

## **Exploring Conflict and Violence in ESOL Writing**

**Laura E. Mendoza**

University of Texas at El Paso

ABSTRACT. Throughout the years, the study of vocabulary, and how words are combined to express particular meanings, has been well-established in language studies. In 1935, Firth argued that the meaning of a word is as much a matter of how it combines with other words in actual use as it is of the meaning it possesses in itself (O’Keeffe et al. 2007). Since then, the study of collocations has been a recurrent area of investigation in corpus linguistics. Interestingly, when writing about sensitive topics, writers, including second language learners, tend to express their opinions very thoughtfully. This study reports on the frequency with which students in an ESOL program refer to sensitive topics in their essays. In addition, specific language patterns used by students are presented. To summarize, the goal here was to explore the connections that exist among writers, language, and texts/discourse when taking into account sensitive topics.

*Keywords:* collocations, second language learners, writing, stance, discourse

**1. INTRODUCTION.** The study reported here was conducted with two main goals in mind: First, to determine patterns of lexical co-occurrence that involve vocabulary related to violence and conflict in written assignments (essays) generated by English language learners, and secondly, to determine any patterns in the use of linguistic markers of stance expressed in relation to the violence and conflict being described by the students. The data used in the study consists of language learner corpus data. In this case, a sub-sample of written texts in the UTEP's Learner Corpus of Academic English (ULCAE) was used in order to analyze the linguistic expression of stance and lexical co-occurrence in topic-specific writing. The analysis of stance focuses on attitudinal and evaluative stance, i.e., the expression of thoughts, opinions, and involvement.

This investigation differs from other studies in the sense that the corpus data used to conduct the analysis relates to a very particular pedagogical situation, since the learners involved share the same L1 and are enrolled in the same institutional program, i.e., the ESOL (English of Speakers of Other Languages) program offered by the Department of Languages and Linguistics at the University of Texas at El Paso. The motivation for conducting the study relates to the fact that many of the students represented in the ULCAE corpus live in Juarez or have family and friends living in Juarez. Given the violence that has affected this city in recent years, as well as my participation as a teaching assistant in the students' program, I considered exploring the extent to which social conflict emerges in their writing, and how they use language in relation to such events, worth doing. This was done in order to understand how students use language in patterned ways to discuss sensitive issues. Even though issues related to social conflict and violence in students' writing have been previously researched, usually to study the benefits of expressive writing after a traumatic event (e.g., Fernández & Páez 2008), the language used by students in L2 writing about these issues has not been investigated. However, studies of L1

discourse have been conducted to investigate the use of formulaic language (including collocations), in emotionally-charged situations, as reported by Wray (2012).

### **1.1. HOW ARE COLLOCATIONS AND STANCE MARKERS DETERMINED OR IDENTIFIED?**

**COLLOCATIONS:** Halliday and Sinclair (1966) were some of the early pioneers in conducting automated studies of collocations, which has shown that not only uncommon words, but also many common words, are highly attracted to one another even in an arbitrary order (O’Keeffe et al. 2007). In fact, collocations tend to be so predictable that collocational patterns indicate a strong representation of mechanics in discourse, even bigger than what was originally expected (Gledhill 2000). A common procedure used while identifying collocations is to retrieve the collocations by using a concordance program. Usually, a frequency list of a specific set of items is created in advance. The next step generally is the analysis of the corpus. Sometimes, sub-corpora are analyzed separately, depending on the purpose of the study. In this study, collocations are identified using MonoConc Pro, and a set of lexical items previously determined in a prior pilot study (Urzúa & Mendoza 2011) was used as a starting point.

**STANCE MARKERS:** Multiple methods of analysis can be employed for investigating stance markers. For example, genre analysis, corpus linguistics, contrastive analysis, quantitative analysis, and translation corpora, among others. Stance markers may be identified clause by clause by analyzing texts, for instance according to frequency. In this case, the study focuses on a relatively small set of stance markers involving first person pronouns, mental verbs (e.g., hope, consider, understand), and complement clauses (e.g., I think [that]..., we believe [that]...,) and by exploring their frequency and functions in context. These stance markers are reported as being commonly found in written discourse (Biber et al. 1999, Hyland 2005).

