... were not initiated for the purpose of teaching... not intended for non-native learners" (p. 176). It is well known that there are varying degrees of authenticity in TESOL literature; however, their pedagogical purpose remains relatively the same.

Within the TESOL field, it is possible to summarize the use of authentic materials in stating that authentic materials are consistently believed to be more effective in the ESL classroom (as opposed to non-authentic materials developed for the purpose of an activity), at the same time providing a more communicative and realistic base for any classroom activity. For example, there may be some correlation between the effectiveness of a task-based activity in the ESL classroom and its degree of authenticity – the underlying explanation for this being that students who are using authentic materials successfully during a task-based activity will be more equipped to perform a particular task outside of the classroom (as the result of the classroom task being duly

representative of what the student may encounter outside the classroom). Another reason a language teacher may choose to use authentic materials in the classroom regards student motivation. Authentic materials, while having a pedagogical rationale, are also more popular among language learners in that their implementation contributes to improved motivation in the language classroom. While it is possible to create suitable non-authentic materials, there is always the risk that those materials may not be representative of real-world encounters with language and/or that students may realize their non-authenticity and experience a decrease in motivation as a result of a lack of incentive for something that is not based on "real-world language". While these are just a few examples of why a language teacher may choose to use authentic materials in the classroom, the benefits of authentic materials clearly subjugate the use of non-authentic materials.

Taking the multiple benefits of using authentic materials in the classroom into account, it would be a shame to ignore the abundancy with which corpora provide authentic materials of many genres to language teachers and learners. Corpus examples are important in language learning as they expose students at an early stage in the learning process to the kinds of sentences and vocabulary that they will encounter in reading genuine texts in the language or in using the language in real communicative settings (McEnery & Wilson, 1996).

Justification & Research Questions

While teachers frequently use authentic materials in the classroom, and may even use corpora for data-driven learning exercises in or outside the classroom, it seems that

they do so in hopes of supplementing a textbook, or providing more communicative input for their students. However, what is overlooked in this case is that it is often textbooks that need supplementing (with authentic materials) and teachers spend endless hours trying to come up with new ways to provide such a medium for their students. While this supplementation[1] has been proven to be effective in language learning classrooms, it is still important to look at the broader picture. That picture states that textbooks can in fact be updated with more authenticity through the use of a corpus-based approach, and if a corpus of a corresponding nature exists, then there is no reason a textbook cannot be based solely on authentic pieces.

One particular study I conducted focused on the dynamics of comparing academic teaching materials to a corpus in order to assess the suitability of materials being taught (What is accurately represented? What is misrepresented?), as well as to look at how a corpus could be used in order to update teaching materials with authentic materials for speakers of English as a second or foreign language. Specifically, this study compared Madden and Rohlck's (2000) textbook *Discussion and Interaction in the Academic Community* (created without the assistance of a corpus) to the Michigan Corpus of Academic Spoken English (MICASE)[2]. Both the textbook and the MICASE were created at the University of Michigan and therefore have some relationship with one another. The particular features that were assessed were the colloquial and idiomatic expressions taught in the textbook, as well as the context in which they occurred. Lastly, this study looked at suggestions for the textbook according to various corpora. The research questions that will be discussed in regards to this study are: How can a corpus be used in order to create better ESL teaching materials (i.e. textbooks)? What is the

frequency of expressions and idioms taught in Madden and Rohlck's (2000) *Discussion & Interaction in the Academic Community* as found in MICASE? How do these frequencies compare to the COBUILD corpus? Are there variations of these expressions and idioms in MICASE? Are there enough variations to be significant? (i.e. Are these expressions outdated?) In what contexts are the expressions being taught and are they 'context-appropriate'? Are there idioms in MICASE that are frequent enough to be taught? What about other idioms "From the News" (COBUILD) that might be applicable? What are some other ways in which the textbook could be improved, according to empirical corpora analysis?

Methodology

As noted previously, the Michigan Corpus of Academic Spoken English (MICASE) was utilized for the purpose of this study. MICASE was established at the University of Michigan in 1999. It is compiled of 20 different speech event types and consists of approximately 1.6 million words overall (Table 1). Moreover, MICASE is representative of various discourse modes, genders, academic roles, and academic divisions (Tables 2 – 4).

Table 1: MICASE speech events

Speech Event Type	**Transcripts**	**Speakers**	**Words**	**% of Total Corpus**	**% Male**	**% Female**	**% Faculty***	**% Students***
Advising	5	20	58,817	3.5	43.5	56.5	14.2	37.2
Colloquia	13	118	151,639	8.9	52.9	47.1	76.9	10.8
Discussion Sections	9	112	74,904	4.4	36.8	63.2	33	66.7
Dissertation Defenses	4	26	56, 837	3.4	55.1	44.9	36.5	62.7

Table 1 (cont'd).

Interviews	3	6	13,015	0.8	82.6	17.4	56.0	44.0
Labs	8	42	73,815	4.4	69.8	30.2	15.1	67.9
Large Lectures	31	217	257,311	15.2	52.6	47.4	93.5	5.9
Small Lectures	31	289	320,893	18.9	43.8	56.2	74.0	22.6
Meetings	6	60	70,038	4.1	65.8	34.2	15.8	61.6
Office Hours	8	79	120,629	7.1	32.1	67.9	26.9	72.8
Seminars	8	79	151,071	8.9	60.2	39.8	58.8	34.9
Study Groups	8	36	129,725	7.7	31.7	68.3	0	100.0
Student Presentations	11	146	143,369	8.5	23.9	76.1	15.4	77.6
Service Encounters	2	90	24,691	1.5	40.6	59.4	.02	60.2
Tours	2	19	21,768	1.3	58.4	41.6	0	60.9
Tutorials	3	18	27,014	1.6	35.4	64.7	15.9	80.9

Table 2: MICASE gender and academic role distribution

Speaker Category		**Total Speakers**	**Total Words**	**% of Total Corpus**
Gender	Male	729	786, 487	46%
	Female	842	909, 053	54%
Academic Role	Faculty	160	825, 829	49%
	Male	84	446, 925	26%
	Female	76	378, 904	22%
	Students	1, 039	742, 348	44%
	Undergraduates	782	368, 433	22%
	Male	336	142, 102	8%
	Female	446	226, 331	13%
	Graduates	257	373, 915	22%
	Male	121	158, 696	9%
	Female	136	215, 219	13%
Language Status	Native Speakers	1, 449	1, 493, 586	88%
	Non-native Speakers	122	201, 954	12%
Totals		1, 571	1, 695, 540	

Table 3: MICASE academic divisions distribution

Academic Division	**Speech Events**	**Speakers**	**Words**	**% of Total Corpus**	**% Male**	**% Female**	**% Faculty**	**% Students**
Humanities & Arts	36	349	434, 669	26	56	44	63	29

Table 3 (cont'd).

Social Sciences & Education	35	452	420, 347	25	25	63	44	55
Biological & Health Sciences	32	257	325, 347	19	19	59	55	42
Physical Sciences & Engineering	36	314	358, 776	21	21	45	44	52
Other/NA	13	199	156, 292	9	9	63	20	41
Totals	152	1, 571	1,695,540					

Table 4: MICASE discourse mode distribution

Primary Discourse Mode	**Speech Events**	**Speakers**	**Words**	**% of Total Corpus**	**% Male**	**% Female**	**% Faculty**	**% Students**
Monologic	61	472	554, 335	33	50	50	84	14
Panel	9	133	141, 505	8	27	73	16	76
Interactive	57	643	715, 333	42	46	54	26	63
Mixed	25	323	284, 367	17	51	49	54	39
Totals	152	1, 571	1,695,540					

*Note: In these tables, percentages for faculty and students do not add up to 100% because of other speaker roles (e.g. staff, researchers, visitors) not included in these counts.

The Wordsmith and MonoConc Pro programs were used almost entirely for the purpose of this research. Within Wordsmith and MonoConc Pro, both concordances and wordlists were created in order to analyze the information from various perspectives. The MICASE online version was also used for a portion of this research, primarily for the examination of contextual representativeness, as the online version allows one to browse the texts easily and look directly at particular occurrences of idiomatic expressions in specific environments. COBUILD online was also examined for a portion of this research.

My methodology for collecting and analyzing data began by first compiling the various lists of words being taught in Madden and Rohlck's *Discussion and Interaction in the Academic Community* (2000), as well as noting the contexts in which they were

being taught. There were two long lists of idioms that I was particularly interested in analyzing, namely one list called "Idioms from Geometry", another list called "Idioms in the News", the latter consisting of baseball idioms that were said to occur with great frequency in various newspapers (See Table 5).

Table 5: Idioms from Madden and Rohlck's *Discussion & Interaction in the Academic Community (2000).*

Idioms in the News (p. 42)	**Idioms from Geometry (p. 94)**
Playing hardball	Give (me) a straight answer
Threw (him) a curve	Going around in circles
Out of left field	Drew a blank
Pinch hit	Gone off on a tangent
Strike out	Get to the point
Hit a home run	Read between the lines
	Start at square one
	To draw a parallel

The next step was to collect other words and colloquial expressions being taught throughout the text. Some of the more specialized expressions can be seen in Table 6 below. (Note: The complete version of all expressions can be found in Appendix B).

Table 6: Sample colloquial expressions from Madden and Rohlck's *Discussion & Interaction in the Academic Community (2000).*

Colloquial Challenges (p. 15)	**Rapport & Complaining (p. 16)**
Come on	That's typical
I don't follow	I don't think so
What are you getting at?	
Are you suggesting...?	

Overall, these idioms and phrases served as the initial base of this research. *Note: The textbook offered more phraseology than what is merely being analyzed in this project. The choice of what was analyzed will be discussed later on.*

Expressions (including all variations thereof) were assessed using a concordance search in MICASE. Once the frequency data were recorded, I then searched the COBUILD in order to examine the frequency of "Idioms in the News". In order to

maintain comparable data, I only examined U.S. newspapers and news in COBUILD, which resulted in a corpus of approximately 15 – 16 million words, inclusive of both spoken and written news. The purpose of this comparative search was to see whether or not the "News" idioms were appropriate representations of common idioms in U.S. newspapers and news broadcasts as the textbook claimed.

The research then separates into different branches as follows: For the idioms that were not found in MICASE, other possible "more modern" or "common" idioms (according to *A Dictionary of American Idioms* (2001)) were then searched in the concordance in order to act as possible suggestions for updating the textbook. Moreover, COBUILD was searched a second time in order to discover what idioms occurred with greatest frequency, possibly illustrating alternatives to the "Idioms in the News" which were not represented in the COBUILD. This approach to searching served the purpose of possibly updating the textbook, while still maintaining the appropriate context of idioms being taught (i.e. "Idioms in the News"). Furthermore, I also looked at contexts of the various colloquial expressions in the MICASE in order to examine what common 2-word, 3-word, 4-word, and 5-word clusters occurred with greatest frequency, in order to reveal common idiomatic expressions (as possible suggestions for updates as well).

Results

My initial prediction regarding overall frequency of occurrences of the selected idioms was that these idioms would be found in MICASE, perhaps illustrating minor variations. It was these variations that I was planning on examining in order to suggest revisions for the textbook. However, when I did search for the idioms (and all variations

thereof), what I found was that all but two of the idioms did *not* occur in the MICASE at all! The two idiomatic expressions that did occur in one form or another were "Gone off on a tangent" and "Get to the point". The occurrences are as follows in Table 7.

Table 7: Frequency results (per 1 million) of the idioms found in MICASE.

Gone off on a tangent	**Occurrences**	**Get to the point**	**Occurrences**
Gonna go off on a tangent	1	Going to the point	1
Go off on a tangent	2	Got to the point	1
Go off on such a tangent	1	Gotten to the point	1
		Getting to the point	1
		Get to the point	2
		Other "to the point" expressions	14

Note: Only occurrences are given, as the frequencies of these idioms were less than 1%.

Since the occurrence of overall idioms was surprisingly low (or non-existent for that matter), it was necessary to search the COBUILD in order to make sure the "Idioms from the News" were, in fact, common, up-to-date idioms from the news. These results showed variation from the MICASE corpus, in that some of the idioms that did not occur in MICASE did occur with some frequency in the COBUILD (See Table 8).

Table 8: Frequency results of "Idioms in the News", as found in COBUILD.

Idioms in the News	**Occurrences**
Playing hardball	23
Threw (him) a curve	0
Out of left field	4
Pinch hit	0
Strike out	50+*
Hit a home run	0

*The most common definition meaning "to strike out in anger", rather than the definition of "to fail", as given in the textbook.

As can be seen in the data above, both "playing hardball" and "strike out" occurred in the COBUILD with some frequency, yet they did not occur in the MICASE. This supports

the textbook in that these may be considered to be common idioms in the news, therefore being duly representative. However, "threw (him) a curve", "pinch hit", and "hit a home run" still did not occur.

In order to suggest alternatives to the idioms that did occur, I used my intuition to come up with more modern versions of particular idioms and searched these expressions in MICASE (See Table 9).

Table 9: List of alternative idiomatic expressions searched

Idiomatic alternative	**Instead of:**
To be frank	To give a straight answer
To go off	To go off on a tangent
Cut to the chase	Get to the point
Start from scratch	Start from square one
Threw (him) for a loop	Threw (him) for a curve
Out of nowhere/out of the blue	Out of left field
To score/To hit the jackpot	To hit a home run

However, of the alternative idioms, the idioms in Table 10 were the only ones that came even remotely close to having alternatives.

Table 10: Alternative idiomatic expression results

Textbook Idiom	**New idiom**	**Frequency of New Idiom**
Start from square one	Start from scratch	4
Threw (him) for a curve	Threw (him) for a loop	3
Out of left field	Out of the blue	2

While the new idioms offer possible revisions of the idioms in the textbook, a new problem arises because these new idioms are of a different context/theme. For example, while "start from scratch" may be a revision of the outdated idiom "start from square one", it does not comply as being an idiom from geometry. This in mind, I then compiled a list of geometry idioms, as found in *A Dictionary of American Idioms* (2001), that could serve to fill the position of deleted idioms (due to infrequency)(See Table 11).

Table 11: List of geometric expressions searched

Geometry Possibilities
Count me out
To come full circle
A vicious circle
Fair and square
A square deal
To square an account
Take care of number one
Safety in numbers
One-sided
Put two and two together
Draw the line
Pluses outweigh the minuses

The search had to be limited to idioms from geometry, as alternative baseball idioms could be not found. However, of all of the above geometric idioms searched, the following are the only idioms that offer possible alternatives (See Table 12).

Table 12: Alternative geometric expressions results

Geometry Idiom	**Occurrence in MICASE**
Safety in numbers	6
One-sided	4
Draw the line	2
"Come full circle" idioms	4

Since the baseball idioms proved to be infrequent in both the MICASE corpus and the COBUILD, and there were no alternative baseball idioms found in either corpus, I then searched for common idioms in the COBUILD in order to satisfy the "Idioms in the News" contextual requirement. However, this proved to be more problematic than originally planned, since the amount of idioms found in the COBUILD was so extensive that it was difficult to narrow the results down to a concise list of practical idioms to include in the textbook. Therefore, this would depend on what the authors of the textbook would be interested in teaching (i.e. instead of a baseball theme, another theme would have to be chosen). Therefore, this portion of the research proved to be indefinite

since anything would be possible, and it was difficult to judge what the "good" themes would be without factual reasoning.

The next step of this study was to examine the various colloquial expressions (as seen in Table 6 and Appendix B). This portion of the research was reassuring on behalf of the textbook in that the majority of the expressions in the textbook were found in the MICASE corpus in some form or another, and with adequate frequency. However, there were discrepancies since the expressions proved to be difficult to search for due to the length of the phrases. However, it was worth doing the searches in order to see the list actually represented in MICASE, as this was a large portion of the textbook's teachings. However, as a sample (and due to the more specific nature of these phrases), we will go back to the Table 6 expressions, and look at the sample colloquial expressions given and their results in detail (See Table 13).

Table 13: Occurrences of sample colloquial expressions from MICASE – results from Table 6

Colloquial Expressions	**Occurrences**
Come on	42
I don't follow	1
What are you getting at?	0
Are you suggesting...?	0
That's typical	0
I don't think so	21

Again, there are discrepancies in these data since the longer phrases are exhibiting zero occurrences. "That's typical" has zero occurrences as well, however, this may be due to being an outdated idiom. What to maintain and what to alter according to these results would be dependent on the author's objectives in creating the textbook, although these results offer a base with which to begin any alterations.

In order to examine how the "Colloquial Expressions" section of the textbook could be further supplemented, I then created wordlists in order to examine the most

commonly used word cluster expressions in order to offer that material as alternatives to certain colloquial expressions, if alternatives were being looked for at any point in the future. In doing so, 2-word, 3-word, 4-word, and 5-word clusters were then filtered out in order to see what counted as colloquial or not. Again, it is important to keep in mind that there will always be discrepancies in these data, as this is a subjective method open to interpretation. The most common 2-word, 3-word, 4-word, and 5-word colloquial expressions in the MICASE are as follows (See Tables 14 – 17).

Table 14: Frequent 2-word clusters in MICASE

2-word clusters	**Occurrences**	**Frequency (per 1 million)**
Kind of	1,996	15%
Sort of	1,672	12%
A lot	1,424	11%
A little	1,108	8%
Lot of	1,060	8%

Table 15: Frequent 3-word clusters in MICASE

3-word clusters	**Occurrences**	**Frequency (per 1 million)**
A lot of	1,024	8%
A little bit	542	4%
The fact that	218	2%
In other words	185	1%
In order to	166	1%
A couple of	147	1%
Point of view	118	<1%
A bunch of	95	<1%

Table 16: Frequent 4-word clusters in MICASE

4-word clusters	**Occurrences**	**Frequency (per 1 million)**
At the same time	114	1%
A little bit of	78	<1%
A lot of the	78	<1%
In terms of the	71	<1%
On the other hand	54	<1%
It turns out that	42	<1%

Table 17: Frequent 5-word clusters in MICASE

5-word clusters	**Occurrences**	**Frequency (per 1 million)**
You know what I mean	54	<1%
As a matter of fact	16	<1%
From the point of view	11	<1%
So on and so on	11	<1%
In a number of ways	6	<1%

Discussion

It is difficult to suggest what should be changed in *Discussion and Interaction in the Academic Community* (2000), as this is an obvious decision only to be made by the authors, according to what they want to work with and what they have had the most success with in teaching in the past. However, these data may be useful in updating the textbook in order to better represent actual speech in the University of Michigan's academic community, if that is in fact an objective of the textbook. In trying to suggest what should remain the same and what should be changed, it is important to take into account the need for a structural format. For example, prior to doing this particular research study, I spoke with Theresa N. Rohlck, co-author of the textbook. Upon my inquiring about what they (the co-authors) chose to go into the textbook, she commented on the notion that students need some type of structural format, or pedagogical norm, for conversation for example, regardless of the words chosen to go into that format. For example, it is necessary to teach a language learner the basic format of conversation (i.e. Hello, How are you? I'm fine, How are you?), regardless of how many times people actually strictly comply with this format and word-usage in everyday conversation. Perhaps while the most common form of greeting around campus may be "Wassup?", it would not necessarily be appropriate to teach this to an academic student, as this may not be the most professional form of the greeting, nor the most effective for the students' academic purposes. Therefore, while we may see great frequencies of certain words or idioms within a corpus, it is still necessary to have a human evaluator who decides what is textbook-appropriate and what is not (based on what students' objectives are for taking an English for Academic Purposes course). This in mind, I would suggest maintaining

the colloquial expressions used throughout the text, as these seem to be general structural-context-appropriate phrases and significantly represented in the MICASE. Furthermore, another suggestion I would make regarding this textbook would be to modernize some of the idioms. This could be done through the analysis of some of the clusters given, or through the modernization of the present idioms in the textbook according to the alternative suggestions given.

Overall, while this study provides one way in which a corpus may be used to evaluate a textbook, what is important are the implications for textbook evaluation in language learning for language teachers and future textbook editors. For example, while I have looked at how a textbook may be updated according to empirical data and only provided suggestions for the textbook, it is important to note that sometimes (despite the lack of empirical data), language researchers may choose to continue to teach particular idioms and expressions if they have been proven effective in their own classrooms. Therefore, how can the continuation of teaching idioms and expressions not found in the MICASE be justified? When should a language teacher/textbook editor choose materials based only on empirical data from a corpus, and when should they teach what they deem intuitively appropriate for their students?

One important and unique characteristic of a corpus-approach is that it incorporates both empirical quantitative data *and* qualitative human analytical techniques. Therefore, despite the credibility of a corpus, I conclude that it is still necessary for language teachers to utilize some type of filtering and selection based on their language teaching goals and on the needs of their students.

Therefore, here are some guided questions to take into consideration for teachers interested in assessing their own textbooks through the comparison of their textbook to a corresponding corpus:

1.) What is the credibility of the materials being taught? Are the chosen materials based on empirical research or teachers 'intuition'?
2.) What is the frequency of idioms and expressions being taught in the textbooks, as found in corpora?
3.) Are there variations of these expressions, and are there enough variations to be significant?
4.) In what contexts are the expressions being taught and are they context-appropriate?
5.) Are there high frequency expressions and idioms that occur in the corpus, but not in the textbook?
6.) What are your language teaching goals? What are the specific needs/interests of your students?
7.) How do you justify teaching expressions that do not occur in any corpus?

Notes:

[1] It is important to note the distinction between authentic materials as texts in the language classroom and authentic materials as a basis for deciding what to teach. In the former case, it is possible to take whole texts and simply use them as authentic reading/comprehension material. However, in the latter case, it is possible to create a corpus or implement an existing corpus, and teach the vocabulary and structures from that corpus, but never actually use the whole texts themselves. This particular study, and other studies in this thesis, focuses solely on the latter use of authentic materials.

[2] While this is the first time that a corpus has been used in order to assess Madden and Rohlck's textbook (2000), the MICASE has been analyzed in the past in order to assess and create teaching materials and textbooks. Data collected for MICASE have already been used for the purpose of creating instructional materials projects. Elizabeth Axelson, University of Michigan English Language Institute (ELI) Lecturer, used transcripts and sound files to develop training materials for International Teaching Assistants, focusing on "linguistic aspects of interactive teaching" (MICASE online: http://www.hti.umich.edu/m/micase). Susan Reinhart of the University of Michigan's ELI has also incorporated MICASE data in a textbook on oral presentations (U-M Press). Furthermore, the testing division, the major sponsor of MICASE, has been using the database as a resource for test development and validation (MICASE online). Word frequency information based on MICASE has been employed in the development of new items for the listening section of the Examination for the Certificate of Proficiency in English (ECPE). Moreover, there have been numerous other studies done and papers published using MICASE for the purposes of analyses ranging over many topics and fields.

CHAPTER 3
TEXTBOOK DEVELOPMENT

Background

While there has been an increase in the number of teachers using authentic materials, one area of instruction that is largely overlooked in ESL/EFL classrooms is the development of pragmatic awareness. Specifically, there is a need to provide students with pragmatically appropriate conversational models, especially in English speaking classes. Tanaka (1997) best explains this need by relating L2 pragmatics to the effectiveness of speaking with native speakers of the target language. For example, Tanaka mentions that although native speakers of a language tend to be tolerant of grammatical errors, they are less tolerant of pragmatic errors. Therefore, there is a present need for pragmatic instruction in our ESL/EFL classrooms for the main reason that students desire to speak with native speakers of their target language (in this case, English) in order to continue to improve their L2 proficiency, and will more likely feel comfortable doing so once they have received some type of input in the classroom. Therefore, the issue at hand regards the best way to teach speaking pragmatic proficiency in an L2.

In an early study, Bardovi-Harlig, Hartford, Mahan-Taylor, Morgan, and Reynolds (1991) explore the concept of the need to increase the role of pragmatics in English-language instruction. Their assumption is based on the lack of "natural" materials provided in many ESL textbooks, as well as the lack of natural examples of "pragmatically-appropriate" conversational models for learners (Bardovi-Harlig et al., 1991). They provide some background of this assumption, stating:

> Language learners interacting with speakers of a target language must be exposed to language samples that observe social, cultural, and discourse conventions – or in other words, which are pragmatically appropriate. Speakers who do not use pragmatically appropriate language run the risk of appearing uncooperative at the least, or, more seriously, rude or insulting (Bardovi-Harlig et al., 1991, p. 4).

This possibility of seeming rude or insulting is the basis of their argument for the need to develop pragmatic awareness in ESL conversation classrooms. Furthermore, they explore the notion that teaching pragmatics in an L2 is easier said than done for the main reason that there are infinite amounts of speech events, and it would be impossible to teach/give input about all situations that students are likely to face. Furthermore, they offer methods to teach pragmatic awareness in conversation classrooms, not only to advanced learners, but to learners at lower levels as well. This research is based on the notion that it is more beneficial to the L2 learners if teachers make students aware of pragmatics in speech, rather than giving them specific input on specific speech acts. There are four steps to the pragmatic-awareness model:

> Four steps to integrating pragmatically appropriate language into the classroom:
> 1. Identification of the speech act
> 2. Data collection and description
> 3. Text and materials evaluation
> 4. Development of new materials (Bardovi-Harlig et al., 1991, p. 5)

It is believed that the reason that it is so important to allow students to come to their own assumptions about pragmatics in L2 speech is because directly teaching pragmatics is based on the intuition of the instructor or the textbook and this intuition may not always be accurate as well.

The main argument that Bardovi-Harlig et al. (1991) maintain is the notion that it is more important for L2 learners to know the *structure* of conversation (as previously

mentioned in Chapter 2 of this thesis), rather than being taught explicitly words or phrases to use in conversation since all conversations differ. They go on to offer specific exercises that L2 learners can benefit from if pragmatic awareness in conversation is their language goal. Some of these exercises are: practice with a native speaker, structured model approaches, classroom role-play, and data collection by students.

The article concludes with the statement that teaching pragmatics "empowers students to experience and experiment with the language at a deeper level" (Bardovi-Harlig et al., 1991, p. 13). Therefore, creating pragmatic awareness in English conversation classrooms will allow L2 learners to overcome cultural boundaries and to equip them with not only the structural aspects of the language, but with the pragmatics as well.

Tanaka (1997) also explores this same issue of pragmatics in the L2. Tanaka begins his argument for the ESL student's need for pragmatic competence by offering his understanding of someone who is "communicatively competent", which is someone who is able to use not only grammar structures effectively, but also someone who has acquired pragmatic knowledge of the sociocultural rules of speaking (Tanaka, 1997, p. 14). However, because this pragmatic competency is most likely to take a very long time to achieve, Tanaka believes that it therefore should be an inclusive part of English conversation curricula so that English learners will obtain the necessary input and interactional opportunities.

The method that Tanaka encourages is the "learners-as-researchers" approach, which is similar to Bardovi-Harlig et al.'s (1991) conclusions about the most beneficial ways to encourage pragmatic awareness in the classroom. Tanaka, however, also

acknowledges the notion that ESL and EFL learners of English often have limited access to native speakers in order to obtain real-world data of speech pragmatics. Therefore, Tanaka offers a solution to this dilemma by providing possible classroom activities and homework assignments for students to interact with native speakers of English and to make these activities optimal opportunities to obtain pragmatic data in order to develop their own pragmatic competence.

However, while Tanaka (1997) and Bardovi-Harlig et al. (1991) offer the background regarding the necessity to teach pragmatic awareness in the classrooms, neither acknowledges the usefulness of teaching idioms in this feat. Fernando (1996) states: "No translator or language-teacher can afford to ignore idioms or idiomaticity if a natural use of the target language is an aim" (p. 234). Furthermore, these sentiments are echoed by Wray (2000), although somewhat less prescriptively.

> Gaining full command of a new language requires the learner to become sensitive to native speakers' preferences for certain sequences of words over others that might appear just as possible. From the bizarre idiom, through the customary collocation, to the turns of phrase that have no apparent linguistic merit than that 'we just say it that way', the subtleties of a language may floor even the proficient non-native, not so much because of a non-alignment between interlanguage and target language forms, as because the learner lacks the necessary sensitivity and experience that will lead him or her unerringly away from all the grammatical ways of expressing a particular idea except the most idiomatic. (p. 463)

There is certainly a great deal of truth to these claims. Learners often have an insatiable desire to learn idioms, as this knowledge is related to a heightened level of fluency in a language. In fact, according to the *Longman Advanced English Dictionary*, the term "idiomatic" is defined as "typical of the way in which someone using their own language speaks and writes." In response to such statements implying native speaker competence (along the same lines as Tanaka's notion of 'pragmatic competence'), many

ESL textbooks dealing with spoken English include idioms and idiomatic expressions in their content, to varying degrees. However, as most language teachers will agree, teaching idioms is by no means an easy task.

Duquette (1995) proposes that idioms, culturally embedded and grouped into units of speech, are helpful in developing communicative and pragmatic competence. According to Duquette, it is necessary to gain knowledge of background cultural information in order to use language skills properly. He explains that language proficiency requires more than just being able to speak the target language. Duquette focuses on three principles of language acquisition:

1.) Language comprehension and language performance requires socio-cultural and "context-defined cues" in order to be meaningful and communicative.
2.) "Comprehensible input" plays a large role in increasing comprehension and promoting the acquisition process.
3.) Target group motivation, integration, acceptance, and identification are important factors in developing communicative competence. (Duquette, 1995, p. 37)

Therefore, one reason that Duquette suggests that the use of idioms by L2 learners should help to increase communicative and pragmatic competence is that the acquisition of idioms and other commonly used expressions give self-confidence to the L2 student (Duquette, 1995). Since cultural meaning may vary from one group to another, and meaning is already 'built-into' these expressions, the L2 student has the opportunity to enable him/her to be well understood in a particularly reliable manner. "Idioms seem to 'bridge the barrier' because they are simple to grasp, frequently used, semantically and culturally loaded, and embedded into the target culture" (Duquette, 1995, p. 37).

Duquette goes on to offer suggestions for language learners and teachers in order to demonstrate ways in which idioms may be taught so that they are more meaningful, or

communicative. One step is to combine idiom teaching with real-life experience. For example, presenting sensory input to the class in the form of a video, a guest speaker, a visit to the supermarket, etc., in order to provide students with concrete background information. This information can then be applied to classroom activities that exercise these task-based idiomatic expressions (Duquette, 1995). These suggestions are directly correlated to the prior background articles on developing pragmatic awareness; therefore, idioms are in fact a possible means of developing competence in L2 speech, as well as developing pragmatic awareness in speech overall.

Justification & Research Questions

What do we as language teachers teach in the classrooms when achieving pragmatic competency is our goal as ESL language teachers? It is well known that students have an insatiable desire to have pragmatic awareness, and a part of that includes a desire for "the idiom". While teaching idioms is by no means an easy task on the part of the language teacher, there are methods that are believed to be more empirically sound than others. For example, many textbooks teach idioms in a special 'idiom' section of the text, but do we know where these idioms were obtained and who decided that these idioms were significant enough to be taught? Furthermore, many instructors teach idioms as an integrative part of their teaching curricula. However, how do the teachers decide which idioms are most frequently used, and how does this account for speaker idiosyncrasies? In response to such inquiries, I believe that it is necessary to go beyond basic teacher intuition and begin to base our teaching on empirical data obtained through research. Therefore, one response to this is to collect such empirical data through a

corpus-based approach to the development of an "idiom" textbook. The research questions that are addressed in this study are: 1.) Do any idioms occur in academic speech for use in academic language learning classrooms/textbooks? 2.) What idioms occur in academic speech (MICASE), and with what frequency? 3.) Are there enough idioms found to compose, supplement, or replace existing "idiom" textbooks based solely on intuition? 4.) How can the idioms found in MICASE be taught communicatively, if their collection is to be used to create a textbook?

