

Teachers of English to Speakers of Other Languages, Inc. (TESOL)

The Most Frequently Used Spoken American English Idioms: A Corpus Analysis and Its Implications

Author(s): Dilin Liu

Source: *TESOL Quarterly*, Vol. 37, No. 4 (Winter, 2003), pp. 671-700

Published by: [Teachers of English to Speakers of Other Languages, Inc. \(TESOL\)](#)

Stable URL: <http://www.jstor.org/stable/3588217>

Accessed: 28/08/2011 18:01

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Teachers of English to Speakers of Other Languages, Inc. (TESOL) is collaborating with JSTOR to digitize, preserve and extend access to *TESOL Quarterly*.

<http://www.jstor.org>

The Most Frequently Used Spoken American English Idioms: A Corpus Analysis and Its Implications

DILIN LIU

Oklahoma City University

Oklahoma City, Oklahoma, United States

Most teaching and reference materials on English idioms are primarily intuition based. As such, they often include seldom-used idioms and incorrect descriptions of the meaning and use of some idioms, hence limiting their usefulness to ESOL students. This article demonstrates how this problem can be addressed through a corpus-based study of the spoken American English idioms used most frequently by college and other professional ESOL students learning American English. The study involved a close concordance search and analysis of the idioms used in three contemporary spoken American English corpora: Corpus of Spoken, Professional American English (Barlow, 2000); Michigan Corpus of Academic Spoken English (Simpson, Briggs, Ovens, & Swales, 2002), and Spoken American Media English (Liu, 2002). According to the search results, four lists of the most frequently used idioms were compiled, with one based on the overall data and the other three on one of the corpora. The study uncovered interesting English idiom use patterns. The results were compared with information in nine major current idiom dictionaries, which revealed inadequacies of the existing idiom teaching and reference materials in terms of item selection, meaning and use explanation, and the appropriateness of the examples provided. The article discusses pedagogical and research implications, including suggestions for improving the development of idiom teaching and reference materials.

Because of their rather rigid structure, quite unpredictable meaning, and fairly extensive use, idioms are “a notoriously difficult” but simultaneously very useful aspect of English for ESOL learners because a grasp of them “can be a great asset to learners in acquiring a new language” (Celce-Murcia & Larsen-Freeman, 1999, p. 39). How to help students acquire idioms has long been a challenge to ESOL educators and researchers alike.

One of the first issues to consider in idiom instruction is which idioms to teach and in what sequence. Many English idiom teaching and reference materials exist for ESOL learners, some of which claim to cover essential idioms. Yet the selection of idioms in these publications often reflects primarily the authors' intuition rather than any empirical data, and a substantial number of them are rarely used. Thus learning these idioms not only is difficult but may also be unhelpful because students rarely encounter and use them. In addition, these materials cover many seldom-used idioms but fail to cover some frequently used ones. Determining the most useful idioms for ESOL students is therefore important. Because idioms are register sensitive, any most useful idiom list must have a specific group of learners and a register in mind. This article reports a corpus study aimed at identifying the most frequently used spoken American English idioms for college and other professional ESOL students learning American English and uncovering some of the idioms' usage patterns.

BACKGROUND

Definitions of Idiom

In any idiom research, an important yet difficult initial question is, What constitutes an idiom? The definition of idiom varies considerably from scholar to scholar and may also depend on context. As Moon (1998) puts it, "Idiom is an ambiguous term, used in conflicting ways" (p. 3). For some scholars, and in a broad sense, the term is rather inclusive, covering, among other things, all fixed phrases, proverbs, formulaic speeches, and, at the extreme, even single polysemic words. For example, scholars such as Cooper (1998) and Katz and Postal (1963) have included as idioms individual words that are used metaphorically, such as *weigh* as in *weigh a decision*. Yet for other scholars, and in a more restrictive use, the term *idiom* is a much narrower concept referring only to those "fixed and semantically opaque or metaphorical" expressions, such as "*kick the bucket* or *spill the beans*" (Moon, 1998, p. 4).

What constitutes an idiom is thus often a decision at the discretion of the researcher. For this reason, Tabossi and Zardon (1993) contend that "idioms are multifaceted objects, whose study requires various viewpoints and different methodological approaches" (p. 145). Therefore, for any researcher, the task of "identifying idioms is simply an attempt to differentiate and label one class of common expressions with specific functions from others on the bases of criteria which strike the analyst as being the most illuminating," and, for that reason, different "analysts will

come up with somewhat different criteria and different identifications” (Fernando, 1996, p. 40).

Whatever definition and criteria one develops and uses in identifying idioms, they must be clear, specific, and systematic. An example is Fernando’s (1996) definition: “conventionalized multi-word expressions often, but not always non-literal” (p. 1). This definition excludes single words as idioms, which, as previously mentioned, some scholars have included. (See Cowie [1998], Stubbs [2001], and Wray [2002] for interesting and rather comprehensive analyses of formulaic language from different perspectives. The scope of their studies is much broader than Fernando’s, however, for they cover almost all types of collocations. As a result, many of the types of phrases in their discussions are not idioms, even in the broadest definition.) Fernando also focuses on the invariant or restricted variant nature of idioms to help distinguish them from other habitual collocations. According to her,

only those expressions which become conventionally fixed in a specific order and lexical form, or have only a restricted set of variants, acquire the status of idioms. Combinations, showing a relatively high degree of variability, especially in the matter of lexical replacement such as *catch a bus*, *catch a train*, etc., are not regarded as idioms, though they exemplify idiomaticity by virtue of habitual co-occurrence: *catch* meaning ‘be in time for’ co-occurs usually with a mode of transport, though *catch the post* is also possible. (p. 31)

While upholding the principles she establishes in distinguishing idioms from nonidioms, Fernando also recognizes the complexity and difficulty of the task. Following previous scholars, she developed a scale system for classifying idiomatic expressions and habitual collocations in which idioms fall into three categories: pure (nonliteral), semiliteral, and literal (see Table 1). Because of its clarity and systematic nature, I adopted Fernando’s approach and criteria for idiom identification in this study. In the Method section, I describe how I applied her theory in deciding what expressions to look for in my concordance search of idioms.

TABLE 1
Three Categories of Idioms

Category	Examples
Pure	<i>kick the bucket</i> , <i>pull someone’s leg</i> , <i>make off with</i> ^a
Semiliteral	<i>fat chance</i> , ^a <i>use something as a step stone</i> , <i>go through</i>
Literal	<i>according to</i> , <i>in sum</i> , ^a <i>throw away</i>

^aFrom Fernando (1996, p. 32).

Idiom Acquisition and Corpus Research

Despite the fact that idioms are difficult for L2 learners, historically idiom acquisition has not received adequate attention in L2 research because of what Ellis (1985) considers to be a traditional emphasis on the acquisition of "grammatical systems" (p. 5) and neglect of the lexis. Even though second language acquisition researchers are paying greater attention to lexis, most idiom-related studies have still focused on L1 (especially children's) idiom comprehension and acquisition (Cacciari, 1993; Cacciari & Levorato, 1989; Cacciari & Tabossi, 1988, 1993; Gibbs, 1986, 1987; Levorato, 1993; Levorato & Cacciari, 1995). However, since the late 1980s, and especially since the late 1990s, a few studies in L2 have appeared (Abdullah & Jackson, 1998; Cooper, 1998, 1999; Irujo, 1986a, 1986b, 1993). These studies examined how ESOL students comprehend, learn, and use idioms, and helped identify some of the special difficulties ESOL students encounter in learning idioms and the distinctive processes they employ in such learning. However, none of these studies has looked at the important question of which idioms ESOL students should learn first, a question whose answers may lie, in part, in the study of the frequency and patterns of use of English idioms. Fortunately, this latter issue has gained some attention in applied linguistics, thanks largely to advances in corpus linguistics.

A few extensive, corpus-based studies (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Francis, Hunston, & Manning, 1996, 1998; Hunston & Francis, 2000; Moon, 1998) have examined partially or exclusively idiom use in English. Based on a thorough analysis of the Longman Spoken and Written English Corpus, which includes over 40 million words, Biber et al.'s work is arguably the most comprehensive single-book corpus study so far on English grammar and usage. It contains small sections on idioms and phrasal verbs and offers, among other things, a brief discussion and short list of the most frequently used idioms. Their analysis shows that idiom use is register sensitive and more common in fiction and conversation. Furthermore, they find pure idioms to be rare in general, fewer than one per million words. Yet because their work is a comprehensive study of grammar, its coverage of idioms is rather limited, and it offers only rather selective information on idiom use. Francis et al.'s (1996, 1998) Cobuild pattern grammars have also uncovered many interesting idiomatic usage patterns, but because they are grammar references, the focus of their discussions is not idioms per se. The same is true of Hunston and Francis's (2000) theoretical treatise of pattern grammar.

