Corpus Linguistics: A Tool for a Critical Approach to Language Teaching?

Jennifer Leigh Johnson

In memory of my great-grandfather, Rodric Jackson (1922-2007), a farmer, with whom I spent my childhood watching the daily evening news, and who helped teach me to pay attention to the small stuff because everything is connected.

Abstract

This research paper attempts to show how corpus linguistics can be employed by language teachers to enrich their linguistic awareness. I first compare and contrast two major approaches to grammar, structural and functional, and then show how corpus linguistics can be used along with the genre approach and critical discourse studies to reveal the "lexicogrammatical" meaning of a word or phrase. I conclude with an exploratory, empirical study using corpus linguistics to understand the context of two politically and socially sensitive phrases, "liberal" and "left wing."

Introduction

Since the 1970s, the communicative language teaching method has been considered the best theoretical backing to syllabus creation in the ESL classroom. After all, it is logical that students should be able to communicate effectively in their target language after a reasonable period of study, and as a result, the field has made a noticeable shift from audiolingualism's focus on practicing grammatical forms in drills to a focus on having students actually produce the desired forms in class activities. The teacher's role subsequently expanded from the distributor of knowledge to that of being responsible for providing opportunities in the language classroom for production; for example, by incorporating group work, oral presentations, and role playing. However, teachers still must provide instruction on the language itself, and many aspiring and practicing ESL teachers may feel trepidation when it comes to teaching grammar and vocabulary in a contextualized manner. Particularly for native speakers, though exposed throughout formal schooling years to the grammar and lexicon of their L1, and possibly for some years to that of another language, there still remains the question of how best to teach the rules of language, and even more challenging, how to introduce the social context surrounding these grammar rules and lexical items. This paper suggests that teachers should employ corpus linguistics as a scholarly exercise to deepen their own linguistic awareness in order to enrich their teaching. When corpus linguistics is combined with an understanding of functional grammar, the genre approach, and critical discourse studies, the result is a more comprehensive framework for decrypting the reasons behind the selection of one grammatical structure or word phrase over another.

The goals of this paper are to provide an overview of the two major approaches to the study of language, to explain corpus linguistics, to show how useful corpus linguistics can be to a language teacher. Finally, I will present an exploratory, empirical study using corpus linguistics to understand the context of two politically and socially charged phrases, "liberal" and "left wing."

Structural Linguistics versus Functional Linguistics

There seems to be a division on how linguists see language based on the social integration it provides and the cognitive faculties it employs. This stretches back to Chomsky's (1965) dichotomy of *competence* and *performance,* and even further to Saussure's (1916) distinction between *langue* and *parole.* Both of these dichotomies claim that there is a system of rules underlying any utterance, and that describing this system should be the focus of linguistics. Like Descartes's split of the mind and body, this distinction proved useful for discussion, but fell short in practice. This problem is elabo-

rated on by Beaugrande (1999) when he criticizes modern linguistics for falling into the same ideological trap of science, which he calls "scientism," a doctrine he defines as "holding that only scientific knowledge is true and valid" (p. 261). He uses a related concept of "idealization" to criticize linguistics study. This ideology, he claims, holds that "language is based upon an ideal mode or order that is not readily evident in the ordinary practices of real discourse within society" (p. 3). While Beaugrande is concerned with the philosophical implications of challenging existing ideologies, a topic outside the scope of this paper, he does point to a newly emerging ideological approach, the "ideology of ecologism" which is more focused on how language is guided by discourse, how it is used in interactions, and what people use it for (p. 262).

Such an ideology would probably include Halliday (1985, 1995) since his functional linguistics theory, in contrast to Chomsky's structural linguistics, focuses on how language is being used to convey meaning. Compare these two quotes:

Linguistic theory is primarily concerned with an *ideal* speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of language in actual performance. (Chomsky, 1965, p. 3, emphasis added)

> Every text—that is, everything that is said or written—unfolds in some context of use; furthermore, it is the uses of language that, over tens of thousands of generations, have shaped the system. Language has evolved to satisfy human needs; and the way it is organized is functional with respect to those needs—it is not arbitrary. A functional grammar is essentially a 'natural' grammar, in the sense that everything in it can be explained, ultimately, by reference to how language is used. (Halliday, 1984, p. xiii)

As expressed in these statements, structuralism and functionalism are similar in their focus on language; however, the crucial difference is what they consider language to be. In the Chomskian approach, cognitivists look for evidence of the Innateness Hypothesis, which states that language is only possible for the human species (Nunan, 1999). In the Hallidayan approach, functionalists look for the meaning of a word in its grammatical and social context. If we apply Beaugrande's (1999) distinction of "standing constraints," or formal grammar, "emergent constraints," or rules for using words related to the topic at hand, and "global constraints," or rules of the social situation (p. 263); we can see why functional linguistics does a lot more than structural linguistics. A structural grammar approach is mainly concerned with the "standing constraints" of grammar, while a functional grammar approach handles those, plus the "emergent constraints" and the "global constraints."

In the broader picture, we could examine how this contrast between the two approaches reflects the major shifts in theoretical paradigms since Plato addressed the Sophists in regard to analyzing speech, writing, and text. Richter (1998) suggested using the ideas of McKeon (1990), who classified thinkers by looking at their basic premises. His categories are *dialectical, operational, problematic, and logistic* (as cited in Richer, 1998, p. 12). While there is significant overlap in these categories, it is useful to think about the fundamental differences in world view: Dialectical and operational thinkers see the world as connected; problematic and logistic thinkers see the world as parts making up a whole. Using these categories, I think that structural grammar fits into the problematic and logistic categories, while functional grammar fits into the dialectical and operational categories. However, while I firmly believe that it is beneficial for a teacher to know the historical paradigms which influenced the current ones, the point of contention in language teaching is how best to develop our students' ability to *use* the target language. Getting a sense of the historical views of language and society and

deciding which thinkers most closely align with personal beliefs will help language teachers better understand their own teaching philosophies.