Methodology

The second study began out of an interest in which idioms occurred in academic speech. The methodology that we followed for finding all of the idioms was to first manually search through the MICASE (Michigan Corpus of Academic English) and read through all of the transcripts (Simpson, Mendis, & Komsic, 2002). The MICASE corpus is made up of over 1.6 million words, and over 20 speech events. Often there would be two readers per transcript. All idioms and idiomatic expressions were then recorded into a Microsoft Access Database based on our guidelines/criteria for the term "idiom". Our rationale/criteria for selecting idioms for our list included:

1. Idioms
    a.) Multiword expressions which cannot be understood by taking the meaning of individual components
    b.) Semantic opacity of the whole
    c.) Compound words which are metaphoric, and are either semantically opaque or cannot be understood by individual components (e.g. oddball, full-fledged)
2. Phrasal Verbs
    a.) 3-part phrasal verbs, e.g. verb+prep+prep or verb+prep+noun (e.g. fall into place)

b.) 2-part phrasal verbs in which the VP is more commonly known as a noun (and thus requires additional semantic processing on the part of the listener/learner) (e.g. breeze through, flesh out)

However, we rejected the following two categories:

1. 2-part phrasal verbs[1] (e.g. get up, wake up)
2. Single verb phrases used metaphorically

Results

After one year of reading through transcripts and recording idiom data, our results showed over 600 occurrences of idiomatic expressions in the MICASE, which constituted over 200 different idiomatic expressions! These idioms should prove to be very useful to language teachers who desire to incorporate idioms into their academic classrooms. Some of the most frequent idioms in the MICASE follow in Table 2 (See Appendix C for a complete list). This list gives all idioms occurring three or more times in the MICASE:

Table 18: Most frequent idiomatic expressions, as found in MICASE

Idiomatic Expressions	**Occurrences**
Bottom line	17
Come into play	16
What the hell	14
Down the line	12
Pros and cons	11
Flip a coin, flip side of a/the same coin	10
Pick up on s.th.	10
Hand in hand	8
Right off the bat (straight = 1)	8
What the heck	8
Draw a/the line (between)	7
On target	7
Out the door	7
The big picture	7
Thumbs up	7
Fall back on	6
On track	6
Rule/s of thumb	6
Take (s.th) at face value	6
Beat to death	5
Fall in love	5

Table 18 (cont'd).

Get around (to doing s.th)	5
Put the heat on	5
Caught up in s.th.	4
Come out of the closet	4
Fall into place	4
Full-fledged	4
Get a handle on	4
Goes to show	4
Line of attack	4
Nitty-gritty	4
On the same page	4
Ring a bell	4
Splitting hairs	4
Take my word for it	4
Truck along on	4
Catch up with	3
Cutting edge	3
Drive me crazy/bananas	3
Fall/throw by the wayside	3
Fine tune	3
For all intents and purposes	3
Get a head start on	3
Give and take	3
Have/has the foggiest notion/idea	3
Hit the wall, hitting a (rock) wall	3
In a nutshell	3
In line with s.th.	3
Lined up in advance	3
On the right track	3
Pick (a) fight	3
Plug and chug	3
Rears its ugly head	3
Run into a wall	3
Steering clear of	3
The kitchen sink	3

Discussion

The above list is just a small portion of the culturally loaded idioms found in the MICASE. These data are significant because they supply not only authentic materials (texts) with which to teach the context of particular idioms in class, but also an empirical collection of idioms that are not solely based on teacher intuition, providing not only the cultural information with which to improve pragmatic awareness, but in the appropriate

register as well. The implications for teaching these idioms are endless since the creativity of language learning teachers may also be non-finite.

However, the issues that are important to address at this stage are whether or not data (in this case, idioms) based on empirical data are worth more than data (idioms) based solely on intuition for textbook development, and if one chooses to use data collected through a 'corpus-based approach', what would the best way be to teach these communicatively? In this first case, there seems to be no question that empirical data should outweigh idiosyncratic data in textbook development if authenticity and true pragmatic awareness are our goals, and therefore I hope this study and others provide support of that stance. In the latter case, there is a question regarding the best way to include idioms in textbooks and how to teach them most effectively. In this case, if one chooses to take a 'corpus-based approach' for textbook development, it will change textbooks in two ways. The first way is that textbooks will be able to contain natural texts (e.g. for reading), which can contain the idiomatic expressions. In this case, there may still be communicative activities that are found in many textbooks, but rather the only thing that will change will be that all of the materials in the textbook are authentic, taken from natural language occurrences. The second way that textbooks will change will be in terms of how they may be updated. In the past, if a particular "idiom" textbook needed to be updated, it would be a matter of starting completely over to create a new textbook with the "newest" expressions. However, by implementing a 'corpus-based approach' for textbook development, it would be possible simply to do concordance searches (similar to those done in Study 1 of this thesis) to validate the credibility of the idioms being taught, and checking which idioms are still frequently used, and which are

in need of updating. This assumption is based, however, on the belief that corpora are kept up to date.

The overall implications that corpora and textbook development have for developing communicative curricula in the ESL classroom will be addressed in the final chapter of this thesis. However, what is important to note about this study is the evidence provided for the abundancy of idioms in academic speech (serving as a substitution for previously developed textbooks based only on intuition). It also acts as a model for future textbook developers wishing to create textbooks based on empirical data.

Notes:

[1]The above list was based on various literature reviews of the term 'idiom'. The reason that we chose to omit regular 2-part phrasal verbs from our study was due to the great frequency of phrasal verbs within the MICASE. On an interesting note, most newer textbooks teach only 2-part phrasal verbs which may solve the problem of having outdated idioms in their textbooks.

CHAPTER 4
ESP MATERIALS DEVELOPMENT

Background

Until this point, I have looked at the ways in which a corpus-based approach may be utilized in order to evaluate and develop textbooks. However, one particularly useful application of a corpus-based approach that should not be overlooked is the use of corpora in the development of ESP (English for Specific Purposes) materials and curricula in general, and ESP for university graduate students in particular, which will be the focus of this chapter.

The creation of a corpus to accommodate a specific population is an emerging application that has only been done in a relatively small number of documented instances, most often in the cases of teaching "content" in language classrooms. According to Kennedy (1998), there are a number of ways the language in a corpus can influence language-teaching pedagogies.

> However valuable may have been the focus on learning language as communication (particularly in increasing language fluency), there has been growing recognition that systemic accuracy is also relevant, and consequently there has been a more recent revival of interest among language teachers in what is being learned: the content of language teaching... Curriculum designers and classroom teachers need to have access to this information through better reference materials and syllabuses... First it can influence the content of language teaching by affecting the selection of what to teach, the sequencing of pedagogy, and the weight given to items or parts of the language being taught, thus contributing directly to the content of instruction. Secondly, through the consciousness-raising of teachers about language and language use, it can show that likelihood of occurrence, or frequency of use, is an important measure of usefulness. (Kennedy, 1998, p. 281)

In this case, it seems that the ESP classroom (e.g. ESP teachers and their curricula) could greatly benefit from the use of corpora, since corpora are a specialized means of

accessing specialized vocabulary in specialized texts. While I also agree that corpora may be used in any general language classroom or EAP classrooms (for the purpose of incorporating authentic materials into the curricula), the fact remains that if a teacher was interested in *creating* corpora (and hence, using that corpora) for their own class objectives (e.g. curricula), it would be an easier feat to do so for the purpose of using it in an ESP classroom rather than in an EAP or general English classroom due to the fact that there would be no confusion as to what texts to input into the corpus, and why those texts should be analyzed.

For example, Sutarsyah and Kennedy (1994) illustrated one study where a corpus was made to accommodate a specific population, thereby examining the usefulness of EAP vocabulary for particular university graduate students. As Sutarsyah and Kennedy have shown, advanced learners of English (university graduate students) wishing to study English only for the purpose of undertaking business studies or economics are unlikely to have received exposure to relevant vocabulary if they are part of a general English for Academic Purposes (EAP) course. Sutarsyah and Kennedy compared a corpus consisting of texts from many academic fields with one of roughly the same size consisting of a single economics text. They found almost two-and-a-half times as many different types of words in the general text as in the specialized text, suggesting that as far as vocabulary is concerned, the learners working through the general academic texts meet new words that would rarely recur, and that EAP courses that go beyond the 3,000 most common words in English may be of little value for learners with specialized needs for learning English. Table 19 shows that learners exposed to the economics specialized text get

about five times as much exposure to the most frequent content words in that text that they would in a corpus consisting of texts from general academic sources.

Table 19: Rank order of the 20 most frequent words used in an economics text compared with their occurrence in general academic English (adapted from Sutarsyah and Kennedy, 1994

	Rank Order		Frequency	
Word	Economics	General academic English	Economics	General academic English
price	9	479	3, 080	90
cost	14	471	2, 251	91
demand	17	411	1, 944	102
curve	21	525	1, 804	83
firm	23	991	1, 743	41
supply	24	509	1, 590	86
quantity	25	807	1, 467	53
margin	27	*	1, 427	24
economy	29	224	1, 353	172
produce	31	234	1, 237	167
income	33	442	1, 183	96
market	36	372	1, 104	110
labour	40	313	1, 004	131
increase	41	113	1, 002	277
consume	42	623	995	70
total	47	362	946	114
change	48	92	927	316
rate	49	104	915	293
capital	51	842	907	50
work	52	58	906	480

* = not in the first 1, 000 words of general academic English

Therefore, the above results show the specificity of language material that an ESP/content-specific corpus may offer, as well as the various ways that students need subject/field-specific vocabulary and language attention.

However, one important thing Kennedy notes following this study is that corpora should be used "judiciously for pedagogical purposes, informing instruction rather than

determining it as to avoid the risk of a return to prescriptivism" (Kennedy, 1998, p. 290). Furthermore, he states:

> Frequency of occurrence in corpora should be only one of the criteria used to influence instruction. Sometimes, according to the goals of the learners, less frequent items or processes in a language may deserve more attention than the most frequent, simply because they are known to be learning problems with a wide range of uses. (Kennedy, 1998, p. 290)

This idea leads to the next topic of EAP, a case in which there is usually a class full of university students, each with individual language needs (e.g. students from different academic departments), and it is not possible to address vocabulary in such a specialized way. This EAP situation is more common at universities in the United States than the idealized environment which Sutarsyah and Kennedy (1994) describe, in which a teacher could simply choose to teach only vocabulary that will most benefit the students based on their field of study. Therefore, it is also necessary to understand how corpora may be used for more general academic populations as well.

Graduate students face a variety of writing tasks as they work toward their chosen degrees, although these tasks will vary from one degree program to another. Swales and Feak (1994) give a general account of how university students at the graduate level may improve their academic writing, focusing on the structure and vocabulary of academic writing. In doing so, they focus on academic writing as a type of *genre,* the expectations that are involved in becoming a successful writer within that genre, and how students should learn more procedural or discourse organizing vocabulary for the aim of writing academic texts in a structured format. The areas that Swales and Feak address are academic organization (e.g. introduction, main body paragraphs, conclusion, etc.), academic style (e.g. vocabulary shift), flow (e.g. general to specific), and presentation

(e.g. paragraph formatting, etc.). The work resembles the multitude of EAP textbooks available, and does not address the use of taking a corpus-based approach in this particular text. However, what is interesting to note about their text is that it addresses the specialized needs of university students, and looks at the ways in which they may work on improving their genre-specific vocabulary. This may be done in conjunction with the MICASE corpus. In this sense, EAP is a type of ESP, since EAP teachers must address the needs of their students not in terms of their academic fields, but as participants in the academic community/genre.

However, taking both the work of Sutarsyah and Kennedy (1994), and Swales and Feak (1994) into account, a potential debate arises regarding the type of curriculum that is most effective for graduate students. Are more common "procedural" and "discourse organizing" vocabulary items most effective, or are more specialized vocabulary items most commonly found in a students' own field most effective? A corpus-based study by Liu and Nesi (1999) addressed this issue by investigating the best type of corpora that would meet the needs of graduate students in engineering - common vocabulary or technical, field-specific vocabulary. The first corpus contains the most common English words (generally taught in EAP classrooms) labeled as "subtechnical" vocabulary items, and the second corpus contains the most common English words in the electrical engineering field (based on engineering texts that were being used in their academic program) labeled as "technical" vocabulary items. While it has been a widely accepted practice that EAP teachers concentrate on vocabulary that is neither technical nor general, Liu and Nesi (1999) speculate that the choice of words to fit this category varies. In their study, they created a list of the most common "technical" vocabulary items taken from

the engineering corpora, as well as a list of the most common "subtechnical" vocabulary items. They chose a group of advanced (upper level) graduate students in engineering and had the students choose which words they found familiar, in hopes of testing to see which words they were most likely to encounter in engineering, therefore illustrating which words should be taught to this select group. What they found was that both the 'technical' and 'subtechnical' words were recognized with the same degree of frequency, and that the 'subtechnical' words that were recognized were procedural or discourse organizing. In their analysis, they noted that it seemed as though the students should have recognized a higher number of common technical words specific to engineering, and that their findings suggested that the subjects had not acquired all the essential technical vocabulary through course work, even though these students were at the end of their study period and were about to return to their countries of origin as qualified engineers! These findings challenge the opinion regarding the type of vocabulary that needs to be taught in EAP classes, since it illustrates that these particular students were not taught (or did not learn) the most important vocabulary for their field. While this particular study did have its weaknesses in terms of the way that the words were tested out of context, ignoring the subjects' ability to produce or recognize words in context, it raised the question of whether graduate university students have the appropriate opportunities (apart from EAP classes) to learn specialized vocabulary for their chosen fields. While the answer to this question is beyond the scope of this particular study, I believe that it is still important to address the possibility of ESP for graduate students, given that opportunity. In other words, while it may not be possible to have ESP classes

at universities for all academic departments, there may be some universities where it is a possibility, at least to some degree.

Justification & Research Questions

Therefore, this particular study looks at the various ways corpora may be used in the ESP language classroom as well as the teaching applications and benefits of taking a corpus-based approach when creating ESP teaching materials. More specifically, this study looks at a small population of M.A. TESOL students at Michigan State University (MSU) and the ways in which corpora may benefit their English writing needs.

To give some background for the justification for this study, non-native speakers of English entering the M.A. TESOL program often complain about the lack of advanced/mastery level EAP courses offered at MSU. Many of these students are far too advanced in their English language skills to take most of the existing EAP courses; however, many have voiced that they would like some means of improving their English skills (both written and oral) in order to help them excel in their TESOL courses, as well as in their futures in teaching English.

One way of improving the writing skills of this select population is through the utilization of a corpus-based approach. The purpose of this research was to find a realistic tool for the improvement of the writing skills of students in the M.A. TESOL program; therefore, I have designed and developed various written corpora based on papers collected from native speakers of English in the M.A. TESOL program at MSU. It was then possible to look at the various teaching applications of the corpora frequency data for both non-native (and native speakers) of English, as well as what these data

imply about the various core TESOL courses. The questions that will be explored in this particular study based on the empirical data found are: 1.) What can be implied from the data frequencies of the corpora? 2.) What are some of the various teaching applications of these corpora in this ESP situation? 3.) How can the corpora be used to help students improve their writing skills? 4.) What do the corpora say about the writing skills of the native speakers? and 5.) What are the benefits and limitations of taking a corpus-based approach in this situation, namely using this approach for ESP curriculum and materials development?

Methodology

As previously mentioned, I designed and developed various written corpora based on papers collected from native speakers of English in the M.A. TESOL program. Papers from M.A. TESOL core courses (See Table 20 below) were collected from 14 native speaker participants.

Table 20: M.A. TESOL core courses at MSU

Course Number	Course Name
LLT 461	Introduction to Second Language Acquisition
LLT 807	Introduction to TESOL
LLT 808	Advanced Studies in TESOL
LLT 841	Special Topics in Language Teaching and Learning
LLT 896	Practicum

The criteria for the papers to be incorporated into the corpus were that they had to have been written in the past four years (Fall 1999- Spring 2003)[1], and to have received a grade of 3.5 or higher (on a 4.0 scale)[2]. A total of 99 papers were included in the corpora, each as a separate text file. The papers which were incorporated into the corpora

included: 461 and 808 summary papers, 461 topic papers, all term/research project papers, project proposals, book reviews, 896 reflective journals, all 841 papers, 807 classroom teaching observation papers, 896 peer reviews, and also various miscellaneous papers/essays written in 461 and 807. Lesson plans were not included, nor were papers from the core course 872[2]. All texts then had to be marked up or tagged accordingly, and submitted into the corpus. Software used for analyses were Wordsmith and MonoConc Pro. One Master TESOL Corpus was developed, as well as various smaller corpora (one for each course).

Results

As noted in Table 21 below, six corpora were initially created overall: 1 corpus per TESOL core course, as well as a Master Corpus that contained all texts combined:

Table 21: Corpora Created

Corpus	Number of Text Files	Number of words	% of Master Corpus
LLT 461	21	50, 132	35.4
LLT 807	21	28, 912	20.4
LLT 808	12	16, 045	11.3
LLT 841	9	15, 622	11.0
LLT 896	36	31, 058	21.9
Total = Master Corpus	99	141, 769	100

The Master Corpus contained 141,769 words based on the above-mentioned 99 text files. This is a relatively large corpus when taking the size of the M.A. TESOL population and the (only) 14 native speaker participants into consideration. At this point it is important to note the word counts and proportions of each of the smaller corpora in the Master

Corpus. For example, the LLT 896 corpus ended up having the most text files, while it was only ranked second in overall word frequency. Similarly, the LLT 461 and LLT 807 corpora contained the same amount of text files, yet a large discrepancy was seen in the comparison of their word tokens. All frequency data results per corpus may be seen in Appendix D. These frequency data results include most frequent word tokens per each corpus[3].

Discussion

There are numerous conclusions that one may make from the data frequency results, all of which are dependent on the interest and needs of the language learner or teacher who is analyzing the corpus. However, it is important to note that the list of possibilities is endless, and the implications and applications discussed below are only some examples of these possibilities. In this case, let us begin by looking at the initial research questions of this particular study.

1.) What can be implied from the data frequencies of the corpora?

Analyzing the smaller corpora, we see that there is a large difference between various core courses. For example, the word "I" is used 629 times in the LLT 896 corpus, whereas it is only used 41 times in the LLT 808 corpus (refer to Appendix D). This may be in part because the LLT 896 corpus included many reflective journals, which were of a more informal nature than the skills summary papers of LLT 808. Can we conclude that native speakers in the M.A. TESOL program are more likely to use the first person in informal writing, but not in formal writing? Can we conclude that LLT

896 has more informal homework assignments than 808? Possibly. Other examples of comparisons between TESOL core courses are between the LLT 807 corpus and the LLT 808 corpus. The two courses are similar, as 808 is a more advanced, theoretical continuation of 807. The frequency lists are similar, both containing words "students", "language", and "vocabulary" in their high frequency words (refer to Table 22 and Appendix D). However, the word "exercises" is a high frequency word in the LLT 807 corpus, yet not in the LLT 808 corpus (refer to Table 22 below). Similarly, the word "pragmatic" is a high frequency word in LLT 808, yet it does not occur *at all* in the LLT 807 corpus (nor do any derivatives of the word).

Table 22: Frequencies

	LLT 807 tokens	LLT 807 % (per 28, 912 words)	LLT 808 tokens	LLT 808 % (per 16, 045 words)
Students	433	1.50%	119	0.74%
Language	144	0.50%	176	1.10%
Vocabulary	156	0.54%	73	0.45%
Exercises	122	0.42%	8	>.01%
Pragmatic	0	0%	58	0.36%

Therefore, depending on one's objectives in analyzing the content of coursework (formality, length of papers, bibliographies, grammar usage, etc.), it can be as simple as comparing two lists of frequencies. While these assumptions from simple word occurrences are speculative, this provides an example of one of the possible analyses one may do with language in the corpora in general.

2.) What are some of the various teaching applications of these corpora in this ESP situation?

In Table 22, we see only frequency lists of extremely common words which students in the M.A. TESOL program are likely to already know. However, if we look more carefully, it would be possible to take the most common words in each course and have students study those particular content words that are specialized for the TESOL field, either on their own or in a highly specialized, advanced ESL course for the non-native speakers in the M.A. TESOL program. In other words, the applications of taking a corpus-based approach in an ESP classroom are similar to those in any language classroom where communicative competence in a particular area is crucial. However, some specific potential teaching applications of the TESOL corpora are:

1.) Using collocations to teach correct word usage and "TESOL" vocabulary for incoming TESOL students (e.g. feedback vs. metalinguistic feedback vs. corrective feedback, etc.)

2.) Creating pre-reading exercises for non-native speakers based on vocabulary in context (of research articles, for example)

3.) Obtaining access to real-world examples of academic writing for the purpose of improving the writing skills (in LLT 841, etc.) of all TESOL students (both native and non-native speakers), and

4.) Analyzing various written assignments and writing styles and looking at how to write effectively in the TESOL field in general (e.g. looking at introductions and conclusions from *TESOL Quarterly*, etc.).

This list is virtually endless, and may depend on the needs of the students. However, in whichever case one chooses to utilize the corpus for the purpose of language materials, the corpora will be available for analyses. Furthermore, as with any language tool, the use of corpora in the classroom will only improve with teacher piloting and time.

3.) How can the corpora be used to help students improve their writing skills?

Some examples of how M.A. students in the program could use the corpora to improve their writing skills on their own are:

1.) Students could examine collocations in order to understand correct word usage and grammar

2.) Students could read other past student essays to understand vocabulary for future classes, and

3.) Students can become familiar with the structure of various writing styles as well as with various written assignments for their courses (e.g. Book reviews, Summary Papers, Reflective Journals, etc.)

In other words, it would be relatively simple to create a specific "Introduction" corpus, as well as a "Book review" corpus (or virtually any type of corpus depending on student interests); however, given the small amount of papers used for this research and the fact that there was such a variety of papers submitted (resulting in only 2 or 3 of each assignment type), this was not possible, but is a future research possibility. With such specialized corpora available, TESOL students could use these corpora to focus on writing styles and techniques that they are expected to master in order to finish the M.A. program[4], even though the writing techniques may not be explicitly taught in the various TESOL core courses. Furthermore, students could examine essays together in class in order to gain insight on how to write effectively (since essentially the essays submitted into the Master Corpus have received a grade of 3.5 or above). Again, these are just a few of the ways that the M.A. TESOL students may use these specific corpora, yet the list is endless.

4.) What do the corpora say about the writing skills of the native speakers in the M.A. TESOL program?

My initial assumptions about the nature of the corpora would be that the frequency lists would contain more complex and intricate vocabulary words which would prove useful for non-native speakers looking to improve their breadth of vocabulary use in their writing. However, looking at the top 1000 most frequent words in each corpus, I do not believe that there are words that current non-native speakers in the M.A. TESOL program do not currently use themselves in their own papers. Therefore, while I would have to create a corpus of non-native speaker papers in the M.A. TESOL program to confirm this supposition, I think it is possible that there is a high correlation between vocabulary usage of native speakers and non-native speakers. Perhaps one way that the lists could prove to be useful would be for incoming students (both native speakers and non-native speakers) who wish to familiarize themselves with the program, as well as the expectations of graduate academic writing. However, it is at this point that it is important to address the limitations encountered in this study.

5.) What are the benefits and limitations of taking a corpus-based approach in this situation, namely using this approach for ESP curriculum and materials development?

One possible limitation of this particular research study is the size of the corpus. One million words are considered to be "standard" for a total size of a corpus. However, while size does not necessarily matter since corpus design is partially defined by research goals, I still would have liked to create a larger corpus (250,000 words). However, the number of native speakers in the M.A. TESOL program, those who were willing to

participate, and the issue of finding where we saved our papers limited this inquiry. Also, in order for the TESOL Master Corpus to be more representative of student writing, there should have been more balance between courses (e.g. each course being 1/5 of the total corpus). However, because of the total number of text files, it was necessary to include all files in the corpus, therefore contributing to the disproportion between courses. Another potential problem, which may arise by making the corpora available to all students, is plagiarism. Therefore, professors would have to implement some type of system to check for this.

Overall, these studies looked at a variety of issues in and applications of a corpus-based approach when creating ESP teaching materials and curricula. This research is relatively new; therefore, I hope to have offered a broad enough background for the reader. Also, I hope that the guidelines provided in the text for the implementation of a corpus in the classroom were supported well enough to have proven how useful a corpus-based approach might be for language teachers. Finally, and most importantly, I hope to have contributed to one possible means of creating some type of language improvement medium that may be used in improving the writing skills of the M.A. TESOL students in the future.

Notes:

[1]Within this particular time frame, the TESOL core courses were taught by the following:

	LLT 461	LLT 807	LLT 808	LLT 841	LLT 896
Fall 1999	Professor 1	Professor 2	N/A	N/A	Professor 4
Spring 2000	N/A	N/A	Professor 1	Professor 1	Professor 1
Fall 2000	Professor 1	Professor 2	N/A	N/A	Professor 4
Spring 2001	N/A	N/A	Professor 1	Professor 2	Professor 2
Fall 2001	Professor 1	Professor 2	N/A	N/A	Professor 4
Spring 2002	N/A	N/A	Professor 1	Professor 1	Professor 1
Fall 2002	Professor 3	Professor 2	N/A	N/A	N/A
Spring 2003	N/A	N/A	Professor 1	Professor 2	Professor 2

[2]Although the course "872: Research Methods" is a core course of the M.A. TESOL program at MSU, I decided not to include these particular papers because the majority of papers from this course were article reviews and research papers whose topics were chosen by the students based on their individual interests, and therefore were not papers which could be deemed as "typical of "an 872 paper since everyone was working on separate topics.

[3]Most frequent word tokens per corpus include any word types that occur in the corpus 5 or more times.

[4]For example, in order to matriculate from the M.A. TESOL program, one must complete either a Masters Thesis or a Teaching Portfolio (which includes various works of writing, including a personal statement of teaching philosophy and a textbook review).

CHAPTER 5
WHY TAKE A 'CORPUS-BASED' APPROACH?

Discussion

This thesis looked at the many ways in which a corpus-based approach can most effectively be utilized in the TESOL field in general, and in ESL textbook and materials development and evaluation in particular.

Chapter 1 gave the reader a background on corpus linguistics and the research that has been done to-date in corpus linguistics' application to language learning. From this background literature, I concluded that corpora could best be utilized for the purpose of textbook and materials development and evaluation, rather than simply data-driven language learning.

Chapter 2 provided evidence for one way that corpora may most effectively be utilized (based on empirical data) for the purpose of textbook evaluation. The study compared an academic speech textbook, *Discussion and Interaction in the Academic Community* (Madden & Rohlck, 2000), with the MICASE. Frequencies of expressions and collocations were compared, and the results suggested that the textbook was not representative of academic speech, as it had previously claimed. The MICASE corpus, as well as the COBUILD corpus, were then utilized in order to update the textbook and provide more realistic data and content for the textbook authors.

Chapter 3 of this thesis focused on the ways in which ESL teachers and researchers may develop textbooks based on information gathered from simple corpus concordance searching in relevant corpora, and how to most effectively transform basic search results into communicative teaching materials. This study specifically looked at

what idioms are most common in academic speech. Findings revealed 600 occurrences of idioms in the MICASE, providing empirical evidence of idioms in academic speech, should any researcher choose to develop a textbook focusing on this topic. This study provided a substantial amount of evidence for language teachers in academic environments who would like to teach idioms, without relying solely on idiosyncratic intuition. Furthermore, this study set a methodological framework as to the most efficient and productive searching techniques (as well as limitations) for language researchers choosing to create textbooks using an empirical corpus-based approach.

Lastly, Chapter 4 focused on the ways in which a corpus-based approach may be used for the purpose of materials development in an ESP classroom. In this study, I created a written corpus of writing taken from 14 native speaker graduate students in the M.A. TESOL program at MSU for the purpose of providing a means with which incoming non-native speakers in the M.A. TESOL program may improve their writing in English in their graduate classes. This study served solely as a prototype model, and introduced the many ways in which corpora may best be utilized to create a specific curriculum that meets the needs of the students, as well as considerations and limitations that need to be taken into account if creating corpora for materials development (e.g. what data to collect, how to collect it, what to do with it, etc.).

However, while I have focused mostly on the advantages of using a corpus-based approach in textbook and materials development and evaluation, there are multiple limitations to take into consideration as well before wholeheartedly accepting this new approach.

One criticism is that corpora can never replace the grammarian's intuition. Many language teachers may say that only the language teacher can know what is appropriate, or not appropriate. Furthermore, many teachers may complain that a corpus-based approach takes too long to find empirical data – they would rather 'make do' with what they already have. However, while some say that this is a limitation to creating language exercises in the classroom, it is not a limitation when creating or assessing textbooks, or when creating corpora to accommodate the content material of a particular population, which is what I have attempted to show in this thesis.

There are some other limitations of using a corpus-based approach when creating teaching materials, including that a computer cannot account for speaker idiosyncrasies. It is time-consuming to create corpora, and also the computer cannot do all of the language analyses; therefore, human qualitative input is always necessary to check for computer errors. While it may be relatively easy for a teacher to skim through and check the corpus materials for suitability (considering the amount of time that they would normally spend creating non-authentic materials or finding authentic materials anyway), the amount of time it takes to create corpus-based materials does seem to remain the biggest downfall of the technique. However, again, while the initial time it may take to evaluate or develop materials using this approach may be relatively great, the teacher/researcher does save time in the long run since they can easily return to the corpus to update their materials, rather than starting "from scratch" (a model for this has been shown in Chapter 2 of this thesis).

Lastly, there does seem to be another limitation of using a corpus-based approach in that it should not be used all of the time, nor should it completely replace

communicative curricula or textbooks that have proven to be effective for any language teacher. Rather, it is more important that the methods described in this thesis be used strategically and with empirical purpose. Owen (1993) has argued that over-dependence on a corpus as a basis for the development of pedagogical grammar can lead to "irrelevance, oversight and misrepresentation" (p.185). However, it could be argued that, at a developmental stage, of course corpus-based approaches to pedagogy may contain inaccuracies and lack of balance, but that it is already believed that "these shortcomings already exist in curricula based on intuition, traditional descriptions, tasks or other bases" (Kennedy, 1998, p. 290).

Conclusion

Overall, this thesis has looked at a variety of issues in and applications of a corpus-based approach when creating and assessing ESL teaching materials. This research is relatively new, and therefore I hope to have offered a broad enough background for the reader. Also, I hope that the guidelines provided in the text for the implementation of a corpus in the classroom were supported well enough to have proven how useful a corpus-based approach might be for language teachers.

In conclusion, what is most important to have gained from this research is the fact that a corpus-based approach allows one to obtain empirical data from language analyses in an efficient and effective manner. I recommend that this be for the purpose of evaluating a textbook, developing a particular textbook in some genre, or creating ESP materials for a select population, instead of simply data-driven language learning practices. In each of these cases, there are both advantages and limitations involved in

whether or not someone should choose to implement a corpus-based technique in the future; however, it should be clear that a corpus-based approach may be the solution to finding unlimited real-world scenarios and vocabulary in context, as well as pragmatically-loaded contexts that may be used in the ESP classroom. The applications of corpora in language teaching should not be any different from other communicative methods that have been proven to be effective in the classroom; however, the main difference in adopting this approach is that it will provide an authentic medium to motivate student learning, provide empirical data (that was not available before) with which to develop textbooks, and provide justification for teaching more specialized vocabulary to university graduate students that they may not receive in their own EAP courses. Lastly, and most importantly, the methods described in this thesis are not reserved simply for the computer-oriented, but rather may be adopted, using the models described, for most teachers without special training, thereby further promoting their accessibility, and leaving no reason that these methods should not be considered.

APPENDICES

APPENDIX A
SAMPLE ACTIVITIES USING A 'CORPUS-BASED APPROACH'

1. (taken from an MSU Colloquium handout presented by L. Zwier and C. Mazak of MSU's ELC (4/11/2002))

Vocabulary Detectives

Learning vocabulary is not just memorizing definitions, or even being able to use the word in a sentence. Really learning vocabulary involves learning the kinds of words that are used with the vocabulary word, and the situations that you are likely to find the vocabulary word in. For example, you might know that "scrutiny" means "looking closely at". But you might not know that you would hardly ever hear a native English speaker saying to a friend 'I watched the TV show with a lot of scrutiny."

So, how do you learn more about how and when to use a new word? Become a vocabulary detective!

Look at the list of sentences [taken from a corpus]. Then, work on these questions in groups.

1. What part of speech is this word?
2. Look at the sentences. What comes before the word?
3. What comes after the word?
4. Make a few formulas for using this word. What kinds of things usually go with the word?
5. Practice making sentences with the word.

Words that come before...	**SCRUTINY**	**Words that come after...**
	scrutiny	
	scrutiny	
	scrutiny	
	scrutiny	

Sample Practice and Test Items Based on Above Class Exercise

*Directions: Look at the sentences below. Circle the one where the word in **bold** is used incorrectly. Then explain WHY it is incorrect.*

1. The rock star is under intense scrutiny.
2. The teacher is under scrutiny for doing a wonderful job.
3. The mayor was under scrutiny for doing a bad job.

_____

_____

Sample response: Number 2 is incorrect because "scrutiny" should be used in a negative context only, and the context of number 2 is positive.

I will give you a situation. You write a sentence using the word.

Situation: The President of MSU has been stealing money from the university.

Scrutiny: _____

_____

2. (taken from an MSU Colloquium handout presented by Dr. R. Simpson of U of M's ELI (1/22/2002))

Focus: so +ADV/ADV +that clause

Directions: Match the sentence fragments on the left with the corresponding that-clause on the right.