Unlike the above studies, Moon's (1998) is devoted exclusively to the use of idioms and fixed expressions in English. Using primarily the Oxford Hector Pilot Corpus, with 18 million words, Moon systematically

and thoroughly analyzed various important aspects of these distinct English expressions, including the definition, frequency, grammatical structure, variation, meaning, and discursual functions of idioms. In addition to finding that pure idioms are very rare across the board, Moon (1998) found that, although idioms of “situational formulae and conventions feature more strongly in spoken discourse” (p. 72), pure idioms are more likely to appear in written discourse. Moon also found surprisingly significant variations in the forms of idioms: “Fixedness is a key property of FEI [fixed expressions and idioms], yet around 40% of database FEIs have lexical variations or strongly institutionalized transformations, and around 14% have two or more variations on their canonical forms” (p. 120). Some of the variations, especially grammatical or structure-dependent variations, are very systematic, whereas others, especially those that are register dependent, are less so. Moon also discusses in detail the different forms of variation in both the systematic and the less systematic categories, such as verb variation (e.g., *up* or *raise the ante*), particle variation (e.g., *by* or *in leaps and bounds*), and truncation (e.g., *a bird in hand* without the rest of the phrase *is worth two in the bush*).

Applying Idiom Research to Teaching

The idiom studies described above focused on general issues regarding idiom use in English, primarily in written British English. They did not investigate the issues of principal concern for English language teaching, that is, the most frequently used idioms in spoken American English and idiom use patterns.

An important reason for developing corpus-based idiom lists is that, based on my research, including brief informal interviews with the authors of some of the existing idiom teaching and reference materials, the idioms in these publications were selected based primarily on the authors’ intuition rather than empirical data. Intuition alone is particularly problematic for identifying idioms because some idioms are regional; even when one’s intuition is correct, the selected idioms may be specific to one region. As a result, these teaching materials and references may include many seldom-used idioms, on the one hand, but leave out some frequently used ones, on the other. For example, some low- and intermediate-level books contain such rarely used idioms as *cop out* and *be on cloud nine* but exclude such frequently used idioms as *come up (with)* and *as of*.

Corpus-based research appears to be a good way to address this issue because, as Biber, Conrad, and Reppen (1994) point out, corpus linguistic analyses “are based on naturally-occurring structures and patterns of [language] use rather than intuitions and perceptions, which

often do not accurately represent actual use" (pp. 169–170). The use of naturally occurring language data is especially helpful and productive in examining use frequencies of language structures and lexical items. The findings of Biber et al. (1999), Coxhead (2000), Francis et al. (1996, 1998), and Moon (1998) provide helpful support in this regard.

Despite their usefulness for teaching, results obtained from corpus-based research cannot be considered the only relevant source of information on what to teach. Although such frequency studies may offer such valuable information as the most accurate count of the use of linguistic items, L2 professionals cannot ignore the importance of teaching—even to low-level students—some of the items that fail to make the list because pure frequency often leaves out some important and useful items in lexical lists. Moreover, interpreters of the results of corpus students should determine whether the corpus employed is representative of the type of language that is relevant for its purpose (Biber, 1993; Coxhead, 2000; Kennedy, 1998; Moon, 1998; Sinclair, 1991). Generally speaking, a corpus needs to contain millions of running words (tokens) to ensure that it has enough data to be sufficiently representative (Sinclair, 1991),¹ but a balanced selection of types and lengths of texts (either spoken or written) is equally important. Linguistic features of texts vary significantly from one register to another (Biber, 1989, Biber et al., 1994, 1998); thus, selecting the register(s) appropriate to one's research interest is crucial (Coxhead, 2000; Simpson & Mendis, 2003). Furthermore, the size and structure of texts chosen must be typical of the register of the researcher's interest (Coxhead, 2000; Sinclair, 1991). A representative corpus should also include as many different texts and as many authors or speakers as possible to avoid data distortion caused by a few individuals' personal styles.

The research reported here sought results that could inform English language teaching, with emphasis on the spoken language, by identifying the most frequently occurring idioms across three large corpora samplings from spoken American English in a variety of situations. Having identified these idioms, I related their frequency, association to registers, variations from the canonical forms, and tense (of idioms that function as verbs) to the findings of previous idiom studies.

¹ *Running words (tokens)* refers to the total number of word forms in a text or corpus; *individual words (types)* refers to each different word in a text regardless of how many times it occurs.

METHOD

The Corpora

In view of my focus on spoken English, I used three corpora containing transcribed spoken language (see Table 2): (a) Barlow’s (2000) Corpus of Spoken, Professional American English (CSPAЕ; hereafter *Professional*); (b) a corpus of spoken American media English (Liu, 2002, compiled with the help of graduate assistants; hereafter *Media*); and (c) Simpson et al.’s (2002) Michigan Corpus of Academic Spoken English (hereafter *MICASE*). The corpora in combination contain about 6 million tokens and 72,402 types and constitute, to my knowledge, the largest available spoken American English corpus to date. I also attempted to include a large number of diverse texts (1,111) and speakers (approximately 4,300) to help ensure the representativeness of the corpus.

The Professional corpus consists of transcripts of discussions at the meetings of various academic institutions and professional organizations and White House press briefings. The Media corpus includes transcripts of spontaneous talk from a variety of TV programs downloaded from the Web sites of the major U.S. networks: ABC, CBS, CNN, Fox News, and NBC. In compiling this corpus, I followed the corpus design principles discussed above and attempted to include as many different TV programs and topics as possible. The corpus contains such diverse TV programs as news reports, debates, interviews, magazine shows, and talk shows, including ABC’s *Nightline* and *20/20*, CNN’s *Larry King Live* and *Your Health*, Fox News’s *Rita Cosby Show*, and NBC’s *Dateline* and *Today*. The last corpus, MICASE, is made up of transcripts of a variety of spoken academic texts, including lectures, advising sessions, office hours, class discussions, and colloquia.

All three corpora are made up of contemporary, everyday, semiformal

TABLE 2
Summary of the Corpora

Corpus	Tokens	Types	Texts	Speakers	Text or transcript types
Professional	2,000,000	25,658	302	400	meetings/news briefings
Media	2,100,000	46,234	657	2,350	variety of TV programs
MICASE	1,848,364	37,975	152	1,571	variety of school functions
Total	6,000,000	72,402 ^a	1,111	4,321	

Note. Some figures are approximate. ^aTotal is not the sum of the number of types in each of the three corpora as some of the types are found in more than one corpus.

spoken American English (not casual or very formal speech; for a sample spoken text, see Appendix A), an important characteristic given that idioms are one of the most time-sensitive aspects of language. I limited my study to spoken American English because idiom use, like other aspects of language, has shown to be language-variety and register sensitive (Biber et al., 1999; Moon, 1998). Idioms common in spoken language may not be so in writing and vice versa. As my resources were limited, I believed that a study with a narrow focus would be more feasible and purposeful, hence maybe more meaningful and productive. The data in the corpora are primarily the type of spoken language students learning American English as an L2 will most likely be exposed to.

The three corpora differ somewhat in the formality of the speech they contain. A comparative analysis of the vocabulary in the three corpora using Heatley, Nation, and Coxhead's (2002) Range and Frequency Programs suggests that MICASE is the most formal of the three in vocabulary use: It contained the highest percentage of tokens found in Coxhead's (2000) Academic Word List (7.2%), followed by the Professional (4.9%) and Media (3.2%) lists. The results are consistent with expectations because MICASE is composed of academic speech events such as lectures and colloquia, and the Professional corpus consists of speeches at professional meetings and White House press conferences. In contrast, the Media corpus involves speakers with diverse social and educational backgrounds.

Idiom Identification

I identified idioms using Fernando's three categories (pure, semiliteral, and literal), as discussed earlier. I also included phrasal verbs as idioms because many of them are fixed in structure and nonliteral or semiliteral in meaning (e.g., *fall through*, *give in*, *put up with*). More importantly, these idiomatic expressions often present great difficulty to ESOL students. However, I excluded verb-plus-particle or verb-plus-preposition structures that most grammarians would not consider phrasal verbs.