Bringing the two major approaches to the study of language, functional linguistics and structural linguistics, to language teaching, Celce-Murcia and Larson Freeman (1999) support Butler's (2004) claim that functional and cognitive approaches compliment one another. While they believe that teachers should have a strong grasp of the language's grammar that they teach, they also encourage teachers to understand grammar's communicative function, emphasizing the need "to take into consideration how grammar operates at three levels: the sub sentential or morphological level, the sentential or syntactic level, and the suprasential or discourse level" (p. 3). More radically, I believe that language teachers can benefit from a move toward a more functional approach to language study by starting with texts rather than by building up from the smaller to the larger units of language. Beaugrande (1999), for example, outlines the ways corpus linguistics could answer to the "idealisation" that is inherent in structural linguistics:

To a significant degree, a language is always in the process of being created and negotiated whilst discourse is in progress; and a different generation of linguistic theories will be needed to explain how. The ideology sustained by corpus linguistics would accordingly be a version of *dynamism* wherein the specification of a theory will be far more actively data-driven and 'bottom-up' than the theories sponsored by lingisticism. Theories will no longer originate 'from the top down' every time some ambitious linguist chooses to fabricate his or her personal idealisation and illustrate it with a handful of fictional sentences. (1999, p. 5)

"Linguisticism," according to Beaugrande (1999) "sustains the doctrine that 'language' is an abstract, uniform, and stable system whose nature and properties can be determined only by the ratiocinations of 'linguistic theory' and not from observing and recording discursive practices" (p. 262), a view transparent in Noam Chomsky's

(1965) statement cited above. A pioneer in starting with texts and speech acts and therefore making the top-down approach to language teaching more concrete is Micheal A.K. Halliday, who published his influential *An Introduction to Functional Grammar* in 1984. In this work, Halliday described language in terms of *ideational*, *interpersonal*, and *textural* levels (p. 5). These levels correlate to three functions of language, and within every sentence all three work together to create meaning (Butt, Fahey, Feez, Spinke & Yallop, 2001). From a functional grammar point of view, "grammar is not just a matter of labeling classes of words—it is a purposeful, constructive and above all social enterprise" (Butt, 2001, *et. al.,* p. vi). This is why this approach is highly useful in conjunction with genre theory and corpus linguistics. To this end, Tsui (2005) remarks that "corpus studies have shown that linguistic forms, contexts and meanings are inextricably linked. The co-occurance of lexical items in different contexts is crucial to the meanings that they take on the pragmatic functions that they perform" (p. 352). She is referring here to the phenomenon of "collocability," the "preferences of some lexical options for appearing with certain others," and "colligability," the "preference of some grammatical options for appearing with certain others" (as defined by Beaugrande, 1999, p. 263). By using corpus linguistics, we can study the way that words are used in context to achieve a certain function. This kind of knowledge would increase a teacher's prowess in presenting genre-specific writing and speaking techniques.

Defining Corpus

Using a corpus means that we are working with "real language data." A *corpus* is a collection of texts to be used a represent-tative sample (Butler, 2004). A corpus is used to study *lexical meanings*, as opposed to "grammatical meanings" (McEnry & Wilson, n.d.). A lexical meaning, according to Webster's dictionary, is "the meaning of the base (as the word *play*) in a paradigm (as *plays, played, playing*);" while grammatical meaning is "the part of meaning that varies from one inflec-

tional form to another (as from *play* to *played* *to playing*)." Corpus analysis, then, is examining the "lexicon-grammatical" sense of a word, since an understanding of both meaning and form are key to language use.

An Overview of Corpus Linguistics

The first computational corpus of American English was compiled in 1967 at Brown University by Henry Kuceera and W. Nelson Francis. Two other major corpora are the British National Corpus and the Cobuild Bank of English. However, using corpora to determine lexical meaning is not a new idea; in fact, the *Oxford English Dictionary* (OED*)* is a project embarked upon first by London's Philological Society, who felt available dictionaries were "deficient." The task of listing and defining all the English words ever used is a daunting one, so the compilers enlisted the help of the public, and from a collection of quotations sent by volunteers, the OED boasts that it "traces the usage of words through 2.5 million quotations from a wide range of international English language sources, from classic literature and specialist periodicals to film scripts and cookery books" (Oxford University Press, 2006). The work continues, now employing technology, and the 34 million pounds sterling spent each year on the project is justified by the claim that the OED is "an irreplaceable part of English" which documents "the continuing development of our society" (Oxford University Press, 2006).

Corpora used in corpus linguistics differ from the OED in that the historicity of the word is not usually a primary consideration (though some studies do include a diachronic analysis—see Orpin, 2005). The types of questions asked by corpus linguistics are resonate with those found in critical discourse analysis, a discipline which is always relating the text to society, and which was delegated as irrelevant by those who think that observing language in use can tell us nothing about the language compared to studying the language explicitly (Beaugrande, 1999).

Application of the potential insights from corpus analysis has increased greatly in the last decade (Tsui, 2005, p. 335). Wolfgang Teuber, editor of the *International Journal of Corpus Linguistics,* stated that almost all linguists use corpora due to its authenticity. In his opinion, "no introspection can claim credence without verification through real language data" (Teubert, 2005, p.1). He also commented on how the method has gone beyond "lexicography and language teaching" and now "offers a perspective on language that sets it apart from received views or the views of cognitive linguistics, both relying heavily on categories gained from introspection rather than from the data itself" (p.2). The entropic growth of corpus linguistic analysis is not surprising if we realize that its usefulness is bound by technological advances. As corpora become larger and search engines more sophisticated, we should expect to see more research into viewing syntax and semantics through the lens of language as a social construct.

