Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi
Mustafa Kemal University Journal of Social Sciences Institute
Yıl/Year: 2013 ♦ Cilt/Volume: 10 ♦ Sayı/Issue: 21, s. 37-54

# GENRE-SPECIFIC SEMANTIC PROSODY: THE CASE OF POSE

**Yrd. Doç. Dr. İhsan ÜNALDI**
Gaziantep Üniversitesi, Eğitim Fakültesi, Yabancı Diller Eğitimi Bölümü
ihsanunaldi@gmail.com

**Abstract**

*This study tries to deal with the semantic prosody of the word pose. The term semantic prosody refers to the fact that lexical items, by nature, tend to occur with positive, negative, or neutral collocations. For example the word cause has a negative semantic prosody (Stubbs, 1995) as it frequently collocates with words such as damage, problem, or pain. In this study, to see whether the target word pose has such a tendency, it was processed in a 464-million-word corpus (Davies, 2008) which is made up of five different genres namely: spoken language, fiction, magazines, newspapers, and academic journals. Collocational variety of pose in academic contexts was compared with the other genres mentioned. The results suggested that it actually has a negative semantic prosody in academic contexts, while in other contexts it tends to collocate mostly with neutral lexical items.*

**Keywords:** *Semantic Prosody, Collocation, Academic Context.*

## TÜRE HAS ANLAMSAL BÜRÜN: İNGİLİZCE *POSE* SÖZCÜĞÜNÜN DURUMU

**Özet**

*Bu çalışma, İngilizce pose sözcüğünü anlamsal bürün açısından çözümleyen derlem tabanlı analizleri kapsamaktadır. Anlamsal bürün terimi sözcüklerin doğaları gereği olumlu, olumsuz yada yansız sözcüklerle birlikte kullanıldıklarını ifade eder. Örneğin, İngilizce cause (neden olmak) sözcüğü sıklıkla hasar, sorun yada acı gibi olumsuz sözcüklerle beraber kullanılır (Stubbs, 1995). Bu çalışmada, İngilizce pose (duruş, durma, poz vermek) sözcüğünün böyle bir eğilimi olup olmadığını belirlemek için, bu sözcük 464 milyon sözcük içeren bir derlemde işlenmiştir. Bu derlem, İngilizce günlük konuşma dilinden, kurgu roman ve öykülerden, popüler dergilerden, gazetelerden ve akademik dergilerden oluşturulmuştur (Davies, 2008). Pose sözcüğü, eşdizim çeşitliliği açısından akademik ve diğer bağlamlarda karşılaştırılmıştır. Sonuçlar, bu sözcüğün akademik bağlamlarda olumsuz sözcüklerle birlikte kullanıldığını, fakat diğer bağlamlarda yansız sözcüklerle kullanıldığını ortaya çıkarmıştır.*

**Anahtar kelimeler**: *Anlamsal Bürün, Eşdizim, Akademik Bağlam.*

## Introduction

### *Corpus Linguistics*

Corpus (plural *corpora*) is generally defined as the digitalized collection of spoken or written registers of a given language. Natural instances of a language are digitalized and stored into a database and retrieved to make calculations. Six types of corpora are mentioned in the related literature— specialized, multilingual, reference, parallel learner, diachronic, and monitor (see Tognini-Bonelli, 2010 for details. Specialized types of corpora include collections that are gathered from a specific group and/or for a specific purpose. The collection of Child Language Data Exchange System (CHILDES) is an example for this type of corpus. Natural instances of child language are collected from a variety of languages which makes it a multilingual corpus as well. Reference corpora also consist of collections of a language, but it tries to deal with the general nature of that language. British National Corpus (BNC) could be counted as an example of this kind of corpora. A parallel corpus is a collection of two or more corpora that were sampled in the same way from different languages for comparison or contrast. A learner corpus (LC), on the other hand, is a mixture of these types of corpora. Language samples are collected from specific language learner groups and from different languages. These samples are compared and contrasted with the samples that are collected from the productions of the native speakers of that language. The International Corpus of Learner English (ICLE) is a well-structured collection of language samples gathered from EFL (English as a Foreign Language) learners from a wide range of languages. A diachronic corpus is used to analyze language change in general, and this sort of corpus makes it possible to track the changes occurring in a language over a period of time. Such a corpus, Helsinki Corpus of English Texts: Diachronic Part, includes 400 text samples written in English from AD 750 to 1700. The last type, monitor corpus builds up in time with new addition to the corpus at regular intervals. Bank of English (BoE), which offers opportunities to analyze recent and ongoing changes in the English Language, is a well-known example of this kind of corpus.