To determine whether a verb-plus-particle structure was a phrasal verb or not, I used criteria agreed upon by many linguists: (a) whether an adverb may be inserted between the verb and the particle (phrasal verbs do not allow such insertion), (b) whether the particle can be fronted in a sentence (phrasal verbs do not allow such fronting), and (c) whether the meaning is completely literal (phrasal verbs are often not completely literal in meaning) (Celce-Murcia & Larsen-Freeman, 1999). The application of all these testing principles excludes as phrasal verbs those extreme literal verb phrases that often contain a directional

particle, such as *come in*, *go out*, *listen to*, *look at*, and *talk about*. It will, however, include most other phrasal verbs, such as *come across*, *pass out*, and *fall apart*.

I identified idioms in four major contemporary English idiom dictionaries and three English phrasal verb dictionaries: *Cambridge International Dictionary of Idioms* (1998) and *Cambridge International Dictionary of Phrasal Verbs* (1997), *Longman American Idioms Dictionary* (1999; no matching Longman phrasal verb dictionary was available), *NTC's American Idioms Dictionary* (Spears, 1994) and *NTC's Dictionary of Phrasal Verbs and Other Idiomatic Verbal Phrases* (Spears, 1993), and *Oxford Idioms Dictionary for Learners of English* (2001) and *Oxford Phrasal Verbs Dictionary for Learners of English* (2001). I selected these dictionaries because they were all rather recent publications from major ESOL publishers and claimed to be comprehensive and contain representative idioms and phrasal verbs.²

An important criterion in identifying an idiom was how difficult the phrase might be for ESOL students, which often hinges on how literal it is. To help reduce subjectivity in determining the difficulty of an idiom, I considered a fairly literal expression to be an idiom if it was listed in two of the four idiom dictionaries or two of the three phrasal verb dictionaries. In total, the idioms identified numbered 9,683.

Concordance Search

I used the concordance computer program MonoConc Pro 2.0 (2000) to search the Professional and Media corpora separately for the 9,683 idioms identified. I searched the MICASE using the search tool provided on the MICASE Web site. I compared and then combined the results of the three corpora to develop the idiom lists and uncover use patterns. I considered the various forms of an idiom as one idiom (e.g., *bring someone up to date/speed*, *in/with respect to*), but searched for the forms one at a time. For example, to search the frequency of the idiom *to bring someone up to date/speed*, I entered the following four separate entries: *bring* * up to date*, *brought * up to date*, *bring* * up to speed*, and *brought * up to speed*. Thus the total number of items searched for would have been much higher if I had counted each form of an idiom separately.

While searching for the most frequently used idioms, I also looked for noticeable usage patterns, especially those that were either not covered or erroneously presented in existing idiom teaching and reference

²None of these references states explicitly the criteria for selecting items, although the publishers of two (Cambridge University Press and Longman) state that a corpus was used in the selection of usage examples. Neither appears to have used a corpus frequency count for idiom selection.

materials. Because the results generated by the concordance search included some expressions that did not exemplify the idiom use I had intended, I often read the results one by one. For example, the search for *kind of* or *sort of* as an idiom expressing *somewhat* or *in a way* also yielded examples of its use as a noun phrase with *of* as a preposition, such as *this kind/sort of book*. Similarly, searching for the idiom *go after*, meaning *pursue someone in order to catch him or her*, also generated examples of the literal meaning *move after someone in sequence*. The analysis of the features and patterns of idiom use in general also demanded a close reading. Finally, because the MICASE online search tool allowed neither Boolean searches with *or* nor the use of truncated wildcard characters (*), the search of this corpus was much more laborious than expected.

FINDINGS

My search resulted in four lists of most frequently used idioms, one for each of the three corpora and one for the combined corpora. In addition, I made observations about the frequency of the idioms relative to the total number of words searched, their association to registers, variations from canonical forms, and the tense of idioms that function as verbs.

Most Frequently Used Idioms

I tabulated four separate lists of the most frequently used English idioms found in the concordance search: one based on the entire data set (see Appendix B) and the other three based on one of the three corpora (see Appendix C). Besides meeting the criteria outlined above, each selected item (following Coxhead, 2000, on frequency and range) occurred at least 12 times in all three corpora combined (i.e., two tokens per million words).

Setting a frequency level of two tokens per million meant that the idioms belonged, at least, to what Moon (1998) classifies as the lowest band of the medium-frequency idioms. I excluded any item that fell into Moon's (1998) two lowest frequency categories: low frequencies (less than one to two tokens per million) and insignificant frequencies (zero to four tokens in the entire corpus).

In terms of range, the four lists included only the 302 items that were listed in at least two of the major idiom dictionaries used to guide the concordance search and occurred in at least two of the three corpora so as to reduce the possibility of inflated results by one speaker, text type, or

topic. Excluded from the lists were 13 items that met the frequency criteria but failed the range test. I classified the 302 idioms into three frequency-of-use bands representing 50 or more, 20–49, and 2–19 tokens per million words (see Table 3 for a comparison of these bands with Moon’s, 1998). These classifications are rather arbitrary and are intended merely as a reference, not a guide, for ESOL teachers and learners to consider in selecting idioms for study.

All the idioms in the three corpus-specific lists in Appendix C also occur in the overall list. To reduce the possibility that the idiom use of individual speakers or texts might inflate the results, I did not include in the sublists idioms that did not meet the criteria for inclusion on the overall list. A comparative analysis of the four lists shows a rather strong convergence in the idiom selection. Of the 302 idioms in the overall list, 283 appeared in all three sublists. Of the 19 that did not, 7 failed to make the Professional list and 12 the MICASE list. All 302 idioms in the overall list occurred in the Media list, suggesting that the Media corpus is the most balanced of the three, which, I believe, is due to the large number of speakers and the broad range of topics it involves.

Frequency

The results of this study support previous findings that pure idioms are rare. Moon (1998, p. 64), for example, found that few such idioms occur with a frequency greater than one per million words. Similarly, Biber et al. (1999, p. 1025) detected a few with a frequency greater than five per million words. I also found few with a frequency of more than two per million words. This finding is most evident in the fact that only a few such idioms made the compiled list, such as *ballpark estimate*, *the ball is in your court*, and *right off the bat*—mostly sports-related idioms. Even these few are all in Band 3, the lowest band, with a frequency of 2–19 per million words.

TABLE 3
Three Bands of the Most Frequently Used Idioms in the Corpora

Band	Number of items	Frequency (tokens per million words)	Comparable frequency band in Moon (1998)
1	47	50 or more	VIII (high)
2	107	11–49	VII (high medium)
3	148	2–19	V and VI (medium)

Variations in Register

The findings also support previous findings that idioms are register sensitive (Biber et al., 1999; Moon, 1998). For example, the multiple-meaning phrasal verb *to come up* (meaning also *to appear* or *be mentioned*) is used far more frequently than *come up with* in the Media corpus (398:159 tokens, 255%) and more in the Professional corpus (392:242, 160%), but the opposite is true in MICASE (153:222, 69%). A plausible explanation is that college students are often asked to “come up with” answers and solutions to various problems presented by their textbooks or by their professors and peers in class.

Another example of the register sensitivity of idioms is the informal idiom *to hang out (with)*, with 8 tokens in the Professional corpus but 26 and 40 respectively in the Media corpus and MICASE. What is more, almost 70% (27:40) of the tokens in MICASE appeared in the 4 least formal (of 16) genres—Tour, Study Groups, Labs, and Office Hours—but none occurred in the 7 more formal categories, such as Presentations, Interviews, and Colloquia.

Variations in Form

An analysis of the idioms identified in spoken American English also corroborates Moon’s findings (1998) about how idioms vary in form. Context-dependent variations are either what Moon (1998) terms “truncation” (p. 131), the shortening of an idiom, or the conversion of the parts of speech of the idiom. Based on my data, the context of an idiom often makes its truncation possible. An example relates to the idiom *to bring someone up to date* (or *up to speed*) *on something*, meaning to update someone about something. At a U.S. White House press briefing, a spokesperson made the following remark:

So this is to bring him [President Clinton] up to speed on all the details of the schedule. Certainly he’s got a number of multilateral meetings to prepare him for the substance of those meetings so he can have thoughtful, meaningful conversations just to continue to *bring him up* [italics added]. (Professional Corpus, Barlow, 2000)

The truncation of the second *to bring him up to speed* to *to bring him up* will not be interpreted as *to rear him* because the context makes such an interpretation impossible.