**1.2. WHAT CAN BE LEARNED FROM DETERMINING PATTERNS OF LEXICAL CO-OCCURRENCE AND STANCE?** Language is never random (Chaudhari et al. 2010); as a consequence, several uses of a word may be learned from determining specific lexical patterns. First of all, determining such patterns may be useful as an indicator of the way that various lexical patterns relate to one another. Besides, word combinations that co-occur recurrently in texts might mirror different underlying word senses (Biber 1993). It has been also proved that the analysis of lexical patterns may be useful to determine how much retention adult learners have about what words tend to occur together. Likewise, having knowledge of frequent collocations between words can serve as testing ground for models of acquisition (Durrant & Schmitt 2010). Interestingly, determining these patterns helps to understand how different units (words) working together may create a specific meaning.

**2. THE STUDY.** In this study, I rely on learner corpus data and corpus-based techniques to explore patterns of lexical co-occurrence, as well as linguistic stance markers commonly used in the students' writings. By focusing on lexical co-occurrences and linguistic stance markers I aimed to discover the core vocabulary that students may use, not only in a descriptive way, but most importantly, how students may use this core vocabulary while trying to express a particular stance (opinions and judgments). In this way, my goal was to explore, in a very small but meaningful way, the connection that exists among writers, language, and texts/discourse.

The expectation is that exploring the connection among those three elements in essays related to violence and conflict may shed some light on the ways in which ESOL students, as members of their discourse community, share some common lexicon and phraseology.

**3. GUIDING/RESEARCH QUESTIONS.** (1) What words tend to co-occur with frequent topic-specific words, i.e., words related to violence and conflict? (2) What do these patterns of co-occurrence tell us about the way students use, in writing, lexicon from a particular semantic field (e.g. violence and conflict)? (3) Are there any patterns in the stance the writers in this study express in relation to the violence and conflict being described?

In summary, the study aims to explore two distinct objectives. One is to determine the vocabulary that students use to write about social conflict and violence as well as the words that tend to co-occur with such topic-specific words. The second objective is to determine the way students express their stance, as writers, using specific linguistic markers in their writing assignments (essays) when these deal with sensitive topics. A preliminary list of target words from a pilot study by Urzúa and Mendoza (2011), extracted from ULCAE data, was used as a starting point to identify essays for analyses and to explore the first objective. This procedure is explained more in detail below.

#### **4. METHODOLOGY.**

**4.1. CORPUS DATA USED IN THE STUDY.** The learner data used in this study comes from the UTEP's Learner Corpus of Academic English (ULCAE). The data for the corpus was collected from learners at The University of Texas at El Paso enrolled in the ESOL (English of Speakers of Other Languages) program. Ninety-five percent of the students in the corpus/program have Spanish as first language and attended Mexican high schools. At present the corpus comprises 1,355,645 words, reflecting a variety of writing genres such as process essays, comparison and contrast essays, evaluative essays, and research reports, among others. All written samples were produced as part of the general course work of the students in their ESOL courses.

All essays are subdivided according to level; ranging from intermediate level (ESOL 1610) to sequential writing intensive courses (ESOL 1309, 1311 and 1312). It is important to mention that the last two levels are equivalent to regular English courses in the core curriculum of the university. For the purpose of the study, two different sub-sets of data have been extracted; data from ESOL 1311 and ESOL 1312, which were the levels included in the analyses. In all cases, only the final version of the essay was taken into account; in other words, the number varied from the original corpus, since the ULCAE comprises not only final versions, but also first drafts. The first sub-set (shown in Table 1, Sub-corpus A) reflects the number of essays included in the corpus to date in those two levels, ESOL 1311 and ESOL 1312. Data from this sub-sample includes a total of 278 files (essays) and 236,418 words.