Tognini-Bonelli (2001) offers valuable insights about the impact of corpus linguistics on language teaching method-ology. She explains that there are two ways that this tool is being used by linguists; one is "corpus-based," which is when people apply corpus data to old theories in hopes of further supporting or expanding them; and the other is "corpus-driven," which is when people are looking at corpus data as an object of study unto itself (pp. 68-77). According to Butler (2004), the latter has resulted in "fascinating insights" (p. 154). Among these, the two that are most important for this paper are collocations and the "idiom principle" (Sinclair, 1991). Sinclair compared the "open-choice principle" with the "idiom principle" and, by using corpus linguistics, came to the conclusion that grammars are not simply puzzles for word classes, but that "a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments" (Sinclair, 1991, p. 110). Significantly, since people choose language phrases in observable patterns, it is clear that our reality is constructed according to the words that we use, and that much of what we say is directly re-

lated to our ideological standpoints. However, before looking at how critical discourse can inform corpus analysis, I will discuss some of the ways that corpus linguistics can be used in language teaching.

Teaching Grammar with Corpus Linguistics

Some ESL textbook writers use corpus linguistics to develop teaching strategies and grammar exercises. One such example is the *Touchstone* series, which was created using corpora. The authors believe that using data from corpora will result in a better representation of how the contemporary society is using language, and how that language is dependent upon the situation (McCarthy, 2004). For example, their corpus analysis yielded patterns such as "In conversation...people use the simple present and simple past more often than any other tense" (McCarthy, McCarten, & Sandidford, 2005, p. 99). Such insights can provide motivation for learning a particular grammatical structure and reveal information about how the language is structured that a traditional grammar approach may not. The methodology backing this book up is evidently communicative, since the emphasis is on language in use.

In Hong Kong, the creators of the online resource for ESL teachers, *Telegram*, also uses a collection of corpora to answer questions from real-world situations. The promulgators recognized that a major problem for English teachers in Hong Kong is that many are themselves former ESL learners, and therefore most likely have in common with their students the obstacle that much of their English language knowledge is from textbooks emphasizing traditional grammar rules (Lock & Tsui, 2000, p. 18). The overall goal of the database is to be a resource for the teachers, and its organization is "strongly influenced by Hallidayan functional grammar" (p. 21), which indicates that *Telenex* is running on the assumption that understanding context is intrinsic with grasping meaning, and that using corpora has the benefit of allowing the learners to compare authentic examples of a particular word or word phrase in current use with

abstract explanations. Tsui (2005) also pointed out that the process of "corpus-based investigations can help teachers to reflect on their knowledge and to make this knowledge explicit" (p. 336). The teacher's questions that she described are not easily answered by a native speaker because they may have never been observed as a pattern before, and this new awareness brought on by comparing corpora samples challenges and deepens the teacher's understanding of how English works.

Teaching Writing with Corpus Linguistics

Corpora have also been employed by writing teachers in the language teaching field (Dinnie & Moran, 2005). From this angle, the corpus is not as much a tool for informing language theory, but one for raising linguistic awareness in the writer, probably for native as well as non-native speakers. One tool suggested by Platinize and Moran (2005) is the Web Vocabulary Profiler, (http://132.208.224.131/vp/), which divides words of an inputted text into four categories: the 1000 most frequently used words, the next 1000 academic words, and the rest of the words. Teachers are advised to check their students' writing as a way to monitor their vocabulary acquisition, and to instruct students on how to take responsibility for their own learning by using the profiler. Research done to test the claims has mixed results, however. For example, Goodfellow, Lemy and Jones (2002) did not find it useful in assessing their students, but did feel that it could be a valuable self-assessment tool.

Returning to Sinclair's (1991) notion of the "idiom principle" mentioned above, we see another area in which corpus linguistics can support teaching with a genre approach, since we can assume that not all language chunks are appropriate in all types of texts. By using a corpus, a teacher can see clearly the difference between closely related words, as Tsui (2005) showed in her research on the *Telenex* website set up for teachers in Hong Kong. The website used the *Modern English Corpus, The South China Morning Post, Cobuild Direct,* and the *British National Corpus*

(Tsui, p. 338). Tsui showed how using the different corpora helped the website maintainers to answer questions about the differences between synonyms such as *big* vs. *large* or *finally* vs. *lastly and* transition phrases such as *less than* vs. *fewer than*, to observe sentences that consistently break prescriptive grammar rules such as *"no sentence initial conjunctions*," and to comment upon the inherent connotative aspect of words and word phrases by exploring the collocates for *in view of* and *based on* (Tsui, pp. 340-352). This critical discourse approach to interpreting collocates was also used by Orpin (2005) in her research reported in a paper entitled *Corpus Linguistics and Critical Discourse Analysis: Examining the Ideology of Sleaze*, which will be discussed in more detail later in this paper.

Thus, teachers can use corpora specifically designed for linguistic analysis when preparing lessons, when searching for a satisfying answer to a question, when using the genre approach, and when thinking about the ideological implications of saying certain things a certain way (Diniz & Moran, 2005; Orpin, 2005). To add another dimension to the study, teachers may use online newspapers or other self-contained databases as corpora (Tsui, 2005). One advantage of using the latter sources is the ability to analyze the ideology of the source, since in the former, the identification of the sources is broader.