An example of the parts-of-speech-conversion type of context-dependent variation is found in the following question in the Professional corpus: “How long would it take, *ballpark* [italics added], for it to work its

way up through to the Supreme Court before we get an answer?” (Barlow, 2000). The word *ballpark* is listed in all dictionaries as either a noun or an adjective in its idiomatic use, but the listener cannot miss its meaning when used as an adverb in this sentence because of the discourse context.

Meaning-related variants can be divided into three subcategories: *meaning dependent*, *meaning extension*, and *hyponym exchange*. The alternate use of *in the long run* and *in the long term* and the counterpart pair, *in the short run* and *in the short term*, provides an example of meaning-dependent variation. The corpus analysis shows that U.S. speakers strongly prefer *in the long run* over *in the long term* (35 vs. 12 tokens, respectively, a ratio of approximately 3:1), but they tend to prefer *in the short term* over *in the short run* (12 vs. 6 tokens, respectively, a ratio of 2:1). Such a variation seems to suggest that to most U.S. speakers, a *run* appears longer than a *term*.

An example of meaning extension is the use of the expression *ballpark idea*: “you can start out with these few observations of conductivity to give you a *ballpark idea* of what, what it, what it could be at the site and how it might vary” (MICASE, Simpson et al., 2002). *Ballpark idea* here means general idea, obviously derived from the idiom *ballpark figures*. In substituting the noun *idea* for *figure* or *number*, the speaker maps the adjective meaning of *ballpark* onto a new conceptual category, *ideas*. Other examples of extension are the remark “Put the ball back in Barak’s court” (Media corpus), which a political commentator used in depicting the battle between the former Israeli leader, Ehud Barak, and the Palestinian leader, Yasir Arafat, and in the question a correspondent asked Dee Dee Myers, White House Press Secretary under President Clinton: “On health care reform it seems like the administration now is being very passive and just *putting the ball in the Senate court*” [italics added] (Professional corpus, Barlow, 2000). The expression *put the ball in someone’s court* has apparently evolved from the idiom *the ball is in someone’s court*. In terms of structure, the variation represents what Moon (1998) calls verb “variation” (p. 124), where an idiom’s verb is replaced by another verb. In meaning, the new form has changed from a passive descriptive state—the ball simply being in someone’s court—to an action that moves the ball from one location into a new, desired position.

For hyponym exchanges, an example is the expression “slip of the lip” (Media corpus), in which the original noun *tongue* has been replaced by another speech organ hyponym, *lip*. This type of variation differs from that in which a word in the idiom is replaced by another word that is the same part of speech but is not a hyponym, as in “pull a highway [for *rabbit*] out of the hat” or “join the Proposition 36 bandwagon” for “jump on the bandwagon” (both from the Media corpus).

Tense of Verbal Idioms

The results suggest that certain phrasal verbs appear predominantly in the present tense, whereas others feature substantially more in the past tense. For example, the idiom *go ahead* is almost always used in the present tense in the data set (635 of 645 tokens, or 98%), as is the idiom *make sure* (1,159 of 1,179 tokens). Yet the verbal idiom *leave out someone or something* registers more past tense than present tense uses (51 vs. 34 tokens). Similarly, the phrasal verbs *work out* and *turn out* register a considerable past tense distribution (about one third of their total use). Such information may help ESOL teachers make more informed decisions about when to teach certain idioms. For instance, *make sure* and *go ahead* may be a good focus of instruction during the introduction of the present tense, whereas *leave out* and *work out* may be best practiced during the instruction of the past tense. On the other hand, the use pattern of *in the long/short run/term* mentioned above may also help students learn to use the phrase more idiomatically.

COMPARISON OF FINDINGS WITH TEACHING AND REFERENCE MATERIALS

As one of the goals of this study is to help improve the development of future idiom teaching and reference materials, I compared the findings of my search with the information in the seven dictionaries I used to guide my concordance search and two essential idiom publications: Dixon's (1994) *Essential Idioms in English* and Spears's (1999) *Essential American Idioms*. None of these nine publications is limited to spoken American English or to British or written idioms. Three of the dictionaries (those with *American idioms* in their titles: *Longman American Idioms Dictionary*, 1999; Spears, 1994, 1999) are devoted to American idioms only, and all nine publications contain idioms from both spoken and written English, ensuring that the comparison is basically fair. The comparative analysis unearths some inadequacies in the idiom selection and meaning and usage explanations in these materials.

Inconsistent Selection

The selection of idioms in these materials seems sometimes inconsistent, with some highly frequent or highly opaque items missing but much less frequent and more transparent ones included. For example, one of the essential idiom dictionaries contains the entries *above par*, *all*

over again, and *all right* but not *according to*.³ Based on my findings, *according to* should be included. In terms of frequency, *according to* occurs far more frequently in my data set than *above par* (533:0). In terms of difficulty level, *according to* appears to be equal to *all over again* and *all right*; all three idioms are rather literal in meaning.

Another example of inconsistent selection is that of the five major idiom dictionaries surveyed (excluding the three phrasal verb dictionaries), only one lists the idiomatic phrase *as of*. The four that omit this phrase include idioms that are either much less common or much more transparent in meaning. Under the heading *as*, one of the four dictionaries lists only *as is*, yet *as is* claims only 16 tokens to the 90 tokens of *as of* in my corpora. Furthermore, *as of* is more opaque in meaning than *as is*. Another dictionary lists under such idioms as *as nutty as a fruitcake* and *as one*. Similar problems are found in the remaining two dictionaries. One records idioms such as *as a duck takes to water* (zero occurrence in the corpora) and *as usual* (very transparent in meaning), and the other lists under *as* only the two idioms *as if* (rather literal despite its high frequency) and *as per usual* (extremely rare with zero tokens in the corpora).

One more example of inconsistency in item inclusion is that only one of the five dictionaries includes the highly frequent (among the top band in my list) idiomatic phrase *with/in regard/respect to*. The fairly transparent meaning of this idiom might be the reason for its exclusion, but many of the listed items, such as *with each passing day* and *with a will*, are even more literal. A final example is the inclusion in one publication of such extremely rare idioms as *cop out* and *cut and dried*.

Inadequate Meaning and Usage Explanations

A more important disparity between the publications and the results from this study is that the primary meaning and typical use of an idiom introduced in these publications are not those found in the corpora. A case in point is the verbal idiom *to bring up*. All the dictionaries that list this phrase give *to rear or educate a person (often a child)* as the primary meaning and present the other meaning—*to mention and start discussion of an issue*—as the second or third entry, or as a subcategory within an entry. For instance, in one dictionary, the second entry for *bring something up* gives two definitions: (a) “to move something up from a lower to a higher position” and (b) “to mention a subject and start to talk about it.”

³ Because it is not the goal of this study to criticize the references, I purposely do not give the titles of the publications referred to in this discussion.

The concordance analysis indicates that this second meaning constitutes more than 90% (287 of 316 tokens) of the uses of the phrase, whereas the meaning *to rear or educate a person* accounts for only about 5% (16 tokens), with the rest of the phrase's meanings, such as *moving something up*, constituting the remaining 5%.

Another example is the phrase *as of*, which, as mentioned earlier, is listed in only one of the dictionaries. This dictionary gives the phrase's meaning and use as "used to indicate the time or date from which something starts: *We shall be at our new address as of mid June*," suggesting that the phrase is used with a future event. Yet 43 (47.8%) of the tokens of this idiom in the corpora are in the past tense (e.g., *as of yesterday*), 45 (50%) are in the present or present perfect tense (e.g., *as of now, today or yet*), and only 2 (2.2%) are in the future tense. Such data demonstrate that the phrase is seldom used in describing future events and is mostly (98%) used with past or ongoing events up to the present. Another example is the idiomatic multimeaning phrase *in place*. Of the two dictionaries that list it as an entry, one does not mention the most frequently used meaning according to the corpora, *prepared/ready or in existence*, giving only the rather literal meaning, "things being in the proper place," and the rarely used figurative meaning, "something being proper and well done."