1. We liked science and engineering **so much**	a. that dealers don't want to touch crack.
2. The bird is nestled **so deep** inside there	b. that even in sending them through the earth the probability of them doing anything on the way was, practically nil.
3. Now, typically you guys are **so close together**	c. that de decided to import those two pillars himself.
4. The faculty in this area are **so good**	d. that in order to get to it you usually have to destroy it.
5. We want to see the penalties **so stiff**	e. that there are many areas that cannot be cultivated at all.
6. They collided **so rarely**	f. that you can't even distinguish between the four of you.
7. He thought it was **so cool**, and, he was **so rich**	g. that we don't want to lose them.
8. In Chiapas, the topography is **so steep**	h. that we were willing to put up with this.

APPENDIX B
COMPLETE LIST OF PHRASES, COLLOQUIAL EXPRESSIONS, AND IDIOMS, FROM MADDEN AND ROHLCK'S *DISCUSSION & INTERACTION IN THE ACADEMIC COMMUNITY* (2000), CATEGORIZED BY CONTEXT

UNIT 1: INTERACTING IN THE ACADEMIC COMMUNITY
Colloquial Challenges (p.15)
Come on
I don't follow
What are you getting at?
Are you suggesting...?
I don't quite understand
Rapport and Complaining (p.16)
That's typical
I don't think so
Negotiating the Office Hour (p. 24)
I'm going to take...
I need to take...
Listening to the News: Idioms in the News (p. 42)
Playing hardball
Threw (him) a curve
Out of left field
Pinch hit
Strike out
Hit a home run
UNIT 2: PARTICIPATING IN THE ACADEMIC COMMUNITY
Opinions and Discussions: Getting Started (pp. 46)
What do you think about...?
How do you feel about...?
Do you agree with...?
Are you opposed to...?
What is your country's position on...?
Giving and Getting Opinions (pp. 47 – 49)
What do you think/feel about...?
Would you agree/say that...?
What's your opinion on/about...?
I believe/think/feel (that)...
I'm convinced (that)...
It seems to me (that)...
I don't feel strongly either way
Actually, I can see both points of view/both sides
I'm not sure, I haven't really thought about it

I don't know
Right/That's right
Exactly/That's true
I think so too...
I kind of agree...
I'm pretty much in agreement with you, but...
Probably you're right...
Well, maybe so, but...
Well, I don't necessarily agree with that
Hmmm, I see it somewhat differently
I see what you're saying, but...
You're wrong
Opinions and Summarizing (pp. 56)
The topic of this article is...
The point is that...
This article presents information about...
Most importantly...
Not surprisingly...
This article focuses on...
More Work on Opinions and Summarizing (pp. 59 – 60)
That's not quite right...
That's not what I meant...
You're close, but what I mean is...
We all agreed that...
We discussed the topic...
Most of us agree that...
Organizing the Discussion (pp. 64 – 66)
Today I would like to...
I want to present...
Today I'm going to talk about...
As we know...
As we have already seen...
As we have all read...
For example...
It's clear that...
Currently, however...
More specifically, I would like to discuss...
Strategies in Discussions (pp. 71)
I have to add two things
I want to say...
Could I interrupt for a moment...?
So is this situation the same as in your country?
Do you agree with what *** just said?
So, what is your opinion of this?
Concluding a Discussion (pp. 72)
So, to conclude...
In conclusion...
Let me summarize what we've talked about

We're just about out of time so I'd like to conclude by saying...
During the Discussion: Feedback (pp. 75 – 76)
Are you saying that...?
I don't quite understand what you are saying
Could you repeat that?
Are you following me?
Is that clear?
Is everyone with me?
Excuse me, but...
Ah, I think we need to move on to...
Perhaps we can get back to your point after we hear from...
UNIT 3: PRESENTING DATA IN THE ACADEMIC COMMUNITY
Idioms from Geometry (pp. 94)
Give (me) a straight answer
Going around in circles
Drew a blank
Gone off on a tangent
Get to the point
Read between the lines
Start at square one
To draw a parallel
Difficult Questions: Asking and Answering (pp. 105 – 109)
In other words...
What I'm saying is...
I mean...
What do you mean...?
I'm sorry, what exactly is your question?
I guess I just don't see your point. Could you clarify?
That's a very interesting question. What I have looked at it is...
That's a good question, but in my case...
I don't have the data (yet), but...
That's our next step
We haven't thought much about it yet, but if you would like to talk afterward we can...
I'd like to hear more about what you're saying
I feel I've answered the question, but we could continue our discussion later

APPENDIX C
COMPLETE LIST OF IDIOMATIC EXPRESSIONS, AS FOUND IN MICASE

Idiomatic Expressions	Occurrences
A dry run	1
All in my head	1
All wet	1
An oddball	1
Bag of tricks	1
Ball game	1
Bang for your buck	1
Beat around the bush	1
Beat into the ground	1
Beat the system	2
Beat to death	5
Bent out of shape	1
Blow by blow	1
Blow the whistle/whistle blower	2
Bottom line	17
Bouncing off of walls	1
Break even	1
Breathing down our neck	1
Breeze through	1
Can't win for losing	1
Catch up with	3
Catch you on (another day...)	2
Catch-22	1
Caught up in s.th.	4
Chicken and egg question	1
Child's play	1
Clean slate	2
Clean-cut	1
Clear cut	2
Come into play	16
Come out of the closet	4
Crash and burn	1
Creek without a paddle	1
Cry wolf	2
Cut a deal with	1
Cut s.o. off	3
Cut to the chase	2
Cutting edge	3
Dancing to that tune	1
Dead end	2
Does the trick	2
Doesn't cut it	2
Down and out	1
Down the line	12
Draw a/the line (between)	7
Drive me crazy/bananas	3
Drop in the bucket	1

Fall back on	6
Fall in love	5
Fall into place	4
Fall/throw by the wayside	3
Feather in their cap	1
Fine tune	3
First-come first-served	1
Flesh out	2
Flip a coin, flip side of a/the same co	10
Flooding the markets	1
Foot in the door	1
For all intents and purposes	3
For good	1
Full-fledged	4
Garden variety	1
Get a grasp of	1
Get a handle on	4
Get a head start on	3
Get around (to doing s.th)	5
Get to the bottom of things	1
Give and take	3
Go back on your word	2
Goes to show	4
Going to town	1
Good heavens	1
Goodness gracious	1
Hand in hand	8
Hand waving	2
Happy camper	1
Have/has the foggiest notion/idea	3
Having the world at his doorstep	1
Hit the wall, hitting a (rock) wall	3
Hit you over the head	1
I see the error of my ways	1
Ice breaker	1
In a nutshell	3
In bed with	1
In limbo	1
In line with s.th.	3
In the dark	2
In your wildest dreams	1
It's all downhill	1
Jump to conclusions	1
Just for the hell of it	2
Just trying to get by	2
Keep an eye on	1
Keep it up	1
Keep tabs (on)	1
Lay in on the line	1
Learning curve	1
Let the cat out of the bag/cat was...	1
Line of attack	4
Line to walk (fine?)	1

Lined up in advance	3
Litmus test	1
Live to tell about it	1
Long road	1
Making out like bandits	1
Million dollar question	1
Neck and neck	1
Nitty-gritty	4
No holds barred	1
No mean feat	1
Odds and ends	1
Off guard	2
Off the deep end	1
Off the wall	1
On a tangent, go off	2
On target	7
On the ball	1
On the fringe	1
On the right track	3
On the same page	4
On the same wavelength	1
On track	6
One catch	1
One fell swoop	1
One nit to pick	1
Open door	1
Out of this world	1
Out of whack	2
Out on a limb	1
Out the door	7
Pain in the neck	2
Pick (a) fight	3
Pick up on s.th.	10
Pick up where the last one took off	2
Pie in the sky	1
Play devil's advocate	2
Plug and chug	3
Pros and cons	11
Pulls the rug out from	2
Put a spin on	2
Put the heat on	5
Put them in their place	1
Rears its ugly head	3
Revolving door	1
Riding piggyback	2
Right off the bat (straight = 1)	8
Right on the dot	1
Right smack between	1
Ring a bell	4
Ripe for the picking	1
Rule/s of thumb	6
Run into a wall	3
Run it by you	2

Run of the mill	1
Scared to death	1
Send chills up and down the spine	1
Set his sights on	1
Shadow hanging over them	1
Shift gears	1
Sick of, sick (to death) of	1
Smack dab in the middle of	1
Smack up against	1
Splitting hairs	4
Stand out like a soar thumb	1
Steering clear of	3
Stick (their) neck out	2
Straight face	2
Strike a happy medium	1
Strike fear into the heart	1
Strike out on his own	1
Struck (such) a chord	1
Take (s.th) at face value	6
Take (something) to heart	1
Take him to task	2
Take it on faith	1
Take my word for it	4
Take on a life of its own	1
Take the plunge	1
Tall order	1
The big picture	7
The ivory tower	2
The kitchen sink	3
The saving grace	1
Thick of something	1
Thorn in your side (side in your thorn)	1
Throw off guard	1
Throw someone for a loop	1
Thumb you nose	1
Thumbs up	7
Time on your/his hands	2
Tipped your hand	1
Tongue in cheek	1
Touch base	2
Truck along on	4
Turn a blind eye	1
Turn a deaf ear	1
Turn a profit from	1
Turn of events	1
Turn of phrase/speech	2
Turn the tables/tables are turned	2
Upper hand (get/got the)	1
What the heck	8
What the hell	14
Which way to turn	2

APPENDIX D
COMPLETE WORD FREQUENCY LIST FOR CREATED CORPORA

Master Corpus
99 files; 141, 769 words

8444	5.9562%	the	438	0.3090%	which
4421	3.1185%	to	433	0.3054%	second
4240	2.9908%	and	423	0.2984%	what
4159	2.9336%	of	419	0.2956%	had
3330	2.3489%	in	416	0.2934%	but
3014	2.1260%	a	397	0.2800%	one
2026	1.4291%	that	393	0.2772%	these
1455	1.0263%	for	390	0.2751%	will
1391	0.9812%	language	378	0.2666%	also
1384	0.9762%	is	378	0.2666%	she
1217	0.8584%	students	369	0.2603%	can
1201	0.8472%	on	353	0.2490%	learners
1087	0.7667%	this	351	0.2476%	grammar
1057	0.7456%	as	348	0.2455%	age
1023	0.7216%	they	346	0.2441%	may
1011	0.7131%	i	346	0.2441%	reading
1006	0.7096%	it	344	0.2426%	use
1003	0.7075%	be	330	0.2328%	vocabulary
1002	0.7068%	with	327	0.2307%	teacher
959	0.6765%	their	298	0.2102%	time
815	0.5749%	are	294	0.2074%	some
801	0.5650%	not	292	0.2060%	acquisition
782	0.5516%	or	290	0.2046%	how
772	0.5445%	have	286	0.2017%	other
729	0.5142%	was	274	0.1933%	if
596	0.4204%	an	272	0.1919%	when
568	0.4007%	more	268	0.1890%	words
543	0.3830%	writing	265	0.1869%	each
498	0.3513%	at	260	0.1834%	l2
492	0.3470%	by	257	0.1813%	who
474	0.3343%	would	251	0.1770%	her
473	0.3336%	learning	251	0.1770%	there
466	0.3287%	were	250	0.1763%	do
453	0.3195%	english	250	0.1763%	so
450	0.3174%	class	248	0.1749%	could
443	0.3125%	them	248	0.1749%	first
442	0.3118%	from	240	0.1693%	he
440	0.3104%	about	236	0.1665%	questions

235	0.1658%	then	165	0.1164%	focus
226	0.1594%	than	163	0.1150%	level
224	0.1580%	used	161	0.1136%	should
212	0.1495%	however	160	0.1129%	since
209	0.1474%	might	157	0.1107%	lesson
208	0.1467%	study	156	0.1100%	skills
205	0.1446%	only	154	0.1086%	students'
204	0.1439%	speakers	153	0.1079%	factors
204	0.1439%	student	153	0.1079%	need
203	0.1432%	motivation	153	0.1079%	order
201	0.1418%	all	152	0.1072%	feedback
200	0.1411%	new	152	0.1072%	important
199	0.1404%	because	147	0.1037%	able
198	0.1397%	different	147	0.1037%	input
198	0.1397%	has	147	0.1037%	make
195	0.1375%	did	145	0.1023%	period
194	0.1368%	activity	144	0.1016%	learn
193	0.1361%	between	144	0.1016%	teachers
193	0.1361%	proficiency	140	0.0988%	children
192	0.1354%	information	139	0.0980%	much
192	0.1354%	my	136	0.0959%	speech
192	0.1354%	such	136	0.0959%	write
191	0.1347%	after	133	0.0938%	read
190	0.1340%	many	132	0.0931%	unit
189	0.1333%	exercises	132	0.0931%	you
189	0.1333%	native	130	0.0917%	asked
188	0.1326%	listening	130	0.0917%	book
186	0.1312%	well	129	0.0910%	textbook
185	0.1305%	we	127	0.0896%	even
184	0.1298%	been	127	0.0896%	test
184	0.1298%	example	127	0.0896%	through
184	0.1298%	into	126	0.0889%	1
182	0.1284%	form	126	0.0889%	part
182	0.1284%	p	126	0.0889%	up
181	0.1277%	very	125	0.0882%	work
177	0.1249%	while	124	0.0875%	program
175	0.1234%	using	124	0.0875%	word
174	0.1227%	most	123	0.0868%	accent
174	0.1227%	teaching	123	0.0868%	like
173	0.1220%	foreign	122	0.0861%	does
173	0.1220%	research	122	0.0861%	found
172	0.1213%	activities	122	0.0861%	same
172	0.1213%	learner	121	0.0854%	both
172	0.1213%	two	121	0.0854%	s
168	0.1185%	his	120	0.0846%	think
166	0.1171%	out	119	0.0839%	before

119	0.0839%	classroom	93	0.0656%	any
119	0.0839%	ideas	93	0.0656%	comprehension
118	0.0832%	group	93	0.0656%	people
118	0.0832%	own	93	0.0656%	where
117	0.0825%	topic	91	0.0642%	context
116	0.0818%	2001	91	0.0642%	given
116	0.0818%	fact	90	0.0635%	researchers
116	0.0818%	meaning	89	0.0628%	ability
115	0.0811%	studies	89	0.0628%	made
115	0.0811%	those	88	0.0621%	although
114	0.0804%	instruction	88	0.0621%	therefore
112	0.0790%	2	87	0.0614%	good
112	0.0790%	over	86	0.0607%	sla
112	0.0790%	text	85	0.0600%	authentic
111	0.0783%	sentences	84	0.0593%	according
110	0.0776%	content	84	0.0593%	course
110	0.0776%	just	84	0.0593%	difficult
109	0.0769%	practice	84	0.0593%	errors
109	0.0769%	speaking	84	0.0593%	few
108	0.0762%	pronunciation	84	0.0593%	its
107	0.0755%	often	84	0.0593%	related
107	0.0755%	target	84	0.0593%	sounds
106	0.0748%	better	83	0.0585%	exercise
106	0.0748%	point	82	0.0578%	error
105	0.0741%	give	82	0.0578%	must
104	0.0734%	during	81	0.0571%	1997
104	0.0734%	rodrigo	81	0.0571%	3
103	0.0727%	possible	81	0.0571%	results
103	0.0727%	process	80	0.0564%	correction
103	0.0727%	take	80	0.0564%	culture
102	0.0719%	materials	80	0.0564%	knowledge
102	0.0719%	section	80	0.0564%	type
101	0.0712%	based	79	0.0557%	appropriate
101	0.0712%	help	79	0.0557%	critical
100	0.0705%	being	79	0.0557%	groups
100	0.0705%	older	78	0.0550%	factor
99	0.0698%	another	78	0.0550%	points
99	0.0698%	correct	78	0.0550%	pragmatic
99	0.0698%	school	78	0.0550%	see
99	0.0698%	way	77	0.0543%	answers
98	0.0691%	know	77	0.0543%	chapter
97	0.0684%	no	77	0.0543%	differences
96	0.0677%	past	77	0.0543%	done
95	0.0670%	answer	77	0.0543%	task
95	0.0670%	esl	76	0.0536%	long
94	0.0663%	academic	76	0.0536%	social

76	0.0536%	three	63	0.0444%	certain
75	0.0529%	article	63	0.0444%	explicit
75	0.0529%	sentence	63	0.0444%	identity
75	0.0529%	t	63	0.0444%	j
74	0.0522%	communicative	63	0.0444%	main
74	0.0522%	hypothesis	62	0.0437%	especially
74	0.0522%	structure	62	0.0437%	me
73	0.0515%	adults	62	0.0437%	northstar
73	0.0515%	e	62	0.0437%	opportunities
73	0.0515%	journal	62	0.0437%	without
73	0.0515%	understand	61	0.0430%	1994
72	0.0508%	development	61	0.0430%	believe
72	0.0508%	high	61	0.0430%	effect
72	0.0508%	still	61	0.0430%	further
71	0.0501%	complete	61	0.0430%	grammatical
71	0.0501%	less	61	0.0430%	instead
71	0.0501%	presented	61	0.0430%	interesting
71	0.0501%	university	61	0.0430%	ll
71	0.0501%	whether	61	0.0430%	rather
70	0.0494%	community	61	0.0430%	short
70	0.0494%	material	61	0.0430%	situation
70	0.0494%	teach	61	0.0430%	within
70	0.0494%	usually	61	0.0430%	working
69	0.0487%	classes	60	0.0423%	amount
69	0.0487%	particular	60	0.0423%	approach
68	0.0480%	case	60	0.0423%	become
68	0.0480%	find	60	0.0423%	conversation
68	0.0480%	readings	60	0.0423%	environment
68	0.0480%	review	60	0.0423%	higher
68	0.0480%	variables	60	0.0423%	learner's
67	0.0473%	american	60	0.0423%	native-like
67	0.0473%	provide	60	0.0423%	spent
67	0.0473%	topics	59	0.0416%	1995
66	0.0466%	data	59	0.0416%	affect
66	0.0466%	present	59	0.0416%	discussion
65	0.0458%	languages	59	0.0416%	essay
65	0.0458%	minutes	59	0.0416%	noticing
65	0.0458%	next	59	0.0416%	several
65	0.0458%	states	59	0.0416%	show
65	0.0458%	want	58	0.0409%	4
64	0.0451%	camila	58	0.0409%	assignment
64	0.0451%	examples	58	0.0409%	attention
64	0.0451%	influence	58	0.0409%	having
64	0.0451%	interaction	58	0.0409%	series
64	0.0451%	question	57	0.0402%	competence
64	0.0451%	tasks	57	0.0402%	down

57	0.0402%	goals	52	0.0367%	younger
57	0.0402%	probably	51	0.0360%	am
57	0.0402%	specific	51	0.0360%	best
57	0.0402%	taught	51	0.0360%	due
56	0.0395%	feel	51	0.0360%	get
56	0.0395%	general	51	0.0360%	japanese
56	0.0395%	others	51	0.0360%	making
56	0.0395%	something	51	0.0360%	maturational
56	0.0395%	written	51	0.0360%	pp
55	0.0388%	1998	51	0.0360%	seem
55	0.0388%	accelerated	51	0.0360%	style
55	0.0388%	ask	51	0.0360%	themselves
55	0.0388%	either	51	0.0360%	types
55	0.0388%	furthermore	50	0.0353%	analysis
55	0.0388%	include	50	0.0353%	develop
55	0.0388%	m	50	0.0353%	terms
55	0.0388%	produce	50	0.0353%	understanding
55	0.0388%	structures	50	0.0353%	various
55	0.0388%	times	49	0.0346%	evidence
55	0.0388%	too	49	0.0346%	focused
54	0.0381%	1991	49	0.0346%	gave
54	0.0381%	aspects	49	0.0346%	involved
54	0.0381%	d	49	0.0346%	likely
54	0.0381%	johnson	49	0.0346%	problems
54	0.0381%	lecture	49	0.0346%	strategies
54	0.0381%	really	49	0.0346%	together
54	0.0381%	said	48	0.0339%	acquire
54	0.0381%	state	48	0.0339%	experience
54	0.0381%	why	48	0.0339%	future
54	0.0381%	years	48	0.0339%	last
53	0.0374%	already	48	0.0339%	puberty
53	0.0374%	awareness	48	0.0339%	success
53	0.0374%	communication	48	0.0339%	united
53	0.0374%	learned	47	0.0332%	1993
53	0.0374%	levels	47	0.0332%	actually
53	0.0374%	life	47	0.0332%	anxiety
53	0.0374%	problem	47	0.0332%	degree
53	0.0374%	simply	47	0.0332%	experiences
52	0.0367%	-	47	0.0332%	forms
52	0.0367%	business	47	0.0332%	meaningful
52	0.0367%	chart	47	0.0332%	necessary
52	0.0367%	developing	47	0.0332%	r
52	0.0367%	enough	47	0.0332%	role
52	0.0367%	helpful	47	0.0332%	say
52	0.0367%	instructor	47	0.0332%	sections
52	0.0367%	participants	46	0.0324%	cited

46	0.0324%	doing	42	0.0296%	speak
46	0.0324%	essays	42	0.0296%	speaker
46	0.0324%	five	42	0.0296%	true
46	0.0324%	him	42	0.0296%	variety
46	0.0324%	objective	42	0.0296%	wanted
46	0.0324%	paper	41	0.0289%	b
46	0.0324%	tesol	41	0.0289%	believes
46	0.0324%	towards	41	0.0289%	consider
46	0.0324%	useful	41	0.0289%	every
45	0.0317%	attitudes	41	0.0289%	heinle
45	0.0317%	effects	41	0.0289%	identify
45	0.0317%	following	41	0.0289%	later
45	0.0317%	linguistic	41	0.0289%	output
45	0.0317%	looking	41	0.0289%	overall
45	0.0317%	person	41	0.0289%	situations
45	0.0317%	relationship	41	0.0289%	subjects
45	0.0317%	theme	40	0.0282%	again
45	0.0317%	wrote	40	0.0282%	back
44	0.0310%	adult	40	0.0282%	french
44	0.0310%	come	40	0.0282%	lot
44	0.0310%	complex	40	0.0282%	natural
44	0.0310%	effort	40	0.0282%	pictures
44	0.0310%	end	40	0.0282%	previous
44	0.0310%	going	40	0.0282%	provides
44	0.0310%	idioms	40	0.0282%	quality
44	0.0310%	items	40	0.0282%	result
44	0.0310%	organization	40	0.0282%	seems
44	0.0310%	someone	40	0.0282%	took
43	0.0303%	authors	39	0.0275%	beginning
43	0.0303%	clear	39	0.0275%	brain
43	0.0303%	g	39	0.0275%	comprehensible
43	0.0303%	homework	39	0.0275%	cultural
43	0.0303%	listen	39	0.0275%	difference
43	0.0303%	objectives	39	0.0275%	exposure
43	0.0303%	once	39	0.0275%	giving
43	0.0303%	picture	39	0.0275%	goal
43	0.0303%	showed	39	0.0275%	little
43	0.0303%	similar	39	0.0275%	notes
43	0.0303%	teacher's	39	0.0275%	regarding
43	0.0303%	thought	39	0.0275%	tests
43	0.0303%	ways	39	0.0275%	your
42	0.0296%	1996	38	0.0268%	2000
42	0.0296%	explicitly	38	0.0268%	al
42	0.0296%	hand	38	0.0268%	basic
42	0.0296%	number	38	0.0268%	board
42	0.0296%	place	38	0.0268%	et

38	0.0268%	final	35	0.0247%	begin
38	0.0268%	go	35	0.0247%	day
38	0.0268%	intermediate	35	0.0247%	discussed
38	0.0268%	methods	35	0.0247%	four
38	0.0268%	negotiation	35	0.0247%	gardner
38	0.0268%	opportunity	35	0.0247%	interest
38	0.0268%	outside	35	0.0247%	interested
38	0.0268%	paragraph	35	0.0247%	introduction
38	0.0268%	partner	35	0.0247%	medical
38	0.0268%	personal	35	0.0247%	never
38	0.0268%	pragmatics	35	0.0247%	phonology
38	0.0268%	provided	35	0.0247%	quarterly
38	0.0268%	reason	35	0.0247%	response
38	0.0268%	worked	35	0.0247%	support
37	0.0261%	achieve	34	0.0240%	abilities
37	0.0261%	always	34	0.0240%	around
37	0.0261%	claims	34	0.0240%	articles
37	0.0261%	details	34	0.0240%	aware
37	0.0261%	dörnyei	34	0.0240%	control
37	0.0261%	effective	34	0.0240%	discuss
37	0.0261%	etc	34	0.0240%	easy
37	0.0261%	flege	34	0.0240%	friends
37	0.0261%	immersion	34	0.0240%	importance
37	0.0261%	lack	34	0.0240%	mind
37	0.0261%	model	34	0.0240%	now
37	0.0261%	passive	34	0.0240%	our
37	0.0261%	peer	34	0.0240%	oxford
37	0.0261%	production	34	0.0240%	taking
37	0.0261%	spanish	34	0.0240%	video
37	0.0261%	strong	34	0.0240%	voice
37	0.0261%	studying	33	0.0233%	advantage
36	0.0254%	1999	33	0.0233%	changes
36	0.0254%	5	33	0.0233%	discourse
36	0.0254%	advanced	33	0.0233%	early
36	0.0254%	change	33	0.0233%	integrative
36	0.0254%	home	33	0.0233%	italian
36	0.0254%	improve	33	0.0233%	keep
36	0.0254%	krashen	33	0.0233%	low
36	0.0254%	least	33	0.0233%	mistakes
36	0.0254%	makes	33	0.0233%	non-native
36	0.0254%	reasons	33	0.0233%	notice
36	0.0254%	sensitive	33	0.0233%	purpose
36	0.0254%	simple	33	0.0233%	rules
36	0.0254%	story	33	0.0233%	small
35	0.0247%	affective	33	0.0233%	taken
35	0.0247%	asking	33	0.0233%	went

33	0.0233%	whole	30	0.0212%	list
32	0.0226%	1989	30	0.0212%	longer
32	0.0226%	achievement	30	0.0212%	lyster
32	0.0226%	describe	30	0.0212%	matter
32	0.0226%	earlier	30	0.0212%	means
32	0.0226%	kind	30	0.0212%	passage
32	0.0226%	korean	30	0.0212%	require
32	0.0226%	look	30	0.0212%	seemed
32	0.0226%	needed	30	0.0212%	self-confidence
32	0.0226%	needs	30	0.0212%	significant
32	0.0226%	newport	30	0.0212%	specifically
32	0.0226%	note	30	0.0212%	talking
32	0.0226%	perhaps	30	0.0212%	though
32	0.0226%	phonological	30	0.0212%	uses
32	0.0226%	providing	29	0.0205%	affects
32	0.0226%	sure	29	0.0205%	background
32	0.0226%	themes	29	0.0205%	cannot
32	0.0226%	trying	29	0.0205%	ed
32	0.0226%	units	29	0.0205%	fill
31	0.0219%	1985	29	0.0205%	lower
31	0.0219%	accuracy	29	0.0205%	phrases
31	0.0219%	arrival	29	0.0205%	positive
31	0.0219%	assignments	29	0.0205%	quickly
31	0.0219%	began	29	0.0205%	student's
31	0.0219%	clearly	29	0.0205%	texts
31	0.0219%	cognitive	29	0.0205%	things
31	0.0219%	expressions	29	0.0205%	truscott
31	0.0219%	features	28	0.0198%	available
31	0.0219%	felt	28	0.0198%	create
31	0.0219%	idea	28	0.0198%	direct
31	0.0219%	included	28	0.0198%	don't
31	0.0219%	including	28	0.0198%	draft
31	0.0219%	individual	28	0.0198%	lead
31	0.0219%	issues	28	0.0198%	letter
31	0.0219%	responses	28	0.0198%	mean
31	0.0219%	right	28	0.0198%	multiple
31	0.0219%	setting	28	0.0198%	notion
31	0.0219%	try	28	0.0198%	raters
31	0.0219%	u	28	0.0198%	required
30	0.0212%	allow	28	0.0198%	rest
30	0.0212%	c	28	0.0198%	status
30	0.0212%	charts	28	0.0198%	stevens
30	0.0212%	choice	28	0.0198%	suggests
30	0.0212%	follow	28	0.0198%	tense
30	0.0212%	gap	28	0.0198%	testing
30	0.0212%	hear	28	0.0198%	thus

27	0.0190%	ages	26	0.0183%	v
27	0.0190%	areas	26	0.0183%	young
27	0.0190%	classrooms	25	0.0176%	addition
27	0.0190%	contains	25	0.0176%	among
27	0.0190%	entire	25	0.0176%	becomes
27	0.0190%	focuses	25	0.0176%	chapters
27	0.0190%	formal	25	0.0176%	child
27	0.0190%	half	25	0.0176%	closely
27	0.0190%	initial	25	0.0176%	considered
27	0.0190%	interactions	25	0.0176%	continuous
27	0.0190%	interlanguage	25	0.0176%	discussions
27	0.0190%	interview	25	0.0176%	divided
27	0.0190%	mentioned	25	0.0176%	encouraged
27	0.0190%	metalinguistic	25	0.0176%	explain
27	0.0190%	negative	25	0.0176%	great
27	0.0190%	primary	25	0.0176%	increase
27	0.0190%	put	25	0.0176%	intended
27	0.0190%	qi	25	0.0176%	nature
27	0.0190%	says	25	0.0176%	pairs
27	0.0190%	seen	25	0.0176%	papers
27	0.0190%	skill	25	0.0176%	patterns
27	0.0190%	snow	25	0.0176%	reformulation
27	0.0190%	successful	25	0.0176%	set
27	0.0190%	tend	25	0.0176%	stated
26	0.0183%	acquired	25	0.0176%	studied
26	0.0183%	audience	25	0.0176%	system
26	0.0183%	bardovi-harlig	25	0.0176%	third
26	0.0183%	comments	25	0.0176%	until
26	0.0183%	constraints	25	0.0176%	upon
26	0.0183%	edition	24	0.0169%	affected
26	0.0183%	express	24	0.0169%	argues
26	0.0183%	family	24	0.0169%	away
26	0.0183%	itself	24	0.0169%	check
26	0.0183%	lapkin	24	0.0169%	college
26	0.0183%	lessons	24	0.0169%	comfortable
26	0.0183%	major	24	0.0169%	country
26	0.0183%	method	24	0.0169%	determine
26	0.0183%	michigan	24	0.0169%	easily
26	0.0183%	one's	24	0.0169%	findings
26	0.0183%	producing	24	0.0169%	haejung
26	0.0183%	sometimes	24	0.0169%	includes
26	0.0183%	suggestions	24	0.0169%	job
26	0.0183%	talk	24	0.0169%	pages
26	0.0183%	textbooks	24	0.0169%	potential
26	0.0183%	told	24	0.0169%	practical
26	0.0183%	understood	24	0.0169%	rate

24	0.0169%	recast	22	0.0155%	didn't
24	0.0169%	relevant	22	0.0155%	difficulty
24	0.0169%	scoring	22	0.0155%	dutch
24	0.0169%	subject	22	0.0155%	examined
24	0.0169%	us	22	0.0155%	exposed
23	0.0162%	10	22	0.0155%	far
23	0.0162%	assume	22	0.0155%	impossible
23	0.0162%	bongaerts	22	0.0155%	move
23	0.0162%	books	22	0.0155%	n
23	0.0162%	called	22	0.0155%	ohp
23	0.0162%	compare	22	0.0155%	plan
23	0.0162%	continue	22	0.0155%	preparing
23	0.0162%	developed	22	0.0155%	progress
23	0.0162%	directly	22	0.0155%	quite
23	0.0162%	efl	22	0.0155%	range
23	0.0162%	explained	22	0.0155%	references
23	0.0162%	key	22	0.0155%	researcher
23	0.0162%	l	22	0.0155%	scores
23	0.0162%	led	22	0.0155%	sense
23	0.0162%	length	22	0.0155%	shows
23	0.0162%	living	22	0.0155%	six
23	0.0162%	pair	22	0.0155%	sound
23	0.0162%	parents	22	0.0155%	tanaka
23	0.0162%	particularly	22	0.0155%	turn
23	0.0162%	performance	22	0.0155%	unable
23	0.0162%	processing	22	0.0155%	york
23	0.0162%	recasts	21	0.0148%	12
23	0.0162%	remote	21	0.0148%	15
23	0.0162%	started	21	0.0148%	1986
23	0.0162%	syntax	21	0.0148%	20
23	0.0162%	talked	21	0.0148%	above
23	0.0162%	ultimate	21	0.0148%	acquiring
23	0.0162%	varied	21	0.0148%	along
22	0.0155%	active	21	0.0148%	assessment
22	0.0155%	almost	21	0.0148%	build
22	0.0155%	area	21	0.0148%	categories
22	0.0155%	argue	21	0.0148%	computer
22	0.0155%	associated	21	0.0148%	conversational
22	0.0155%	benefit	21	0.0148%	corrections
22	0.0155%	came	21	0.0148%	deal
22	0.0155%	comes	21	0.0148%	decline
22	0.0155%	common	21	0.0148%	familiar
22	0.0155%	comparing	21	0.0148%	greater
22	0.0155%	contexts	21	0.0148%	immediate
22	0.0155%	days	21	0.0148%	introduce
22	0.0155%	desire	21	0.0148%	involving