Bringing a Critical Perspective to Corpus-oriented Language Teaching

As Beaugrande (1999) warned, ideology is "so deeply anchored at the base of human awareness" that people will have a difficult time accepting that any other ideology is valid: "People may well defend and cling to their ideology even when they perceive symptoms of disorientation, fearing that they might lose whatever orientation they still have" (p. 5). This captures the problem Araffin (2006) mentioned in her discussion of culture in the ESL classroom, namely the Telekinetic (1993) study on Muslim students' perceptions of the American saying, "A dog is a man's best friend." Such a fundamental cultural/ideological difference is the tip of the iceberg when it comes to communication breakdowns. To see a different possibility of looking at the world would require a study of how the world can be represented in language in various ways. Gathering corpus data on a particular word or word phrase can help a teacher/student to step back from personal persuasions on what particular words or phrases mean or do. Such implications for corpus linguistics go far beyond defining a word or describing functional grammar; they allow us to understand how words work in their natural habitat—context of use.

Take, for example, the construction of a genre. A genre is first of all a communication act that has connections to society, and Nunan (1999) pointed out that genres are best studied in a functional grammar approach: "Such [social] events generally result in spoken and written texts that can be differentiated according to their generic structure and grammatical features" (p. 230). For example, the passive voice is used to take attention off individuals and focus it on the action, so it is commonly found in scientific writing where the emphasis is on the study and not on the researcher. The first person is often used in fiction or personal essays to close the distance between the reader and the author. Such implications mean that not only do language teachers need to have a deeper under-standing of the grammar of a language, they also need to be able to meet a myriad other student needs. Students, business associates, health care providers, laborers, and immigrants all need to be capable of working with and utilizing different genres. Language teachers need to be familiar with the characteristics of different genres, or at least have a ready plan to learn, if they want to run a needs-based classroom, since students are afforded a greater opportunity for success if they are presented not only with content vocabulary, but genre structure and commonly used grammatical structures as well. Butler (2004) postulated that we may find it more useful to have grammars that are area-specific rather than general, an idea which is concurrent with the genre approach. Swales (1990), for example, was concerned mainly with

teaching people how to write a research article. He deconstructed articles and described them in terms of "moves" and "steps." After Swales' (1990) influential work, people researching within the parameters of genre analysis have started using sociology methods and terms, for example, "discourse community," and/or work at identifying grammatical forms which signify a particular genre (Evans, 1995).

When critical discourse theory is applied to genre-based approaches, it becomes clear that "certain literacies yield more power than others" and it follows that being able to produce in a powerful genre will enable the writer access to some of the power (Auerbach, 2005, p. 2). Critical discourse studies attempt to link language use to particular ideological frameworks. ESL teachers are aware of the cultural implications of learning (and teaching) English, and know to be aware of the identity construction that necessarily accompanies the gain of communicative competence. Canagarajah (1999) pointed out that "in a post-modern world, education has lost its innocence. The realization that education may involve the propagation of knowledge and ideologies held by dominant social groups has inspired a critical orientation to pedagogical paradigms" (p. 3). *Ideology,* a word that Beaugrande (1999) found absent in his readings of linguistics, is now an integrated philosophy. In other words, that we say what we say how we say it has been somewhat accepted as an indicator as to the social role we wish to take.

This idea is supported by the corpus data cited in Tsui (2005) as mentioned above. One teacher's inquiry concerned the difference between *in view of* and *based on.* She found that "the former means that you take into consideration facts that have just been mentioned or about to be mentioned, and the latter means that the first thing develops from the second" (p. 353). Going beyond this definition and showing how this linguistic device has a functional and semantic meaning will help a non-native speaker understand, for instance, what is coming when a supervisor says or writes during a performance review, "In view of..."

rather than "Based on..." Communication is a social act, and if the "idiom principle" (Sinclair, 1991) is correct, we are somewhat bound to certain phrases, which have semantic consequences. Orpin (2005) succinctly summarized the benefit of combining functional grammar and critical discourse analysis and concluded:

> One of the strengths of CDA is that by marrying a Hallidayan approach to linguistic analysis (an approach that sees language as firmly rooted in its socio-linguistic context) with theories relating to the mediation of ideology and its relation to power structures in society, the researcher can make insightful statements about the socio-political implications of the instances of language use. (p. 38).

Orpin's (1995) corpus research into the nouns *sleaze* and *corruption* had shown a more negative connotation attached to *corruption*. However, in view of criticism from Stubbs (1997), she expanded her search to include more synonyms and used a larger corpus, then compared the meaning of the words in context to the linguistic norms. As a result, she was able to discern the subtleties between synonyms and to identify collates that synonyms shared and collocates which were exclusive. Studies like these reveal that the affordance of technology is helping to build the bridge between knowing about a language and using that language in an effective, communicative manner.

In the remainder of this paper, I present an exploratory corpus linguistics study on two socially and politically charged words *left wing* and *liberal.* This was inspired by the theoretical foundation of functional linguistics, the power of corpus linguistics, and the perspectives of critical discourse analysis. The results, though preliminary, are very revealing and helpful for the teaching of English as a second language.

Research Question
In this exploratory research, I aim to answer the following questions: What collocates differentiate the socially and politically charged words *left wing* and *liberal*? How can

analyzing these collocates enhance our understanding of their meaning in recent usage? What impact will this study have on teaching practices?