One more discrepancy between my findings and the materials is that, as discussed earlier, the materials typically do not introduce the variations of many of the idioms. Even when the variations are included, their frequency is not mentioned. Yet often one of the variations is the dominant form. For example, three dictionaries list both *with regard to* and *in regard to* but give no information on their frequency of use. According to the corpora, *with regard to* (92 tokens) occurs almost seven times as frequently as *in regard to* (14 tokens). The difference between *with respect to* and *in respect to* is even more striking, with a ratio of 382:3. A related point is that *with respect to* and *with regard to* are synonymous, yet the number of tokens of the former (382) is more than four times that of the latter (92). Having such information in teaching and reference materials would be helpful to ESOL teachers and students.

CONCLUSION

This corpus study of idiom use has resulted in (a) the development of four lists of the most frequently used idioms in spoken American English, (b) a comparison of idiom use patterns in spoken American English with those found in previous research, and (c) the identification of discrepancies between the findings and the presentation of idioms in idiom teaching and reference materials. Based on these results, five areas

for improvement in the teaching of idioms for ESL learners can be identified.

First, items in the teaching and reference materials, especially in those so-called essential idiom publications, need to be selected in a more rigorous, systematic way, and should be based on authentic language rather than on intuition in order to increase their content representativeness. Similarly, rather than relying on made-up sentences for idiom use illustrations, writers of such materials should use examples from a corpus, as some publishers have reportedly done in some of their dictionaries. Second, these publications should strive to provide more accurate descriptions of the meanings and uses of their items. Third, the publications should include additional descriptive information, such as an idiom's distribution and use frequency, because such information may help students develop a more complete grasp of the idioms or decide to what extent they want to learn and use those idioms. Fourth, ESOL teachers, especially those of low-level students, might want to consult corpus-based lists of most frequently used idioms in selecting idioms to teach, particularly when more objective data on frequency become available. Such consultation may help decrease the chance of having students work on idioms not useful to them at the time of instruction. Of course, frequency and range of idioms should not be the only selection criterion in lexical instruction decisions, a point I elaborate on below. Fifth, teachers may want to include information on idiom variations in their instruction. This will help make students' learning of idioms more complete.

Although this study, and corpus research more generally, contributes to the understanding and selection of idioms for learning and teaching in ESOL or the acquisition of other L2s, it has some limitations. For example, an idiom textbook or dictionary whose item selection is based entirely on a frequency count from a corpus study may not include low-frequency idioms that could be important to some students, such as *call it a day*, an utterance that some instructors use to signal the end of a class or meeting, or *out of the question* (both with fewer than one token in a million words in the corpora). Not understanding the first may result in a student's failure to perform a routine yet important speech act; failing to understand the second may have serious consequences (e.g., mistaking the meaning as *no question* or *no problem*, the opposite of its actual meaning). Compilers of a textbook or reference that is useful for a specific group of learners thus may have to resort to additional methods to look for item candidates. What may further limit the value of a corpus-based study is the difficulty of finding a corpus that is truly representative of the language use that is the focus of an investigation. Concerning the present study, as stated earlier, the corpora employed may not have been large enough and the criteria for identifying idioms may not have been

rigorous enough. The results of the study therefore need to be interpreted cautiously.

Validating the findings of the present study will require studies of substantially larger corpora of spoken American English when such corpora become available. Investigations of written corpora would enable some meaningful comparative analyses between the written and spoken registers. In the meantime, however, the results demonstrate the particular advantages of corpus research in revealing valuable information about American English idiom use that might not be unearthed otherwise. The findings of the study also indicate the need in TESOL to develop more informed and, it is hoped, more effective idiom teaching and reference materials.

ACKNOWLEDGMENTS

This research was partially funded by the Faculty Development Grant of the Petree College of Arts and Sciences of Oklahoma City University. I thank Jill Coley, Cody Lowry, David Moody, and Patrick Wright for assistance in compiling the Media spoken English corpus used in the study; four anonymous reviewers and the *TESOL Quarterly* editor for their valuable comments and suggestions; and Michael Barlow of Athelstan and R. C. Simpson, S. L. Briggs, J. Ovens, and J. M. Swales of the English Language Institute of the University of Michigan for making their respective corpora available.

THE AUTHOR

Dilin Liu is professor and director of TESOL at Oklahoma City University. He has published in the areas of language acquisition, English grammar and usage, sociolinguistics, and pedagogy.

REFERENCES

- Abdullah, K., & Jackson, H. (1998). Idioms and the language learner: Contrasting English and Syrian Arabic. *Language in Contrast*, 1, 84–107.
- Barlow, M. (2000). Corpus of spoken, professional American English [CD-ROM]. Houston, TX: Athelstan.
- Biber, D. (1989). A typology of English texts. *Linguistics*, 27, 3–43.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8, 243–257.
- Biber, D., Conrad, S., & Reppen, R. (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics*, 15, 169–189.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, England: Longman.
- Cambridge international dictionary of idioms*. (1998). Cambridge: Cambridge University Press.

- Cambridge international dictionary of phrasal verbs*. (1997). Cambridge: Cambridge University Press.
- Cacciari, C. (1993). The place of idioms in a literal and metaphorical world. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 27–55). Hillsdale, NJ: Erlbaum.
- Cacciari, C., & Levorato, M. C. (1989). How children understand idioms in discourse. *Journal of Child Language*, 16, 387–405.
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27, 668–683.
- Cacciari, C., & Tabossi, P. (Eds.). (1993). *Idioms: Processing, structure, and interpretation*. Hillsdale, NJ: Erlbaum.
- Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). Boston: Heinle & Heinle.
- Cooper, T. C. (1998). Teaching idioms. *Foreign Language Annals*, 31, 255–266.
- Cooper, T. C. (1999). Processing of idioms by L2 learners of English. *TESOL Quarterly*, 33, 233–262.
- Cowie, A. P. (Ed.). (1998). *Phraseology: Theory, analysis, and applications*. Oxford: Oxford University Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213–238.
- Dixon, R. J. (1994). *Essential idioms in English: New Edition*. White Plains, NY: Pearson Education.
- Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Fernando, C. (1996). *Idioms and idiomaticity*. Oxford: Oxford University Press.
- Francis, G., Hunston, S., & Manning, E. (1996). *Collins Cobuild grammar 1: Verbs*. London: HarperCollins.
- Francis, G., Hunston, S., & Manning, E. (1998). *Collins Cobuild grammar II: Nouns*. London: HarperCollins.
- Gibbs, R. W., Jr. (1986). Skating on thin ice: Literal meaning and understanding of idioms in conversation. *Discourse Processes*, 9, 17–30.
- Gibbs, R. W., Jr. (1987). Linguistic factors in children's understanding of idioms. *Journal of Child Language*, 14, 569–86.
- Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). *Range and frequency programs* [Software program]. Available from Victoria University, School of Linguistics and Applied Language Studies Web site, http://www.vuw.ac.nz/lals/staff/Paul_Nation
- Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Irujo, S. (1986a). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly*, 20, 287–304.
- Irujo, S. (1986b). A piece of cake: Learning and teaching idioms. *English Language Teaching*, 40, 236–242.
- Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. *IRAL*, 31, 205–219.
- Katz, J. J., & Postal, P. M. (1963). Semantic interpretations of idioms and sentences containing them. *MIT Research Laboratory of Electronics Quarterly Progress Report*, 70, 275–282.
- Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Addison Wesley Longman.
- Levorato, M. C. (1993). The acquisition of idioms and the development of figurative competence. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 101–128). Hillsdale, NJ: Erlbaum.
- Levorato, M. C., & Cacciari, C. (1995). The effects of different tasks on the

- comprehension and production of idioms in children. *Journal of Experimental Child Psychology*, 60, 261–283.
- Longman American idioms dictionary*. (1999). Harlow, England: Pearson Education.
- MonoConc Pro (Version 2.0) [Computer software]. (2000). Houston, TX: Athelstan.
- Moon, R. (1998). *Fixed expressions and idioms in English*. Oxford: Clarendon Press.
- Oxford idioms dictionary for learners of English*. (2001). Oxford: Oxford University Press.
- Oxford phrasal verbs dictionary for learners of English*. (2001). Oxford: Oxford University Press.
- Simpson, R. C., Briggs, S. L., Ovens, J., & Swales, J. M. (2002). Michigan corpus of academic spoken English [Data file]. Available from University of Michigan Web site, <http://www.hti.umich.edu/m/micase>
- Simpson, R., & Mendis, D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly*, 37, 419–442.
- Sinclair, J. (1991). *Corpus, concordance, and collocation*. Oxford: Oxford University Press.
- Spears, R. A. (1993). *NTC's dictionary of phrasal verbs and other idiomatic verbal phrases*. Lincolnwood, IL: NTC.
- Spears, R. A. (1994). *NTC's American idioms dictionary* (2nd ed.). Lincolnwood, IL: NTC.
- Spears, R. A. (1999). *Essential American idioms* (2nd ed.). Lincolnwood, IL: NTC.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.
- Tabossi, P., & Zardon, F. (1993). The activation of idiomatic meaning in spoken language comprehension. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 27–55). Hillsdale, NJ: Erlbaum.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

APPENDIX A

Excerpt From the Corpora

White House Press Briefing by Dee Dee Myers, June 23, 1994, from Barlow's (2000) Corpus of Spoken Professional American English

MYERS: Following up on a couple of things from this morning—first of all, President Clinton will meet with President Carlos Menem of Argentina at the White House on Friday, June 24th, to discuss a broad range of bilateral and international issues *with a view toward* continuing the close collaboration between the United States and Argentina. The two Presidents last met at the White House in June of '93. There was also great interest in the menu at tonight's Congressional picnic. It is hard-shell crabs.