21	0.0148%	kinds	20	0.0141%	tried
21	0.0148%	large	20	0.0141%	verb
21	0.0148%	limited	20	0.0141%	verbs
21	0.0148%	majority	20	0.0141%	week
21	0.0148%	offer	20	0.0141%	world
21	0.0148%	page	20	0.0141%	year
21	0.0148%	possibly	20	0.0141%	yet
21	0.0148%	press	19	0.0134%	1990
21	0.0148%	received	19	0.0134%	7
21	0.0148%	start	19	0.0134%	access
21	0.0148%	techniques	19	0.0134%	account
21	0.0148%	toward	19	0.0134%	accurately
20	0.0141%	0	19	0.0134%	advertisement
20	0.0141%	6	19	0.0134%	applied
20	0.0141%	basis	19	0.0134%	biological
20	0.0141%	believed	19	0.0134%	challenge
20	0.0141%	benefits	19	0.0134%	consideration
20	0.0141%	building	19	0.0134%	encourage
20	0.0141%	chile	19	0.0134%	field
20	0.0141%	choose	19	0.0134%	function
20	0.0141%	compared	19	0.0134%	getting
20	0.0141%	concerned	19	0.0134%	hour
20	0.0141%	creating	19	0.0134%	k
20	0.0141%	curriculum	19	0.0134%	lectures
20	0.0141%	depending	19	0.0134%	manual
20	0.0141%	duquette	19	0.0134%	meanings
20	0.0141%	easier	19	0.0134%	months
20	0.0141%	eds	19	0.0134%	observation
20	0.0141%	education	19	0.0134%	off
20	0.0141%	essential	19	0.0134%	optimal
20	0.0141%	explanation	19	0.0134%	programs
20	0.0141%	follows	19	0.0134%	real
20	0.0141%	guess	19	0.0134%	rhetorical
20	0.0141%	instrumental	19	0.0134%	showing
20	0.0141%	interviews	19	0.0134%	stage
20	0.0141%	issue	19	0.0134%	statements
20	0.0141%	learners'	19	0.0134%	styles
20	0.0141%	motivated	19	0.0134%	suggested
20	0.0141%	noticed	19	0.0134%	thinking
20	0.0141%	opinion	19	0.0134%	training
20	0.0141%	presentation	19	0.0134%	view
20	0.0141%	proficient	18	0.0127%	accents
20	0.0141%	receive	18	0.0127%	additional
20	0.0141%	stress	18	0.0127%	affecting
20	0.0141%	suggest	18	0.0127%	apply
20	0.0141%	today	18	0.0127%	argument

18	0.0127%	assumptions	17	0.0120%	became
18	0.0127%	beyond	17	0.0120%	begins
18	0.0127%	bring	17	0.0120%	camila's
18	0.0127%	calts	17	0.0120%	cases
18	0.0127%	cambridge	17	0.0120%	cause
18	0.0127%	chinese	17	0.0120%	celce-murcia
18	0.0127%	claim	17	0.0120%	clément
18	0.0127%	completed	17	0.0120%	closing
18	0.0127%	consists	17	0.0120%	comparison
18	0.0127%	contrast	17	0.0120%	conclusion
18	0.0127%	describing	17	0.0120%	corpus
18	0.0127%	directions	17	0.0120%	correlation
18	0.0127%	draw	17	0.0120%	countries
18	0.0127%	elicited	17	0.0120%	daily
18	0.0127%	everything	17	0.0120%	definition
18	0.0127%	examine	17	0.0120%	elicit
18	0.0127%	false	17	0.0120%	encourages
18	0.0127%	fluent	17	0.0120%	everyone
18	0.0127%	helped	17	0.0120%	expression
18	0.0127%	intensive	17	0.0120%	games
18	0.0127%	interact	17	0.0120%	gass
18	0.0127%	involvement	17	0.0120%	global
18	0.0127%	john	17	0.0120%	h
18	0.0127%	larsen-freeman	17	0.0120%	house
18	0.0127%	lenneberg	17	0.0120%	increased
18	0.0127%	lexical	17	0.0120%	looked
18	0.0127%	modern	17	0.0120%	markers
18	0.0127%	observed	17	0.0120%	murder
18	0.0127%	occur	17	0.0120%	newspaper
18	0.0127%	organize	17	0.0120%	pick
18	0.0127%	outline	17	0.0120%	play
18	0.0127%	prepare	17	0.0120%	plus
18	0.0127%	problematic	17	0.0120%	project
18	0.0127%	regards	17	0.0120%	qualitative
18	0.0127%	relationships	17	0.0120%	rated
18	0.0127%	room	17	0.0120%	remember
18	0.0127%	sheet	17	0.0120%	rich
18	0.0127%	stages	17	0.0120%	shown
18	0.0127%	ten	17	0.0120%	swain
18	0.0127%	worksheet	17	0.0120%	tested
18	0.0127%	wrong	17	0.0120%	utterance
17	0.0120%	analyze	17	0.0120%	wu
17	0.0120%	assess	16	0.0113%	address
17	0.0120%	atmosphere	16	0.0113%	ahead
17	0.0120%	attached	16	0.0113%	allows
17	0.0120%	average	16	0.0113%	appears

16	0.0113%	arrived	16	0.0113%	spend
16	0.0113%	asher	16	0.0113%	stoller
16	0.0113%	beneficial	16	0.0113%	ultimately
16	0.0113%	bit	16	0.0113%	variable
16	0.0113%	call	15	0.0106%	25
16	0.0113%	clarification	15	0.0106%	8
16	0.0113%	compositions	15	0.0106%	against
16	0.0113%	correcting	15	0.0106%	agree
16	0.0113%	corrective	15	0.0106%	americans
16	0.0113%	created	15	0.0106%	analyzing
16	0.0113%	description	15	0.0106%	attainment
16	0.0113%	designed	15	0.0106%	attempt
16	0.0113%	encounter	15	0.0106%	brainstorming
16	0.0113%	feels	15	0.0106%	challenging
16	0.0113%	fluency	15	0.0106%	concerns
16	0.0113%	garcia	15	0.0106%	conversations
16	0.0113%	helps	15	0.0106%	correctly
16	0.0113%	here	15	0.0106%	courses
16	0.0113%	ideal	15	0.0106%	cover
16	0.0113%	implications	15	0.0106%	cues
16	0.0113%	improved	15	0.0106%	determines
16	0.0113%	indicate	15	0.0106%	emphasis
16	0.0113%	international	15	0.0106%	encountered
16	0.0113%	introduced	15	0.0106%	evaluate
16	0.0113%	involve	15	0.0106%	exploring
16	0.0113%	late	15	0.0106%	extra
16	0.0113%	laufer	15	0.0106%	extremely
16	0.0113%	mistake	15	0.0106%	face
16	0.0113%	mostly	15	0.0106%	frequency
16	0.0113%	msu	15	0.0106%	functions
16	0.0113%	oyama	15	0.0106%	grabe
16	0.0113%	peers	15	0.0106%	highly
16	0.0113%	progressive	15	0.0106%	idiomatic
16	0.0113%	promote	15	0.0106%	immigration
16	0.0113%	prompt	15	0.0106%	individually
16	0.0113%	protocols	15	0.0106%	involves
16	0.0113%	proved	15	0.0106%	it's
16	0.0113%	real-life	15	0.0106%	leads
16	0.0113%	recall	15	0.0106%	limitations
16	0.0113%	repeat	15	0.0106%	linguistics
16	0.0113%	requires	15	0.0106%	lives
16	0.0113%	retain	15	0.0106%	original
16	0.0113%	samples	15	0.0106%	passages
16	0.0113%	semester	15	0.0106%	primarily
16	0.0113%	settings	15	0.0106%	ranta
16	0.0113%	sources	15	0.0106%	respond

15	0.0106%	responded	14	0.0099%	guidelines
15	0.0106%	retention	14	0.0099%	immediately
15	0.0106%	rule	14	0.0099%	immigrants
15	0.0106%	saying	14	0.0099%	incorrect
15	0.0106%	scale	14	0.0099%	influences
15	0.0106%	share	14	0.0099%	instructions
15	0.0106%	somewhat	14	0.0099%	instructors
15	0.0106%	spoken	14	0.0099%	interactive
15	0.0106%	step	14	0.0099%	jobs
15	0.0106%	takahashi	14	0.0099%	letters
15	0.0106%	technique	14	0.0099%	meant
15	0.0106%	term	14	0.0099%	meet
15	0.0106%	theories	14	0.0099%	morphology
15	0.0106%	thing	14	0.0099%	noels
15	0.0106%	utterances	14	0.0099%	old
15	0.0106%	version	14	0.0099%	participant
15	0.0106%	writers	14	0.0099%	participate
14	0.0099%	16	14	0.0099%	pica
14	0.0099%	1984	14	0.0099%	piece
14	0.0099%	24	14	0.0099%	principles
14	0.0099%	30	14	0.0099%	prior
14	0.0099%	acts	14	0.0099%	quiz
14	0.0099%	aloud	14	0.0099%	realize
14	0.0099%	approaches	14	0.0099%	slavoff
14	0.0099%	asks	14	0.0099%	spelling
14	0.0099%	association	14	0.0099%	statement
14	0.0099%	beebe	14	0.0099%	stories
14	0.0099%	center	14	0.0099%	sub-section
14	0.0099%	certainly	14	0.0099%	sufficient
14	0.0099%	children's	14	0.0099%	supplement
14	0.0099%	chose	14	0.0099%	supports
14	0.0099%	clues	14	0.0099%	takes
14	0.0099%	coady	14	0.0099%	theoretical
14	0.0099%	communicate	14	0.0099%	think-aloud
14	0.0099%	completing	14	0.0099%	throughout
14	0.0099%	concluded	14	0.0099%	travel
14	0.0099%	corrected	14	0.0099%	turns
14	0.0099%	eisenstein	14	0.0099%	visual
14	0.0099%	emphasized	13	0.0092%	17
14	0.0099%	examining	13	0.0092%	19
14	0.0099%	explore	13	0.0092%	1981
14	0.0099%	followed	13	0.0092%	ad
14	0.0099%	format	13	0.0092%	addressed
14	0.0099%	genie	13	0.0092%	advance
14	0.0099%	goes	13	0.0092%	alone
14	0.0099%	got	13	0.0092%	answered

13	0.0092%	answering	13	0.0092%	purple
13	0.0092%	appear	13	0.0092%	purposes
13	0.0092%	approaching	13	0.0092%	quantity
13	0.0092%	chance	13	0.0092%	raise
13	0.0092%	classmates	13	0.0092%	recasting
13	0.0092%	component	13	0.0092%	recognize
13	0.0092%	debate	13	0.0092%	relatively
13	0.0092%	dimensions	13	0.0092%	report
13	0.0092%	discussing	13	0.0092%	resources
13	0.0092%	english-only	13	0.0092%	reviewing
13	0.0092%	established	13	0.0092%	roles
13	0.0092%	evaluated	13	0.0092%	schmidt
13	0.0092%	evaluation	13	0.0092%	search
13	0.0092%	events	13	0.0092%	shearin
13	0.0092%	expected	13	0.0092%	spada
13	0.0092%	feelings	13	0.0092%	successfully
13	0.0092%	finally	13	0.0092%	suited
13	0.0092%	finished	13	0.0092%	summary
13	0.0092%	free	13	0.0092%	superior
13	0.0092%	frequently	13	0.0092%	supplemented
13	0.0092%	full	13	0.0092%	supported
13	0.0092%	fully	13	0.0092%	today's
13	0.0092%	fun	13	0.0092%	top
13	0.0092%	generally	13	0.0092%	trait
13	0.0092%	hulstijn	13	0.0092%	usage
13	0.0092%	identification	12	0.0085%	1976
13	0.0092%	ii	12	0.0085%	1980
13	0.0092%	improving	12	0.0085%	9
13	0.0092%	instances	12	0.0085%	act
13	0.0092%	integrativeness	12	0.0085%	arm
13	0.0092%	intonation	12	0.0085%	assigned
13	0.0092%	kasper	12	0.0085%	attitude
13	0.0092%	let	12	0.0085%	behind
13	0.0092%	lines	12	0.0085%	bibliography
13	0.0092%	listed	12	0.0085%	bodman
13	0.0092%	literacy	12	0.0085%	break
13	0.0092%	longman	12	0.0085%	brought
13	0.0092%	naturally	12	0.0085%	category
13	0.0092%	necessarily	12	0.0085%	challenges
13	0.0092%	night	12	0.0085%	cohesion
13	0.0092%	obtain	12	0.0085%	constraint
13	0.0092%	obvious	12	0.0085%	contain
13	0.0092%	onto	12	0.0085%	continued
13	0.0092%	orientation	12	0.0085%	deeper
13	0.0092%	perform	12	0.0085%	design
13	0.0092%	perspectives	12	0.0085%	determining

12	0.0085%	difficulties	12	0.0085%	regardless
12	0.0085%	doctor	12	0.0085%	repairs
12	0.0085%	enhanced	12	0.0085%	requests
12	0.0085%	enjoyed	12	0.0085%	revise
12	0.0085%	eventually	12	0.0085%	robb
12	0.0085%	exceptions	12	0.0085%	ross
12	0.0085%	explains	12	0.0085%	segment
12	0.0085%	explanations	12	0.0085%	shortreed
12	0.0085%	expressing	12	0.0085%	stronger
12	0.0085%	extensive	12	0.0085%	su
12	0.0085%	fll	12	0.0085%	technology
12	0.0085%	front	12	0.0085%	tell
12	0.0085%	german	12	0.0085%	theory
12	0.0085%	implicit	12	0.0085%	varonis
12	0.0085%	influenced	12	0.0085%	vi
12	0.0085%	initially	12	0.0085%	vot
12	0.0085%	instructed	12	0.0085%	w
12	0.0085%	instructional	12	0.0085%	whereas
12	0.0085%	investment	12	0.0085%	willing
12	0.0085%	knew	12	0.0085%	works
12	0.0085%	leading	11	0.0078%	--
12	0.0085%	leave	11	0.0078%	1977
12	0.0085%	lost	11	0.0078%	achieved
12	0.0085%	measure	11	0.0078%	advantages
12	0.0085%	measures	11	0.0078%	aid
12	0.0085%	models	11	0.0078%	anything
12	0.0085%	motivational	11	0.0078%	body
12	0.0085%	moved	11	0.0078%	bottom
12	0.0085%	neufeld	11	0.0078%	checking
12	0.0085%	noted	11	0.0078%	cities
12	0.0085%	o	11	0.0078%	classic
12	0.0085%	online	11	0.0078%	clauses
12	0.0085%	opinions	11	0.0078%	comic
12	0.0085%	oral	11	0.0078%	connection
12	0.0085%	organized	11	0.0078%	consequently
12	0.0085%	outlining	11	0.0078%	construct
12	0.0085%	paragraphs	11	0.0078%	creative
12	0.0085%	paraphrasing	11	0.0078%	current
12	0.0085%	plans	11	0.0078%	dealing
12	0.0085%	plays	11	0.0078%	depends
12	0.0085%	population	11	0.0078%	describes
12	0.0085%	practicum	11	0.0078%	descriptive
12	0.0085%	principle	11	0.0078%	devices
12	0.0085%	professor	11	0.0078%	distinguish
12	0.0085%	recent	11	0.0078%	drawing
12	0.0085%	refer	11	0.0078%	educational

11	0.0078%	f	11	0.0078%	proper
11	0.0078%	failure	11	0.0078%	psychological
11	0.0078%	fieldwork	11	0.0078%	ready
11	0.0078%	figure	11	0.0078%	reminded
11	0.0078%	finding	11	0.0078%	repair
11	0.0078%	floor	11	0.0078%	repetition
11	0.0078%	focusing	11	0.0078%	requirements
11	0.0078%	frequent	11	0.0078%	reviewed
11	0.0078%	gain	11	0.0078%	rodrigo's
11	0.0078%	gender	11	0.0078%	sample
11	0.0078%	gratitude	11	0.0078%	saw
11	0.0078%	holistic	11	0.0078%	scene
11	0.0078%	imitate	11	0.0078%	schools
11	0.0078%	incorporate	11	0.0078%	schumann
11	0.0078%	indirect	11	0.0078%	self-study
11	0.0078%	internet	11	0.0078%	society
11	0.0078%	kept	11	0.0078%	ssla
11	0.0078%	larger	11	0.0078%	standard
11	0.0078%	maintain	11	0.0078%	student-centered
11	0.0078%	manner	11	0.0078%	surface
11	0.0078%	maps	11	0.0078%	syllabus
11	0.0078%	maybe	11	0.0078%	thoughts
11	0.0078%	moving	11	0.0078%	tl
11	0.0078%	negotiate	11	0.0078%	total
11	0.0078%	news	11	0.0078%	transition
11	0.0078%	normal	11	0.0078%	uptake
11	0.0078%	normally	11	0.0078%	variations
11	0.0078%	north	11	0.0078%	vary
11	0.0078%	notetaking	11	0.0078%	weeks
11	0.0078%	offers	11	0.0078%	writes
11	0.0078%	participants'	10	0.0071%	1969
11	0.0078%	partners	10	0.0071%	45
11	0.0078%	parts	10	0.0071%	accurate
11	0.0078%	pedagogical	10	0.0071%	actual
11	0.0078%	perfect	10	0.0071%	addresses
11	0.0078%	physical	10	0.0071%	afterwards
11	0.0078%	planned	10	0.0071%	aim
11	0.0078%	possibility	10	0.0071%	analytical
11	0.0078%	practiced	10	0.0071%	arrivals
11	0.0078%	prediction	10	0.0071%	aspect
11	0.0078%	presents	10	0.0071%	associations
11	0.0078%	pre-writing	10	0.0071%	boston
11	0.0078%	processes	10	0.0071%	built
11	0.0078%	produced	10	0.0071%	canada
11	0.0078%	productive	10	0.0071%	canadian
11	0.0078%	promotes	10	0.0071%	cats

10	0.0071%	caused	10	0.0071%	members
10	0.0071%	chances	10	0.0071%	met
10	0.0071%	clarify	10	0.0071%	morley
10	0.0071%	closer	10	0.0071%	numerous
10	0.0071%	collected	10	0.0071%	person's
10	0.0071%	communities	10	0.0071%	planning
10	0.0071%	completely	10	0.0071%	practicing
10	0.0071%	composition	10	0.0071%	predicting
10	0.0071%	considering	10	0.0071%	prompted
10	0.0071%	crime	10	0.0071%	proposed
10	0.0071%	culturally	10	0.0071%	questionnaire
10	0.0071%	dealt	10	0.0071%	radio
10	0.0071%	decide	10	0.0071%	reach
10	0.0071%	demonstrate	10	0.0071%	reality
10	0.0071%	detail	10	0.0071%	reference
10	0.0071%	doughty	10	0.0071%	relative
10	0.0071%	dynamic	10	0.0071%	reveal
10	0.0071%	editor	10	0.0071%	risks
10	0.0071%	effectively	10	0.0071%	rose
10	0.0071%	elc	10	0.0071%	run
10	0.0071%	elements	10	0.0071%	scanning
10	0.0071%	else	10	0.0071%	schema
10	0.0071%	emotional	10	0.0071%	segments
10	0.0071%	enhance	10	0.0071%	slightly
10	0.0071%	entitled	10	0.0071%	sorts
10	0.0071%	evident	10	0.0071%	store
10	0.0071%	expect	10	0.0071%	subsequent
10	0.0071%	experiment	10	0.0071%	survey
10	0.0071%	facilitate	10	0.0071%	taker
10	0.0071%	forming	10	0.0071%	tape
10	0.0071%	genuine	10	0.0071%	tapes
10	0.0071%	gets	10	0.0071%	thinks
10	0.0071%	grade	10	0.0071%	tool
10	0.0071%	guide	10	0.0071%	varying
10	0.0071%	history	10	0.0071%	whose
10	0.0071%	hoefnagel-höhle	10	0.0071%	wishes
10	0.0071%	hope	10	0.0071%	women
10	0.0071%	hours	10	0.0071%	yes
10	0.0071%	identifying	10	0.0071%	zamel
10	0.0071%	incidental	9	0.0063%	13
10	0.0071%	integrate	9	0.0063%	1992
10	0.0071%	intellectual	9	0.0063%	28
10	0.0071%	interpretation	9	0.0063%	abstract
10	0.0071%	investigate	9	0.0063%	acceptable
10	0.0071%	lists	9	0.0063%	accepted
10	0.0071%	master	9	0.0063%	across

9	0.0063%	addison	9	0.0063%	intend
9	0.0063%	additionally	9	0.0063%	interference
9	0.0063%	analyzed	9	0.0063%	introduces
9	0.0063%	aptitude	9	0.0063%	introducing
9	0.0063%	assuming	9	0.0063%	item
9	0.0063%	assumption	9	0.0063%	knowing
9	0.0063%	becoming	9	0.0063%	lansing
9	0.0063%	brainstorm	9	0.0063%	largely
9	0.0063%	broken	9	0.0063%	lee
9	0.0063%	capacity	9	0.0063%	links
9	0.0063%	coding	9	0.0063%	literature
9	0.0063%	collect	9	0.0063%	man
9	0.0063%	components	9	0.0063%	maturation
9	0.0063%	compound	9	0.0063%	media
9	0.0063%	concepts	9	0.0063%	meetings
9	0.0063%	conclude	9	0.0063%	mode
9	0.0063%	connected	9	0.0063%	mystery
9	0.0063%	contrasting	9	0.0063%	name
9	0.0063%	corpora	9	0.0063%	near
9	0.0063%	currently	9	0.0063%	neither
9	0.0063%	decarrico	9	0.0063%	neurobiological
9	0.0063%	defined	9	0.0063%	neurological
9	0.0063%	definitions	9	0.0063%	newspapers
9	0.0063%	demonstrated	9	0.0063%	non-academic
9	0.0063%	distinction	9	0.0063%	obviously
9	0.0063%	drafts	9	0.0063%	ones
9	0.0063%	dunkel	9	0.0063%	organizers
9	0.0063%	eap	9	0.0063%	oteíza
9	0.0063%	ego	9	0.0063%	outcome
9	0.0063%	exactly	9	0.0063%	pay
9	0.0063%	excellent	9	0.0063%	peirce
9	0.0063%	exist	9	0.0063%	perceived
9	0.0063%	follow-up	9	0.0063%	personality
9	0.0063%	formation	9	0.0063%	personally
9	0.0063%	gives	9	0.0063%	phrase
9	0.0063%	gradual	9	0.0063%	played
9	0.0063%	grasp	9	0.0063%	pointed
9	0.0063%	iep	9	0.0063%	position
9	0.0063%	iii	9	0.0063%	preparation
9	0.0063%	i'm	9	0.0063%	prepositions
9	0.0063%	impression	9	0.0063%	presenting
9	0.0063%	improvement	9	0.0063%	psychology
9	0.0063%	indeed	9	0.0063%	public
9	0.0063%	individuals	9	0.0063%	putting
9	0.0063%	infer	9	0.0063%	quizzes
9	0.0063%	inference	9	0.0063%	quotes

9	0.0063%	readers	8	0.0056%	allowed
9	0.0063%	realistic	8	0.0056%	amsterdam
9	0.0063%	record	8	0.0056%	apart
9	0.0063%	regular	8	0.0056%	apartments
9	0.0063%	repeated	8	0.0056%	apparently
9	0.0063%	representative	8	0.0056%	appendix
9	0.0063%	represented	8	0.0056%	arguments
9	0.0063%	residence	8	0.0056%	assessing
9	0.0063%	schils	8	0.0056%	audio
9	0.0063%	scored	8	0.0056%	backgrounds
9	0.0063%	significantly	8	0.0056%	brings
9	0.0063%	single	8	0.0056%	close
9	0.0063%	solely	8	0.0056%	cloze
9	0.0063%	songs	8	0.0056%	coming
9	0.0063%	space	8	0.0056%	comment
9	0.0063%	spoke	8	0.0056%	commonly
9	0.0063%	star	8	0.0056%	complexity
9	0.0063%	starting	8	0.0056%	comprehend
9	0.0063%	stopped	8	0.0056%	concept
9	0.0063%	strategy	8	0.0056%	concerning
9	0.0063%	stressed	8	0.0056%	conditions
9	0.0063%	targeted	8	0.0056%	consistent
9	0.0063%	teacher-centered	8	0.0056%	contained
9	0.0063%	tenses	8	0.0056%	content-based
9	0.0063%	thematic	8	0.0056%	contrasts
9	0.0063%	transfer	8	0.0056%	covered
9	0.0063%	under	8	0.0056%	crucial
9	0.0063%	users	8	0.0056%	cultures
9	0.0063%	value	8	0.0056%	davis
9	0.0063%	versus	8	0.0056%	decision
9	0.0063%	walk	8	0.0056%	despite
9	0.0063%	wants	8	0.0056%	disagree
9	0.0063%	wasn't	8	0.0056%	doesn't
9	0.0063%	wesley	8	0.0056%	editing
9	0.0063%	white	8	0.0056%	editors
9	0.0063%	williams	8	0.0056%	encouraging
9	0.0063%	wish	8	0.0056%	engage
8	0.0056%	1978	8	0.0056%	english-speaking
8	0.0056%	1988	8	0.0056%	enjoy
8	0.0056%	84	8	0.0056%	entirely
8	0.0056%	acceptance	8	0.0056%	exists
8	0.0056%	accompany	8	0.0056%	expand
8	0.0056%	addressing	8	0.0056%	experienced
8	0.0056%	adjectives	8	0.0056%	fairly
8	0.0056%	adjustment	8	0.0056%	faster
8	0.0056%	adulthood	8	0.0056%	finish

8	0.0056%	fl	8	0.0056%	per
8	0.0056%	fourth	8	0.0056%	performed
8	0.0056%	framework	8	0.0056%	phonetic
8	0.0056%	game	8	0.0056%	poem
8	0.0056%	gardner's	8	0.0056%	porter
8	0.0056%	gone	8	0.0056%	portion
8	0.0056%	guesses	8	0.0056%	predictor
8	0.0056%	guides	8	0.0056%	printed
8	0.0056%	handed	8	0.0056%	procedure
8	0.0056%	helping	8	0.0056%	quick
8	0.0056%	hill	8	0.0056%	recently
8	0.0056%	hopefully	8	0.0056%	recognition
8	0.0056%	identified	8	0.0056%	reinforce
8	0.0056%	illustrate	8	0.0056%	relevance
8	0.0056%	immigrant	8	0.0056%	reported
8	0.0056%	inc	8	0.0056%	represent
8	0.0056%	intellectually	8	0.0056%	requiring
8	0.0056%	interpret	8	0.0056%	revising
8	0.0056%	intuitions	8	0.0056%	revision
8	0.0056%	investigated	8	0.0056%	rewritten
8	0.0056%	journals	8	0.0056%	score
8	0.0056%	lenneberg's	8	0.0056%	seek
8	0.0056%	line	8	0.0056%	self-consciousness
8	0.0056%	logical	8	0.0056%	semantic
8	0.0056%	magazine	8	0.0056%	shorter
8	0.0056%	mainly	8	0.0056%	smaller
8	0.0056%	male	8	0.0056%	sophisticated
8	0.0056%	member	8	0.0056%	speaker's
8	0.0056%	memory	8	0.0056%	stand
8	0.0056%	methodology	8	0.0056%	sub-sections
8	0.0056%	monitor	8	0.0056%	susan
8	0.0056%	motivate	8	0.0056%	syllable
8	0.0056%	myself	8	0.0056%	thematically
8	0.0056%	noun	8	0.0056%	traditional
8	0.0056%	ny	8	0.0056%	trilingual
8	0.0056%	nyikos	8	0.0056%	truly
8	0.0056%	occurred	8	0.0056%	unfortunately
8	0.0056%	offered	8	0.0056%	valuable
8	0.0056%	orally	8	0.0056%	visuals
8	0.0056%	otherwise	8	0.0056%	vs
8	0.0056%	outlines	8	0.0056%	waste
8	0.0056%	pace	8	0.0056%	worth
8	0.0056%	pain	7	0.0049%	1975
8	0.0056%	paraphrase	7	0.0049%	1982
8	0.0056%	partner's	7	0.0049%	1983
8	0.0056%	peace	7	0.0049%	29

7	0.0049%	300	7	0.0049%	considerable
7	0.0049%	34	7	0.0049%	consistently
7	0.0049%	47	7	0.0049%	contribute
7	0.0049%	accepting	7	0.0049%	count
7	0.0049%	accomplish	7	0.0049%	couple
7	0.0049%	acculturation	7	0.0049%	cross-cultural
7	0.0049%	achieving	7	0.0049%	deep
7	0.0049%	actions	7	0.0049%	depend
7	0.0049%	advertising	7	0.0049%	described
7	0.0049%	advocates	7	0.0049%	desired
7	0.0049%	amounts	7	0.0049%	detailed
7	0.0049%	anyway	7	0.0049%	detectives
7	0.0049%	apartment	7	0.0049%	diagnostic
7	0.0049%	appeared	7	0.0049%	dictionary
7	0.0049%	appropriately	7	0.0049%	dinner
7	0.0049%	approximately	7	0.0049%	discovered
7	0.0049%	arise	7	0.0049%	discriminate
7	0.0049%	arises	7	0.0049%	distracters
7	0.0049%	assessed	7	0.0049%	effectiveness
7	0.0049%	assist	7	0.0049%	eighteen
7	0.0049%	assistant	7	0.0049%	ellis
7	0.0049%	assumes	7	0.0049%	enjoyable
7	0.0049%	attempted	7	0.0049%	ensure
7	0.0049%	author	7	0.0049%	entails
7	0.0049%	avoid	7	0.0049%	enter
7	0.0049%	bad	7	0.0049%	environments
7	0.0049%	besides	7	0.0049%	essentially
7	0.0049%	bilingual	7	0.0049%	evaluating
7	0.0049%	brief	7	0.0049%	examination
7	0.0049%	california	7	0.0049%	expectations
7	0.0049%	calt	7	0.0049%	fall
7	0.0049%	can't	7	0.0049%	fear
7	0.0049%	capable	7	0.0049%	fewer
7	0.0049%	catch	7	0.0049%	flawed
7	0.0049%	changed	7	0.0049%	flexibility
7	0.0049%	characteristics	7	0.0049%	flexible
7	0.0049%	childhood	7	0.0049%	flow
7	0.0049%	choosing	7	0.0049%	force
7	0.0049%	chosen	7	0.0049%	formed
7	0.0049%	columns	7	0.0049%	formulate
7	0.0049%	competency	7	0.0049%	gradually
7	0.0049%	composing	7	0.0049%	graduate
7	0.0049%	conducive	7	0.0049%	harder
7	0.0049%	conference	7	0.0049%	heard
7	0.0049%	conferences	7	0.0049%	himself
7	0.0049%	confidence	7	0.0049%	hold

7	0.0049%	human	7	0.0049%	plasticity
7	0.0049%	idiom	7	0.0049%	plenty
7	0.0049%	imitating	7	0.0049%	pragmatically
7	0.0049%	immigrated	7	0.0049%	prepared
7	0.0049%	incorporated	7	0.0049%	presence
7	0.0049%	inhibit	7	0.0049%	prove
7	0.0049%	insight	7	0.0049%	punctuation
7	0.0049%	integrated	7	0.0049%	questionable
7	0.0049%	isn't	7	0.0049%	raising
7	0.0049%	layout	7	0.0049%	rarely
7	0.0049%	left	7	0.0049%	reader
7	0.0049%	limit	7	0.0049%	realized
7	0.0049%	linked	7	0.0049%	reliability
7	0.0049%	listeners	7	0.0049%	reluctant
7	0.0049%	looks	7	0.0049%	rely
7	0.0049%	magazines	7	0.0049%	requirement
7	0.0049%	majors	7	0.0049%	responding
7	0.0049%	map	7	0.0049%	responsible
7	0.0049%	marianne	7	0.0049%	russian
7	0.0049%	mark	7	0.0049%	second-language
7	0.0049%	mental	7	0.0049%	seeing
7	0.0049%	merely	7	0.0049%	sequence
7	0.0049%	metalanguage	7	0.0049%	sequences
7	0.0049%	modified	7	0.0049%	sharing
7	0.0049%	necessity	7	0.0049%	side
7	0.0049%	negatively	7	0.0049%	signaling
7	0.0049%	nine	7	0.0049%	soon
7	0.0049%	norton-pierce	7	0.0049%	source
7	0.0049%	nouns	7	0.0049%	specialized
7	0.0049%	occasionally	7	0.0049%	speculate
7	0.0049%	ok	7	0.0049%	stating
7	0.0049%	onset	7	0.0049%	steps
7	0.0049%	opposite	7	0.0049%	strongest
7	0.0049%	overhead	7	0.0049%	strongly
7	0.0049%	partially	7	0.0049%	structured
7	0.0049%	participation	7	0.0049%	subjects'
7	0.0049%	pass	7	0.0049%	subtle
7	0.0049%	patients	7	0.0049%	summarize
7	0.0049%	perception	7	0.0049%	superiority
7	0.0049%	perspective	7	0.0049%	suspects
7	0.0049%	phenomenon	7	0.0049%	systems
7	0.0049%	philosophy	7	0.0049%	technical
7	0.0049%	photos	7	0.0049%	translation
7	0.0049%	placed	7	0.0049%	typical
7	0.0049%	places	7	0.0049%	unfamiliar
7	0.0049%	planken	7	0.0049%	unknown