Methodology

Six different corpora were used: *Googlescholar*, the search engine's answer to academic cries for a reliable, scholarly database; *The New York Times*, a respected newspaper in a major US city criticized by Herman and Chomsky (1988) for having a corporate bias; *National Public Radio*, a respected news source accused by Republicans for catering to the educated elite and for being liberal; *Adbusters*, a political magazine associated with Marxism and which receives complex criticism; *Fox News*, a respected news source, accused by Democrats of having a conservative bias; and *Cobuild*, a corpus of language used in general media and literature. The spectrum provides examples of the words used by different social groups for different purposes. Also, *Cobuild* and *NPR* provide that spoken language samples also be included. In these corpora, I did not count instances of *liberal* with affixes (e.g., *liberalism*). Representative examples from the corpora examined are presented in Appendix A.

To determine what words collocate with *left wing* and *liberal*, I examined each entry and noted the four "content" words that were available on either side of the term in question. I then tallied the occurrences, and then sorted their collocates with my own data-generated categorization. Based on my interpretation of the types of collocation associated with each phrase, I identified a number of categories to group these collocations (see Appendix B).

Findings

The Occurrence of Left Wing and Liberal in Different Sources

Table 1 shows the number of samples collected from each source. This varies according to availability and relevance. For example, the term *left wing* might be found describing an aircraft, so all references of this kind had to be discarded. With closer study, this data reveal interesting aspects about the words. That *liberal* (27 items) was used more often than *left wing* (19 items) in results from *Google Scholar* could mean that the latter term is not widely used in the academic lexicon. *Adbusters* offered only one true example of *left wing*, while providing seven examples of *liberal*. This makes sense if we think of their association with Marxism, a social group which knows too well the ambiguity of the term *left wing*, as it is at once identified with and rejected by other ideologies associated with being *left wing*. The *New York Times* uses the word *liberal* (18) far more often than *left wing* (7), which may support Herman and Chomsky's (1988) criticism that they have a corporate bias, since corporations would be more concerned with liberal policy makers than with left wing anything. *NPR, Fox,* and *Cobuild* all offer near equal examples of both terms, with 11 and 12 from *Fox*, 40 and 41 from *Cobuild* and 19 and 19 for *NPR*. Since these are three very different kinds of sources, closer look at the collocations can reveal more about their usage of the terms.

Table 1

The Number of Samples Collected from Each Source

	Occurrences of leftwing	Occurrences of liberal
GoogleScholar	19	27
National Public Radio	19	19
Adbusters	2	7
New York Times	7	18
Fox News	11	12
CoBuild	40	4

Collocates and Parts of Speech

As shown in Appendix B, for *left wing,* the most frequent collocate was *party,* with 14 matches. *Democratic Party* was found to co-occur with *left wing* 10 times, and *right/right wing* to co-occur with it eight times. *Liberal* co-occurred with *left wing* in six instances. *Activists, political,* and *demonstrated* co-occurred with *left wing* four times each, and *union, government, opposition, blog,* and *extreme* each co-occurred with *left wing* three times. Those co-occurring twice with *left wing* were *anti-Bush, critics, media, women, guerrillas, Republican, leaders, model, authoritarianism, communism* and *populism.*

For the word *liberal,* the most frequent collocate noted was *democracy,* with five co-occurrences. When this value is collapsed with the other forms of the word (*Democrats* and *Democratic*), a very strong connection, 17 instances, emerges. *State* co-occurred with *liberal* six times; and *party, conservative, ruling, criticizes,* and *social* each co-occurred with *liberal* three times. Those found to co-occur with *liberal* were *circles, activist, Christians, church, bourgeoisies, judge, American, nationalism, corporate, bloggers, California, perhaps, alternative, impossibility, political, sexual, progressive, norms, view, bias, peace, reputation, college degree,* and *justice.*

Words that were collocates with both left wing and liberal include: party, organizations, students, activists, media, Democratic Party, government, blog/ bloggers, socialist, nationalism/nationalization, young, radical, bias, Clinton, and labor. Interestingly, in one instance, left wing was changed into a noun with the addition of an affix, while liberal is commonly used with different affixes.

The following example was taken from a paragraph in *Adbusters*:

Liberal hawk. War dove. Academic circles may prefer *liberal* interventionist. The term applies to left-wingers who are committed to using military force to preserve and promote human rights and…

The addition of the suffix –*er* has a stronger impact than the phrase "those who are left-wing," and designates *left wingers* as particular members of a social group.

This leads to the discussion of what word class each item belongs in. In all sources, both are used predominantly as adjectives. Here, I will highlight the use of *left wing* and *liberal* as nouns, because this shows the clearest distinction. Interestingly, as a noun, *left wing* tends to describe a group, while when *liberal* is used as a noun, it tends

to describe a single person. This in part explains the one aberration in my data of *Adbusters* calling someone a *leftwinger*. A political party, a candidate, or a news source can be *left wing*, but not a single person, unless the use of a "slang" term is employed. This implies that being *left wing* means being associated with a group of people, and connotatively, dissenters. Being called *liberal* in comparison is a slap on the wrist, a comment on one's social values. This subtle distinction has led to those who oppose the current President Bush and his administration's invasion of Iraq to use the slogan "Dissent is NOT unpatriotic." All anti-war people are called *left wing*, thus associated with radical revolutionary groups worldwide. The following includes examples of the terms used as nouns from each source (no example will be provided when a corpus only used the term as an adjective).

From the *Googlescholar* corpus examined, I found one example of *liberal* being used as a noun, and none for *left wing*:

The Impossibility of a Paretian Liberal...

However, in *The New York Times* corpus, all uses of *left wing* were nouns (used to describe the non-right wing of the Democratic Party), while all uses of *liberal* were adjectives

…left wing of the Democratic Party…

…frustrated Democrats and liberal bloggers, who relentlessly dogged…

From the NPR corpus, two terms were used as nouns with almost equal counts (seven and eight, respectively):

…What Clinton did clearly silenced the left wing of the party…

…Bob Casey Jr. is a pro-life, pro-gun, social conservative and economic liberal...