....

VOICE: We have a new poll out today that shows that 41 percent only of Americans approve of the way the President is handling health care reform; 50 percent of them disapprove of it. What do you think is the problem?

MYERS: Well, I think if you ask people specifically do they support universal coverage, overwhelmingly they do; if they support an employer-based system, the system that we currently have, overwhelmingly they do; if they think the Congress should act now to produce some kind of comprehensive health care reform, overwhelmingly they support that. So I think they

Excerpt used with permission. For samples from the other two corpora, see the respective MICASE (Simpson et al., 2002) and TV network Web sites.

certainly support elements of the President's plan. The President is going to continue to fight very hard to get universal health care passed this year.

VOICE: Well, why do you think they don't seem to have confidence in the way he is leading the fight on this?

MYERS: I think millions and millions of dollars have been spent providing misinformation about the President's plan. That's unfortunate. The President, the First Lady, members of the Cabinet and others in the administration have worked hard to try to counter that, to *put out* correct information, to work with Congress to pass elements of the President's package. I think *things are moving in the right direction* in Congress. We're encouraged by discussions that are happening. *We're a long way from* getting it done, but the President remains confident that he will get a bill on his desk this year that will provide guaranteed private insurance for every American.

VOICE: Senator Dole accused the President of *throwing in the towel* on North Korea. And Senator McCain today said that the President would become a co-conspirator with Kim Il Sung if they continue to reprocess. What is your reaction to those Republican criticisms?

MYERS: Well, I think it's unfortunate. The message we got from North Korea yesterday was encouraging. As the President said, it was a bit of good news. The North Koreans agreed to freeze their nuclear program during a third round of talks. There will be nuclear inspectors there on the ground in North Korea to assure that they meet those commitments. We're moving forward now in planning for a third round of high level discussions with the North Korea, *with an eye toward* resolving the outstanding nuclear issues between the United States, the rest of the international community, and North Korea. Certainly, we're going to *make sure* that we verify along the way. We *look forward to* that dialogue. We hope that it happens.

VOICE: Dee Dee, Senator Dole and the other Senate Republicans in the Senate sent a letter to the President today asking him to join them in repudiating remarks by some in his own party that they say used terms like "fire-breathing," "Christian radical right," and they cheapen our democracy through religious bigotry. Is he going to join them?

MYERS: I'm unaware of the letter. I haven't seen it yet. I mean, certainly the President supports very strongly the principle of religious freedom. But beyond that I haven't seen the letter, so I can't comment.

VOICE: Dee Dee, what do you make of *The New York Times* report that the Russians have been secretly working on chemical weapons? Chernomyrdin didn't know anything about it, but—

MYERS: We're *following up on* that. We signed a memorandum of understanding on this. They were—in which they agreed to provide us with information. We don't believe they provided all the information that they *set forward*—that we need for that MOU. And the memorandum also provides for us to ask for additional information should we think what we receive is insufficient. We're in the process of discussing that with them now.

VOICE: Do we think they're hiding something?

MYERS: We're in the process—we want more information, and so we're *following up*. We're not going to draw any conclusions until we have all the information we think we need.

VOICE: Dee Dee, the crime bill seems to be stalled at the moment, or at least it was when I left the office. Do you guys have anything to say about that?

MYERS: As you know, the President met earlier this week with House and Senate leaders to discuss the crime bill. He's certainly been working very hard. I think 95 percent of that has been resolved between the Houses, which is substantial progress. There are *a few stumbling blocks*, which as I understand it are being worked out among the members of the Congress right now.

VOICE: Well, they say it's—

MYERS: Well, *it's not up to me, it's up to* the members of Congress to decide what the *stumbling blocks* are. I think certainly there are a couple of outstanding issues which they're addressing. I think the President hopes that they reach agreement soon and pass a crime bill and send it to his desk.

VOICE: Some of them suggested that it's *up to* him at this point to *come up with* a solution to the whole racial justice issue and remaining *stumbling blocks*, that it won't happen among them.

MYERS: He discussed it on Friday. I think they're aware of his position. We've certainly worked hard on this crime bill and we'll continue to do that. But I think the *ball is now with the members*

of Congress who are *working out* some of the final details. And I just don't have any more on it at this point.

VOICE: Their meeting yesterday fell apart because they couldn't reach a conclusion and said it won't happen without him.

MYERS: I think there are discussions *going on*, on the Hill today, and again the President will remain engaged in this and would like to see something passed.

VOICE: What is he doing today besides—he did some ambassadors fairly early this morning that wasn't on his schedule.

MYERS: It wasn't on the public schedule.

VOICE: Another question on the racial justice issue. What is the President's position on the racial justice issue?

MYERS: He hasn't taken one.

VOICE: Why doesn't he have a position on the racial justice issue?

MYERS: Throughout the discussion of this crime bill, he's laid out specifically what it was he wanted to see in that bill—100,000 new police officers on the street, things for people to say yes to *as well as* to say no to. *Three strikes and you're out* proposal with specific language. All of those things, all of the provisions that he *laid out* have been included in the bill. Those are the things he fought for, those are the things that have been included. There are other things that members of Congress have added and other things that they've *taken out*. He has not taken—he did not take a position on every item that *came up* throughout this debate.

....

APPENDIX B

Most Frequently Used Idioms Across Three Corpora of Spoken American English (in Order of Frequency)

Band 1

kind of (meaning *somewhat*)
sort of (meaning *somewhat*)
of course
in terms of
in fact
deal with
at all
as well
make sure
go through
come up
look for
find out
go on (*with* + gerund or present participle)
as well as
in a/some way
go ahead

in order to/that
get into
first of all
come up with
figure out
put on
in other words
end up (*with* + gerund or present participle/ adjective)
according to
as/so far as
in a/some sense (*of/that*)
so far
point out
by the way
take place (*of*)

pick up
make sense
turn out
as to (+ *wh*- clause/noun phrase)
set up
with respect to/in respect to
be/have (*something*) *in place*
used to (verb)
as long as/so long as
work out
have something/nothing to do
with
bring up
have/keep in mind
call for
in general

Band 2

take care of
go over
on the other hand
put out
go with
to the/some extent (of/that)
as soon as
put together
show up
get on
get rid of
as if
get out of
give up
no matter (+ any *wh*- clause)
look forward to
once again

in time
get back to
take on
go for
be about to
after all
follow up on
on behalf of
in effect
go along (with)
come on
be used to something/doing
something
get through
take out
go off
for sure/certain

call (up)on
make up (of)
back and forth
go after
carry out
make a difference
have/play a part/role (in)
take off
right away
take advantage of
run out (of)
in favor of
all of a sudden
(be) in charge (of)
break down
put up
take over