The different essays that ESOL 1311 comprises are: Evaluation Essay, Problem/Solution Essay, a proposal, and a final exam. And, the essays that students write in ESOL 1312 are: Genre Analysis, Research Proposal, Literature Review, Research Paper, Synthesis Paper, and a Final Exam. The number of texts in the ESOL 1311 and ESOL 1312 sections of the ULCAE, including only final drafts, by level/semester is shown below:

<b>Sub-corpus A</b>				
<b>Course</b>	<b>Semester / Year</b>	<b>Number of</b>		
		<b>Sections</b>	<b>Texts</b>	<b>Words</b>
<b>ESOL 1311</b>	Fall 2010	3	41	<b>22,718</b>
<b>ESOL 1311</b>	Spring 2011	4	75	<b>47,271</b>
<b>ESOL 1312</b>	Spring 2010	4	122	<b>129,679</b>
<b>ESOL 1312</b>	Spring 2011	3	40	<b>36,750</b>
	<b>TOTALS</b>	<b>14</b>	<b>278</b>	<b>236,418</b>

TABLE 1. Number of texts in ESOL 1311 and 1312 in the ULCAE

In order to create the second sub-set of data, i.e., the data for analysis, henceforth referred to as the 'V' sub-corpus, the essays that include words related to violence and conflict, extracted from the first sub-corpus (sub-corpus A) mentioned above, were identified and extracted, using the procedure described below.

**4.2. IDENTIFYING ESSAYS FOR ANALYSIS ('V' SUB-CORPUS).** In order to identify the essays to be included in the 'V' sub-corpus, the automated text search feature of the concordance software MonoConc Pro was used to locate the presence of any of the target items (from Urzúa & Mendoza 2011) in the original data (sub-corpus A). The data yielded by MonoConc was thus used to identify essays to include in the sample to be used in the analyses ('V' sub-corpus). In addition, once a preliminary sample of essays was determined, frequency lists of all the words found in this preliminary sample were manually explored in order to identify any potentially relevant lexical items not included in the original list used in Urzúa and Mendoza (2011). As a consequence, after locating the new lexical items in the list, another search was performed in order to identify any missing essays and include them in the corpus for analysis.

As shown in Table 2, this second sub-set includes a total of 51 files (essays) and 50, 145 words. This sample constitutes the focus of the analyses. The remaining essays, that is, those essays in the ESOL 1311 and ESOL 1312 sections of the ULCAE corpus that do not include terms (vocabulary) related to social conflict and violence constitute a third sample of essays, which was used here for comparison purposes. This third sample, henceforth referred to as the 'Non-V' sub-corpus is shown in Table 3.

<b>Sub-corpus ‘V’</b>				
<b>Course</b>	<b>Semester / Year</b>	<b>Number of</b>		
		<b>Sections</b>	<b>Texts</b>	<b>Words</b>
<b>ESOL 1311</b>	Fall 2010	3	7	<b>4,079</b>
<b>ESOL 1311</b>	Spring 2011	4	19	<b>11,363</b>
<b>ESOL 1312</b>	Spring 2010	4	16	<b>23,975</b>
<b>ESOL 1312</b>	Spring 2011	3	9	<b>10,728</b>
	<b>TOTALS</b>	<b>14</b>	<b>51</b>	<b>50,145</b>

TABLE 2. Number or texts included in the ‘V’ sub-corpus

<b>Sub-corpus ‘Non-V’</b>				
<b>Course</b>	<b>Semester / Year</b>	<b>Number of</b>		
		<b>Sections</b>	<b>Texts</b>	<b>Words</b>
<b>ESOL 1311</b>	Fall 2010	3	34	18,639
<b>ESOL 1311</b>	Spring 2011	4	51	32,765
<b>ESOL 1312</b>	Spring 2010	4	108	109,750
<b>ESOL 1312</b>	Spring 2011	3	31	26,022
	<b>TOTALS</b>	<b>14</b>	<b>224</b>	<b>187,176</b>

TABLE 3. Number of texts included in the ‘non-V’ sub-corpus

## 5. ANALYSES AND RESULTS.

### 5.1. PERCENTAGE OF ESSAYS DEALING WITH TOPICS RELATED TO SOCIAL CONFLICT AND VIOLENCE.

The first stage in the analysis aimed at determining the percentage of essays from the original ESOL 1311/ESOL 1312 sub-sections of the ULCAE that deal with topics related to social conflict and violence. Table 4 below shows how often students’ writings reflect such topics.