From the *Adbusters* corpus examined, there was only the aforementioned example of *left wingers* as a noun, and one use of *liberal* as a noun:

…Increases the likelihood that liberals who do so will be the contemporary version of the boy who cried wolf…

Examination of the *Fox* data reveals that the terms were used only as adjectives.

…Unions and left-wing opposition parties fear privatization will erode job security…

…We are in danger of our country falling under socialist liberal progressive politicians…

In the *Cobuild* data, *left wing* was used as a noun six times, but *liberal* was only used as an adjective in the four examples that I found.

…In 1952 the liberal non-Communist left wing had managed to remain powerful…

…to flee Germany, his homeland, because of his liberal political views…

In the single *Googlescholar* example, the only example of "liberal" being used as a noun is in reference to Vilfredo Pareto, an Italian economist whose work concerned finding the best solution possible when solving a problem with more than one, seemingly unrelated, criteria (Rodrigue, 2007). This is reflective of the site's commitment to providing scholarly articles, and not opinion pieces.

However, the *New York Times*, with its reputation for being a Democratic newspaper is careful to use "left wing" as part of a phrase which marks political distinctions, especially within the Democratic party, and reserving "liberal" to use as an adjective describing people's societal leanings.

NPR, which has been criticized as leaning left, or at least promoting the Democratic Party over the Republican, is like the *New York Times* in terms of the usage of "left wing" in regards to divisions within political parties, and "liberal" as an adjective with a questionable connotation. Both of these sources seem to be trying to avoid the negative connotations of "left wing."

Adbusters, which is defiantly not promoting the mainstream viewpoint, avoids using the term of "left wing" almost completely, yet seems to adopt the common usage of "liberal" as one with a particular views on society and government.

Fox, known for its exclusively right-wing stance, uses both words in ways that give a negative connotation.

Cobuild, which represents the widest variety of printed sources, is also using "left wing" to denote a part of a larger group, such as a union or political party, while using "liberal" as an adjective. From this analysis, it seems that the term "left wing" is one which can be employed as a useful demarcation between opinions held in a larger group, or as marking a dangerous deviation from the norm, while "liberal" is widely used to describe someone who wants a stronger tie between government and the individual, for example, in health care or education.

It is interesting to note that the *Fox* news network, which unquestionably supports the Bush administration, uses both terms in a negative sense, while the *New York Times* and *NPR* use the more neutral definitions. This must have an influence on the consumer's understanding of these politically charged words, and the members of groups being described as such. I hope the results of my study promote people to get their news from a wide variety of sources, and to pay closer attention to how words are being used.

Discussion

One point that became clear to me is that the word *liberal* is used in more combinations than *left wing*. The two words have clear semantic overlap, as both are found in the domains of politics, religion, theory, and popular culture, but the examination of their collocates indicates that *left wing* has a more negative connotation of the two. After all, it was found near *fascists, idiots, bastards, terrorism, whining, arguing, hostile, vehement, childish, whacko,* and *freak show*.

I also noticed that liberal is frequently found near genre-specific compound nouns. For example, <u>*liberal*</u> *social justice,* <u>*liberal*</u> *international organizations,* <u>*liberal*</u> *minority rights,* and <u>*liberal*</u> *social conservative*. Liberal is also part of the proper names of multiple recognized political parties, further complicating the situation. This vividly illustrates the "idiom principle" (Sinclair, 1991), mentioned above.

Finally, I find it very interesting that in some instances the two words are interchangeable, especially when talking about "celebrity politicians" like the Clintons and Leiberman, or when discussing the internal divisions in political parties, when in fact the two have quite different connotations, based on the high occurrence of negative collocates for *left wing*. I think it is cause for concern that the "social groups" which were found in the same environment of *left wing* include *rebels, anti-Bush, fascists, intellectuals, idiots, faculty members, women, guerrillas, film writers, etc.* Those found for *liberal* were much more neutral, such as *societies, circles, people* and the possibly charged *elites, pro-gun, Christian*. Also, the use of "anti" with *left wing* versus *pro* with *liberal* indicates a certain position. These types of results point to *left wing* as meaning someone or something that is "against the norms of society" while *liberal* remains a rather vague idea that correlates with people's views of how society should be organized. We might want to be concerned that *left wing* is associated with *Hamas, Red Army, the Democratic Party, Fidel Castro, Clinton* and *Leiberman*.

Conclusion

This research project was undertaken to accomplish two goals: namely, to explore the semantic meaning of *left wing* and *liberal* through interpreting corpora data, and secondly, to make my first attempt at participating in the research genre.

I believe that the rigorous scholarly exercise of performing corpora searches can benefit any language teacher. By exploring the intricacies of lexicon and grammar, teachers are focusing on the language itself, and to interpret the data is a process of reflection on assumptions about the nature of language, and in this case, the structure of English. This research could be expanded to include more information about the distribution of the key words in the various sources, which would be considered critical discourse since the object of such an endeavor would be to uncover the ideological backdrops.

References

Araffin, S. (2006). Culture in EFL teaching: Issues and solutions. *HPU TESL Working Paper Series 4*(1), 75-78.

Auerbach, E. 2005). The power of writing, the writing of power: Approaches to adult ESOL writing instruction. National center for the study of adult learning and *literacy.* Retrieved from http://www.ncsall.net/?id=341 on October 12, 2005.

Beaugrande, R. (1999). Discourse studies and ideology: On 'liberalism' and 'liberalisation' in three large corpora of English. *Discourse Studies 1*(3), 259-295.

Butler, C.S. (2004). Corpus studies and functional linguistic theories. *Functions of Language, 11*(2), 147-186.