Band 3

be open to ideas
rule out
as for
fill in something (or someone
on something)
be up to somebody/something
(meaning depending on
somebody/something)
hold on
regardless of
account for
in advance
in public
with regard to
break up (transitive/
intransitive verb)
in case
in someone's view
take up
in someone's/the interest (of)
take steps
throw out
as of
run into
wind up with/in/gerund (+
participle)
stick (be stuck) with
go/move/be too far
look up something (in)
as a matter of fact
more or less
leave out something/someone
be/keep in touch with
down the road
turn on something
make it

do one's best
turn in something/somebody
all along
on time
turn around (something/
someone)
(be) better off
keep up (with)
come by (meaning to visit)
get away with
hang out (with)
put forward
take into account
in light of
go wrong
for someone's/the sake (of)
count on
get around (something)
set out
shut down (something)
hand out
live with (meaning accept/
exist with)
run through (something)
(meaning to go over)
touch on a topic/issue
in the long/short run/term
hold someone/something
accountable
pass out
stick to
in common
(be) under way
come across (meaning to
encounter)
in turn

up front (about)
at stake
by hand
to the/that/this effect (that/of)
first and foremost
in the/a/some fashion (that)
so to speak
come about (meaning to
happen)
come off
have/get a clue
bring about (meaning to make
happen)
hold up (meaning to endure
or stand testing)
in essence
chances are/were
cut down (meaning cut down
the cost)
(get) in the/someone's way
pay off
in the wake of
buy into
by and large
out of control
have (something be) on one's
mind
keep/be on track
make good on something
throw away something
fall apart
get to the point
in detail
on and off (or off and on)
come/go/bring into effect
(can't) get over something

bring out
 crack down (on)
 hold on to
 turn up
 by far
 to date
 cope with
 give someone a break
 shut up
 up to date
 at large
 in control
 follow through
 for that matter
 shed/cast light on something
 sign off
 take part (in)
 be/put on hold
 (something as) a fair game
 after the fact
 above all
 drop off
 up in the air
 all out
 come to mind
 in private
 in the (somebody's) eyes of
 in the wrong
 live up to
 as usual
 by virtue of
 the big picture
 on the verge of
 ballpark (e.g., figure, estimate)
 keep an eye on
 on the whole
 screw up (something)
 at the (somebody's) expense (of)
 be in for (meaning to
 experience)
 draw the line

get/grab hold of somebody/
 something
 be over one's head
 get/have a handle on
 something
 go for it
 in (good/bad) shape
 make fun of
 hold up (meaning to delay or
 hold as hostage)
 in line with
 in the fore of
 in keeping with
 (a) level playing field
 to the contrary
 at issue
 call something into question
 for good
 in good faith
 get/have hands on something
 off the top of my head
 put something to rest
 take its toll
 all over again
 make up one's mind
 the ball is in your/their court
 in the event of/that
 so on and so forth
 get something across
 in place of/in someone's place
 by no means
 have/get a say/voice (in
 something)
 give away
 leave something/someone alone
 on the horizon
 take effect
 the other way around
 to somebody's credit
 to the (or somebody's) best
 knowledge of
 at somebody's disposal

hang in there
 make up for
 put up with
 to somebody's advantage
 come across as (meaning to
 appear as)
 for the time being
 bring forward
 give rise to
 make out
 right off the bat
 sell out (meaning to
 compromise)
 something/things are up for
 grabs
 take issue with
 a rule of thumb
 bits and pieces
 do away with something
 err on the side of
 fall short
 for real
 in due course
 in no way
 in practice
 in the works
 quid pro quo
 take something/someone for
 granted
 in order (in need, get/put
 house in order)
 break off/break off (something)
 beg the question
 from scratch
 hit home
 in the pipeline
 make/catch/hit headlines
 once and for all
 push the envelope
 with (keep) one's eye on
 something

APPENDIX C

Most Frequently Used Idioms in Each of Three Corpora of Spoken American English (in Order of Frequency)

Professional Corpus

in terms of
deal with
sort of
kind of
in fact
make sure
go on
of course
as well
come up
go through
at all
as well as
look for
with/in respect to
get into
come up with
in other words
find out
first of all
in order to/that
take (the) place (of)
in a/some way
as/so far as
figure out
be (have something) in place
go ahead
put on
in a/any sense of/that
have (something/nothing) to do with
point out
work out
keep/have something in mind
as to (wh- clause/noun phrase)
so far
to the/some extent
look forward to
follow up on
in general
make sense
bring up
set up
on behalf of
put together
as soon as
on the other hand
by the way

go with
end up
call for
as (so) long as
pick up
get back to
put out
according to
get on
go over
turn out
go along (with)
on one's own
in effect
back and forth
in time
with/in regard to
call (up)on
show up
used to (+ verb)
be open to (ideas)
take on
fill in
rule out
in advance
get rid of
put forward
get out of
in favor of
as if
be used to (+ noun phrase/gerund/present participle)
play/have a role/part in
take out
in light of
take steps
be up to somebody/something
take advantage of
take care of
get through
down the road
carry out
go for
in the (or somebody's) interest (of)
leave out
make up
as of

put up
regardless
give up
in touch with
in someone's view
make a difference
no matter (wh- clause)
in charge of
once again
touch on
go off
all along
in case
for sure/certain
take into account
as a matter of fact
take up
first and foremost
up front
run into
in public
on time
set out
do one's best
break down
come on
wind up
as for
stick (be stuck) with
buy into
turn around
better off
get away with
hold someone accountable
account for
throw out
in the/a fashion (that)
keep up with/doing
more or less
live with (meaning to accept)
get around
pass out
run out of
after all
make it
a fair game
sign off
be about to
take off

go after
 run through (go over)
 in the wake of
 break up
 right away
 all of a sudden
 come/go into effect
 in the long/short run/term
 hand out
 by and large
 have/get a clue
 keep on track
 in turn
 take over
 ballpark (e.g., figure)
 the ball is in your court
 hold on
 in good faith
 after the fact
 to somebody's best knowledge
 on and off
 stick to
 bring about
 in essence
 for somebody's/the sake (of)
 at stake
 the big picture
 by virtue of
 so to speak
 in keeping with
 live up to
 draw the line
 to the contrary
 in line with
 off the top of my head
 follow through
 up to date
 hold up (meaning to delay)
 to date
 come across
 by hand
 hold up (to a test)
 in common
 in somebody's/the way (of)
 so on and so forth
 come about (happen)
 shed/cast light on
 in somebody's/the eyes (of)
 get/have a handle on
 something
 turn in
 under way
 in the fore of

put something to rest
 in due course
 bring forward
 err on the side of
 fall short
 turn on
 go wrong
 have somebody/something on
 one's mind
 in detail
 pay off
 fall apart
 go for it
 up in the air
 in the event of/that
 get something across
 take part in
 by no means
 hang out (with)
 in the wrong
 level playing field
 keep an eye on
 on the whole
 take effect
 bring out
 chances are
 crack down
 come to mind
 in private
 call something into question
 to somebody's credit
 on the verge of
 all out
 turn up
 by far
 get to the point
 on the horizon
 quid pro quo
 screw up
 come off
 come by (meaning to visit)
 cut down
 make good on something
 throw away
 above all
 in control
 at somebody's/the expense (of)
 make up for
 in the pipeline
 in practice
 as usual
 be in for (meaning to
 experience)

be over one's head
 get/have hands on
 cope with
 make up one's mind
 the other way around
 in order (meaning in
 sequence)
 push the envelope
 once and for all
 a rule of thumb
 for the time being
 (can't) get over something
 at large
 drop off
 look up something (in)
 gerund/ grab hold of
 at issue
 all over again
 in place of/in somebody's place
 have a say/voice in
 be/put on hold
 leave somebody/something alone
 do away with
 give rise to
 in no way
 from scratch
 take somebody/something for
 granted
 in the works
 in (good/bad) shape
 come across as (meaning to
 appear as)
 bits and pieces
 for good
 for real
 for that matter
 at somebody's disposal
 hang in there
 give somebody a break
 right off the bat
 put up with
 take issue with
 beg the question
 break off
 come across as (meaning to
 appear as)
 give away
 hold on to
 out of control
 shut up
 make fun of
 make/hit headlines
 take its toll