The first, naturally primitive, corpus appeared when it was first realized that language is something observable. This point of view claiming that we can observe language and make deductions about it has been around for a relatively long period of time with certain changes in its paradigm. The famous distinction between *parole* and *langue* which was made by Saussure (Finch, 2005) was one of the turning points in the perception of language in general. The *real* language that is being used was referred to as *Parole;* and it was seen as the outgrowth of *langue* which is a much more complex system.

Similarly, with a cognitive-algorithmic model, Noam Chomsky (1965) tried to attribute humans' knowledge and ability concerning language to an unconscious system which could enable production in any. In other words, humans are endowed with universal on-off switches for grammar whereby any human

language is acquired; the focus is on language competence. On the other side, there are supporters of a performance model which takes into consideration the actual psychological and physical processes involved during language production.

From a Chomskyan perspective, "...the corpus linguist aims for descriptive adequacy (a lower level of adequacy), and it is arguable whether explanatory adequacy is even achievable through corpus analysis." (Meyer, 2004). What is meant here is that from the three types of adequacy levels, i.e., observational, descriptive, and explanatory (Chomsky, 1965) corpus linguistics lacks the means to reach higher levels of adequacy and cannot go beyond descriptive adequacy level which can be reached by means of generative grammar. The underlying reason for such an approach is that the real language is actually quite *messy* with false starts, hesitations, and speech fragments. At this point, corpus linguists claim that "a word in or on itself does not carry any meaning, but the meaning is often made through several words in a sequence" (Sinclair, 1991); and "[t]he aim is not to study idiosyncratic details of performance which are, by chance, recorded in a corpus. On the contrary, a corpus reveals what frequently recurs, sometimes hundreds or thousands of times, and cannot possibly be due to chance" (Stubbs, 2004, p. 111). Actually, this point is noteworthy because discussions concerning the representativeness of language samples focus on whether or not we could make deductions about language in general solely based on frequency.

Today, technological developments have made digital data storage easier and compiled language data has become more accessible. In 1960's the first electronic corpus, Brown Corpus, was compiled. It contained digitalized documents comprising of a million words and it is still in use today. Applications from different areas of study derived from corpus linguistics; among them the most notable ones are lexicography, translation, stylistics, grammar, gender studies, forensic linguistics, computational linguistics and language teaching (Tognini Bonelli, 2001).

***Collocations***

The concept of collocation can be traced back to Firth (1935). He claimed that it is hard for a lexical item to create meaning on its own but meaning is realized through the connections that the lexical item creates. For instance, the word *plane* is almost free of meaning until it combines with another word like *delay*. Upon hearing the statement 'Our *plane* is *delayed*.', we deduce that the word *plane* refers to a flying vehicle. However, the same lexical item creates a different meaning when used in a different context like the following: They were on completely different intellectual *planes*. In this sentence, the same word, *plane,* refers to a level or standard of thought, dignity, or character. As one might guess, this is actually a probabilistic approach to lexicon not a deterministic one. Therefore, the term collocation can basically be defined as a habitual co-occurrence of lexical items.

However, collocations are generally regarded as lexical items displaying immediate occurrences. For example, you might hear people getting *almost killed* in accidents, people who are *happily married*, or things that don't *make sense*. These fixed expressions occur naturally both in spoken and written registers, and they are so frequent that it would be unrealistic to consider them as peculiarities. From this perspective, language is made up of such prefabricated units bound together with grammar. Some researchers even claim that these structures actually make language fluency possible by reducing the processing effort (Aitchison, 1987 and Partington, 1996), and despite its probabilistic nature, language consists of phraseological recurrent patterns (*see Altenberg, 1998; Sinclair, 1991, and Stubbs 2001*)

When it comes to learning a second language, It could be claimed that dealing with collocations is one of the most challenging areas. This is particularly valid for English as it is full of collocations. In English, knowing the meaning of a word does not only require knowing its dictionary definition; learners should also know the type of words with which it is often related or connected. Collocations, either fixed or more flexible, are the result of many years of habitual use by fluent speakers of the English (Prodromou, 2004).