7	0.0049%	variation	6	0.0042%	boxes
7	0.0049%	verbal	6	0.0042%	capacities
7	0.0049%	via	6	0.0042%	capitalize
7	0.0049%	videos	6	0.0042%	career
7	0.0049%	virtually	6	0.0042%	carefully
7	0.0049%	volunteer	6	0.0042%	cars
7	0.0049%	warm-up	6	0.0042%	cd-rom
7	0.0049%	web	6	0.0042%	challenged
7	0.0049%	weekly	6	0.0042%	changing
7	0.0049%	workbook	6	0.0042%	checked
7	0.0049%	worthwhile	6	0.0042%	code
6	0.0042%	11	6	0.0042%	cohen
6	0.0042%	14	6	0.0042%	collection
6	0.0042%	18	6	0.0042%	color
6	0.0042%	1973	6	0.0042%	combination
6	0.0042%	1989a	6	0.0042%	combine
6	0.0042%	21	6	0.0042%	combined
6	0.0042%	22	6	0.0042%	communicatively
6	0.0042%	23	6	0.0042%	comparative
6	0.0042%	80	6	0.0042%	comparisons
6	0.0042%	accented	6	0.0042%	computerized
6	0.0042%	accept	6	0.0042%	concentrate
6	0.0042%	accounting	6	0.0042%	concern
6	0.0042%	accustomed	6	0.0042%	conclusions
6	0.0042%	adequately	6	0.0042%	conducted
6	0.0042%	administered	6	0.0042%	confident
6	0.0042%	aids	6	0.0042%	connect
6	0.0042%	alcohol	6	0.0042%	conscious
6	0.0042%	algebra	6	0.0042%	conservation
6	0.0042%	allowing	6	0.0042%	considerations
6	0.0042%	alternatives	6	0.0042%	contact
6	0.0042%	anxious	6	0.0042%	conventions
6	0.0042%	anymore	6	0.0042%	copying
6	0.0042%	anyone	6	0.0042%	core
6	0.0042%	assisted	6	0.0042%	corps
6	0.0042%	attain	6	0.0042%	counterparts
6	0.0042%	attending	6	0.0042%	creates
6	0.0042%	audiocassettes	6	0.0042%	criticism
6	0.0042%	authority	6	0.0042%	croatian
6	0.0042%	basically	6	0.0042%	cycle
6	0.0042%	begun	6	0.0042%	david
6	0.0042%	bilingualism	6	0.0042%	decided
6	0.0042%	bill	6	0.0042%	declined
6	0.0042%	biology	6	0.0042%	define
6	0.0042%	blank	6	0.0042%	degrees
6	0.0042%	bottom-up	6	0.0042%	dependent

6	0.0042%	dialogues	6	0.0042%	homosexual
6	0.0042%	dictation	6	0.0042%	hoping
6	0.0042%	disadvantage	6	0.0042%	housework
6	0.0042%	disagreement	6	0.0042%	ibid
6	0.0042%	discriminating	6	0.0042%	ignore
6	0.0042%	drawbacks	6	0.0042%	imagine
6	0.0042%	earliest	6	0.0042%	imitation
6	0.0042%	east	6	0.0042%	imposed
6	0.0042%	eight	6	0.0042%	increases
6	0.0042%	elt	6	0.0042%	increasingly
6	0.0042%	emerged	6	0.0042%	indicates
6	0.0042%	enable	6	0.0042%	indistinguishable
6	0.0042%	ended	6	0.0042%	influencing
6	0.0042%	equally	6	0.0042%	institutions
6	0.0042%	exact	6	0.0042%	instrumentality
6	0.0042%	examines	6	0.0042%	intentions
6	0.0042%	existing	6	0.0042%	interactional
6	0.0042%	expanding	6	0.0042%	interestingly
6	0.0042%	expert	6	0.0042%	intriguing
6	0.0042%	explaining	6	0.0042%	i've
6	0.0042%	explores	6	0.0042%	james
6	0.0042%	extend	6	0.0042%	keeping
6	0.0042%	extent	6	0.0042%	knows
6	0.0042%	families	6	0.0042%	korea
6	0.0042%	famous	6	0.0042%	krashen's
6	0.0042%	fashion	6	0.0042%	kroll
6	0.0042%	fathman	6	0.0042%	lacks
6	0.0042%	fifteen	6	0.0042%	learners-as-
6	0.0042%	figures	researchers		
6	0.0042%	filling	6	0.0042%	liked
6	0.0042%	finds	6	0.0042%	limitation
6	0.0042%	fits	6	0.0042%	limits
6	0.0042%	food	6	0.0042%	linking
6	0.0042%	foreign-born	6	0.0042%	live
6	0.0042%	fragments	6	0.0042%	longitudinal
6	0.0042%	fulfill	6	0.0042%	loss
6	0.0042%	gained	6	0.0042%	mackay
6	0.0042%	gathered	6	0.0042%	marinova-todd
6	0.0042%	georges	6	0.0042%	medicine
6	0.0042%	grace	6	0.0042%	memorization
6	0.0042%	graded	6	0.0042%	memorize
6	0.0042%	halloween	6	0.0042%	middle
6	0.0042%	happened	6	0.0042%	minor
6	0.0042%	hard	6	0.0042%	missed
6	0.0042%	henry	6	0.0042%	missing
6	0.0042%	highlighted	6	0.0042%	modals

6	0.0042%	modify	6	0.0042%	secondary
6	0.0042%	moreover	6	0.0042%	september
6	0.0042%	mother	6	0.0042%	serious
6	0.0042%	motivations	6	0.0042%	serve
6	0.0042%	motives	6	0.0042%	service
6	0.0042%	names	6	0.0042%	seven
6	0.0042%	natives	6	0.0042%	significance
6	0.0042%	nearly	6	0.0042%	solution
6	0.0042%	nevertheless	6	0.0042%	solving
6	0.0042%	newbury	6	0.0042%	sparked
6	0.0042%	none	6	0.0042%	speed
6	0.0042%	nor	6	0.0042%	spending
6	0.0042%	noticeable	6	0.0042%	sports
6	0.0042%	one-word	6	0.0042%	stimulation
6	0.0042%	ongoing	6	0.0042%	strip
6	0.0042%	opening	6	0.0042%	structural
6	0.0042%	party	6	0.0042%	stylistic
6	0.0042%	perceive	6	0.0042%	substantial
6	0.0042%	peterson	6	0.0042%	summarized
6	0.0042%	please	6	0.0042%	supplemental
6	0.0042%	polio	6	0.0042%	surveys
6	0.0042%	power	6	0.0042%	teachers'
6	0.0042%	preferring	6	0.0042%	tended
6	0.0042%	proactive	6	0.0042%	thorough
6	0.0042%	professors	6	0.0042%	thousands
6	0.0042%	prompts	6	0.0042%	three-stage
6	0.0042%	protocol	6	0.0042%	time-consuming
6	0.0042%	published	6	0.0042%	top-down
6	0.0042%	purely	6	0.0042%	totally
6	0.0042%	quantitative	6	0.0042%	trends
6	0.0042%	questionnaires	6	0.0042%	trouble
6	0.0042%	rating	6	0.0042%	turned
6	0.0042%	react	6	0.0042%	twice
6	0.0042%	readily	6	0.0042%	ukraine
6	0.0042%	recipe	6	0.0042%	universal
6	0.0042%	recognized	6	0.0042%	unless
6	0.0042%	reflect	6	0.0042%	van
6	0.0042%	regard	6	0.0042%	varies
6	0.0042%	relations	6	0.0042%	vermicomposting
6	0.0042%	remembers	6	0.0042%	walks
6	0.0042%	revealed	6	0.0042%	watch
6	0.0042%	reveals	6	0.0042%	wesche
6	0.0042%	rude	6	0.0042%	whom
6	0.0042%	salient	6	0.0042%	wide
6	0.0042%	savignon	6	0.0042%	worry
6	0.0042%	schedule	6	0.0042%	x

6	0.0042%	z	5	0.0035%	buy
5	0.0035%	'	5	0.0035%	car
5	0.0035%	00	5	0.0035%	cards
5	0.0035%	100	5	0.0035%	care
5	0.0035%	1967	5	0.0035%	chased
5	0.0035%	1972	5	0.0035%	chaudron
5	0.0035%	1987	5	0.0035%	chilean
5	0.0035%	2002	5	0.0035%	chronological
5	0.0035%	75	5	0.0035%	circumstances
5	0.0035%	82	5	0.0035%	cite
5	0.0035%	9-31	5	0.0035%	college-level
5	0.0035%	activate	5	0.0035%	communicated
5	0.0035%	adapted	5	0.0035%	competent
5	0.0035%	added	5	0.0035%	completion
5	0.0035%	adjective	5	0.0035%	comprehensive
5	0.0035%	advice	5	0.0035%	conduct
5	0.0035%	agreed	5	0.0035%	confusion
5	0.0035%	alike	5	0.0035%	coniam
5	0.0035%	altogether	5	0.0035%	consisted
5	0.0035%	america	5	0.0035%	containing
5	0.0035%	ample	5	0.0035%	contextual
5	0.0035%	analyses	5	0.0035%	controversial
5	0.0035%	ann	5	0.0035%	copies
5	0.0035%	application	5	0.0035%	correlated
5	0.0035%	applying	5	0.0035%	criteria
5	0.0035%	approached	5	0.0035%	critically
5	0.0035%	approximate	5	0.0035%	davenport
5	0.0035%	assign	5	0.0035%	day's
5	0.0035%	assimilation	5	0.0035%	designing
5	0.0035%	attachment	5	0.0035%	determined
5	0.0035%	attempts	5	0.0035%	developmental
5	0.0035%	attended	5	0.0035%	develops
5	0.0035%	automatically	5	0.0035%	dialogue
5	0.0035%	auxiliary	5	0.0035%	diane
5	0.0035%	awkward	5	0.0035%	differently
5	0.0035%	bear	5	0.0035%	disasters
5	0.0035%	belief	5	0.0035%	dishes
5	0.0035%	below	5	0.0035%	distribution
5	0.0035%	benjamins	5	0.0035%	drawer
5	0.0035%	bibliographies	5	0.0035%	dream
5	0.0035%	big	5	0.0035%	dynamics
5	0.0035%	black	5	0.0035%	ease
5	0.0035%	blanks	5	0.0035%	efforts
5	0.0035%	bored	5	0.0035%	elaborate
5	0.0035%	boys	5	0.0035%	eliciting
5	0.0035%	businessmen	5	0.0035%	embarrassing

5	0.0035%	embedded	5	0.0035%	independent
5	0.0035%	emphasizes	5	0.0035%	inferences
5	0.0035%	empirical	5	0.0035%	influential
5	0.0035%	ends	5	0.0035%	insects
5	0.0035%	eric	5	0.0035%	inside
5	0.0035%	erlbaum	5	0.0035%	integral
5	0.0035%	establish	5	0.0035%	integration
5	0.0035%	evaluators	5	0.0035%	intention
5	0.0035%	except	5	0.0035%	interests
5	0.0035%	exchange	5	0.0035%	interviewed
5	0.0035%	existence	5	0.0035%	introspective
5	0.0035%	extended	5	0.0035%	intuitively
5	0.0035%	facts	5	0.0035%	isbn
5	0.0035%	favor	5	0.0035%	isolated
5	0.0035%	fernando	5	0.0035%	isolation
5	0.0035%	ferris	5	0.0035%	italics
5	0.0035%	filter	5	0.0035%	justice
5	0.0035%	fine	5	0.0035%	lambert
5	0.0035%	forced	5	0.0035%	latent
5	0.0035%	formally	5	0.0035%	latvia
5	0.0035%	formats	5	0.0035%	learning's
5	0.0035%	form-focused	5	0.0035%	lies
5	0.0035%	formulated	5	0.0035%	light
5	0.0035%	foundations	5	0.0035%	likewise
5	0.0035%	friend	5	0.0035%	link
5	0.0035%	genre	5	0.0035%	listener
5	0.0035%	gist	5	0.0035%	lived
5	0.0035%	gotten	5	0.0035%	london
5	0.0035%	grades	5	0.0035%	low-frequency
5	0.0035%	grammaticality	5	0.0035%	macintyre
5	0.0035%	green	5	0.0035%	mahan-taylor
5	0.0035%	grouped	5	0.0035%	mahwah
5	0.0035%	hall	5	0.0035%	makino
5	0.0035%	handout	5	0.0035%	marriage
5	0.0035%	happen	5	0.0035%	mastery
5	0.0035%	harley	5	0.0035%	matching
5	0.0035%	hart	5	0.0035%	meaningfully
5	0.0035%	hartford	5	0.0035%	mechanics
5	0.0035%	head	5	0.0035%	medium
5	0.0035%	hearing	5	0.0035%	meets
5	0.0035%	hinder	5	0.0035%	men
5	0.0035%	hoffman	5	0.0035%	mention
5	0.0035%	homes	5	0.0035%	mentor
5	0.0035%	housing	5	0.0035%	meunier
5	0.0035%	identifies	5	0.0035%	miami
5	0.0035%	il	5	0.0035%	mice

5	0.0035%	million	5	0.0035%	remaining
5	0.0035%	misuse	5	0.0035%	remains
5	0.0035%	modes	5	0.0035%	repeating
5	0.0035%	money	5	0.0035%	request
5	0.0035%	monolingual	5	0.0035%	resolved
5	0.0035%	morgan	5	0.0035%	resource
5	0.0035%	nakajima	5	0.0035%	retrospection
5	0.0035%	negotiations	5	0.0035%	rewards
5	0.0035%	nj	5	0.0035%	reynolds
5	0.0035%	non-evident	5	0.0035%	richards
5	0.0035%	non-interactive	5	0.0035%	role-plays
5	0.0035%	northern	5	0.0035%	salience
5	0.0035%	numbers	5	0.0035%	secondly
5	0.0035%	observe	5	0.0035%	seeking
5	0.0035%	oh	5	0.0035%	self-efficacy
5	0.0035%	olds	5	0.0035%	separate
5	0.0035%	one-on-one	5	0.0035%	serves
5	0.0035%	open	5	0.0035%	sets
5	0.0035%	out-of-class	5	0.0035%	sexual
5	0.0035%	overview	5	0.0035%	shelly
5	0.0035%	paribakht	5	0.0035%	short-term
5	0.0035%	pause	5	0.0035%	siblings
5	0.0035%	periods	5	0.0035%	site
5	0.0035%	perseverance	5	0.0035%	skilled
5	0.0035%	philadelphia	5	0.0035%	skimming
5	0.0035%	phonemes	5	0.0035%	slowly
5	0.0035%	phrasal	5	0.0035%	sociocultural
5	0.0035%	pieces	5	0.0035%	socio-economic
5	0.0035%	placement	5	0.0035%	softeners
5	0.0035%	pleasure	5	0.0035%	sort
5	0.0035%	portuguese	5	0.0035%	speakers'
5	0.0035%	potentially	5	0.0035%	spellings
5	0.0035%	practices	5	0.0035%	standards
5	0.0035%	prefer	5	0.0035%	stimulating
5	0.0035%	preliminary	5	0.0035%	stop
5	0.0035%	previously	5	0.0035%	street
5	0.0035%	product	5	0.0035%	struggle
5	0.0035%	progression	5	0.0035%	succeed
5	0.0035%	purse	5	0.0035%	supporting
5	0.0035%	puzzles	5	0.0035%	suprasegmental
5	0.0035%	reached	5	0.0035%	synonym
5	0.0035%	reactions	5	0.0035%	synonyms
5	0.0035%	reactive	5	0.0035%	table
5	0.0035%	real-world	5	0.0035%	taped
5	0.0035%	regional	5	0.0035%	targets
5	0.0035%	relation	5	0.0035%	task-based

5	0.0035%	teaches
5	0.0035%	tendency
5	0.0035%	terminology
5	0.0035%	theme-based
5	0.0035%	thesis
5	0.0035%	time-management
5	0.0035%	timing
5	0.0035%	toefl
5	0.0035%	tone
5	0.0035%	transparency
5	0.0035%	tv
5	0.0035%	twenty-four
5	0.0035%	typically
5	0.0035%	unclear
5	0.0035%	universities
5	0.0035%	unlike
5	0.0035%	vague
5	0.0035%	vanpatten
5	0.0035%	vansummeren
5	0.0035%	viewed
5	0.0035%	viewpoint
5	0.0035%	weekend
5	0.0035%	wh-
5	0.0035%	whatever
5	0.0035%	window
5	0.0035%	worthless
5	0.0035%	wray
5	0.0035%	writer
5	0.0035%	xi

461 Corpus
21 files; 50, 132 words

2566	5.1185%	the	148	0.2952%	factors
1715	3.4210%	of	142	0.2833%	speakers
1460	2.9123%	and	140	0.2793%	will
1434	2.8604%	to	139	0.2773%	foreign
1369	2.7308%	in	137	0.2733%	but
1111	2.2161%	a	134	0.2673%	children
944	1.8830%	language	134	0.2673%	these
874	1.7434%	that	128	0.2553%	i
572	1.1410%	is	126	0.2513%	one
438	0.8737%	for	123	0.2454%	learner
393	0.7839%	it	121	0.2414%	accent
378	0.7540%	as	121	0.2414%	might
361	0.7201%	be	121	0.2414%	native
357	0.7121%	not	120	0.2394%	had
356	0.7101%	learning	119	0.2374%	proficiency
356	0.7101%	on	118	0.2354%	period
348	0.6942%	age	115	0.2294%	were
336	0.6702%	have	114	0.2274%	he
323	0.6443%	this	114	0.2274%	his
315	0.6283%	with	113	0.2254%	other
297	0.5924%	their	111	0.2214%	time
294	0.5865%	second	109	0.2174%	some
276	0.5505%	or	105	0.2094%	input
255	0.5087%	they	104	0.2075%	rodrigo
240	0.4787%	english	98	0.1955%	between
236	0.4708%	are	98	0.1955%	older
229	0.4568%	acquisition	97	0.1935%	how
223	0.4448%	learners	97	0.1935%	studies
216	0.4309%	an	95	0.1895%	than
207	0.4129%	more	92	0.1835%	first
204	0.4069%	may	91	0.1815%	what
186	0.3710%	motivation	86	0.1715%	has
183	0.3650%	at	86	0.1715%	sla
175	0.3491%	was	85	0.1696%	learn
171	0.3411%	by	85	0.1696%	when
164	0.3271%	l2	84	0.1676%	been
160	0.3192%	from	84	0.1676%	however
157	0.3132%	would	83	0.1656%	different
156	0.3112%	which	83	0.1656%	program
156	0.3112%	who	80	0.1596%	only
154	0.3072%	can	79	0.1576%	pronunciation
149	0.2972%	study	78	0.1556%	because
148	0.2952%	also	78	0.1556%	new

78	0.1556%	sounds	57	0.1137%	languages
78	0.1556%	there	56	0.1117%	culture
77	0.1536%	if	56	0.1117%	form
77	0.1536%	level	55	0.1097%	accelerated
76	0.1516%	research	55	0.1097%	used
76	0.1516%	while	55	0.1097%	well
75	0.1496%	important	54	0.1077%	native-like
74	0.1476%	factor	54	0.1077%	them
74	0.1476%	her	53	0.1057%	according
72	0.1436%	could	53	0.1057%	johnson
71	0.1416%	critical	53	0.1057%	over
70	0.1396%	social	52	0.1037%	learner's
69	0.1376%	about	52	0.1037%	possible
69	0.1376%	adults	52	0.1037%	younger
69	0.1376%	most	51	0.1017%	maturational
69	0.1376%	such	51	0.1017%	much
67	0.1336%	all	50	0.0997%	its
66	0.1317%	do	49	0.0977%	feedback
66	0.1317%	hypothesis	48	0.0957%	although
66	0.1317%	those	48	0.0957%	content
65	0.1297%	found	48	0.0957%	puberty
65	0.1297%	since	47	0.0938%	environment
65	0.1297%	so	47	0.0938%	interaction
65	0.1297%	speech	47	0.0938%	out
64	0.1277%	after	47	0.0938%	process
64	0.1277%	camila	47	0.0938%	results
64	0.1277%	many	46	0.0918%	affect
64	0.1277%	p	46	0.0918%	up
63	0.1257%	use	45	0.0898%	amount
62	0.1237%	fact	45	0.0898%	does
62	0.1237%	researchers	45	0.0898%	through
62	0.1237%	school	44	0.0878%	just
62	0.1237%	target	44	0.0878%	take
61	0.1217%	ability	43	0.0858%	anxiety
61	0.1217%	both	43	0.0858%	effect
61	0.1217%	community	43	0.0858%	type
61	0.1217%	did	42	0.0838%	before
61	0.1217%	example	42	0.0838%	degree
60	0.1197%	identity	42	0.0838%	into
60	0.1197%	variables	42	0.0838%	then
59	0.1177%	influence	42	0.0838%	two
59	0.1177%	she	41	0.0818%	adult
58	0.1157%	order	41	0.0818%	attitudes
58	0.1157%	s	41	0.0818%	better
58	0.1157%	students	41	0.0818%	data
57	0.1137%	even	41	0.0818%	difficult

41	0.0818%	ll	34	0.0678%	1994
41	0.0818%	my	34	0.0678%	evidence
41	0.0818%	words	34	0.0678%	probably
40	0.0798%	become	34	0.0678%	success
40	0.0798%	case	34	0.0678%	united
40	0.0798%	development	34	0.0678%	we
40	0.0798%	e	34	0.0678%	whether
40	0.0798%	like	33	0.0658%	claims
40	0.0798%	very	33	0.0658%	error
39	0.0778%	acquire	33	0.0658%	integrative
39	0.0778%	brain	33	0.0658%	italian
39	0.0778%	no	33	0.0658%	long
39	0.0778%	series	33	0.0658%	opportunities
39	0.0778%	teacher	33	0.0658%	same
39	0.0778%	years	33	0.0658%	without
38	0.0758%	able	32	0.0638%	differences
38	0.0758%	each	32	0.0638%	immersion
38	0.0758%	french	32	0.0638%	life
38	0.0758%	make	32	0.0638%	need
38	0.0758%	produce	32	0.0638%	newport
38	0.0758%	teachers	32	0.0638%	oxford
38	0.0758%	way	32	0.0638%	sensitive
37	0.0738%	any	32	0.0638%	spanish
37	0.0738%	dörnyei	32	0.0638%	towards
37	0.0738%	effects	32	0.0638%	vocabulary
37	0.0738%	flege	31	0.0618%	arrival
37	0.0738%	part	31	0.0618%	being
37	0.0738%	people	31	0.0618%	complex
37	0.0738%	related	31	0.0618%	less
37	0.0738%	student	31	0.0618%	negotiation
37	0.0738%	writing	31	0.0618%	skills
36	0.0718%	certain	30	0.0598%	2001
36	0.0718%	experiences	30	0.0598%	achieve
36	0.0718%	exposure	30	0.0598%	achievement
36	0.0718%	group	30	0.0598%	high
36	0.0718%	information	30	0.0598%	linguistic
36	0.0718%	r	30	0.0598%	lyster
36	0.0718%	states	30	0.0598%	point
36	0.0718%	therefore	30	0.0598%	should
35	0.0698%	1997	30	0.0598%	still
35	0.0698%	based	29	0.0578%	1995
35	0.0698%	gardner	29	0.0578%	1996
35	0.0698%	j	29	0.0578%	aspects
35	0.0698%	phonology	29	0.0578%	effort
35	0.0698%	relationship	29	0.0578%	general
35	0.0698%	situation	29	0.0578%	often

29	0.0578%	output	24	0.0479%	focus
29	0.0578%	participants	24	0.0479%	furthermore
29	0.0578%	show	24	0.0479%	japanese
29	0.0578%	want	24	0.0479%	levels
28	0.0559%	1	24	0.0479%	must
28	0.0559%	consider	24	0.0479%	scoring
28	0.0559%	know	24	0.0479%	seems
28	0.0559%	likely	24	0.0479%	speaking
28	0.0559%	others	24	0.0479%	test
28	0.0559%	phonological	24	0.0479%	within
28	0.0559%	raters	24	0.0479%	young
28	0.0559%	self-confidence	23	0.0459%	actually
28	0.0559%	stevens	23	0.0459%	al
27	0.0539%	1993	23	0.0459%	bongaerts
27	0.0539%	advantage	23	0.0459%	et
27	0.0539%	ages	23	0.0459%	g
27	0.0539%	errors	23	0.0459%	instead
27	0.0539%	own	23	0.0459%	involved
27	0.0539%	result	23	0.0459%	place
27	0.0539%	strong	23	0.0459%	questions
27	0.0539%	three	23	0.0459%	remote
27	0.0539%	types	23	0.0459%	speak
26	0.0519%	1999	23	0.0459%	u
26	0.0519%	affective	23	0.0459%	using
26	0.0519%	affects	22	0.0439%	1989
26	0.0519%	another	22	0.0439%	2
26	0.0519%	attention	22	0.0439%	american
26	0.0519%	experience	22	0.0439%	believe
26	0.0519%	friends	22	0.0439%	cognitive
26	0.0519%	him	22	0.0439%	difference
26	0.0519%	particular	22	0.0439%	done
26	0.0519%	simply	22	0.0439%	dutch
26	0.0519%	subjects	22	0.0439%	family
26	0.0519%	terms	22	0.0439%	give
25	0.0499%	child	22	0.0439%	given
25	0.0499%	communication	22	0.0439%	higher
25	0.0499%	comprehensible	22	0.0439%	made
25	0.0499%	context	22	0.0439%	one's
25	0.0499%	correct	22	0.0439%	significant
25	0.0499%	d	22	0.0439%	ultimate
25	0.0499%	early	21	0.0419%	1998
25	0.0499%	knowledge	21	0.0419%	becomes
25	0.0499%	learned	21	0.0419%	changes
25	0.0499%	role	21	0.0419%	decline
24	0.0479%	classroom	21	0.0419%	develop
24	0.0479%	constraints	21	0.0419%	due

21	0.0419%	during	19	0.0379%	task
21	0.0419%	especially	19	0.0379%	toward
21	0.0419%	find	19	0.0379%	true
21	0.0419%	journal	19	0.0379%	work
21	0.0419%	natural	18	0.0359%	1991
21	0.0419%	opportunity	18	0.0359%	4
21	0.0419%	rather	18	0.0359%	acquiring
21	0.0419%	see	18	0.0359%	affecting
21	0.0419%	support	18	0.0359%	analysis
21	0.0419%	syntax	18	0.0359%	argues
21	0.0419%	unable	18	0.0359%	believed
20	0.0399%	affected	18	0.0359%	biological
20	0.0399%	chile	18	0.0359%	change
20	0.0399%	cited	18	0.0359%	comes
20	0.0399%	competence	18	0.0359%	examined
20	0.0399%	control	18	0.0359%	goal
20	0.0399%	correction	18	0.0359%	goals
20	0.0399%	further	18	0.0359%	help
20	0.0399%	initial	18	0.0359%	identify
20	0.0399%	instrumental	18	0.0359%	lenneberg
20	0.0399%	lack	18	0.0359%	parents
20	0.0399%	m	18	0.0359%	performance
20	0.0399%	number	18	0.0359%	provide
20	0.0399%	person	18	0.0359%	quality
20	0.0399%	pragmatic	18	0.0359%	question
20	0.0399%	primary	18	0.0359%	six
20	0.0399%	producing	18	0.0359%	word
20	0.0399%	rate	17	0.0339%	accents
20	0.0399%	understanding	17	0.0339%	acquired
19	0.0379%	abilities	17	0.0339%	already
19	0.0379%	course	17	0.0339%	argue
19	0.0379%	earlier	17	0.0339%	associated
19	0.0379%	feel	17	0.0339%	camila's
19	0.0379%	five	17	0.0339%	claim
19	0.0379%	groups	17	0.0339%	clément
19	0.0379%	importance	17	0.0339%	desire
19	0.0379%	instruction	17	0.0339%	enough
19	0.0379%	interactions	17	0.0339%	findings
19	0.0379%	krashen	17	0.0339%	model
19	0.0379%	low	17	0.0339%	motivated
19	0.0379%	past	17	0.0339%	negative
19	0.0379%	several	17	0.0339%	practice
19	0.0379%	situations	17	0.0339%	production
19	0.0379%	speaker	17	0.0339%	rated
19	0.0379%	state	17	0.0339%	similar
19	0.0379%	students'	17	0.0339%	something

17	0.0339%	strategies	15	0.0299%	exposed
17	0.0339%	studying	15	0.0299%	features
17	0.0339%	understand	15	0.0299%	global
17	0.0339%	why	15	0.0299%	hand
17	0.0339%	you	15	0.0299%	having
16	0.0319%	-	15	0.0299%	impossible
16	0.0319%	2000	15	0.0299%	length
16	0.0319%	accuracy	15	0.0299%	means
16	0.0319%	approach	15	0.0299%	mind
16	0.0319%	appropriate	15	0.0299%	ranta
16	0.0319%	arrived	15	0.0299%	say
16	0.0319%	asher	15	0.0299%	seem
16	0.0319%	away	15	0.0299%	setting
16	0.0319%	basic	15	0.0299%	someone
16	0.0319%	categories	15	0.0299%	thus
16	0.0319%	clear	15	0.0299%	upon
16	0.0319%	depending	15	0.0299%	uses
16	0.0319%	forms	14	0.0279%	account
16	0.0319%	games	14	0.0279%	among
16	0.0319%	garcia	14	0.0279%	areas
16	0.0319%	late	14	0.0279%	association
16	0.0319%	making	14	0.0279%	awareness
16	0.0319%	months	14	0.0279%	b
16	0.0319%	nature	14	0.0279%	closely
16	0.0319%	optimal	14	0.0279%	complete
16	0.0319%	oyama	14	0.0279%	concluded
16	0.0319%	pp	14	0.0279%	considered
16	0.0319%	quarterly	14	0.0279%	days
16	0.0319%	recast	14	0.0279%	determines
16	0.0319%	regarding	14	0.0279%	examining
16	0.0319%	showed	14	0.0279%	fluent
16	0.0319%	specific	14	0.0279%	genie
16	0.0319%	specifically	14	0.0279%	grammar
16	0.0319%	spent	14	0.0279%	grammatical
16	0.0319%	teaching	14	0.0279%	influences
16	0.0319%	tesol	14	0.0279%	keep
16	0.0319%	themselves	14	0.0279%	least
16	0.0319%	ultimately	14	0.0279%	meaning
15	0.0299%	1985	14	0.0279%	methods
15	0.0299%	attainment	14	0.0279%	morphology
15	0.0299%	cannot	14	0.0279%	noels
15	0.0299%	communicative	14	0.0279%	positive
15	0.0299%	conversational	14	0.0279%	proficient
15	0.0299%	developed	14	0.0279%	relationships
15	0.0299%	every	14	0.0279%	seen
15	0.0299%	examine	14	0.0279%	sense