Media Corpus

kind of
 of course
 in fact
 deal with
 at all
 sort of
 as well
 come up
 find out
 look for
 go on
 go through
 according to
 go ahead
 make sure
 in a/some way
 in terms of
 as well as
 put on
 first of all
 by the way
 pick up
 so far
 call for
 point out
 end up
 get into
 take care of
 as (so) far as
 come up with
 turn out
 in order to/that
 as to (+ *wh*-clause/noun
 phrase)
 take place (of)
 used to (+ verb)
 as (so) long as
 give up
 in a/some sense
 set up
 once again
 get out of
 after all
 figure out
 show up
 be about to
 no matter (+ *wh*-clause)
 go after
 bring up
 in other words
 get on
 work out
 be/have something in place
 on the other hand
 put out
 as if

have something/nothing to do
 with
 as soon as
 run out of
 go for
 take on
 all of a sudden
 get rid of
 go over
 put together
 right away
 in charge of
 call (up) on
 get through
 go off
 as for
 in time
 take off
 take over
 be used to (+ noun phrase/
 gerund/present
 participle)
 have/keep something in mind
 in effect
 in public
 look forward to
 make a difference
 hold on
 throw out
 come by (meaning to visit)
 go with
 to the/some extent
 for sure/sure
 carry out
 get back to
 take advantage of
 in favor of
 take out
 in somebody's/the view (of)
 go wrong
 go/move/be too far
 make sense
 hold somebody/something
 accountable
 under way
 shut down
 put up
 play/have a role/part in
 in general
 break up
 make up (of)
 on behalf of
 in somebody's/the interest (of)
 regardless of
 come on
 go along (with)

do one's best
 turn in something
 get away with
 for somebody's/the sake (of)
 at stake
 out of control
 in case
 take steps
 as of
 stick (be stuck) with
 account for
 break down
 back and forth
 in the long/short run/term
 all along
 count on
 keep up
 hold up (meaning to endure/
 be tested)
 pay off
 turn on
 be up to somebody/something
 with/in respect to
 in advance
 turn around
 cut down
 crack down (on)
 hang out (with)
 set out
 down the road
 have something on one's mind
 give someone a break
 live with something (meaning
 to accept)
 by hand
 be open to ideas
 follow up on
 in touch with
 with/in regard to
 in somebody's/the interest (of)
 more or less
 better off
 hand out
 stick to
 bring about
 fall apart
 in the wake of
 get in somebody's/the way (of)
 throw away
 get to the point
 hold on to
 cope with
 (can't) get over with
 shut up
 at large
 in control

take part in
take its toll
on time
make it
leave out somebody/something
rule out
fill in
run into
in common
come about
chances are
make good on something
run through something
in light of
in essence
so to speak
come/go into effect
in turn
by far
fill in
keep/be on track
get around
pass out
in a/the fashion (that)
have/get a clue
take into account
come across
first and foremost
by and large
on and off
for that matter
be in for (meaning to
experience)
above all
be over one's head
in private
up to date
shed/cast light on
keep an eye on
follow through
come off
on the verge of
put forward
bring out
to date
take something/somebody for
granted
take part in

at somebody's/the expense (of)
after the fact
all over again
the ball is in somebody's court
make fun of
level playing field
at issue
make up one's mind
hang in there
leave somebody/something alone
in somebody's/the eyes (of)
get/grab hold of
as usual
all out
up in the air
above all
drop off
get/have one's hands on
in the fore of
come across as (meaning to
appear as)
put up with
sell out (meaning to
compromise)
up for grabs
in (good/bad) shape
(get/be) in somebody's/the way
screw up
sign off
put something to rest
at somebody's disposal
hold up (meaning to endure)
to the/this effect that
touch on
up front
in detail
buy into
look up (something) in
to somebody's advantage
take issue with
hit home
make/hit headlines
keep/have an eye on
for good
in the event of/that
have a say/voice in
on the horizon
take effect

call something into question
in keeping with
live up to
by no means
to somebody's credit
a fair game
by no means
come to mind
go for it
to the contrary
for the time being
for real
in the works
beg the question
push the envelope
in the wrong
the big picture
in line with
in the pipeline
make out
right off the bat
break off
in order/sequence
bits and pieces
do away with
give away
in good faith
make up for
on the whole
once and for all
rule of thumb
draw the line
fall short
in no way
quid pro quo
(the) ball is in your court
ballpark (e.g., figure)
bring forward
by virtue of
err on the side of
get/have a hand on
give rise to
in due course
in place of/in somebody's place
in practice
off the top of one's head
so on and so forth
the other way around

MICASE

sort of
 kind of
 go on
 of course
 in terms of
 in fact
 go through
 at all
 as well
 deal with
 make sure
 in order to/that
 figure out
 look for
 end up
 make sense
 in other words
 get into
 come up with
 in a/any sense of/that
 find out
 turn out
 in a/some way
 come up
 first of all
 as/so far as
 set up
 used to (+ verb)
 go over
 go ahead
 put on
 in general
 according to
 pick up
 by the way
 point out
 as well as
 so far
 get rid of
 have/keep in mind
 as/so long as
 come on
 as if
 with/in respect to
 ring up
 go with
 as to (+ wh- clause/noun
 phrase)
 look up something in
 take place
 no matter (+ wh- clause)
 put out
 take care of
 work out
 have something/nothing to do
 with

make up (of)
 on the other hand
 for sure/certain
 in time
 get out of
 show up
 to the/an extent (that)
 go for
 as soon as
 be about to
 break down
 more or less
 put together
 take out
 go along (with)
 get on
 account for
 take off
 give up
 go off
 no matter (+ wh- clause)
 be used to (+ noun phrase/
 gerund/present
 participle)
 run into
 break up
 turn on
 after all
 make it
 carry out
 hold on
 take over
 hang out (with)
 back and forth
 get back to
 be up to somebody/something
 play/have a role/part in
 make a difference
 right away
 get through
 take on
 in case
 once again
 stick (be stuck) with
 take up
 be (have something) in place
 turn in
 in effect
 regardless of
 all of a sudden
 fill in
 in somebody's/the interest (of)
 rule out
 take into account
 come across
 come off

in common
 get around
 on time
 as a matter of fact
 better off
 hand out
 run through (meaning to go
 over something)
 turn around
 as for
 be open to (e.g., ideas)
 in detail
 take advantage of
 in favor of
 rule out
 keep up with (+ gerund/
 present participle)
 in touch with
 leave out somebody/something
 pass out
 for somebody's/the sake (of)
 so to speak
 stick to
 be/go/move too far
 go wrong
 in somebody's/the view (of)
 screw up
 touch on
 come about
 in public
 in turn
 bring out
 chances are
 do ones' best
 up front
 shut down
 by hand
 call for
 come by (meaning to visit)
 count on
 get away with
 have/get a clue of
 hold on to
 in advance
 in the long/short run/term
 make good on something
 throw out
 drop off
 in essence
 in the wrong
 live with (meaning to accept)
 look forward to
 all along
 as of
 (can't) get over with
 come to mind

go after
 set out
 throw away
 for that matter
 give away
 in charge (of)
 by far
 by virtue of
 down the road
 follow up on
 get to the point
 hold on to
 in somebody's/the way (of)
 in (good/bad) shape
 on the whole
 call (up) on
 cut down
 get/have a handle on
 the other way around
 on behalf of
 as usual
 bring about
 buy into
 for good
 give rise to
 make fun of
 make out
 to date
 with/in regard to
 after the fact
 all out
 get/grab hold of
 in the/a fashion (of/that)
 on and off
 right off the bat
 shut up
 take part in
 take steps
 up in the air
 above all
 the big picture
 bits and pieces
 break off
 by and large
 cope with

draw the line
 go for it
 in place of/in somebody's place
 take something/somebody for granted
 to somebody's credit
 at large
 ballpark figure
 call something into question
 first and foremost
 get/have hands on
 get something across
 in practice
 keep/be on track
 live up to
 make up for
 shed/cast light on
 at somebody's disposal
 at somebody's/the expense (of)
 have/get a say/voice in
 in control
 in no way
 off the top of one's head
 rule of thumb
 so on and so forth
 up to date
 at issue
 come/go into effect
 fall apart
 from scratch
 give somebody a break
 in somebody's/the eyes (of)
 pay off
 put up with
 to the/this effect (that)
 turn up
 at stake
 beg the question
 come across as (meaning to appear as)
 do away with
 follow through
 for real
 have something/be on somebody's mind

hit home
 in light of
 in line with
 in private
 on the verge of
 once and for all
 quid pro quo
 sell out (meaning to comprise)
 take issue with
 up for grabs
 a fair game
 all over again
 be in for
 be over one's head
 be/put on hold
 bring forward
 by no means
 for the time being
 in the wake of
 in the works
 leave somebody/something alone
 make/hit the headlines
 make up one's mind
 on the horizon
 out of control
 put forward
 to somebody's credit
 err on the side of
 hang in there
 hold somebody/something accountable
 hold up (meaning to delay)
 in keeping with
 in the event of/that
 in the fore of
 in order (meaning in sequence)
 keep an eye on
 push the envelope (meaning to move beyond the usual or normally accepted practice)
 sign off
 to somebody's best knowledge
 to the contrary
 with (keep) an eye on