Collocations have been identified as one of the ways that differentiate native speakers and second language learners. Complex ideas are hard to express unless one can use simple vocabulary in a range of collocations (Lewis, 1993). Wray (2002) highlights that collocations are particularly important for learners striving for a high degree of competence in a second language, because they enhance not only accuracy but also fluency. Nesselhauf (2003) states that "Collocations are of particular importance for learners striving for a high degree of competence in the second language but they are also of importance for learners with less ambitious aspirations, as they not only enhance accuracy but also fluency" (p.223).

***Collocation learning & teaching***

The importance of collocations for successful language learning was recognized over seventy years ago (Palmer, 1933). Nation (2001) abridges that;

- language knowledge is collocational knowledge;
- collocational knowledge is important for developing both fluency and accuracy;
- knowing a word involves knowing its set of collocates.

Collocation knowledge is difficult to acquire simply because there is so much of it. Native speakers carry hundreds of thousands—possibly millions—of lexical chunks in their heads, ready to draw upon in order to produce fluent, accurate and

meaningful language (Lewis, 1997). This generally causes a discouraging challenge to language learners. Collocation learning is a collective process that involves great attempt more than rote memorization. Collocation dictionaries and concordancers therefore potentially supply a useful tool for the learning and teaching of collocations.

Although it is widely recognized that collocations in language learning is crucial, there are still discussions about how they should be taught. The general recognition includes three aspects:

1. awareness raising;
2. collocation selection;
3. learning strategies.

Many researchers believe that collocations should be learned consciously. The first step should be to draw students' attention to their existence. Nation (2001) suggests that teachers should encourage students to split text containing familiar items into chunks and seek patterns in them. Lewis (1997) proposes that important collocations are presented in the classroom and students trained to learn them in their entirety and break them into parts later. Conzett (2000) promotes selecting books that include many collocations and training students to observe and note as many as possible through reading, and reinforce them in their writing.

Collocation learning is challenging, and to develop efficient learning strategies learners generally need help and support. In the classroom, collocation dictionaries can be used and examples can be recorded in their notebook while exploring text or preparing essays. Additionally, computer concordancers can expose learners to collocations in natural recurring contexts. Hoey (2000) suggests using concordancers to study the same collocations in different texts, and to find keywords in a text and learn how they combine with other words in context.

Teaching collocations in the classroom is likely to help students overcome problems of vocabulary, style and usage (Leed & Nakhimovsky, 1979). Leed and Nakhimovsky (1979) argue that vocabulary exercises should be based on the findings of a well-structured lexical analysis, in the same way that pronunciation exercises are based on phonology. Such an approach would help foreign language learners with problems of vocabulary, style and usage, and give teachers a method to produce and carry out lexical exercises in the classroom, as well as concentrate on the teaching of restricted collocations such as 'heavy drinker', 'heavy smoker', 'deep trouble', etc., (Leed & Nakhimovsky, 1979: p. 109).

***Semantic Prosody***

The term *semantic prosody* (SP) is generally attributed to Bill Louw (1993). However, before the term was coined, the *prosody* part of it had already been in use. Firth (1957) used *prosody* in terms of phonology to refer to the semantic coloring that spreads beyond segmental boundaries. He claimed that, during the natural course of speaking, the rhythm, speakers voice and the choice of pitch level are not idiosyncrasies. Utterances of native speakers contained certain patterns in these terms. For example, in the word *animal* the sound æ is endowed with a nasal quality because of the nasal sound '*n*' following it. Through these prosodic properties of speech, infants, in the very early stages of acquisition process, are able to discriminate one language from another (Guasti, 2002). Similarly, in semantics terms, lexical items compose "a consistent aura of meaning" Louw (1993). For example the word 'cause' has a strong tendency to co-occur with words that usually have negative meanings like damage, problems, pain etc. (Stubbs, 1995). In the natural course of discourse, this word and many others like it are followed by undesirable lexical items. According to Sinclair (1996) "The initial choice of SP is the functional choice which links meaning to purpose, all subsequent choices within the lexical item relate back to the prosody." (p. 86). In this statement, a clear emphasis is laid on the functionality of SP. Therefore, we could assume that SP actually operates by linking meaning to a given purpose playing a leading role in the integration of a lexical item with its context (ibid.).

Prosodies are generally either positive or negative; it means that a lexical item is either surrounded by some words with negative meanings or it has words which have positive associations as neighbors. Naturally, if a lexical item has no positive or negative associations in its contexts, it could be labeled as neutral. From this perspective, Stubbs (1996) classified semantic prosody into three main categories positive, negative, and neutral. In this study, he revealed negative semantic prosodies in causation. After the analysis of a collection of 40,000 examples of the lemma *'cause'*, he claimed that it collocates mostly with negative concepts such as *accident, concern, damage, death and trouble*.