Course	Semester / Year	Number of Texts		
		Sub-corpus A	Sub-corpus 'V'	%
<b>ESOL 1311</b>	Fall 2010	41	7	<b>17</b>
<b>ESOL 1311</b>	Spring 2011	75	19	<b>25</b>
<b>ESOL 1312</b>	Spring 2010	122	16	<b>13</b>
<b>ESOL 1312</b>	Spring 2011	40	9	<b>23</b>
	<b>TOTALS</b>	<b>278</b>	<b>51</b>	

TABLE 4. Texts that deal with topics related to violence

Tables 5 and 6 show the percentage of essays that were written in response to a prompt (assigned by the teacher or selected by the student) that dealt with topics directly linked to conflict and violence, and the percentage of essays that were composed in response to a prompt that did not deal with social conflict or violence, but in which students themselves shifted the topic from non-violent to violent events in their texts. In other words, essays with a non-sensitive topic (e.g. the use of Facebook among college students) in which the student-writer brings up an incident, anecdote or example that relates to conflict or violence (e.g. a person becoming a potential kidnapping victim because of private information posted in Facebook).

Course	Semester / Year	Number of Texts		
		Sub-corpus A	With topics related to violence	%
<b>ESOL 1311</b>	Fall 2010	41	7	<b>17</b>
<b>ESOL 1311</b>	Spring 2011	75	13	<b>17</b>
<b>ESOL 1312</b>	Spring 2010	122	8	<b>7</b>
<b>ESOL 1312</b>	Spring 2011	40	5	<b>13</b>
	<b>TOTALS</b>	<b>278</b>	<b>33</b>	

TABLE 5. Texts with topics directly related to violence

Course	Semester / Year	Number of Texts		
		Sub-corpus A	With terms related to violence	%
<b>ESOL 1311</b>	Fall 2010	41	0	<b>0</b>
<b>ESOL 1311</b>	Spring 2011	75	6	<b>8</b>
<b>ESOL 1312</b>	Spring 2010	122	8	<b>7</b>
<b>ESOL 1312</b>	Spring 2011	40	4	<b>10</b>
	<b>TOTALS</b>	<b>278</b>	<b>18</b>	

TABLE 6. Texts with terms directly related to violence

**5.2. FREQUENCY OF OCCURRENCE OF TERMS RELATED TO VIOLENCE AND CONFLICT IN ‘V’ SUB-CORPUS.** Once the sample of essays to be used in the analysis (‘V’ sub-corpus) was determined, the frequency with which each item in the list of terms related to violence and conflict was obtained, recording frequencies of occurrence in each individual essay, also by using the MonoConc Pro concordance program.

**5.3. SEARCHING TARGET LEXICAL ITEMS IN ‘V’ SUB-CORPUS.** As described above, a pre-determined list of lexical items from a previous pilot study (Urzúa & Mendoza 2011) was first used to identify as many essays as possible in sub-corpus A that included any explicit references to violence or conflict. This list was later expanded with relevant items used in the essays being identified for inclusion in the ‘V’ sub-corpus. Therefore, the procedure used to determine the list of lexical items to be used in the study was a recursive one: a preliminary list used to identify essays to be analyzed, and these essays used in turn to compile a more complete list of relevant vocabulary items to include in the lexical analysis.

MonoConc Pro allows the identification of the exact location of the lexical item within each specific essay. It also allows the researcher to look just for specific parts of a word, as a

consequence, special characters (such as ‘\*’) were frequently used while conducting the searches in order to avoid, first of all, misspellings of the words, and secondly, different variants of the target items. For example, when looking for the word BELIEVE, the given string of letters was BELIE\*, which gives as a result a search for all the different possibilities that those string of letters may offer (e.g., BELIEF(S), BELIEVE and BELIEVED).