14	0.0279%	short	13	0.0259%	think
14	0.0279%	slavoff	13	0.0259%	trait
14	0.0279%	snow	13	0.0259%	utterance
14	0.0279%	successful	12	0.0239%	24
14	0.0279%	system	12	0.0239%	access
14	0.0279%	taking	12	0.0239%	beginning
14	0.0279%	tested	12	0.0239%	class
14	0.0279%	whole	12	0.0239%	come
13	0.0259%	1990	12	0.0239%	consideration
13	0.0259%	3	12	0.0239%	constraint
13	0.0259%	accurately	12	0.0239%	correlation
13	0.0259%	around	12	0.0239%	direct
13	0.0259%	available	12	0.0239%	far
13	0.0259%	begin	12	0.0239%	fll
13	0.0259%	c	12	0.0239%	focused
13	0.0259%	challenge	12	0.0239%	gass
13	0.0259%	children's	12	0.0239%	highly
13	0.0259%	chinese	12	0.0239%	individual
13	0.0259%	english-only	12	0.0239%	investment
13	0.0259%	explicit	12	0.0239%	kasper
13	0.0259%	few	12	0.0239%	later
13	0.0259%	function	12	0.0239%	living
13	0.0259%	greater	12	0.0239%	looked
13	0.0259%	h	12	0.0239%	makes
13	0.0259%	immigrants	12	0.0239%	modern
13	0.0259%	include	12	0.0239%	n
13	0.0259%	integrativeness	12	0.0239%	neufeld
13	0.0259%	interesting	12	0.0239%	off
13	0.0259%	interlanguage	12	0.0239%	pragmatics
13	0.0259%	involvement	12	0.0239%	repairs
13	0.0259%	l	12	0.0239%	samples
13	0.0259%	lead	12	0.0239%	scale
13	0.0259%	little	12	0.0239%	shows
13	0.0259%	major	12	0.0239%	sound
13	0.0259%	matter	12	0.0239%	t
13	0.0259%	move	12	0.0239%	vot
13	0.0259%	non-native	12	0.0239%	wanted
13	0.0259%	occur	11	0.0219%	addition
13	0.0259%	orientation	11	0.0219%	allow
13	0.0259%	overall	11	0.0219%	area
13	0.0259%	phrases	11	0.0219%	assume
13	0.0259%	points	11	0.0219%	aware
13	0.0259%	programs	11	0.0219%	began
13	0.0259%	rules	11	0.0219%	best
13	0.0259%	shearin	11	0.0219%	called
13	0.0259%	shown	11	0.0219%	choice

11	0.0219%	clearly	10	0.0199%	above
11	0.0219%	cohesion	10	0.0199%	activities
11	0.0219%	compared	10	0.0199%	arrivals
11	0.0219%	debate	10	0.0199%	ask
11	0.0219%	developing	10	0.0199%	assessment
11	0.0219%	either	10	0.0199%	attitude
11	0.0219%	essential	10	0.0199%	became
11	0.0219%	established	10	0.0199%	chances
11	0.0219%	extremely	10	0.0199%	classic
11	0.0219%	going	10	0.0199%	communities
11	0.0219%	helpful	10	0.0199%	eds
11	0.0219%	ideas	10	0.0199%	end
11	0.0219%	imitate	10	0.0199%	exceptions
11	0.0219%	immediate	10	0.0199%	field
11	0.0219%	including	10	0.0199%	focuses
11	0.0219%	indicate	10	0.0199%	gender
11	0.0219%	interact	10	0.0199%	hear
11	0.0219%	interest	10	0.0199%	hoefnagel-höhle
11	0.0219%	involving	10	0.0199%	holistic
11	0.0219%	issues	10	0.0199%	idea
11	0.0219%	me	10	0.0199%	implications
11	0.0219%	notion	10	0.0199%	improve
11	0.0219%	patterns	10	0.0199%	increased
11	0.0219%	personal	10	0.0199%	influenced
11	0.0219%	present	10	0.0199%	initially
11	0.0219%	provided	10	0.0199%	intended
11	0.0219%	qualitative	10	0.0199%	intensive
11	0.0219%	reasons	10	0.0199%	itself
11	0.0219%	repair	10	0.0199%	kind
11	0.0219%	rodrigo's	10	0.0199%	limited
11	0.0219%	schumann	10	0.0199%	meaningful
11	0.0219%	scores	10	0.0199%	measures
11	0.0219%	stages	10	0.0199%	mentioned
11	0.0219%	suggests	10	0.0199%	necessary
11	0.0219%	though	10	0.0199%	note
11	0.0219%	university	10	0.0199%	once
11	0.0219%	various	10	0.0199%	paper
11	0.0219%	varonis	10	0.0199%	peers
11	0.0219%	where	10	0.0199%	presented
10	0.0199%	1969	10	0.0199%	processing
10	0.0199%	1976	10	0.0199%	psychological
10	0.0199%	1977	10	0.0199%	quantity
10	0.0199%	1981	10	0.0199%	quickly
10	0.0199%	20	10	0.0199%	reading
10	0.0199%	6	10	0.0199%	real
10	0.0199%	7	10	0.0199%	recasts

10	0.0199%	researcher	9	0.0180%	limitations
10	0.0199%	reveal	9	0.0180%	look
10	0.0199%	rose	9	0.0180%	materials
10	0.0199%	schmidt	9	0.0180%	maturation
10	0.0199%	showing	9	0.0180%	mean
10	0.0199%	ssla	9	0.0180%	motivational
10	0.0199%	start	9	0.0180%	multiple
10	0.0199%	suggest	9	0.0180%	neurobiological
10	0.0199%	supports	9	0.0180%	neurological
10	0.0199%	tasks	9	0.0180%	oteíza
10	0.0199%	things	9	0.0180%	outside
10	0.0199%	training	9	0.0180%	particularly
10	0.0199%	uptake	9	0.0180%	peirce
10	0.0199%	video	9	0.0180%	physical
10	0.0199%	view	9	0.0180%	plays
9	0.0180%	12	9	0.0180%	practical
9	0.0180%	additional	9	0.0180%	regardless
9	0.0180%	advertisement	9	0.0180%	relative
9	0.0180%	average	9	0.0180%	risks
9	0.0180%	basis	9	0.0180%	schils
9	0.0180%	component	9	0.0180%	settings
9	0.0180%	consequently	9	0.0180%	society
9	0.0180%	contexts	9	0.0180%	somewhat
9	0.0180%	conversations	9	0.0180%	stronger
9	0.0180%	corrective	9	0.0180%	studied
9	0.0180%	determine	9	0.0180%	supported
9	0.0180%	determining	9	0.0180%	tend
9	0.0180%	difficulty	9	0.0180%	tests
9	0.0180%	dynamic	9	0.0180%	theories
9	0.0180%	effective	9	0.0180%	thought
9	0.0180%	ego	9	0.0180%	times
9	0.0180%	essays	9	0.0180%	too
9	0.0180%	face	9	0.0180%	travel
9	0.0180%	fluency	9	0.0180%	users
9	0.0180%	follow	9	0.0180%	variable
9	0.0180%	following	9	0.0180%	variety
9	0.0180%	formal	9	0.0180%	york
9	0.0180%	german	8	0.0160%	19
9	0.0180%	home	8	0.0160%	5
9	0.0180%	interested	8	0.0160%	achieved
9	0.0180%	involves	8	0.0160%	adulthood
9	0.0180%	k	8	0.0160%	analytical
9	0.0180%	larsen-freeman	8	0.0160%	aptitude
9	0.0180%	leave	8	0.0160%	article
9	0.0180%	led	8	0.0160%	asked
9	0.0180%	lexical	8	0.0160%	assuming

8	0.0160%	beyond	8	0.0160%	possibly
8	0.0160%	bring	8	0.0160%	predictor
8	0.0160%	canadian	8	0.0160%	problem
8	0.0160%	capacity	8	0.0160%	progress
8	0.0160%	category	8	0.0160%	raise
8	0.0160%	concerned	8	0.0160%	receive
8	0.0160%	contrast	8	0.0160%	references
8	0.0160%	conversation	8	0.0160%	regards
8	0.0160%	daily	8	0.0160%	relatively
8	0.0160%	day	8	0.0160%	relevant
8	0.0160%	despite	8	0.0160%	repetition
8	0.0160%	directly	8	0.0160%	require
8	0.0160%	easier	8	0.0160%	requires
8	0.0160%	emotional	8	0.0160%	residence
8	0.0160%	emphasis	8	0.0160%	response
8	0.0160%	entirely	8	0.0160%	right
8	0.0160%	esl	8	0.0160%	segments
8	0.0160%	failure	8	0.0160%	self-consciousness
8	0.0160%	familiar	8	0.0160%	sentences
8	0.0160%	felt	8	0.0160%	set
8	0.0160%	fl	8	0.0160%	simple
8	0.0160%	frequent	8	0.0160%	small
8	0.0160%	fully	8	0.0160%	spoke
8	0.0160%	gardner's	8	0.0160%	superior
8	0.0160%	get	8	0.0160%	taken
8	0.0160%	gradual	8	0.0160%	tapes
8	0.0160%	half	8	0.0160%	theory
8	0.0160%	hill	8	0.0160%	trilingual
8	0.0160%	immediately	8	0.0160%	turn
8	0.0160%	instructed	8	0.0160%	ways
8	0.0160%	instructional	8	0.0160%	yet
8	0.0160%	intonation	7	0.0140%	1975
8	0.0160%	john	7	0.0140%	1984
8	0.0160%	kinds	7	0.0140%	acculturation
8	0.0160%	learners'	7	0.0140%	achieving
8	0.0160%	lenneberg's	7	0.0140%	adjustment
8	0.0160%	lost	7	0.0140%	against
8	0.0160%	mistakes	7	0.0140%	always
8	0.0160%	moved	7	0.0140%	business
8	0.0160%	notice	7	0.0140%	cases
8	0.0160%	nyikos	7	0.0140%	cause
8	0.0160%	old	7	0.0140%	center
8	0.0160%	perceived	7	0.0140%	challenges
8	0.0160%	person's	7	0.0140%	childhood
8	0.0160%	phonetic	7	0.0140%	comfortable
8	0.0160%	pick	7	0.0140%	common

7	0.0140%	comprehension	7	0.0140%	planken
7	0.0140%	conditions	7	0.0140%	plasticity
7	0.0140%	connection	7	0.0140%	play
7	0.0140%	contains	7	0.0140%	primarily
7	0.0140%	country	7	0.0140%	problematic
7	0.0140%	depends	7	0.0140%	psychology
7	0.0140%	distinguish	7	0.0140%	public
7	0.0140%	easily	7	0.0140%	reach
7	0.0140%	easy	7	0.0140%	received
7	0.0140%	eighteen	7	0.0140%	requests
7	0.0140%	encourages	7	0.0140%	responses
7	0.0140%	english-speaking	7	0.0140%	retain
7	0.0140%	enhanced	7	0.0140%	russian
7	0.0140%	entails	7	0.0140%	says
7	0.0140%	etc	7	0.0140%	scored
7	0.0140%	explains	7	0.0140%	seemed
7	0.0140%	extra	7	0.0140%	significantly
7	0.0140%	fear	7	0.0140%	single
7	0.0140%	feelings	7	0.0140%	student's
7	0.0140%	go	7	0.0140%	subsequent
7	0.0140%	good	7	0.0140%	sufficient
7	0.0140%	gradually	7	0.0140%	suggested
7	0.0140%	imitating	7	0.0140%	suited
7	0.0140%	immigrated	7	0.0140%	topic
7	0.0140%	immigration	7	0.0140%	trying
7	0.0140%	inhibit	7	0.0140%	v
7	0.0140%	intellectual	7	0.0140%	videos
7	0.0140%	interference	7	0.0140%	women
7	0.0140%	intermediate	7	0.0140%	world
7	0.0140%	investigate	7	0.0140%	year
7	0.0140%	job	6	0.0120%	13
7	0.0140%	last	6	0.0120%	1978
7	0.0140%	looking	6	0.0120%	1980
7	0.0140%	lower	6	0.0120%	1982
7	0.0140%	maps	6	0.0120%	accounting
7	0.0140%	master	6	0.0120%	act
7	0.0140%	members	6	0.0120%	actual
7	0.0140%	negatively	6	0.0120%	ad
7	0.0140%	never	6	0.0120%	additionally
7	0.0140%	norton-pierce	6	0.0120%	address
7	0.0140%	noticing	6	0.0120%	alcohol
7	0.0140%	numerous	6	0.0120%	alternatives
7	0.0140%	obvious	6	0.0120%	analyzed
7	0.0140%	onset	6	0.0120%	answer
7	0.0140%	our	6	0.0120%	assess
7	0.0140%	perception	6	0.0120%	assessed

6	0.0120%	assessing	6	0.0120%	functions
6	0.0120%	associations	6	0.0120%	generally
6	0.0120%	atmosphere	6	0.0120%	giving
6	0.0120%	attain	6	0.0120%	harder
6	0.0120%	audiocassettes	6	0.0120%	ibid
6	0.0120%	authentic	6	0.0120%	identification
6	0.0120%	bilingual	6	0.0120%	immigrant
6	0.0120%	bilingualism	6	0.0120%	imposed
6	0.0120%	bodman	6	0.0120%	impression
6	0.0120%	break	6	0.0120%	included
6	0.0120%	brings	6	0.0120%	includes
6	0.0120%	california	6	0.0120%	indistinguishable
6	0.0120%	capacities	6	0.0120%	instrumentality
6	0.0120%	classrooms	6	0.0120%	international
6	0.0120%	collect	6	0.0120%	intriguing
6	0.0120%	college	6	0.0120%	leads
6	0.0120%	combined	6	0.0120%	let
6	0.0120%	conclude	6	0.0120%	limitation
6	0.0120%	conclusion	6	0.0120%	longitudinal
6	0.0120%	connected	6	0.0120%	mackay
6	0.0120%	considerable	6	0.0120%	marinova-todd
6	0.0120%	contact	6	0.0120%	measure
6	0.0120%	counterparts	6	0.0120%	merely
6	0.0120%	croatian	6	0.0120%	metalinguistic
6	0.0120%	cultures	6	0.0120%	motives
6	0.0120%	declined	6	0.0120%	name
6	0.0120%	desired	6	0.0120%	necessarily
6	0.0120%	distinction	6	0.0120%	neither
6	0.0120%	earliest	6	0.0120%	nine
6	0.0120%	eisenstein	6	0.0120%	noticeable
6	0.0120%	emerged	6	0.0120%	organization
6	0.0120%	entitled	6	0.0120%	otherwise
6	0.0120%	evaluate	6	0.0120%	participants'
6	0.0120%	evaluated	6	0.0120%	perceive
6	0.0120%	evaluation	6	0.0120%	perhaps
6	0.0120%	eventually	6	0.0120%	personality
6	0.0120%	everything	6	0.0120%	piece
6	0.0120%	examines	6	0.0120%	polio
6	0.0120%	exist	6	0.0120%	possibility
6	0.0120%	faster	6	0.0120%	press
6	0.0120%	fathman	6	0.0120%	prior
6	0.0120%	flexible	6	0.0120%	produced
6	0.0120%	foreign-born	6	0.0120%	provides
6	0.0120%	formation	6	0.0120%	put
6	0.0120%	frequently	6	0.0120%	quantitative
6	0.0120%	full	6	0.0120%	raising

6	0.0120%	rating	5	0.0100%	again
6	0.0120%	read	5	0.0100%	agreed
6	0.0120%	recasting	5	0.0100%	along
6	0.0120%	recent	5	0.0100%	altogether
6	0.0120%	recognize	5	0.0100%	appear
6	0.0120%	reluctant	5	0.0100%	appears
6	0.0120%	responded	5	0.0100%	apply
6	0.0120%	reveals	5	0.0100%	approaches
6	0.0120%	roles	5	0.0100%	approximate
6	0.0120%	said	5	0.0100%	assimilation
6	0.0120%	score	5	0.0100%	attempted
6	0.0120%	seek	5	0.0100%	audio
6	0.0120%	significance	5	0.0100%	bardovi-harlig
6	0.0120%	solely	5	0.0100%	basically
6	0.0120%	spoken	5	0.0100%	becoming
6	0.0120%	started	5	0.0100%	benefit
6	0.0120%	strongest	5	0.0100%	boys
6	0.0120%	strongly	5	0.0100%	canada
6	0.0120%	subtle	5	0.0100%	cards
6	0.0120%	teach	5	0.0100%	certainly
6	0.0120%	thousands	5	0.0100%	chilean
6	0.0120%	together	5	0.0100%	circumstances
6	0.0120%	told	5	0.0100%	classmates
6	0.0120%	universal	5	0.0100%	close
6	0.0120%	us	5	0.0100%	closer
6	0.0120%	value	5	0.0100%	comparison
6	0.0120%	varying	5	0.0100%	conscious
6	0.0120%	waste	5	0.0100%	contain
6	0.0120%	whereas	5	0.0100%	continue
6	0.0120%	willing	5	0.0100%	contribute
6	0.0120%	write	5	0.0100%	corrections
6	0.0120%	z	5	0.0100%	cover
5	0.0100%	15	5	0.0100%	create
5	0.0100%	1967	5	0.0100%	crucial
5	0.0100%	1986	5	0.0100%	cultural
5	0.0100%	1992	5	0.0100%	cycle
5	0.0100%	2002	5	0.0100%	davis
5	0.0100%	21	5	0.0100%	disadvantage
5	0.0100%	25	5	0.0100%	discourse
5	0.0100%	28	5	0.0100%	discuss
5	0.0100%	29	5	0.0100%	discussion
5	0.0100%	9-31	5	0.0100%	doing
5	0.0100%	accented	5	0.0100%	ed
5	0.0100%	accomplish	5	0.0100%	education
5	0.0100%	advanced	5	0.0100%	educational
5	0.0100%	advantages	5	0.0100%	eight

5	0.0100%	emphasized	5	0.0100%	lived
5	0.0100%	enter	5	0.0100%	longer
5	0.0100%	environments	5	0.0100%	loss
5	0.0100%	essay	5	0.0100%	macintyre
5	0.0100%	evaluators	5	0.0100%	method
5	0.0100%	exact	5	0.0100%	motivations
5	0.0100%	exactly	5	0.0100%	needed
5	0.0100%	existence	5	0.0100%	negotiate
5	0.0100%	expect	5	0.0100%	northern
5	0.0100%	expert	5	0.0100%	noted
5	0.0100%	explicitly	5	0.0100%	o
5	0.0100%	explore	5	0.0100%	objective
5	0.0100%	expressing	5	0.0100%	organize
5	0.0100%	force	5	0.0100%	outcome
5	0.0100%	formulated	5	0.0100%	partially
5	0.0100%	foundations	5	0.0100%	participation
5	0.0100%	fourth	5	0.0100%	pass
5	0.0100%	frequency	5	0.0100%	perform
5	0.0100%	fun	5	0.0100%	perseverance
5	0.0100%	gap	5	0.0100%	perspectives
5	0.0100%	gets	5	0.0100%	phenomenon
5	0.0100%	getting	5	0.0100%	picture
5	0.0100%	grammaticality	5	0.0100%	placed
5	0.0100%	gratitude	5	0.0100%	power
5	0.0100%	harley	5	0.0100%	preferring
5	0.0100%	hart	5	0.0100%	problems
5	0.0100%	hoffman	5	0.0100%	prompt
5	0.0100%	house	5	0.0100%	proposed
5	0.0100%	human	5	0.0100%	purposes
5	0.0100%	imitation	5	0.0100%	range
5	0.0100%	improved	5	0.0100%	realistic
5	0.0100%	increase	5	0.0100%	reality
5	0.0100%	influencing	5	0.0100%	record
5	0.0100%	influential	5	0.0100%	refer
5	0.0100%	institutions	5	0.0100%	represented
5	0.0100%	interpret	5	0.0100%	responsible
5	0.0100%	interview	5	0.0100%	rewards
5	0.0100%	interviews	5	0.0100%	schools
5	0.0100%	items	5	0.0100%	second-language
5	0.0100%	jobs	5	0.0100%	self-efficacy
5	0.0100%	korean	5	0.0100%	semantic
5	0.0100%	lambert	5	0.0100%	short-term
5	0.0100%	latent	5	0.0100%	siblings
5	0.0100%	leading	5	0.0100%	sometimes
5	0.0100%	learning's	5	0.0100%	sorts
5	0.0100%	limits	5	0.0100%	sources

5	0.0100%	speaker's
5	0.0100%	starting
5	0.0100%	stated
5	0.0100%	stimulation
5	0.0100%	styles
5	0.0100%	subjects'
5	0.0100%	sure
5	0.0100%	talk
5	0.0100%	techniques
5	0.0100%	total
5	0.0100%	totally
5	0.0100%	transfer
5	0.0100%	tried
5	0.0100%	under
5	0.0100%	until
5	0.0100%	useful
5	0.0100%	usually
5	0.0100%	valuable
5	0.0100%	van
5	0.0100%	vansummeren
5	0.0100%	whose
5	0.0100%	wishes
5	0.0100%	working
5	0.0100%	worthless

807 Corpus
21 files; 28, 912 words

2165	7.4882%	the	93	0.3217%	textbook
949	3.2824%	and	91	0.3147%	more
916	3.1682%	to	87	0.3009%	words
717	2.4799%	of	86	0.2975%	information
628	2.1721%	in	85	0.2940%	what
585	2.0234%	a	84	0.2905%	at
433	1.4976%	students	84	0.2905%	each
369	1.2763%	for	83	0.2871%	also
328	1.1345%	that	80	0.2767%	one
290	1.0030%	be	76	0.2629%	unit
289	0.9996%	is	72	0.2490%	them
247	0.8543%	are	71	0.2456%	focus
246	0.8509%	on	70	0.2421%	but
242	0.8370%	as	68	0.2352%	activities
224	0.7748%	this	68	0.2352%	section
220	0.7609%	with	63	0.2179%	activity
219	0.7575%	they	63	0.2179%	questions
207	0.7160%	or	61	0.2110%	may
185	0.6399%	their	60	0.2075%	book
168	0.5811%	it	60	0.2075%	example
156	0.5396%	vocabulary	60	0.2075%	lesson
152	0.5257%	grammar	59	0.2041%	can
148	0.5119%	would	59	0.2041%	form
147	0.5084%	i	58	0.2006%	skills
145	0.5015%	reading	57	0.1971%	if
144	0.4981%	language	55	0.1902%	she
139	0.4808%	was	55	0.1902%	students'
133	0.4600%	which	53	0.1833%	lecture
132	0.4566%	an	53	0.1833%	topic
128	0.4427%	listening	52	0.1799%	do
126	0.4358%	class	52	0.1799%	english
122	0.4220%	exercises	52	0.1799%	such
121	0.4185%	will	52	0.1799%	there
120	0.4151%	have	52	0.1799%	when
110	0.3805%	teacher	51	0.1764%	could
109	0.3770%	from	51	0.1764%	so
106	0.3666%	these	50	0.1729%	how
105	0.3632%	not	50	0.1729%	text
105	0.3632%	use	50	0.1729%	then
104	0.3597%	about	50	0.1729%	used
101	0.3493%	were	49	0.1695%	new
97	0.3355%	by	49	0.1695%	some
94	0.3251%	writing	48	0.1660%	into

48	0.1660%	other	34	0.1176%	level
48	0.1660%	using	33	0.1141%	both
47	0.1626%	learners	33	0.1141%	comprehension
47	0.1626%	p	33	0.1141%	different
47	0.1626%	time	33	0.1141%	northstar
47	0.1626%	two	33	0.1141%	only
46	0.1591%	authentic	33	0.1141%	well
46	0.1591%	you	32	0.1107%	classroom
45	0.1556%	instruction	32	0.1107%	through
45	0.1556%	need	31	0.1072%	context
45	0.1556%	student	31	0.1072%	listen
45	0.1556%	word	31	0.1072%	medical
44	0.1522%	answer	31	0.1072%	since
44	0.1522%	answers	31	0.1072%	think
44	0.1522%	first	31	0.1072%	while
44	0.1522%	instructor	30	0.1038%	authors
44	0.1522%	than	30	0.1038%	main
43	0.1487%	meaning	30	0.1038%	sentences
43	0.1487%	second	29	0.1003%	between
43	0.1487%	teaching	29	0.1003%	had
43	0.1487%	usually	29	0.1003%	ideas
42	0.1453%	topics	29	0.1003%	order
41	0.1418%	however	29	0.1003%	out
41	0.1418%	learning	29	0.1003%	work
41	0.1418%	material	28	0.0968%	2
41	0.1418%	very	28	0.0968%	correct
40	0.1384%	after	28	0.0968%	important
40	0.1384%	chapter	28	0.0968%	partner
40	0.1384%	during	28	0.0968%	readings
39	0.1349%	materials	28	0.0968%	same
39	0.1349%	should	28	0.0968%	sections
38	0.1314%	academic	28	0.0968%	structures
38	0.1314%	read	27	0.0934%	complete
38	0.1314%	research	27	0.0934%	course
36	0.1245%	2001	27	0.0934%	down
36	0.1245%	able	27	0.0934%	help
36	0.1245%	all	26	0.0899%	discussion
36	0.1245%	practice	26	0.0899%	give
36	0.1245%	where	26	0.0899%	make
35	0.1211%	given	26	0.0899%	must
35	0.1211%	part	26	0.0899%	point
35	0.1211%	past	26	0.0899%	structure
35	0.1211%	speaking	26	0.0899%	theme
34	0.1176%	esl	25	0.0865%	appropriate
34	0.1176%	exercise	25	0.0865%	classes
34	0.1176%	has	25	0.0865%	examples

25	0.0865%	group	20	0.0692%	units
25	0.0865%	many	19	0.0657%	acquisition
25	0.0865%	meaningful	19	0.0657%	believes
25	0.0865%	often	19	0.0657%	chapters
25	0.0865%	teach	19	0.0657%	chart
25	0.0865%	teacher's	19	0.0657%	details
25	0.0865%	working	19	0.0657%	discourse
24	0.0830%	asked	19	0.0657%	further
24	0.0830%	being	19	0.0657%	goals
24	0.0830%	communication	19	0.0657%	lectures
24	0.0830%	differences	19	0.0657%	much
24	0.0830%	does	19	0.0657%	next
24	0.0830%	good	19	0.0657%	see
24	0.0830%	native	19	0.0657%	states
24	0.0830%	notes	19	0.0657%	therefore
24	0.0830%	own	19	0.0657%	university
24	0.0830%	pp	19	0.0657%	who
23	0.0796%	1	18	0.0623%	3
23	0.0796%	ask	18	0.0623%	although
23	0.0796%	communicative	18	0.0623%	content
23	0.0796%	minutes	18	0.0623%	having
23	0.0796%	most	18	0.0623%	include
23	0.0796%	objectives	18	0.0623%	knowledge
23	0.0796%	provide	18	0.0623%	passive
23	0.0796%	within	18	0.0623%	presented
23	0.0796%	your	18	0.0623%	related
22	0.0761%	another	18	0.0623%	speech
22	0.0761%	because	18	0.0623%	spent
22	0.0761%	did	18	0.0623%	task
22	0.0761%	explicitly	18	0.0623%	up
22	0.0761%	passage	18	0.0623%	voice
22	0.0761%	picture	18	0.0623%	we
22	0.0761%	style	17	0.0588%	done
22	0.0761%	take	17	0.0588%	gap
22	0.0761%	tasks	17	0.0588%	intermediate
22	0.0761%	taught	17	0.0588%	key
22	0.0761%	teachers	17	0.0588%	learner
21	0.0726%	any	17	0.0588%	sentence
21	0.0726%	based	17	0.0588%	study
21	0.0726%	heinle	17	0.0588%	three
21	0.0726%	might	17	0.0588%	whether
21	0.0726%	outside	17	0.0588%	worked
21	0.0726%	write	16	0.0553%	1998
20	0.0692%	believe	16	0.0553%	before
20	0.0692%	edition	16	0.0553%	charts
20	0.0692%	interesting	16	0.0553%	culture

16	0.0553%	discuss	13	0.0450%	find
16	0.0553%	focused	13	0.0450%	foreign
16	0.0553%	four	13	0.0450%	found
16	0.0553%	high	13	0.0450%	gave
16	0.0553%	introduction	13	0.0450%	homework
16	0.0553%	observation	13	0.0450%	made
16	0.0553%	speakers	13	0.0450%	majority
16	0.0553%	times	13	0.0450%	making
15	0.0519%	american	13	0.0450%	markers
15	0.0519%	back	13	0.0450%	needed
15	0.0519%	error	13	0.0450%	opportunities
15	0.0519%	every	13	0.0450%	pages
15	0.0519%	few	13	0.0450%	pairs
15	0.0519%	five	13	0.0450%	period
15	0.0519%	future	13	0.0450%	present
15	0.0519%	grammatical	13	0.0450%	purpose
15	0.0519%	like	13	0.0450%	rather
15	0.0519%	review	13	0.0450%	results
15	0.0519%	short	13	0.0450%	several
15	0.0519%	worksheet	13	0.0450%	types
14	0.0484%	active	13	0.0450%	understand
14	0.0484%	asking	13	0.0450%	variety
14	0.0484%	been	12	0.0415%	already
14	0.0484%	continuous	12	0.0415%	approach
14	0.0484%	doing	12	0.0415%	arm
14	0.0484%	e	12	0.0415%	background
14	0.0484%	etc	12	0.0415%	draw
14	0.0484%	fill	12	0.0415%	essays
14	0.0484%	giving	12	0.0415%	hand
14	0.0484%	groups	12	0.0415%	learn
14	0.0484%	know	12	0.0415%	murder
14	0.0484%	necessary	12	0.0415%	needs
14	0.0484%	points	12	0.0415%	over
14	0.0484%	previous	12	0.0415%	people
14	0.0484%	proved	12	0.0415%	problem
14	0.0484%	sub-section	12	0.0415%	program
14	0.0484%	texts	12	0.0415%	providing
14	0.0484%	type	12	0.0415%	question
14	0.0484%	useful	12	0.0415%	recall
13	0.0450%	4	12	0.0415%	s
13	0.0450%	best	12	0.0415%	school
13	0.0450%	better	12	0.0415%	series
13	0.0450%	consists	12	0.0415%	t
13	0.0450%	cultural	12	0.0415%	themes
13	0.0450%	either	12	0.0415%	themselves
13	0.0450%	explicit	12	0.0415%	third

12	0.0415%	way	10	0.0346%	interaction
11	0.0380%	5	10	0.0346%	interested
11	0.0380%	attention	10	0.0346%	introduce
11	0.0380%	come	10	0.0346%	involved
11	0.0380%	contains	10	0.0346%	items
11	0.0380%	cues	10	0.0346%	manual
11	0.0380%	describe	10	0.0346%	morley
11	0.0380%	determine	10	0.0346%	multiple
11	0.0380%	difficult	10	0.0346%	my
11	0.0380%	discussions	10	0.0346%	north
11	0.0380%	due	10	0.0346%	objective
11	0.0380%	ed	10	0.0346%	personal
11	0.0380%	end	10	0.0346%	potential
11	0.0380%	feedback	10	0.0346%	principle
11	0.0380%	follow	10	0.0346%	progress
11	0.0380%	long	10	0.0346%	pronunciation
11	0.0380%	notetaking	10	0.0346%	put
11	0.0380%	opinion	10	0.0346%	real-life
11	0.0380%	opportunity	10	0.0346%	rhetorical
11	0.0380%	overall	10	0.0346%	similar
11	0.0380%	paragraph	10	0.0346%	situations
11	0.0380%	particular	10	0.0346%	state
11	0.0380%	provides	10	0.0346%	stated
11	0.0380%	rest	10	0.0346%	strategies
11	0.0380%	situation	10	0.0346%	student-centered
11	0.0380%	target	10	0.0346%	studying
11	0.0380%	ten	10	0.0346%	tests
11	0.0380%	those	10	0.0346%	thought
11	0.0380%	us	10	0.0346%	verb
11	0.0380%	varied	10	0.0346%	vi
10	0.0346%	advanced	10	0.0346%	want
10	0.0346%	attached	10	0.0346%	written
10	0.0346%	beginning	9	0.0311%	analysis
10	0.0346%	choose	9	0.0311%	aware
10	0.0346%	correction	9	0.0311%	books
10	0.0346%	directions	9	0.0311%	choice
10	0.0346%	doughty	9	0.0311%	competence
10	0.0346%	effect	9	0.0311%	conversation
10	0.0346%	encouraged	9	0.0311%	divided
10	0.0346%	especially	9	0.0311%	easy
10	0.0346%	even	9	0.0311%	elicit
10	0.0346%	explained	9	0.0311%	floor
10	0.0346%	false	9	0.0311%	following
10	0.0346%	fieldwork	9	0.0311%	forms
10	0.0346%	follows	9	0.0311%	furthermore
10	0.0346%	instructions	9	0.0311%	g