Partington (1998), in a similar study, claims that the adjective *impressive* collocates with items such as *achievement*, *talent* and *dignity*, and has a positive SP. On the other hand the word *rife* appears to collocate with negative words like *crime*, *misery*, or *disease*. With a slightly different point of view from Stubbs, he defines SP as "the spreading of connotational coloring beyond single word boundaries" (p. 68). From this perspective, a bolder emphasis is placed on the association between SP and connotations.

According to Tribble (2000), SP could be both examined in terms of the whole language, or it could be subject to a specific genre or context. The analysis of the word *experience* revealed that for a given word a *local semantic prosody* was possible in a specific genre. Similarly, Hunston and Francis (2000) state that lexical items might be said to have particular SPs if they can be shown in particular

semantic sets. Interestingly, if a lexical item is used with odd collocations it becomes a subject of humor or irony. For example, the famous quotation from Peter Ustinov "Comedy is simply a funny way of being serious." is humorous because the word *funny* collocates with being serious in a single clause, which makes the statement sound humorous.

From the studies mentioned so far, about 20 lexical items in English have been reported to have certain types of semantic prosodies -negative, positive, or positive. The following table illustrates some of them.

**Table 1.** *Some of the lexical items with negative & positive prosodies*

| | **Negative SP** | **Positive SP** |
|---|---|---|
| **Sinclair (1991)** | break out, happen, set in, bent on | |
| **Stubbs (1995, 1996)** | accost, cause, signs of | provide, career |
| **Partington (1998, 2004)** | commit, peddle/peddler, rife | impressive |

As can be seen from the table, in these studies only three words, *provide*, career and impressive, are reported to have positive SPs; all the other words which were studied have negative SPs. It means that the lexical items with negative SPs tend to co-occur mostly with negative words. Let's try to make sense out of the following sentences.

1) Children *cause* happiness at home.

2) After an *impressive* accident on the 5$^{th}$ avenue, the traffic didn't move for hours.

3) A frequently dry mouth and deteriorating vision might be the *signs of* diabetes.

It should be obvious that the first two sentences don't sound as natural as the last one because, statistically speaking, the word *cause* (1) needs negative neighboring words and the word *impressive* (2) fits in mostly with positive collocations. The phrase *signs of* (3), on the other hand, reflects negativity collocating with a disease, diabetes.

***The case of 'pose'***

In the Oxford Dictionary, the word 'pose' is defined as *to present or constitute (a problem or danger), or assume a particular position in order to be photographed, painted, or drawn*. From my personal observations, the word *pose* seems to differ in terms of collocations from context to context. In academic contexts, something poses a threat or danger in a specific situation. The following

statement, for example, is taken from an academic context (Longman Dictionary of Contemporary English, 2005)

4) *The discrepancies between these different analyses pose a number of problems.*

The second example is probably from a magazine or a fiction (ibid.):

5) *Winning meant standing on a podium, smiling for cameras and posing for pictures.*

In statement (4), the word pose is a verb and followed by a noun phrase *a number of problems* while in statement (5), it is still a verb but followed by a prepositional phrase *for pictures*. Although in both statements the verb *pose* constructs similar verb phrase structures (syntax), the words it collocates with seem to differ in terms of meaning (semantics). The question to be asked here is whether this situation is an idiosyncrasy, which means that it is specific to these contexts, or a general characteristics of the verb *pose* which tend to show different SPs in different contexts. This is actually the rationale behind the research question of this study:

Does the word *pose* have a specific semantic prosody that depends on the context it is being used in?

**Method**

This is a corpus-based study trying to analyze the potential SP of a single word, *pose*. Traditionally, the words to be analyzed are processed in a corpus made up of thousands or millions of tokens. From time to time, different corpora are compared and/or contrasted. In this study, since the concern is to determine whether a specific word has a SP specific to academic contexts, corpora representing different genres were used. Data was gathered from The Corpus of Contemporary American English (COCA) which is one of the largest corpora of English available online. The corpus is composed of more than 450 million words taken from 189,431 texts, including 20 million words each year from 1990 to 2012. Table 2 exhibits the five main genres and some related parameters.