When investigating the target lexical items, such a procedure was used. For example, when exploring the lexical item GUN with MonoConc Pro, the software allows you to provide just a string of characters, like \*GUN\*, which the program analyzes in order to provide all the possible cases in which those characters appear. To illustrate, \*GUN\* provides lexical items such as GUN, GUNS, HANDGUN, and HANDGUNS, among others. As mentioned before, the software allows the researcher to locate the specific item within a certain text providing the context in which the item occurs. For instance, the string of characters \*GUN\* provides information of the immediate context in which the words occur in each essay, as in ...THE EFFECTIVENESS OF GUN LAWS REMAINS UNPROVEN... and identified the corresponding essay (Essay ID Number: 1311-2-11-0266-4.2). This kind of search was performed for each of the target items.

**5.4. MOST FREQUENT LEXICAL ITEMS RELATED TO CONFLICT AND VIOLENCE.** The results show the frequency of occurrence of the five most frequent lexical items (i.e. the number of tokens) first in ‘V’ sub-corpus and then, for comparative purposes, in the ‘Non-V’ sub-corpus.

Table 7 shows the five most frequently occurring target items in the ESOL 1311 essays included in the ‘V’ sub-corpus. As shown, the most frequent lexical item was the word drug (including the plural form, DRUGS, as well as its use as a modifier, as in DRUG DEALER (S), DRUG DEALING, DRUG LORDS, and DRUG BUSINESS), with 142 occurrences. The second most-frequent word was CRIME including variants such as CRIMES, CRIMINAL (S), and CRIMINALIZE with 97

occurrences. The third one in the ranking is GUN, with variants GUNS and HANDGUN, with 85 occurrences, followed by VIOLENCE/-VIOLENT, with 47 occurrences and by government(s), with 30 occurrences.

Table 8 shows the five most frequently occurring target items in the ESOL 1312 essays included in the ‘V’ sub-corpus. As shown, the most frequent lexical items are: drug and its variants, with 423 occurrences, followed by violence/violent, with 158 occurrences. The third lexical item in ‘V’ sub-corpus refers to Mexico with 101 occurrences, and the fourth one is marijuana with 88 occurrences. The last item among the top five is government(s), with 57 occurrences.

<b>ESOL 1311</b>		
<b>Rank</b>	<b>Lexical Items</b>	<b>Number of Occurrences</b>
<b>1</b>	drug (s)	142
<b>2</b>	crime (s) [criminal(s), criminalize]	97
<b>3</b>	gun (s) [handgun]	85
<b>4</b>	violence/violent	47
<b>5</b>	government (s)	30

TABLE 7. Number of occurrences of target lexical items in ‘V’ sub-corpus

<b>ESOL 1312</b>		
<b>Rank</b>	<b>Lexical Items</b>	<b>Number of Occurrences</b>
<b>1</b>	drug (s)	423
<b>2</b>	violence/violent	158
<b>3</b>	Mexico	101
<b>4</b>	marijuana	88
<b>5</b>	government (s)	57

TABLE 8. Number of occurrences of target lexical items in ‘V’ sub-corpus

As can be seen in tables 7 and 8, three out of the five most used lexical items in both sections, ESOL 1311 and ESOL 1312, in 'V' sub-corpus are the same: drug(s), violence/violent, and government(s). More interestingly, in both cases the most frequent lexical item is drug(s), with 142 (ESOL 1311) and 423 (ESOL 1312) instances. In both cases, government(s) appears as the fifth most frequently used lexical item.

**6. PATTERNS OF CO-OCCURRENCE AND COLLOCATIONAL ANALYSES.** Analyses of word combinations were conducted in order to explore any existing patterns of co-occurrence and collocational frames involving the most frequent target words. The purpose for doing this was to gain a better understanding of how these frequent words combined with other words and what these could tell us about the way students refer to social conflict and violence in their writing.