9	0.0311%	goal	8	0.0277%	environment
9	0.0311%	intended	8	0.0277%	essential
9	0.0311%	interviews	8	0.0277%	express
9	0.0311%	its	8	0.0277%	focuses
9	0.0311%	just	8	0.0277%	format
9	0.0311%	last	8	0.0277%	get
9	0.0311%	never	8	0.0277%	he
9	0.0311%	non-academic	8	0.0277%	identify
9	0.0311%	observed	8	0.0277%	input
9	0.0311%	outline	8	0.0277%	interactive
9	0.0311%	pictures	8	0.0277%	jobs
9	0.0311%	plus	8	0.0277%	lessons
9	0.0311%	prepositions	8	0.0277%	m
9	0.0311%	process	8	0.0277%	matter
9	0.0311%	require	8	0.0277%	note
9	0.0311%	required	8	0.0277%	now
9	0.0311%	seemed	8	0.0277%	number
9	0.0311%	showed	8	0.0277%	offer
9	0.0311%	small	8	0.0277%	once
9	0.0311%	student's	8	0.0277%	organization
9	0.0311%	subject	8	0.0277%	plan
9	0.0311%	suggestions	8	0.0277%	portion
9	0.0311%	supplement	8	0.0277%	preparing
9	0.0311%	supplemented	8	0.0277%	principles
9	0.0311%	test	8	0.0277%	progressive
9	0.0311%	together	8	0.0277%	relevant
9	0.0311%	took	8	0.0277%	reviewing
9	0.0311%	true	8	0.0277%	scene
9	0.0311%	understanding	8	0.0277%	snow
9	0.0311%	united	8	0.0277%	speaker
9	0.0311%	uses	8	0.0277%	specific
9	0.0311%	williams	8	0.0277%	star
9	0.0311%	without	8	0.0277%	subjects
8	0.0277%	1994	8	0.0277%	sub-sections
8	0.0277%	always	8	0.0277%	sure
8	0.0277%	apartments	8	0.0277%	teacher-centered
8	0.0277%	assignment	8	0.0277%	tense
8	0.0277%	celce-murcia	8	0.0277%	thus
8	0.0277%	check	8	0.0277%	top
8	0.0277%	clear	8	0.0277%	various
8	0.0277%	curriculum	8	0.0277%	ways
8	0.0277%	definition	7	0.0242%	1991
8	0.0277%	describing	7	0.0242%	1995
8	0.0277%	earlier	7	0.0242%	accuracy
8	0.0277%	elicited	7	0.0242%	addison
8	0.0277%	entire	7	0.0242%	aid

7	0.0242%	allow	7	0.0242%	predicting
7	0.0242%	answering	7	0.0242%	prediction
7	0.0242%	apartment	7	0.0242%	prepare
7	0.0242%	assumptions	7	0.0242%	presentation
7	0.0242%	available	7	0.0242%	problems
7	0.0242%	begin	7	0.0242%	prompt
7	0.0242%	begins	7	0.0242%	rule
7	0.0242%	board	7	0.0242%	seem
7	0.0242%	certain	7	0.0242%	setting
7	0.0242%	comprehensible	7	0.0242%	share
7	0.0242%	create	7	0.0242%	sharing
7	0.0242%	current	7	0.0242%	show
7	0.0242%	developing	7	0.0242%	signaling
7	0.0242%	effective	7	0.0242%	skill
7	0.0242%	encountered	7	0.0242%	successfully
7	0.0242%	everyone	7	0.0242%	swain
7	0.0242%	explain	7	0.0242%	talking
7	0.0242%	follow-up	7	0.0242%	tenses
7	0.0242%	going	7	0.0242%	term
7	0.0242%	hear	7	0.0242%	terms
7	0.0242%	helpful	7	0.0242%	textbooks
7	0.0242%	home	7	0.0242%	thematic
7	0.0242%	ii	7	0.0242%	thematically
7	0.0242%	implicit	7	0.0242%	v
7	0.0242%	includes	7	0.0242%	vary
7	0.0242%	individual	7	0.0242%	visual
7	0.0242%	inference	7	0.0242%	wesley
7	0.0242%	instead	7	0.0242%	why
7	0.0242%	interest	6	0.0208%	0
7	0.0242%	krashen	6	0.0208%	15
7	0.0242%	ll	6	0.0208%	1985
7	0.0242%	larsen-freeman	6	0.0208%	abilities
7	0.0242%	later	6	0.0208%	addressed
7	0.0242%	led	6	0.0208%	affective
7	0.0242%	life	6	0.0208%	again
7	0.0242%	lines	6	0.0208%	algebra
7	0.0242%	list	6	0.0208%	approximately
7	0.0242%	listed	6	0.0208%	assess
7	0.0242%	listeners	6	0.0208%	attempt
7	0.0242%	living	6	0.0208%	audience
7	0.0242%	marianne	6	0.0208%	authority
7	0.0242%	mistake	6	0.0208%	bit
7	0.0242%	msu	6	0.0208%	bottom
7	0.0242%	organized	6	0.0208%	bottom-up
7	0.0242%	pica	6	0.0208%	broken
7	0.0242%	possible	6	0.0208%	c

6	0.0208%	called	6	0.0208%	peterson
6	0.0208%	clues	6	0.0208%	practiced
6	0.0208%	communicate	6	0.0208%	proactive
6	0.0208%	community	6	0.0208%	production
6	0.0208%	completed	6	0.0208%	purple
6	0.0208%	creating	6	0.0208%	quality
6	0.0208%	crime	6	0.0208%	quotes
6	0.0208%	demonstrated	6	0.0208%	real
6	0.0208%	effort	6	0.0208%	recasting
6	0.0208%	encounter	6	0.0208%	recipe
6	0.0208%	enough	6	0.0208%	regarding
6	0.0208%	explanation	6	0.0208%	regards
6	0.0208%	expression	6	0.0208%	reminded
6	0.0208%	extra	6	0.0208%	repeat
6	0.0208%	features	6	0.0208%	retain
6	0.0208%	focusing	6	0.0208%	savignon
6	0.0208%	followed	6	0.0208%	self-study
6	0.0208%	functions	6	0.0208%	speak
6	0.0208%	genuine	6	0.0208%	stress
6	0.0208%	halloween	6	0.0208%	studies
6	0.0208%	her	6	0.0208%	suggested
6	0.0208%	homosexual	6	0.0208%	superiority
6	0.0208%	housework	6	0.0208%	until
6	0.0208%	idea	6	0.0208%	walk
6	0.0208%	identifying	6	0.0208%	walks
6	0.0208%	included	6	0.0208%	wrote
6	0.0208%	involving	6	0.0208%	year
6	0.0208%	job	5	0.0173%	-
6	0.0208%	kinds	5	0.0173%	10
6	0.0208%	learned	5	0.0173%	25
6	0.0208%	lives	5	0.0173%	8
6	0.0208%	longman	5	0.0173%	ability
6	0.0208%	male	5	0.0173%	above
6	0.0208%	media	5	0.0173%	across
6	0.0208%	mentioned	5	0.0173%	addition
6	0.0208%	mostly	5	0.0173%	affect
6	0.0208%	non-native	5	0.0173%	aim
6	0.0208%	o	5	0.0173%	allows
6	0.0208%	ohp	5	0.0173%	along
6	0.0208%	page	5	0.0173%	answered
6	0.0208%	pair	5	0.0173%	approaching
6	0.0208%	paper	5	0.0173%	appropriately
6	0.0208%	party	5	0.0173%	build
6	0.0208%	passages	5	0.0173%	carefully
6	0.0208%	patterns	5	0.0173%	cases
6	0.0208%	perhaps	5	0.0173%	cause

5	0.0173%	challenging	5	0.0173%	meanings
5	0.0173%	change	5	0.0173%	medicine
5	0.0173%	chaudron	5	0.0173%	mentor
5	0.0173%	checking	5	0.0173%	method
5	0.0173%	classmates	5	0.0173%	modals
5	0.0173%	clearly	5	0.0173%	motivate
5	0.0173%	closely	5	0.0173%	mystery
5	0.0173%	cloze	5	0.0173%	negotiation
5	0.0173%	completing	5	0.0173%	no
5	0.0173%	continued	5	0.0173%	non-evident
5	0.0173%	contrasts	5	0.0173%	non-interactive
5	0.0173%	corpus	5	0.0173%	occasionally
5	0.0173%	countries	5	0.0173%	our
5	0.0173%	country	5	0.0173%	out-of-class
5	0.0173%	dealt	5	0.0173%	particularly
5	0.0173%	development	5	0.0173%	patients
5	0.0173%	difference	5	0.0173%	probably
5	0.0173%	dimensions	5	0.0173%	produce
5	0.0173%	easily	5	0.0173%	protocols
5	0.0173%	education	5	0.0173%	quite
5	0.0173%	efl	5	0.0173%	reactive
5	0.0173%	encourage	5	0.0173%	really
5	0.0173%	evidence	5	0.0173%	reinforce
5	0.0173%	experiment	5	0.0173%	requirements
5	0.0173%	fact	5	0.0173%	retrospection
5	0.0173%	fairly	5	0.0173%	richards
5	0.0173%	field	5	0.0173%	role
5	0.0173%	friends	5	0.0173%	rules
5	0.0173%	go	5	0.0173%	shows
5	0.0173%	increase	5	0.0173%	simple
5	0.0173%	instances	5	0.0173%	simply
5	0.0173%	introduced	5	0.0173%	something
5	0.0173%	introduces	5	0.0173%	space
5	0.0173%	introducing	5	0.0173%	stories
5	0.0173%	isbn	5	0.0173%	studied
5	0.0173%	issue	5	0.0173%	taken
5	0.0173%	issues	5	0.0173%	talk
5	0.0173%	john	5	0.0173%	talked
5	0.0173%	journal	5	0.0173%	technique
5	0.0173%	latvia	5	0.0173%	tell
5	0.0173%	lead	5	0.0173%	tend
5	0.0173%	length	5	0.0173%	tool
5	0.0173%	levels	5	0.0173%	top-down
5	0.0173%	little	5	0.0173%	trends
5	0.0173%	low-frequency	5	0.0173%	turn
5	0.0173%	man	5	0.0173%	turns

5	0.0173%	understood
5	0.0173%	verbs
5	0.0173%	wanted
5	0.0173%	warm-up
5	0.0173%	went
5	0.0173%	whereas
5	0.0173%	white
5	0.0173%	willing
5	0.0173%	york

808 Corpus
12 files; 16, 045 words

743	4.6307%	the	40	0.2493%	business
557	3.4715%	of	39	0.2431%	idioms
473	2.9480%	to	39	0.2431%	there
438	2.7298%	and	38	0.2368%	from
374	2.3309%	in	38	0.2368%	speech
282	1.7576%	a	37	0.2306%	acquisition
257	1.6017%	that	37	0.2306%	also
176	1.0969%	language	37	0.2306%	but
169	1.0533%	is	37	0.2306%	correction
148	0.9224%	for	36	0.2244%	learning
136	0.8476%	on	36	0.2244%	one
124	0.7728%	their	36	0.2244%	p
123	0.7666%	be	36	0.2244%	research
120	0.7479%	as	36	0.2244%	second
119	0.7417%	students	35	0.2181%	may
110	0.6856%	are	35	0.2181%	need
110	0.6856%	with	35	0.2181%	reading
98	0.6108%	this	35	0.2181%	were
93	0.5796%	not	35	0.2181%	what
84	0.5235%	or	34	0.2119%	esl
75	0.4674%	test	34	0.2119%	speakers
73	0.4550%	more	34	0.2119%	use
73	0.4550%	vocabulary	33	0.2057%	some
71	0.4425%	english	32	0.1994%	1997
71	0.4425%	they	32	0.1994%	conversation
65	0.4051%	by	32	0.1994%	errors
62	0.3864%	teachers	30	0.1870%	1
59	0.3677%	an	30	0.1870%	all
58	0.3615%	it	30	0.1870%	awareness
58	0.3615%	pragmatic	30	0.1870%	than
57	0.3553%	have	29	0.1807%	classroom
55	0.3428%	l2	29	0.1807%	developing
55	0.3428%	learners	29	0.1807%	students'
51	0.3179%	teaching	29	0.1807%	truscott
50	0.3116%	at	29	0.1807%	used
50	0.3116%	can	29	0.1807%	using
49	0.3054%	these	28	0.1745%	grammar
48	0.2992%	writing	28	0.1745%	input
46	0.2867%	words	28	0.1745%	many
43	0.2680%	which	28	0.1745%	native
42	0.2618%	will	28	0.1745%	order
41	0.2555%	feedback	28	0.1745%	time
41	0.2555%	i	27	0.1683%	competence

27	0.1683%	if	20	0.1246%	error
27	0.1683%	such	20	0.1246%	his
27	0.1683%	would	20	0.1246%	we
26	0.1620%	because	20	0.1246%	when
26	0.1620%	make	19	0.1184%	based
26	0.1620%	proficiency	19	0.1184%	development
26	0.1620%	should	19	0.1184%	different
26	0.1620%	student	19	0.1184%	her
25	0.1558%	2	19	0.1184%	less
25	0.1558%	do	19	0.1184%	listening
25	0.1558%	my	19	0.1184%	someone
25	0.1558%	pragmatics	19	0.1184%	therefore
25	0.1558%	them	19	0.1184%	university
24	0.1496%	about	18	0.1122%	activities
24	0.1496%	article	18	0.1122%	calts
24	0.1496%	expressions	18	0.1122%	context
24	0.1496%	instruction	18	0.1122%	explicit
24	0.1496%	journal	18	0.1122%	focus
24	0.1496%	materials	18	0.1122%	higher
24	0.1496%	most	18	0.1122%	possible
24	0.1496%	while	18	0.1122%	testing
23	0.1433%	1991	17	0.1060%	1993
23	0.1433%	communicative	17	0.1060%	classrooms
23	0.1433%	how	17	0.1060%	does
23	0.1433%	however	17	0.1060%	even
23	0.1433%	japanese	17	0.1060%	important
23	0.1433%	often	17	0.1060%	knowledge
23	0.1433%	only	17	0.1060%	might
23	0.1433%	s	17	0.1060%	out
23	0.1433%	status	17	0.1060%	own
23	0.1433%	target	17	0.1060%	textbooks
22	0.1371%	between	16	0.0997%	2000
22	0.1371%	d	16	0.0997%	approach
22	0.1371%	j	16	0.0997%	being
22	0.1371%	tanaka	16	0.0997%	best
22	0.1371%	well	16	0.0997%	form
22	0.1371%	who	16	0.0997%	has
21	0.1309%	1995	16	0.0997%	help
21	0.1309%	bardovi-harlig	16	0.0997%	meaning
21	0.1309%	into	16	0.0997%	notion
21	0.1309%	must	16	0.0997%	tesol
21	0.1309%	other	16	0.0997%	through
21	0.1309%	she	16	0.0997%	word
21	0.1309%	so	15	0.0935%	1998
21	0.1309%	teach	15	0.0935%	al
20	0.1246%	duquette	15	0.0935%	appropriate

15	0.0935%	correct	13	0.0810%	written
15	0.0935%	et	12	0.0748%	3
15	0.0935%	example	12	0.0748%	argument
15	0.0935%	grammatical	12	0.0748%	case
15	0.0935%	lower	12	0.0748%	computer
15	0.0935%	m	12	0.0748%	corpus
15	0.0935%	no	12	0.0748%	data
15	0.0935%	reformulation	12	0.0748%	especially
15	0.0935%	speaking	12	0.0748%	had
15	0.0935%	specific	12	0.0748%	idiomatic
15	0.0935%	strategies	12	0.0748%	increase
15	0.0935%	takahashi	12	0.0748%	noticing
15	0.0935%	very	12	0.0748%	paper
15	0.0935%	was	12	0.0748%	question
14	0.0873%	2001	12	0.0748%	since
14	0.0873%	academic	12	0.0748%	text
14	0.0873%	advanced	12	0.0748%	then
14	0.0873%	beebe	11	0.0686%	-
14	0.0873%	coady	11	0.0686%	4
14	0.0873%	could	11	0.0686%	able
14	0.0873%	fact	11	0.0686%	according
14	0.0873%	learn	11	0.0686%	been
14	0.0873%	level	11	0.0686%	concerns
14	0.0873%	methods	11	0.0686%	efl
14	0.0873%	provide	11	0.0686%	experience
14	0.0873%	researchers	11	0.0686%	found
14	0.0873%	study	11	0.0686%	know
14	0.0873%	taught	11	0.0686%	learner
14	0.0873%	tests	11	0.0686%	linguistic
13	0.0810%	ability	11	0.0686%	made
13	0.0810%	b	11	0.0686%	people
13	0.0810%	closing	11	0.0686%	r
13	0.0810%	comprehension	11	0.0686%	rather
13	0.0810%	course	11	0.0686%	results
13	0.0810%	give	11	0.0686%	suggests
13	0.0810%	he	11	0.0686%	t
13	0.0810%	items	11	0.0686%	understand
13	0.0810%	new	10	0.0623%	0
13	0.0810%	problems	10	0.0623%	1986
13	0.0810%	project	10	0.0623%	1994
13	0.0810%	providing	10	0.0623%	1996
13	0.0810%	questions	10	0.0623%	acts
13	0.0810%	same	10	0.0623%	american
13	0.0810%	various	10	0.0623%	another
13	0.0810%	way	10	0.0623%	background
13	0.0810%	where	10	0.0623%	become

10	0.0623%	classes	9	0.0561%	studies
10	0.0623%	evidence	9	0.0561%	task
10	0.0623%	great	9	0.0561%	two
10	0.0623%	group	9	0.0561%	type
10	0.0623%	groups	9	0.0561%	write
10	0.0623%	information	8	0.0499%	again
10	0.0623%	krashen	8	0.0499%	americans
10	0.0623%	levels	8	0.0499%	any
10	0.0623%	likely	8	0.0499%	aware
10	0.0623%	offer	8	0.0499%	cambridge
10	0.0623%	process	8	0.0499%	cited
10	0.0623%	professor	8	0.0499%	cultural
10	0.0623%	robb	8	0.0499%	curriculum
10	0.0623%	role	8	0.0499%	develop
10	0.0623%	ross	8	0.0499%	did
10	0.0623%	shortreed	8	0.0499%	eisenstein
10	0.0623%	show	8	0.0499%	elements
10	0.0623%	skills	8	0.0499%	exercises
10	0.0623%	still	8	0.0499%	feel
10	0.0623%	take	8	0.0499%	first
10	0.0623%	taker	8	0.0499%	four
10	0.0623%	ways	8	0.0499%	furthermore
9	0.0561%	1989	8	0.0499%	g
9	0.0561%	analysis	8	0.0499%	good
9	0.0561%	applied	8	0.0499%	large
9	0.0561%	aspects	8	0.0499%	laufer
9	0.0561%	compositions	8	0.0499%	long
9	0.0561%	corpora	8	0.0499%	meanings
9	0.0561%	e	8	0.0499%	michigan
9	0.0561%	enough	8	0.0499%	necessary
9	0.0561%	helpful	8	0.0499%	non-native
9	0.0561%	interlanguage	8	0.0499%	over
9	0.0561%	issue	8	0.0499%	person
9	0.0561%	just	8	0.0499%	point
9	0.0561%	linguistics	8	0.0499%	points
9	0.0561%	meetings	8	0.0499%	practice
9	0.0561%	much	8	0.0499%	range
9	0.0561%	natural	8	0.0499%	reasons
9	0.0561%	our	8	0.0499%	response
9	0.0561%	positive	8	0.0499%	scores
9	0.0561%	read	8	0.0499%	see
9	0.0561%	reason	8	0.0499%	situation
9	0.0561%	sections	8	0.0499%	suggestions
9	0.0561%	several	8	0.0499%	tend
9	0.0561%	situations	8	0.0499%	texts
9	0.0561%	speaker	8	0.0499%	theoretical

8	0.0499%	up	7	0.0436%	useful
8	0.0499%	zamel	7	0.0436%	you
7	0.0436%	advocates	6	0.0374%	1989a
7	0.0436%	against	6	0.0374%	1999
7	0.0436%	although	6	0.0374%	45
7	0.0436%	authentic	6	0.0374%	agree
7	0.0436%	begins	6	0.0374%	allow
7	0.0436%	benefits	6	0.0374%	answer
7	0.0436%	both	6	0.0374%	argues
7	0.0436%	build	6	0.0374%	articles
7	0.0436%	c	6	0.0374%	assumption
7	0.0436%	calt	6	0.0374%	attention
7	0.0436%	certain	6	0.0374%	basic
7	0.0436%	content	6	0.0374%	basis
7	0.0436%	contexts	6	0.0374%	believe
7	0.0436%	corrections	6	0.0374%	better
7	0.0436%	create	6	0.0374%	bodman
7	0.0436%	differences	6	0.0374%	cannot
7	0.0436%	distracters	6	0.0374%	clear
7	0.0436%	each	6	0.0374%	comments
7	0.0436%	ed	6	0.0374%	computerized
7	0.0436%	few	6	0.0374%	corps
7	0.0436%	foreign	6	0.0374%	decarrico
7	0.0436%	high	6	0.0374%	difficult
7	0.0436%	improve	6	0.0374%	difficulty
7	0.0436%	item	6	0.0374%	done
7	0.0436%	k	6	0.0374%	dunkel
7	0.0436%	like	6	0.0374%	eds
7	0.0436%	meaningful	6	0.0374%	effective
7	0.0436%	needs	6	0.0374%	either
7	0.0436%	negative	6	0.0374%	elt
7	0.0436%	notice	6	0.0374%	evaluated
7	0.0436%	overall	6	0.0374%	evaluation
7	0.0436%	part	6	0.0374%	frequency
7	0.0436%	past	6	0.0374%	general
7	0.0436%	plus	6	0.0374%	georges
7	0.0436%	pp	6	0.0374%	given
7	0.0436%	practical	6	0.0374%	gratitude
7	0.0436%	pragmatically	6	0.0374%	guidelines
7	0.0436%	press	6	0.0374%	hand
7	0.0436%	quarterly	6	0.0374%	heinle
7	0.0436%	rules	6	0.0374%	him
7	0.0436%	seem	6	0.0374%	hypothesis
7	0.0436%	setting	6	0.0374%	ideal
7	0.0436%	textbook	6	0.0374%	ideas
7	0.0436%	three	6	0.0374%	idiom

6	0.0374%	incidental	5	0.0312%	6
6	0.0374%	include	5	0.0312%	abilities
6	0.0374%	intuitions	5	0.0312%	acceptable
6	0.0374%	its	5	0.0312%	acquire
6	0.0374%	lack	5	0.0312%	acquired
6	0.0374%	lapkin	5	0.0312%	alone
6	0.0374%	learned	5	0.0312%	among
6	0.0374%	learners'	5	0.0312%	amount
6	0.0374%	learners-as-	5	0.0312%	amsterdam
researchers			5	0.0312%	assignments
6	0.0374%	main	5	0.0312%	beginning
6	0.0374%	making	5	0.0312%	behind
6	0.0374%	mean	5	0.0312%	beneficial
6	0.0374%	means	5	0.0312%	benefit
6	0.0374%	method	5	0.0312%	building
6	0.0374%	mind	5	0.0312%	cats
6	0.0374%	needed	5	0.0312%	cohen
6	0.0374%	obtain	5	0.0312%	common
6	0.0374%	others	5	0.0312%	coniam
6	0.0374%	papers	5	0.0312%	correlation
6	0.0374%	particular	5	0.0312%	creating
6	0.0374%	peace	5	0.0312%	dictionary
6	0.0374%	position	5	0.0312%	direct
6	0.0374%	program	5	0.0312%	directly
6	0.0374%	quality	5	0.0312%	discourse
6	0.0374%	questionnaire	5	0.0312%	distribution
6	0.0374%	received	5	0.0312%	effectiveness
6	0.0374%	references	5	0.0312%	embarrassing
6	0.0374%	regarding	5	0.0312%	evaluate
6	0.0374%	requires	5	0.0312%	exercise
6	0.0374%	rude	5	0.0312%	extensive
6	0.0374%	said	5	0.0312%	fernando
6	0.0374%	similar	5	0.0312%	forms
6	0.0374%	simple	5	0.0312%	further
6	0.0374%	states	5	0.0312%	giving
6	0.0374%	student's	5	0.0312%	grace
6	0.0374%	system	5	0.0312%	hartford
6	0.0374%	themselves	5	0.0312%	hulstijn
6	0.0374%	towards	5	0.0312%	impossible
6	0.0374%	try	5	0.0312%	including
6	0.0374%	ukraine	5	0.0312%	indirect
6	0.0374%	w	5	0.0312%	individual
6	0.0374%	work	5	0.0312%	instead
5	0.0312%	17	5	0.0312%	interaction
5	0.0312%	1980	5	0.0312%	interest
5	0.0312%	1985	5	0.0312%	issues

5	0.0312%	john	5	0.0312%	video
5	0.0312%	l	5	0.0312%	why
5	0.0312%	ll	5	0.0312%	within
5	0.0312%	least	5	0.0312%	working
5	0.0312%	low	5	0.0312%	wray
5	0.0312%	mahan-taylor	5	0.0312%	yet
5	0.0312%	makes			
5	0.0312%	makino			
5	0.0312%	meunier			
5	0.0312%	model			
5	0.0312%	morgan			
5	0.0312%	nakajima			
5	0.0312%	number			
5	0.0312%	once			
5	0.0312%	opportunities			
5	0.0312%	particularly			
5	0.0312%	phrase			
5	0.0312%	pleasure			
5	0.0312%	present			
5	0.0312%	presented			
5	0.0312%	promote			
5	0.0312%	provided			
5	0.0312%	purpose			
5	0.0312%	qi			
5	0.0312%	reported			
5	0.0312%	responding			
5	0.0312%	result			
5	0.0312%	retention			
5	0.0312%	rewritten			
5	0.0312%	reynolds			
5	0.0312%	search			
5	0.0312%	set			
5	0.0312%	significant			
5	0.0312%	softeners			
5	0.0312%	speak			
5	0.0312%	spend			
5	0.0312%	state			
5	0.0312%	strategy			
5	0.0312%	structure			
5	0.0312%	sufficient			
5	0.0312%	suggest			
5	0.0312%	taken			
5	0.0312%	tasks			
5	0.0312%	teacher			
5	0.0312%	technique			
5	0.0312%	vague			

841 Corpus
9 files; 15, 622 words

953	6.1004%	the	47	0.3009%	one
503	3.2198%	to	46	0.2945%	english
437	2.7973%	and	46	0.2945%	how
405	2.5925%	of	43	0.2753%	student
369	2.3621%	a	42	0.2689%	would
366	2.3428%	in	41	0.2625%	from
247	1.5811%	writing	41	0.2625%	has
237	1.5171%	that	40	0.2560%	if
227	1.4531%	she	40	0.2560%	proficiency
161	1.0306%	for	40	0.2560%	reading
158	1.0114%	is	39	0.2496%	assignment
155	0.9922%	students	39	0.2496%	each
142	0.9090%	her	39	0.2496%	other
133	0.8514%	on	38	0.2432%	then
130	0.8322%	it	36	0.2304%	also
122	0.7809%	they	36	0.2304%	noticing
118	0.7553%	their	36	0.2304%	when
113	0.7233%	this	35	0.2240%	l2
104	0.6657%	was	35	0.2240%	these
103	0.6593%	with	35	0.2240%	will
96	0.6145%	as	34	0.2176%	asked
82	0.5249%	not	34	0.2176%	process
80	0.5121%	be	34	0.2176%	use
77	0.4929%	teacher	33	0.2112%	feedback
76	0.4865%	what	33	0.2112%	two
71	0.4545%	are	32	0.2048%	can
68	0.4353%	had	32	0.2048%	chapter
66	0.4225%	i	32	0.2048%	which
66	0.4225%	or	30	0.1920%	essay
66	0.4225%	them	29	0.1856%	may
61	0.3905%	by	29	0.1856%	only
60	0.3841%	an	29	0.1856%	text
59	0.3777%	at	28	0.1792%	academic
58	0.3713%	were	28	0.1792%	questions
57	0.3649%	more	28	0.1792%	read
55	0.3521%	but	27	0.1728%	first
54	0.3457%	class	27	0.1728%	ideas
53	0.3393%	grammar	27	0.1728%	peer
53	0.3393%	have	26	0.1664%	content
52	0.3329%	language	26	0.1664%	into
51	0.3265%	about	26	0.1664%	second
49	0.3137%	write	26	0.1664%	there
47	0.3009%	could	25	0.1600%	been

25	0.1600%	did	18	0.1152%	out
25	0.1600%	draft	18	0.1152%	reason
24	0.1536%	do	18	0.1152%	said
24	0.1536%	haejung	18	0.1152%	sentence
24	0.1536%	he	18	0.1152%	since
24	0.1536%	learners	18	0.1152%	who
24	0.1536%	review	17	0.1088%	good
24	0.1536%	should	17	0.1088%	instruction
24	0.1536%	study	17	0.1088%	like
23	0.1472%	korean	17	0.1088%	participants
23	0.1472%	my	17	0.1088%	research
23	0.1472%	skills	17	0.1088%	those
23	0.1472%	students'	17	0.1088%	way
23	0.1472%	well	17	0.1088%	wu
22	0.1408%	1	16	0.1024%	examples
22	0.1408%	assignments	16	0.1024%	made
22	0.1408%	different	16	0.1024%	might
22	0.1408%	p	16	0.1024%	most
22	0.1408%	qi	16	0.1024%	task
22	0.1408%	same	16	0.1024%	time
22	0.1408%	so	16	0.1024%	work
22	0.1408%	some	15	0.0960%	able
22	0.1408%	used	15	0.0960%	activities
21	0.1344%	better	15	0.0960%	done
21	0.1344%	between	15	0.0960%	example
21	0.1344%	information	15	0.0960%	interview
21	0.1344%	own	15	0.0960%	learning
21	0.1344%	using	15	0.0960%	much
20	0.1280%	after	15	0.0960%	notice
20	0.1280%	because	15	0.0960%	objective
20	0.1280%	before	15	0.0960%	order
20	0.1280%	lapkin	15	0.0960%	response
20	0.1280%	learner	15	0.0960%	stage
20	0.1280%	many	15	0.0960%	teachers
20	0.1280%	school	14	0.0896%	ahead
19	0.1216%	classes	14	0.0896%	another
19	0.1216%	final	14	0.0896%	book
19	0.1216%	give	14	0.0896%	comments
19	0.1216%	says	14	0.0896%	errors
19	0.1216%	teaching	14	0.0896%	fact
18	0.1152%	does	14	0.0896%	given
18	0.1152%	help	14	0.0896%	his
18	0.1152%	however	14	0.0896%	new
18	0.1152%	looking	14	0.0896%	over
18	0.1152%	make	14	0.0896%	paper
18	0.1152%	no	14	0.0896%	particular

14	0.0896%	people	11	0.0704%	know
14	0.0896%	present	11	0.0704%	native
14	0.0896%	tasks	11	0.0704%	next
14	0.0896%	tense	11	0.0704%	papers
14	0.0896%	than	11	0.0704%	point
14	0.0896%	vocabulary	11	0.0704%	researchers
14	0.0896%	words	11	0.0704%	sheet
14	0.0896%	written	11	0.0704%	think
13	0.0832%	2	11	0.0704%	topic
13	0.0832%	analysis	11	0.0704%	while
13	0.0832%	essays	11	0.0704%	why
13	0.0832%	even	11	0.0704%	wrote
13	0.0832%	gave	10	0.0640%	according
13	0.0832%	just	10	0.0640%	activity
13	0.0832%	level	10	0.0640%	appropriate
13	0.0832%	need	10	0.0640%	both
13	0.0832%	past	10	0.0640%	correct
13	0.0832%	provides	10	0.0640%	discussion
13	0.0832%	say	10	0.0640%	doing
13	0.0832%	short	10	0.0640%	during
13	0.0832%	something	10	0.0640%	esl
13	0.0832%	structure	10	0.0640%	higher
13	0.0832%	word	10	0.0640%	include
12	0.0768%	3	10	0.0640%	last
12	0.0768%	am	10	0.0640%	less
12	0.0768%	based	10	0.0640%	literacy
12	0.0768%	believes	10	0.0640%	mistakes
12	0.0768%	error	10	0.0640%	organization
12	0.0768%	participant	10	0.0640%	presented
12	0.0768%	practice	10	0.0640%	problems
12	0.0768%	researcher	10	0.0640%	program
12	0.0768%	su	10	0.0640%	protocols
12	0.0768%	three	10	0.0640%	readings
12	0.0768%	through	10	0.0640%	reformulation
12	0.0768%	times	10	0.0640%	related
12	0.0768%	university	10	0.0640%	several
12	0.0768%	very	10	0.0640%	such
12	0.0768%	you	10	0.0640%	teacher's
11	0.0704%	actually	10	0.0640%	test
11	0.0704%	all	10	0.0640%	think-aloud
11	0.0704%	although	10	0.0640%	up
11	0.0704%	data	9	0.0576%	analyzing
11	0.0704%	develop	9	0.0576%	authentic
11	0.0704%	explained	9	0.0576%	being
11	0.0704%	feels	9	0.0576%	came
11	0.0704%	form	9	0.0576%	discussed