Butt, D., Fahey, R., Feez, S., Spinke, S., & Yallop, C. (2001). *Using functional grammar: An explorer's guide* (2nd ed.). Sydney: National Centre for English Language Teaching and Research.

Canagarajah, A. (1999). *Resisting linguistic imperialism in English teaching.* Oxford: Oxford University Press.

Celce-Mercia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). Boston, MA:Heinle & Heinle.

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

Evans, T.D. (1995). Common-core and specific approaches to the teaching of academic writing. *Academic Writing in a Second Language: Essays on Research and Pedagogy.* Norwood, N.J.: Ablex Pub. 294-312

Goodfellow, R., Lemy, M., & Jones, G. (2002). Assessing learners' writing using lexical frequency. *Cambridge Journals, 14*(1), 133-145.

Halliday, M.A.K. (1984). *An introduction to functional grammar.* London: Edward Arnold.

Halliday, M.A.K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.

Herman, E., & Chomsky, N. (1988). Manufacturing consent: *The political economy of the Mass Media.* New York: Pantheon Press.

Lock, G., & Tsui, A.B.M. (2000). Customizing linguistics: Developing an electronic grammar database for tea-chers. *Language Awareness, 9*(1), 17-33.

McCarthy, M. 2004). *Touchstone: From corpus to course book.* Cambridge: Cambridge University Press.

McCarthy, M., McCarten, J, & Sandiford, H. (2005). *Touchstone: Student book 1.* Cambridge: Cambridge University Press.

McEnry, T., Wilson, A. (n.d.). *Corpus Linguistics.* Department of Linguistics: Lancaster University. Retrieved March 29, 2006 from http://bowland-files. lancs.ac.uk/monkey/ihe/linguistics/co rpus1/1fra1.htm.

McKeon, R. Freedom and history and other essays. From Richter, D., Ed. (1998). *The Critical Tradition: Classic texts and contemporary trends.* Boston: Bedford Books.

Orpin, D. (2005). Corpus linguistics and critical discourse analysis: Examining the ideology of sleaze. *International Journal of Corpus Linguistics*, 10(1), 37-61.

Oxford University Press (2006). Oxford English Dictionary: The definitive record of the English language. Retrieved March 30, 2006 from http://www.oed.com/about/

Rodrigue, J.P. (2007). Pareto optimality tradeoff curve. Retrieved April 4, 2007 from http://people.hofstra.edu/ geo trans/eng/ch1en/appl1en/pareto.html

Saussure, F. (1916). Nature of the linguistic sign. From Richter, D., Ed. (1998). *The Critical Tradition: Classic texts and contemporary trends.* Boston: Bedford Books.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings.* Cambridge: Cambridge University Press.

Tsui, A.B.M. (2005). ESL teachers' questions and corpus evidence. *International Journal of Corpus Linguistics,* 10(3), 335-356.

Tueber, W. (2005). My version of corpus linguistics. *International Journal of Corpus* Linguistics, 10(1),1-13.

Appendix A

Examples of *leftwing* and *liberal* in the Examined Corpora

Examples of *leftwing*

From *Google Scholar*:
> Why have Women Become Left-Wing? The Political Gender Gap and the Decline in Marriage
> The myth of left-wing authoritarianism
> The left-wing opposition in Denmark, Norway and Sweden: cases of Europhobia?
> the party has moved since 1977 from a right-wing to a slightly left-wing position
> US colleges would not have purged their left-wing faculty members during the McCarthy era

From *The New York Times*:
> the right and left wing of the Democratic Party over
> in with the left wing of the Democratic Party and
> her party's left wing. She returned to Washington
> and an Democratic left wing invigorated by Mr. Lieberman's
> to the extreme left wing to combat the Republicans

From *NPR:*
> Decades later, Pinochet faced charges for his role in a plan to eliminate left-wing opponents.
> presidential nomination in 2008. What Clinton did clearly silenced the left wing of the party and
> New York Times-Reading, Body-Piercing, Hollywood-Loving, Left-Wing Freak show. That blog know, is a left-wing communist, right-wing fascist, anti-Bush propagandist, pro-Bush toady
> the left-wing Peace Now group believes the army should do what it takes to prevent Hezbellah

From *Adbuster:s*
> By contrast, Japans remaining left-wing activists and politicians were, increasingly, older and out

From *Fox News:*
> Delay is not the problem. It was the left-wing media and the idealess Democrats who concocted
> destruction of this country: abortion, crime rampant immorality, massive debt, and Left-wing champion.
> Unions and left-wing opposition parties fear privatization will erode job security
> Chomsky is a linguistics professor at MIT, and an avowed left-wing political activist, who since
> of this campaign liberal bloggers, left-wing bloggers are going to get Democrats in trouble again

From *Cobuild*:
> hell of a fight with 'em. My union's got a left wing and a right wing. I'm with the left. When
> (1969) Deak, Istvan, Weimar Germany's Left-wing Intellectuals: A politial history of the
> the military coup of 1987 was a left-wing government, bent on nationalization and income
> Methodists were bitterly split over the left-wing activities of the church-affiliated Federation
> in jail terms for Alger Hiss and the left-wing scriptwriters knows as the Hollywood Ten

Examples of *liberal*

From *Google Scholar:*
    Multicultural Citizenship: A Liberal Theory of Minority Rights
    Preferences and Power in the European Community: A Liberal Intergovernmentalist Approach
    Modes of immigration politics in liberal democratic states.
    [Book] The Corporate Ideal in the Liberal State, 1900-1918
    The radical future of liberal feminism