**6.1. PROCEDURE USED TO IDENTIFY EXISTING COLLOCATIONAL FRAMES.** Once the frequencies of occurrence were determined, co-occurrence and collocational analyses were conducted. MonoConc Pro provides collocational information, such as the most frequent collocates to the left and to the right of specific lexical items. In other words, the software provides lists of the items that tend to co-occur with relative frequency with target words.

To illustrate, the chart below shows how MonoConc presents the most frequent collocates for a particular target word. For this example, the lexical item explored was DRUG(S). As shown in figure 1, the most frequent collocates of DRUG(S) are the preposition OF (18 occurrences) and the definite article THE (17 occurrences) (one collocate to the left), and forms of the verb BE, ARE and WERE (7 occurrences each) (one collocate to the right).

3-Left	2-Left	1-Left	1-Right	2-Right	3-Right
13 the	12 of	18 of	7 are	7 legal	6 a
9 to	7 re-legalization	17 the	7 were	7 have	6 the
7 of	6 medications	9 illicit	6 in	5 not	6 to
5 and	5 to	7 legalizing	5 is	5 be	5 many
4 in	5 that	7 if	5 would	4 the	4 if
4 that	4 these	6 as	4 dealers	4 to	4 be
3 be	4 and	5 that	4 to	4 as	4 that
3 means	4 by	4 new	4 store	4 can	4 of
2 for	3 use	4 and	4 will	3 is	3 can
2 the	3 a	4 with	4 the	3 very	3 this
2 money	3 get	3 these	4 and	3 and	2 would
2 not	3 if	3 by	4 this	3 in	2 has
2 would	2 conclusion	3 a	4 stores	3 are	2 there
2 is	2 instance	3 any	3 there	2 that	2 in
2 also	2 for	2 illegal	3 I	2 will	2 sell
2 use	2 legalization	2 because	3 as	2 all	2 have
2 they	2 need	2 be	3 they	2 people	2 available
2 get	2 as	2 medical	3 users	2 means	2 cocaine
1 leads	1 affected	2 selling	2 but	2 problem	2 legal
1 legalization	1 pain	2 in	2 at	2 legalized	2 1001
1 muscular	1 sell	2 legal	2 should	2 would	2 opinion
1 only	1 know	2 drugs	2 abuse	2 a	2 new
1 any	1 problem	2 about	2 dealer	2 solution	1 arguments
1 pros	1 ridiculous	2 legalize	2 can	2 cause	1 while
1 medication	1 down	1 but	2 also	2 my	1 no
1 president	1 law	1 consuming	2 one	1 could	1 best
1 may	1 legalizing	1 big	2 companies	1 change	1 other
1 medications	1 front	1 use	2 cannot	1 summarize	1 dramatically
1 adults	1 the	1 forcing	2 have	1 won't	1 tiber
1 producing	1 be	1 activity	2 such	1 if	1 allow

FIGURE 1. Collocational frame for drug(s)

7. STANCE. As mentioned in the Background section of this paper, stance is understood as the speaker/writer's expression of thoughts, attitudes, value judgments, or assessments (Biber et al. 1999). For instance, the tone used by a writer may be more or less formal, authors may write in more personal or impersonal ways (distancing themselves from their texts), agents may become focalized or defocalized, and so on. A clear example of making a text more personal and taking a more involved stance as writers is the frequent use of first person pronouns, so writers can position themselves in particular or strategic ways toward their texts. Pronouns then serve not only exclusively as referents, but also have illocutionary and perlocutionary effects, in terms of locating agency, moral responsibilities, duties, points of view, etc. (Yates & Hiles 2010).

Given the importance of first person reference in the explicit expression of stance, a first stage in this part of the analysis was to focus on first person reference (I, WE, ME, OUR). After that, the use of first person pronouns with complement clauses headed by mental or attitudinal verbs (e.g. THINK, HOPE, BELIEVE, AGREE, DOUBT, UNDERSTAND) is analyzed; and finally, nominal

expressions referring to attitude or thoughts/ideas, also in combination with first person reference are examined (e.g. IN MY OPINION, OUR BELIEFS).