9	0.0576%	exercises	8	0.0512%	understand
9	0.0576%	focus	8	0.0512%	version
9	0.0576%	formal	8	0.0512%	where
9	0.0576%	later	7	0.0448%	2001
9	0.0576%	learn	7	0.0448%	4
9	0.0576%	lot	7	0.0448%	accepting
9	0.0576%	michigan	7	0.0448%	ad
9	0.0576%	noticed	7	0.0448%	believe
9	0.0576%	possible	7	0.0448%	case
9	0.0576%	previous	7	0.0448%	change
9	0.0576%	problem	7	0.0448%	classroom
9	0.0576%	provided	7	0.0448%	coding
9	0.0576%	quality	7	0.0448%	composition
9	0.0576%	question	7	0.0448%	compositions
9	0.0576%	speakers	7	0.0448%	conclusion
9	0.0576%	speaking	7	0.0448%	considered
9	0.0576%	state	7	0.0448%	features
9	0.0576%	things	7	0.0448%	felt
9	0.0576%	true	7	0.0448%	find
9	0.0576%	writers	7	0.0448%	further
8	0.0512%	ability	7	0.0448%	group
8	0.0512%	approach	7	0.0448%	including
8	0.0512%	began	7	0.0448%	its
8	0.0512%	college	7	0.0448%	itself
8	0.0512%	difficult	7	0.0448%	look
8	0.0512%	due	7	0.0448%	materials
8	0.0512%	education	7	0.0448%	mean
8	0.0512%	express	7	0.0448%	needs
8	0.0512%	follow	7	0.0448%	never
8	0.0512%	free	7	0.0448%	original
8	0.0512%	giving	7	0.0448%	preparing
8	0.0512%	going	7	0.0448%	reasons
8	0.0512%	having	7	0.0448%	receive
8	0.0512%	improve	7	0.0448%	required
8	0.0512%	introduction	7	0.0448%	revise
8	0.0512%	me	7	0.0448%	showed
8	0.0512%	model	7	0.0448%	simple
8	0.0512%	now	7	0.0448%	specifically
8	0.0512%	organize	7	0.0448%	thoughts
8	0.0512%	part	7	0.0448%	tl
8	0.0512%	perhaps	7	0.0448%	variable
8	0.0512%	rhetorical	7	0.0448%	variety
8	0.0512%	see	7	0.0448%	we
8	0.0512%	structures	7	0.0448%	year
8	0.0512%	take	7	0.0448%	years
8	0.0512%	type	6	0.0384%	-

6	0.0384%	12	6	0.0384%	points
6	0.0384%	1991	6	0.0384%	prepare
6	0.0384%	accuracy	6	0.0384%	problematic
6	0.0384%	aloud	6	0.0384%	pronunciation
6	0.0384%	answer	6	0.0384%	protocol
6	0.0384%	appendix	6	0.0384%	put
6	0.0384%	audience	6	0.0384%	quarterly
6	0.0384%	authors	6	0.0384%	relationship
6	0.0384%	basic	6	0.0384%	require
6	0.0384%	benefits	6	0.0384%	seem
6	0.0384%	changes	6	0.0384%	sentences
6	0.0384%	come	6	0.0384%	series
6	0.0384%	continuous	6	0.0384%	still
6	0.0384%	course	6	0.0384%	studied
6	0.0384%	courses	6	0.0384%	style
6	0.0384%	day	6	0.0384%	taking
6	0.0384%	discussions	6	0.0384%	tesol
6	0.0384%	either	6	0.0384%	therefore
6	0.0384%	end	6	0.0384%	together
6	0.0384%	focuses	6	0.0384%	told
6	0.0384%	following	6	0.0384%	took
6	0.0384%	foreign	6	0.0384%	topics
6	0.0384%	four	6	0.0384%	trying
6	0.0384%	heinle	6	0.0384%	understanding
6	0.0384%	helpful	6	0.0384%	until
6	0.0384%	helps	6	0.0384%	various
6	0.0384%	iep	6	0.0384%	without
6	0.0384%	importance	6	0.0384%	writes
6	0.0384%	includes	6	0.0384%	yes
6	0.0384%	incorporated	6	0.0384%	york
6	0.0384%	instead	5	0.0320%	30
6	0.0384%	interested	5	0.0320%	already
6	0.0384%	letters	5	0.0320%	arise
6	0.0384%	levels	5	0.0320%	articles
6	0.0384%	lower	5	0.0320%	awareness
6	0.0384%	makes	5	0.0320%	b
6	0.0384%	meet	5	0.0320%	begin
6	0.0384%	non-native	5	0.0320%	best
6	0.0384%	note	5	0.0320%	body
6	0.0384%	often	5	0.0320%	books
6	0.0384%	ohp	5	0.0320%	clear
6	0.0384%	others	5	0.0320%	columns
6	0.0384%	page	5	0.0320%	comprehension
6	0.0384%	period	5	0.0320%	corrected
6	0.0384%	personal	5	0.0320%	davenport
6	0.0384%	place	5	0.0320%	describing

5	0.0320%	designed	5	0.0320%	peers
5	0.0320%	directly	5	0.0320%	person
5	0.0320%	discuss	5	0.0320%	population
5	0.0320%	drawer	5	0.0320%	practical
5	0.0320%	early	5	0.0320%	pre-writing
5	0.0320%	experiences	5	0.0320%	purpose
5	0.0320%	explain	5	0.0320%	really
5	0.0320%	explanation	5	0.0320%	remembers
5	0.0320%	few	5	0.0320%	rest
5	0.0320%	focused	5	0.0320%	s
5	0.0320%	formed	5	0.0320%	section
5	0.0320%	get	5	0.0320%	seems
5	0.0320%	goals	5	0.0320%	similar
5	0.0320%	grammatical	5	0.0320%	skill
5	0.0320%	guide	5	0.0320%	spent
5	0.0320%	hand	5	0.0320%	stopped
5	0.0320%	home	5	0.0320%	teach
5	0.0320%	immediate	5	0.0320%	texts
5	0.0320%	important	5	0.0320%	third
5	0.0320%	incorrect	5	0.0320%	though
5	0.0320%	involved	5	0.0320%	throughout
5	0.0320%	journal	5	0.0320%	try
5	0.0320%	knowledge	5	0.0320%	types
5	0.0320%	korea	5	0.0320%	usually
5	0.0320%	ll	5	0.0320%	verbal
5	0.0320%	lead	5	0.0320%	went
5	0.0320%	leads	5	0.0320%	wishes
5	0.0320%	learned	5	0.0320%	within
5	0.0320%	learners'			
5	0.0320%	least			
5	0.0320%	lesson			
5	0.0320%	letter			
5	0.0320%	little			
5	0.0320%	logical			
5	0.0320%	main			
5	0.0320%	metalanguage			
5	0.0320%	method			
5	0.0320%	minutes			
5	0.0320%	mode			
5	0.0320%	naturally			
5	0.0320%	night			
5	0.0320%	objectives			
5	0.0320%	opportunities			
5	0.0320%	oral			
5	0.0320%	paragraph			
5	0.0320%	participate			

896 Corpus
36 files; 31, 058 words

2017	6.4943%	the
1095	3.5257%	to
956	3.0781%	and
765	2.4631%	of
667	2.1476%	a
629	2.0252%	i
593	1.9093%	in
452	1.4553%	students
356	1.1462%	they
339	1.0915%	for
330	1.0625%	on
330	1.0625%	that
329	1.0593%	this
296	0.9531%	was
257	0.8275%	it
255	0.8210%	class
254	0.8178%	with
235	0.7566%	their
226	0.7277%	them
221	0.7116%	as
206	0.6633%	have
196	0.6311%	is
192	0.6182%	about
190	0.6118%	had
164	0.5280%	not
157	0.5055%	were
151	0.4862%	are
149	0.4797%	be
149	0.4797%	or
140	0.4508%	more
136	0.4379%	what
129	0.4154%	an
122	0.3928%	at
119	0.3832%	activity
117	0.3767%	but
117	0.3767%	writing
116	0.3735%	reading
109	0.3510%	questions
108	0.3477%	one
108	0.3477%	use
106	0.3413%	we
104	0.3349%	grammar
100	0.3220%	would
98	0.3155%	by
97	0.3123%	each
96	0.3091%	teacher
96	0.3091%	time
94	0.3027%	from
93	0.2994%	my
93	0.2994%	then
91	0.2930%	so
88	0.2833%	lesson
83	0.2672%	do
81	0.2608%	he
81	0.2608%	some
80	0.2576%	words
79	0.2544%	did
79	0.2544%	when
77	0.2479%	first
75	0.2415%	language
74	0.2383%	also
74	0.2383%	can
74	0.2383%	how
74	0.2383%	which
73	0.2350%	if
73	0.2350%	very
69	0.2222%	these
68	0.2189%	used
65	0.2093%	other
65	0.2093%	sentences
64	0.2061%	could
63	0.2028%	after
63	0.2028%	asked
63	0.2028%	think
61	0.1964%	activities
57	0.1835%	all
56	0.1803%	there
55	0.1771%	out
55	0.1771%	vocabulary
55	0.1771%	work
54	0.1739%	using
53	0.1706%	because
53	0.1706%	book
53	0.1706%	many
53	0.1706%	student
53	0.1706%	well

52	0.1674%	read	35	0.1127%	while
52	0.1674%	unit	34	0.1095%	know
52	0.1674%	will	34	0.1095%	might
51	0.1642%	write	34	0.1095%	second
50	0.1610%	been	34	0.1095%	since
50	0.1610%	you	34	0.1095%	skills
49	0.1578%	exercises	34	0.1095%	such
47	0.1513%	able	33	0.1063%	answer
47	0.1513%	into	33	0.1063%	example
46	0.1481%	however	33	0.1063%	sentence
46	0.1481%	ideas	32	0.1030%	groups
46	0.1481%	new	32	0.1030%	next
45	0.1449%	much	32	0.1030%	word
45	0.1449%	teaching	31	0.0998%	good
44	0.1417%	english	31	0.0998%	point
44	0.1417%	few	30	0.0966%	board
44	0.1417%	like	30	0.0966%	even
44	0.1417%	topic	29	0.0934%	2001
44	0.1417%	up	29	0.0934%	chart
43	0.1385%	focus	29	0.0934%	during
43	0.1385%	than	29	0.0934%	found
42	0.1352%	most	29	0.0934%	northstar
42	0.1352%	should	29	0.0934%	own
42	0.1352%	who	29	0.0934%	readings
41	0.1320%	different	28	0.0902%	am
41	0.1320%	me	28	0.0902%	get
41	0.1320%	really	28	0.0902%	need
41	0.1320%	two	28	0.0902%	presented
40	0.1288%	form	28	0.0902%	structure
40	0.1288%	group	28	0.0902%	students'
40	0.1288%	meaning	28	0.0902%	textbook
40	0.1288%	only	28	0.0902%	wrote
40	0.1288%	t	27	0.0869%	another
39	0.1256%	exercise	27	0.0869%	classroom
39	0.1256%	information	27	0.0869%	important
39	0.1256%	make	27	0.0869%	made
39	0.1256%	part	26	0.0837%	3
37	0.1191%	before	26	0.0837%	answers
37	0.1191%	minutes	26	0.0837%	interesting
37	0.1191%	points	26	0.0837%	same
36	0.1159%	practice	26	0.0837%	section
36	0.1159%	too	26	0.0837%	speaking
35	0.1127%	article	25	0.0805%	better
35	0.1127%	comprehension	25	0.0805%	feel
35	0.1127%	just	25	0.0805%	future
35	0.1127%	listening	25	0.0805%	give

25	0.0805%	learning	20	0.0644%	no
25	0.0805%	level	20	0.0644%	problem
25	0.0805%	over	20	0.0644%	spent
25	0.0805%	story	20	0.0644%	style
25	0.0805%	together	20	0.0644%	themes
25	0.0805%	where	20	0.0644%	thought
24	0.0773%	2	20	0.0644%	want
24	0.0773%	go	20	0.0644%	wanted
24	0.0773%	learn	19	0.0612%	course
24	0.0773%	often	19	0.0612%	day
24	0.0773%	understand	19	0.0612%	didn't
23	0.0741%	1	19	0.0612%	discussed
23	0.0741%	any	19	0.0612%	either
23	0.0741%	between	19	0.0612%	essay
23	0.0741%	complete	19	0.0612%	helpful
23	0.0741%	down	19	0.0612%	his
23	0.0741%	find	19	0.0612%	main
23	0.0741%	homework	19	0.0612%	once
23	0.0741%	lot	19	0.0612%	organization
23	0.0741%	materials	19	0.0612%	people
23	0.0741%	order	19	0.0612%	take
23	0.0741%	pictures	19	0.0612%	way
23	0.0741%	present	18	0.0580%	difficult
23	0.0741%	review	18	0.0580%	discussion
23	0.0741%	s	18	0.0580%	does
23	0.0741%	took	18	0.0580%	enough
23	0.0741%	working	18	0.0580%	feedback
22	0.0708%	gave	18	0.0580%	going
22	0.0708%	help	18	0.0580%	journal
22	0.0708%	material	18	0.0580%	letter
22	0.0708%	past	18	0.0580%	passive
22	0.0708%	see	18	0.0580%	something
22	0.0708%	still	18	0.0580%	useful
22	0.0708%	through	17	0.0547%	always
21	0.0676%	again	17	0.0547%	articles
21	0.0676%	correct	17	0.0547%	authentic
21	0.0676%	fact	17	0.0547%	done
21	0.0676%	has	17	0.0547%	especially
21	0.0676%	paragraph	17	0.0547%	examples
21	0.0676%	said	17	0.0547%	following
21	0.0676%	went	17	0.0547%	may
20	0.0644%	being	17	0.0547%	pair
20	0.0644%	don't	17	0.0547%	possible
20	0.0644%	instead	17	0.0547%	simply
20	0.0644%	last	17	0.0547%	talked
20	0.0644%	long	17	0.0547%	text

17	0.0547%	those	14	0.0451%	right
17	0.0547%	worked	14	0.0451%	specific
16	0.0515%	already	14	0.0451%	speech
16	0.0515%	american	14	0.0451%	started
16	0.0515%	ask	14	0.0451%	stoller
16	0.0515%	back	14	0.0451%	taught
16	0.0515%	explicit	14	0.0451%	theme
16	0.0515%	hour	14	0.0451%	topics
16	0.0515%	longer	14	0.0451%	try
16	0.0515%	objective	14	0.0451%	why
16	0.0515%	related	13	0.0419%	appropriate
16	0.0515%	seem	13	0.0419%	focused
16	0.0515%	she	13	0.0419%	grabe
16	0.0515%	short	13	0.0419%	guess
16	0.0515%	talking	13	0.0419%	having
15	0.0483%	asking	13	0.0419%	hear
15	0.0483%	effective	13	0.0419%	items
15	0.0483%	entire	13	0.0419%	lessons
15	0.0483%	forms	13	0.0419%	looking
15	0.0483%	general	13	0.0419%	making
15	0.0483%	high	13	0.0419%	p
15	0.0483%	home	13	0.0419%	picture
15	0.0483%	knowledge	13	0.0419%	previous
15	0.0483%	talk	13	0.0419%	probably
15	0.0483%	task	13	0.0419%	program
15	0.0483%	times	13	0.0419%	pronunciation
15	0.0483%	ways	13	0.0419%	question
15	0.0483%	written	13	0.0419%	say
14	0.0451%	-	13	0.0419%	sure
14	0.0451%	around	13	0.0419%	tasks
14	0.0451%	based	13	0.0419%	teach
14	0.0451%	charts	13	0.0419%	themselves
14	0.0451%	check	13	0.0419%	three
14	0.0451%	context	13	0.0419%	trying
14	0.0451%	describe	13	0.0419%	usually
14	0.0451%	doing	13	0.0419%	verbs
14	0.0451%	easy	13	0.0419%	video
14	0.0451%	end	13	0.0419%	without
14	0.0451%	felt	12	0.0386%	5
14	0.0451%	given	12	0.0386%	almost
14	0.0451%	goals	12	0.0386%	attention
14	0.0451%	little	12	0.0386%	audience
14	0.0451%	newspaper	12	0.0386%	brainstorming
14	0.0451%	now	12	0.0386%	classes
14	0.0451%	problems	12	0.0386%	communicative
14	0.0451%	responses	12	0.0386%	compare

12	0.0386%	correction	11	0.0354%	seemed
12	0.0386%	cultural	11	0.0354%	showed
12	0.0386%	details	11	0.0354%	state
12	0.0386%	grammatical	11	0.0354%	structures
12	0.0386%	included	11	0.0354%	though
12	0.0386%	interest	11	0.0354%	today
12	0.0386%	its	11	0.0354%	understanding
12	0.0386%	list	11	0.0354%	understood
12	0.0386%	metalinguistic	11	0.0354%	variety
12	0.0386%	necessary	11	0.0354%	voice
12	0.0386%	others	11	0.0354%	week
12	0.0386%	particular	10	0.0322%	academic
12	0.0386%	practicum	10	0.0322%	actually
12	0.0386%	rather	10	0.0322%	began
12	0.0386%	sometimes	10	0.0322%	both
12	0.0386%	taken	10	0.0322%	building
12	0.0386%	terms	10	0.0322%	comfortable
12	0.0386%	various	10	0.0322%	completed
12	0.0386%	whether	10	0.0322%	correcting
12	0.0386%	whole	10	0.0322%	countries
12	0.0386%	your	10	0.0322%	differences
11	0.0354%	--	10	0.0322%	due
11	0.0354%	10	10	0.0322%	easily
11	0.0354%	1997	10	0.0322%	enjoyed
11	0.0354%	certain	10	0.0322%	finished
11	0.0354%	cited	10	0.0322%	further
11	0.0354%	come	10	0.0322%	furthermore
11	0.0354%	content	10	0.0322%	helped
11	0.0354%	etc	10	0.0322%	her
11	0.0354%	fill	10	0.0322%	here
11	0.0354%	half	10	0.0322%	house
11	0.0354%	idea	10	0.0322%	individually
11	0.0354%	identify	10	0.0322%	interested
11	0.0354%	kind	10	0.0322%	it's
11	0.0354%	later	10	0.0322%	keep
11	0.0354%	learned	10	0.0322%	life
11	0.0354%	look	10	0.0322%	motivation
11	0.0354%	mistakes	10	0.0322%	never
11	0.0354%	objectives	10	0.0322%	news
11	0.0354%	online	10	0.0322%	ohp
11	0.0354%	our	10	0.0322%	planned
11	0.0354%	paraphrasing	10	0.0322%	production
11	0.0354%	presentation	10	0.0322%	quiz
11	0.0354%	quickly	10	0.0322%	simple
11	0.0354%	recasts	10	0.0322%	skill
11	0.0354%	room	10	0.0322%	spelling

10	0.0322%	told	9	0.0290%	test
10	0.0322%	transition	9	0.0290%	thinking
10	0.0322%	tried	9	0.0290%	today's
10	0.0322%	units	9	0.0290%	verb
10	0.0322%	university	8	0.0258%	adjectives
10	0.0322%	utterances	8	0.0258%	along
9	0.0290%	4	8	0.0258%	approach
9	0.0290%	aspects	8	0.0258%	atmosphere
9	0.0290%	assignment	8	0.0258%	bit
9	0.0290%	beginning	8	0.0258%	brought
9	0.0290%	bibliography	8	0.0258%	call
9	0.0290%	cities	8	0.0258%	came
9	0.0290%	clearly	8	0.0258%	chance
9	0.0290%	comic	8	0.0258%	clarification
9	0.0290%	comparing	8	0.0258%	clear
9	0.0290%	country	8	0.0258%	continue
9	0.0290%	doctor	8	0.0258%	control
9	0.0290%	esl	8	0.0258%	conversation
9	0.0290%	essays	8	0.0258%	create
9	0.0290%	everyone	8	0.0258%	created
9	0.0290%	familiar	8	0.0258%	difference
9	0.0290%	final	8	0.0258%	e
9	0.0290%	getting	8	0.0258%	elc
9	0.0290%	instruction	8	0.0258%	encourage
9	0.0290%	intermediate	8	0.0258%	errors
9	0.0290%	introduce	8	0.0258%	everything
9	0.0290%	involved	8	0.0258%	explain
9	0.0290%	makes	8	0.0258%	explicitly
9	0.0290%	mind	8	0.0258%	five
9	0.0290%	onto	8	0.0258%	foreign
9	0.0290%	page	8	0.0258%	fun
9	0.0290%	pages	8	0.0258%	got
9	0.0290%	paragraphs	8	0.0258%	handed
9	0.0290%	person	8	0.0258%	heinle
9	0.0290%	personal	8	0.0258%	improve
9	0.0290%	plans	8	0.0258%	include
9	0.0290%	provide	8	0.0258%	internet
9	0.0290%	provided	8	0.0258%	least
9	0.0290%	rest	8	0.0258%	letters
9	0.0290%	semester	8	0.0258%	levels
9	0.0290%	show	8	0.0258%	listen
9	0.0290%	small	8	0.0258%	m
9	0.0290%	spada	8	0.0258%	must
9	0.0290%	stress	8	0.0258%	overall
9	0.0290%	successful	8	0.0258%	pain
9	0.0290%	taking	8	0.0258%	pairs

8	0.0258%	perhaps	7	0.0225%	less
8	0.0258%	phrases	7	0.0225%	move
8	0.0258%	plan	7	0.0225%	multiple
8	0.0258%	progressive	7	0.0225%	natural
8	0.0258%	quite	7	0.0225%	newspapers
8	0.0258%	regarding	7	0.0225%	notes
8	0.0258%	saying	7	0.0225%	ok
8	0.0258%	scanning	7	0.0225%	opinions
8	0.0258%	sections	7	0.0225%	outside
8	0.0258%	segment	7	0.0225%	partner
8	0.0258%	several	7	0.0225%	period
8	0.0258%	spend	7	0.0225%	place
8	0.0258%	subject	7	0.0225%	played
8	0.0258%	therefore	7	0.0225%	practicing
8	0.0258%	things	7	0.0225%	purple
8	0.0258%	thinks	7	0.0225%	quizzes
8	0.0258%	v	7	0.0225%	radio
7	0.0225%	1994	7	0.0225%	reasons
7	0.0225%	30	7	0.0225%	required
7	0.0225%	analyze	7	0.0225%	sample
7	0.0225%	assigned	7	0.0225%	share
7	0.0225%	assume	7	0.0225%	start
7	0.0225%	basic	7	0.0225%	stressed
7	0.0225%	books	7	0.0225%	suggestions
7	0.0225%	case	7	0.0225%	summary
7	0.0225%	challenging	7	0.0225%	survey
7	0.0225%	choice	7	0.0225%	target
7	0.0225%	chose	7	0.0225%	teachers
7	0.0225%	clauses	7	0.0225%	teacher's
7	0.0225%	completing	7	0.0225%	thing
7	0.0225%	complex	7	0.0225%	until
7	0.0225%	compound	7	0.0225%	varied
7	0.0225%	concerned	7	0.0225%	wasn't
7	0.0225%	creating	7	0.0225%	weekly
7	0.0225%	creative	7	0.0225%	wrong
7	0.0225%	develop	6	0.0193%	8
7	0.0225%	development	6	0.0193%	according
7	0.0225%	discussing	6	0.0193%	advertisement
7	0.0225%	discussions	6	0.0193%	amount
7	0.0225%	encouraged	6	0.0193%	anymore
7	0.0225%	game	6	0.0193%	apply
7	0.0225%	introduced	6	0.0193%	asks
7	0.0225%	introduction	6	0.0193%	assumptions
7	0.0225%	issues	6	0.0193%	begin
7	0.0225%	kept	6	0.0193%	believe
7	0.0225%	led	6	0.0193%	best

6	0.0193%	boston	6	0.0193%	range
6	0.0193%	build	6	0.0193%	realize
6	0.0193%	change	6	0.0193%	reason
6	0.0193%	checking	6	0.0193%	require
6	0.0193%	comparative	6	0.0193%	research
6	0.0193%	consider	6	0.0193%	result
6	0.0193%	corrective	6	0.0193%	results
6	0.0193%	days	6	0.0193%	rich
6	0.0193%	deal	6	0.0193%	rules
6	0.0193%	description	6	0.0193%	seen
6	0.0193%	descriptive	6	0.0193%	set
6	0.0193%	detectives	6	0.0193%	speak
6	0.0193%	discuss	6	0.0193%	stated
6	0.0193%	divided	6	0.0193%	statements
6	0.0193%	edition	6	0.0193%	studying
6	0.0193%	expected	6	0.0193%	styles
6	0.0193%	express	6	0.0193%	success
6	0.0193%	far	6	0.0193%	support
6	0.0193%	finish	6	0.0193%	suspects
6	0.0193%	giving	6	0.0193%	techniques
6	0.0193%	great	6	0.0193%	tense
6	0.0193%	higher	6	0.0193%	third
6	0.0193%	him	6	0.0193%	type
6	0.0193%	hulstijn	6	0.0193%	u
6	0.0193%	i'm	6	0.0193%	worry
6	0.0193%	job	5	0.0161%	1985
6	0.0193%	laufer	5	0.0161%	active
6	0.0193%	likely	5	0.0161%	allows
6	0.0193%	listed	5	0.0161%	answered
6	0.0193%	lives	5	0.0161%	answering
6	0.0193%	majority	5	0.0161%	author
6	0.0193%	maybe	5	0.0161%	authors
6	0.0193%	msu	5	0.0161%	auxiliary
6	0.0193%	noticed	5	0.0161%	available
6	0.0193%	number	5	0.0161%	away
6	0.0193%	one-word	5	0.0161%	b
6	0.0193%	opportunities	5	0.0161%	bad
6	0.0193%	partners	5	0.0161%	bibliographies
6	0.0193%	peer	5	0.0161%	brainstorm
6	0.0193%	poem	5	0.0161%	cars
6	0.0193%	potential	5	0.0161%	catch
6	0.0193%	pre-writing	5	0.0161%	cats
6	0.0193%	printed	5	0.0161%	chased
6	0.0193%	produce	5	0.0161%	choose
6	0.0193%	provides	5	0.0161%	choosing
6	0.0193%	purpose	5	0.0161%	college

5	0.0161%	computer	5	0.0161%	michigan
5	0.0161%	contains	5	0.0161%	model
5	0.0161%	continued	5	0.0161%	mostly
5	0.0161%	contrasting	5	0.0161%	murder
5	0.0161%	couple	5	0.0161%	native
5	0.0161%	decide	5	0.0161%	needed
5	0.0161%	deeper	5	0.0161%	negotiate
5	0.0161%	detail	5	0.0161%	normally
5	0.0161%	detailed	5	0.0161%	note
5	0.0161%	directions	5	0.0161%	noun
5	0.0161%	dream	5	0.0161%	nouns
5	0.0161%	easier	5	0.0161%	obvious
5	0.0161%	editors	5	0.0161%	off
5	0.0161%	elicit	5	0.0161%	old
5	0.0161%	elicited	5	0.0161%	outline
5	0.0161%	else	5	0.0161%	outlines
5	0.0161%	encountered	5	0.0161%	parts
5	0.0161%	ended	5	0.0161%	pause
5	0.0161%	events	5	0.0161%	perfect
5	0.0161%	excellent	5	0.0161%	photos
5	0.0161%	experiences	5	0.0161%	pica
5	0.0161%	explanation	5	0.0161%	pick
5	0.0161%	expression	5	0.0161%	planning
5	0.0161%	facilitate	5	0.0161%	possibly
5	0.0161%	fashion	5	0.0161%	preliminary
5	0.0161%	finally	5	0.0161%	promote
5	0.0161%	flow	5	0.0161%	purse
5	0.0161%	food	5	0.0161%	put
5	0.0161%	format	5	0.0161%	received
5	0.0161%	form-focused	5	0.0161%	regular
5	0.0161%	forming	5	0.0161%	remember
5	0.0161%	fragments	5	0.0161%	report
5	0.0161%	hope	5	0.0161%	request
5	0.0161%	improved	5	0.0161%	requests
5	0.0161%	insects	5	0.0161%	reviewing
5	0.0161%	interactions	5	0.0161%	search
5	0.0161%	involve	5	0.0161%	seems
5	0.0161%	itself	5	0.0161%	self-study
5	0.0161%	knew	5	0.0161%	shelly
5	0.0161%	let	5	0.0161%	similar
5	0.0161%	matter	5	0.0161%	situation
5	0.0161%	meaningful	5	0.0161%	skimming
5	0.0161%	meanings	5	0.0161%	snow
5	0.0161%	mentioned	5	0.0161%	sound
5	0.0161%	method	5	0.0161%	spellings
5	0.0161%	mice	5	0.0161%	stories

5	0.0161%	strip	5	0.0161%	weekend
5	0.0161%	technology	5	0.0161%	weeks
5	0.0161%	testing	5	0.0161%	wish
5	0.0161%	time-management	5	0.0161%	workbook
5	0.0161%	timing	5	0.0161%	x
5	0.0161%	turned	5	0.0161%	years

REFERENCES

REFERENCES

Aston, G. (1997). Involving learners in developing learning methods: exploiting text corpora in self-access. In P. Benson and P. Voller (eds.), *Autonomy and the Independence in Language Learning*. London: Longman.

Bardovi-Harlig, K. Hartford, B.A.S., Mahan-Taylor, R., Morgan, M.J., & Reynolds, D.W. (1991). Developing pragmatic awareness: Closing the conversation. *ELT Journal* 45, 5-15.

Barlow, M. (1996). Corpora for theory and practice. *International Journal of Corpus Linguistics* 1/1, 1 –37.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use.* Cambridge: Cambridge University Press.

COBUILD Direct, Bank of English. *http://www.ldc.upenn.edu/ldc/online/index/html*

Duquette, G. (1995). Developing comprehension and interaction skills with idiomatic expressions. *Second Language Practice: Classroom Strategies for Developing Communicative Competence.* Ed. Georges Duquette. Cromwell Press: Great Britain.

Johns, T. (1991a). Should you be persuaded: two examples of data-driven learning. In T. Johns and P. Kind (eds.), *Classroom Concordancing*, ELR Journal Vol. 4. Birmingham: Birmingham University Press, 1 – 16.

---------- (1991b). From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. In T. Johns and P. Kind (eds.), *Classroom Concordancing*, ELR Journal Vol. 4. Birmingham: Birmingham University Press, 27 – 46.

Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. London: Longman Publishing.

Kirk, J. M., ed. (1998). *Corpora Galore: Analyses and Techniques in Describing English.* Papers from the Nineteenth International Conference on English Language Research on Computerized Corpora (ICAME 1998). Amsterdam – Atlanta, GA: Rodopi, The Queens University of Belfast.

Liu, J., and Nesi, H. (1999). Are we teaching the right words? A study of

students' receptive knowledge of two types of vocabulary: 'subtechnical' and 'technical'. In H. Book and P. Luford (eds.), *Academic Standards and Expectations: The role of EAP.* Nottingham: Nottingham University Press.

Madden, C. G., and Rohlck, T. N. (2000). *Discussion & Interaction in the Academic Community.* Michigan Series in English for Academic Purposes & Professional Purposes. Madden, Carolyn G., and Swales, John M., ed. Ann Arbor: The University of Michigan Press.

Makkai, A. (2001). *A Dictionary of American Idioms.* Third Edition. Barron's Educational Series, Incorporated.

McEnery, T., and Wilson, A. (1996). *Corpus Linguistics.* Manchester: Edinburgh University Press.

MICASE Online. http://www.hti.umich.edu/m/micase

Owen, C. (1993). Corpus-based grammar and the Heineken effect: Lexico-grammatical description for language learners. *Applied Linguistics 14:* 167-187.

Partington, A. (1998). *Patterns and Meanings: Using Corpora for English Language Research and Teaching.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Porter, D. and Roberts, J. (1987). Authentic listening activities. In *Methodology in TESOL,* edited by M. Long and J. Richards. Rowley, MA: Newbury House.

R.C. Simpson, S.L. Briggs, J. Ovens, and J.M. Swales. (1999). The Michigan Corpus of Academic Spoken English. Ann Arbor, MI: The Regents of the University of Michigan.

Rogers, C.V., and Medley, F.W. (1998). Language with a purpose: Using authentic materials in the foreign language classroom". *Foreign Language Annals* 21, 467 – 478.

Simpson, R., and Swales, J. (2001). *Corpus Linguistics in North America: Selections from the 1999 Symposium.* Ann Arbor: The University of Michigan Press.

Simpson, R., Mendis, D., and Komsic, A. (2002). Corpus-based study of idioms in academic speech. Conference paper given at AAAL, Salt Lake City, UT.

Sutarsyah, C., Nation, I.S.P., and Kennedy, G. (1994). How useful is EAP vocabulary for ESP? – A corpus-based case study. *RELC Journal* 25/ 2, 34 – 50.

Swales, J., and Feak, C. (1994). *Academic Writing for Graduate Students.* Ann

Arbor, MI: The University of Michigan Press.

Tanaka, K. (1997) Developing pragmatic competence: A learners-as-researchers approach. *TESOL Journal* 6, 14 –18.

Tribble, C. and Jones, G. (1990). *Concordances in the Classroom*. London: Longman.