From *The New York Times:*
    frustrated Democrats and liberal bloggers, who relentlessly dogged
    has a more liberal reputation while Jonker and Maloney
    of the urban liberal bourgeoisie. Some familiar dysfunctions
    a California liberal who is the current Democratic
    is a pervasive liberal bias in the news media

From *NPR:*
    with a radio spot that linked Shuler to Nancy Pellosi and the quote, unquote, "Pellosi liberals."
    candidate Bob Casey Jr. is a pro-life, pro-gun, social conservative and economic liberal. I
    tax-exempt status in jeopardy. One liberal church in Pasadena, California, finds itself smack-dab
    come in for a lot of criticism over the years that he's a classic liberal activist judge who stretches
    with a liberal George Clooney and with the African-American Don Cheadle as he goes into

From *Adbusters:*
    One of the basic principles of a democratic liberal country is the decentralization of power.
    heard the terms. Liberal hawk. War dove. Academic circles may prefer liberal interventionist
    The play ethic, at least in the West, is about keeping societies liberal and complex. We need
    are a core issue in the election for Koizumis successor in the ruling Liberal Democratic Party
    As the late Liberal Democratic peer, Roy Jenkings put it:
    My view is that the prime minister, far

From *Fox News:*
    the Democrats have in their corner is a left-leaning liberal media that jumps up and airs GOP
    ecisions of the court will start to show a greater tendency to be less liberal."
    or Williams) she also scored higher in words deemed negative (including liberal, biased, bad
    liberal smackdown! Ex-Clinton advisers James Cerville and Paul Begala on Clinton-Wallace
    Jesus was very liberal for 30 AD
    your comment is proof of your misguided liberalism get Bush

From *Cobuild*:
    flee Germany, his homeland, because of his liberal political view. He was also a follower of
    be much happier, a happiness reflected in more liberal ratios. Second, instead of &dollar
    of the symposium. Her thesis was that in liberal bourgeois society teaching must be seen
    subversive of the intellectual foundation of liberal democratic society. Indeed, Professor Fish

Appendix B

Categories of Collocates for *left wing* and *liberal*

Categories	left wing		liberal	
Proper names	Clinton (2) Bertrand Russell Salih Zidan Lieberman	Pinochet Fidel Castro Rendell Powell	Bill Clinton Hillary Clinton Clinton Alito Koizumi Jim Jeffords Kerry Charles Schumer	George Clooney Ruie Tauscher Nacy Pelosi Shuler Fowler Jesus John Roberts Studds
Places	Hollywood (2) Washington MIT Denmark Germany (2)	Japan Panama Columbia Israel	California (2) Pasenda San Francisco Deep South New York	
Time	60's and 70's 1987 old days		1950-1985 future ruling (3) current	30 AD 1900-1918 contemporary
Actions	attack (2) demonstrated (4) bashes eliminate Purged	taking over combat counterattack treatment	attend dogged criticize (3) annihilate retiring governing	normalizes refusing programming assessing (2) dancing
Adjectives	young radical whining arguing hostile vehement	whacko childish freak show Italian clandestine liberal (6) body piercing	young radical negative interventionist frustrated pervasive mainstream urban multicultural Paretian independent left leaning complex classic alternative (2) watchdog critical familiar	leaning reflected impossibility (2) economic late (as in dead) Democratic (8) political (2) rhetorical sexual (2) treasonous erudite intellectual international institutional moderate decentralization progressive (2) interdependence social (3)

Category				
Modifiers	extreme (3) ultra predominantly	Exclusively variant variety slightly	relentlessly advanced more	perhaps (2) less **politically**
Direct references to technology	blog (3) net wired		bloggers (2) Web sites	
Political and social structures	Labor human rights government (3) political (4) president chief executive leaders (2) chieftains power model (2)	leadership presidential nomination opponents candidate military coups debt gender gap	Labor minority rights citizenship justice (2) essays literature arts (4) dailies news reputation (2) member conflict imagination	war peace (2) college degree (2) democracy (5) norms (2) tradition legacies goods foreign affairs smackdown firestorm trap self
Political and social trends	nationalization socialist right/right wing (8) authoritarianism (2) communism (2) McCarthyism populism (2)	terrorism movements activities modernist collectivization counterculture	nationalism (2) socialist conservative (3) secularism institutionalism	feminism schooling immigration leave policy corporate (2)

Social groups	party (14)	fascists	party (3)	realist
	Democratic	guerrillas (2)	Democrats (4)	Christians (2)
	Party(10)	idiots	media (5)	Protestant
	media (2)	corporations	activist (2)	European Parlia-
	activists (4)	stalwarts	organizations	ment
	organizations	Hamas	students	group (4)
	students	Red Army	academics	Government
	intellectuals	Republican (2)	school	people
	colleges	Peace Now	think tank	bloc
	faculty members	Shining Path	dove	artist
	anti-war	bastards	bourgeoisies (2)	peer
	workforce	newscasters	network	societies (4)
	union (3)	film makers	patriot	state (6)
	rebels	script writers	MoveOn.org	circles (2)
	anti-Bush (2)	publicist	judge (2)	successor
	guardians	critics (2)	man	pro-gun
	women (2)	mass culture	hawk	church (2)
			politicians	community
			dupes	country
			elites	district
				American (2)
Other abstract concepts	bias (2)	base	bias (2)	elements
	sentiment (2)	religion	view (2)	content
	observation (2)	freedoms	version	rhetoric
	notions (2)	reprisals	theory	label
	behavior (2)	exemplar	approach	goal
	complaints	immorality	critique	interest
			foundation	thesis
			virtue	ratios
			ideal	terms
			ambivalence	purposes
			sticking point	actions
			dysfunctions	debate
			techniques	validity
			methods	case

[This page is intentionally blank.]