**7.1. USE OF FIRST PERSON SINGULAR AND PLURAL PRONOUNS.** A frequency analysis of first person subject pronouns and possessives indicates noticeable differences between the way these elements are used in ESOL 1311 and ESOL 1312, as well as between its use in the ‘V’ corpus and ‘Non-V’ corpus.

ESOL 1311				
	Sub-corpus 'V'		Sub-corpus 'Non-V'	
	Number of		Number of	
	Essays	Words	Essays	Words
	25	14,846	86	52,000
Pronouns	Tokens	Occurrences per 1000 words	Tokens	Occurrences per 1000 words
<b>I</b>	45	3.03	310	5.96
<b>MY</b>	12	0.8	103	1.98
Sub-total	<b>57</b>	<b>3.83</b>	<b>413</b>	<b>7.94</b>
<b>WE</b>	75	5.05	507	9.75
<b>OUR</b>	36	2.42	189	3.63
Sub-total	<b>111</b>	<b>7.47</b>	<b>696</b>	<b>13.38</b>

TABLE 9. Number of occurrences of first-person reference in ESOL 1311

Table 9 shows the differences in first-person reference in each of the sub-samples in ESOL 1311. It indicates that when students in ESOL 1311 refer to themselves in their writing, they tend to use first person plural forms (we, our) twice as often in relation to their use of first person singular forms (I, my). In addition, when writing about violence/conflict, students seem to adopt a more distanced (impersonal) stance, as they use less first person reference than when writing about other topics. A similar pattern was also observed in ESOL 1312 essays, as shown in table 10.

ESOL 1312				
	Sub-corpus 'V'		Sub-corpus 'Non-V'	
	Number of		Number of	
	Essays	Words	Essays	Words
	25	34,703	139	135,772
Pronouns	Tokens	Occurrences per 1000 words	Tokens	Occurrences per 1000 words
<b>I</b>	46	1.3	269	1.98
<b>MY</b>	20	0.57	91	0.67
Sub-total	<b>66</b>	<b>1.87</b>	<b>360</b>	<b>2.65</b>
<b>WE</b>	80	2.31	500	3.7
<b>OUR</b>	31	0.89	224	1.65
Sub-total	<b>111</b>	<b>3.2</b>	<b>724</b>	<b>5.35</b>

TABLE 10. Number of occurrences of first-person reference in ESOL 1312

**8. CONCLUSION.** The results obtained in this exploratory study show that students in ESOL 1311 and ESOL 1312 often write about situations involving social conflict and violence. In fact, these topics may be found in as much as 25% of the essays generated in one of these courses in a given semester. The examination of 'V' sub-corpus by means of frequency analyses, using the text retrieval program MonoConc Pro made possible to identify those lexical items that represent the core vocabulary students use when writing about social conflict and violence. This core vocabulary used in the identified texts includes lexical items such as: DRUGS, CRIMES, GUNS, MARIJUANA, MEXICO, CARTEL, and POWER, among many others. Three of these items, DRUGS, VIOLENCE, and GOVERNMENT, were ranked among the five top most frequent items for both ESOL 1311 and ESOL 1312. These items served as starting point for the collocational analyses.

The patterns of co-occurrence identified and illustrated in this paper tend to occur with relative frequency, as evidenced by their percentage of use in relation to the total number of instance a lexical items occurred in the data, and in comparison to the uncountable number of possible combinations that any target item given may have. These patterns show, for example,

that drugs are often discussed in relation to their legal status, and that whenever the word ‘crime’ is used, it often tends to be preceded by the adjective ‘violent.’ The patterns found also indicate that students often write about drugs in terms of what should, or must, or has to be done about them, either to try to reduce the number of drug-related violence, or to solve the many problems associated with it.

In relation to the analysis of personal stance, it was found that, if students refer to themselves in their writing, they tend to use first person plural forms (we, our), which occur twice as often as first person singular forms (I, my) in the data. Furthermore, when writing about violence/conflict, students seem to adopt a more distanced (impersonal) stance. In addition, students tend to use we-statements to express need, possibility, or conditionals, and I-statements to express agreement or disagreement.

What these results suggest, even when they are limited to a particular semantic field and a small set of stance markers here, is that studies like this one can help us understand how specific units of language may become more salient to learners. In other words, this type of study can enable us to understand how smaller (or bigger) units can be combined and recombined into larger collocations based upon different points of the acquisition process.

**9. LIMITATIONS AND FURTHER RESEARCH.** A limitation of this study was that data used in the analyses were extracted from only two sub-sets of students’ essays in the ULCAE corpus. Another limitation is the fact that a limited sub-set of stance markers was examined here. A first possible step to expand the study would be to include more of the course levels in the ESOL program represented in the ULCAE corpus. This can be of interest given that the vocabulary and level of writing proficiency of students in the beginning levels of ESOL (classes such as ESOL

1610) have, without a doubt, different linguistic resources than those students enrolled in higher-level intensive writing courses (ESOL 1311 and ESOL 1312). Additional research can also be conducted to explore potential differences in the marking of stance depending on the gender of the writer. As the literature suggests, women express greater sadness or depression than men when using expressive writing to cope with an unpleasant situation (e.g. Fernández et al. 2008). To do that, it may be necessary, though, to examine a larger set of linguistic markers of stance. In sum, much remains to be done. This paper is a first exploration into how essays composed by ESOL students at UTEP write about situations dealing with social conflict and violence. These are topics that, unfortunately, touch them in direct and indirect ways. As an instructor, former ESOL and UTEP student, and research assistant, they have become very important to me as well.

## REFERENCES

- BIBER, DOUGLAS. 1993. Co-occurrence patterns among collocations: A tool for corpus-based lexical knowledge acquisition. *Computational Linguistics*.19 (3). 531-538.
- BIBER, DOUGLAS, STIG JOHANSSON, GEOFFREY LEECH, SUSAN CONRAD, AND EDWARD FINEGAN. 1999. *Grammar of Spoken and Written English*. Harlow: Pearson Education.
- CHAUDHARI, DIPAK, L. DAMANI, P. OM, AND SRIVATSAN LAXMAN. 2010. Lexical co-occurrence, statistical significance, and word association. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2011.
- DURRANT, PHILIP AND NORBERT SCHMITT. 2010. Adult learners' retention of collocations from exposure. *Second Language Research*. 26 (2), 163-188.
- FERNÁNDEZ, ITZIAR, DARIO PÁEZ, AND JAMES, W. PENNEBAKER, ET AL. 2008. Comparison of expressive writing after the terrorist attacks of September 11<sup>th</sup> and March 11<sup>th</sup>. *International Journal of Clinical and Health Psychology*. 9 (1), 89-103.
- FERNÁNDEZ, ITZIAR AND PÁEZ, DARIO. 2008. The benefits of expressive writing after the Madrid terrorist attack: implications for emotional activation and positive affect. *British Journal of Health Psychology* 13.31-34.
- GLEDHILL, CHRIS. 2000. The discourse function of collocation in research article introductions. *English for Specific Purposes*. 19, 115-135.
- HYLAND, KEN. 2005. Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192.
- O' KEEFFE, ANNE, MICHAEL MCCARTHY, AND RONALD CARTER. 2007. *From corpus to classroom*. Cambridge University Press. UK.

- SINCLAIR, JOHN. 1966. Beginning the study of lexis. In Bazell, C.E. et al. (Eds.). In Memory of J.R. Firth. Longman, London, 410- 31.
- URZÚA, ALFREDO AND LAURA MENDOZA, E. 2011. Border Violence and Social Conflict in Students' Writing at UTEP. Center for Inter-American and Border Studies Research Seminar, El Paso, TX.
- WRAY, ALISON. 2012. What do we think (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics*, 32, 231-254.
- YATES, SCOTT AND DAVID HILES. 2010. 'You can't' but 'I do': Rules, ethics, and the significance of shifts in pronominal forms for self-positioning in talk. *Discourse Studies*. 12(4), 535-551.