**Table 2.** *Five Main Genres in COCA and Related Parameters*

| | Content | Sources | Total no. of words |
|---|---|---|---|
| **Spoken** | Conversation from more than 150 different TV and radio programs | All Things Considered (NPR), Newshour (PBS), Good Morning America (ABC), Today Show (NBC), 60 Minutes (CBS) | 95,385,672 |
| **Fiction** | Short stories and plays | Literary magazines, children's magazines, popular magazines, first chapters of first edition books 1990-present, and movie scripts | 90,344,134 |
| **Magazines** | Nearly 100 different magazines | Time, Men's Health, Good Housekeeping, Cosmopolitan, Fortune, Christian Century | 95,564,706 |
| **Newspapers** | Newspaper articles from the US | USA Today, New York Times, Atlanta Journal Constitution, San Francisco Chronicle, etc. | 91,680,966 |
| **Journals** | Articles from peer-reviewed academic journals | Nearly 100 different peer-reviewed journals from a variety of fields | 91,044,778 |
| **Total** | | | 464,020,256 |

In Table 2, the first row represents the name of the five genres composing the corpus. The spoken corpus includes conversations taken from authentic TV and radio shows and has more than 95 million words. The second genre, fiction, is composed of short stories and plays from a variety of magazines, books and movie scripts and has more than 90 million words in total. The corpus collected from magazines, with more than 95 million tokens, is the third genre and includes data from about 100 magazines published in the USA. Newspapers is the fourth genre and has almost 92 million words from a variety of US newspapers. The final one, journals, consists of academic articles with more than 91 million words. In total, the collection reaches to an impressive 464 million words with new additions each year.

Data analyses were carried out by using the corpora mentioned above. In order to see whether the word *pose* has a genre-specific SP, its collocations in spoken language, fiction, magazines and newspapers were compared with the collocations in the academic contexts. In this kind of analysis, certain traditional measurements are made to determine the collocational tendencies. Frequency refers to the number of occurrences of words and it can represent "raw data e.g. there are 58,860 occurrences of the word man in the British National Corpus (BNC)" (Baker et al., 2006), or we can use them to refer to percentages or

proportions, which would make comparisons more meaningful. Another statistical measure which is frequently used while trying to deal with collocations is the mutual information score (MI). Since this score is critical to this study, McEnery & Wilson's (2001) definition of the term should be well understood.

> Mutual information is a formula borrowed from the area of theoretical computer science known as information theory. The mutual information score between any given pair of words - or indeed any pair of other items such as, for example part-of-speech categories - compares the probability that the two items occur together as a joint event (i.e. because they belong together) with the probability that they occur individually and that their co-occurrences are simply result of chance. For example. the words riding and boots may occur as a joint event by reason of their belonging to the same multiword unit (riding boots) whereas the words formula and borrowed in the sentence above simply occur together in a relatively one-off juxtaposition: they do not have any special relationship to each other. The more strongly connected two items are, the higher will be the mutual information score. On the other hand, if the two items have a very low level of co-occurrence, that is, they occur more often in isolation than together, then the mutual information score will be a negative number. And if the co-occurrence of item 1 and item 2 is largely due to chance, then the mutual information score will be close to zero. In other words, pairs of items with high positive mutual information scores are more likely to constitute characteristic collocations than pairs with much lower mutual information scores.

From the detailed definition above, we can easily deduce that the higher the MI score is, the more related the two lexical items are. Typically, scores of about 3.0 or above show a "semantic bonding" between the two words (Davies, 2008). However, MI score is not the only measure employed in collocational computations. Asymptotic hypothesis tests such as z-score, t-score and chi-square are also frequently used in such computations. Their choice mainly depends on their relative ease of application (Schilk, 2011).

**Results and Discussions**

In the analysis process, first of all, the word *pose* was rendered in the COCA database in order to check its collocation candidates. The term *collocation candidate,* in this context, refers to the probabilistic paradigm mentioned earlier. To put in another way, if a certain lexical item co-occurs with another frequently, it

means that they are collocation candidates as it is always possible for them to collocate with other lexical items as well.

To be able to make sound deductions concerning collocational tendencies of *pose*, the following criterion were taken: (1) frequency was the first criteria. Only the collocations with high frequency were taken into consideration; (2) in terms of frequency, eight or more words which co-occur repeatedly with the target word were analyzed; (3) analyses were carried out within a word span of 5:5, which means that collocations within a range of five words from the left and five words from the right to the target word were taken into account, and (4) collocations with MI scores higher than 3.0 were regarded statistically significant, as is the general tendency. With these criterion in mind, the target word, pose, was processed in the academic database of the COCA, and the relevant results are exhibited in Table 3.

**Table 3.** *Collocation Candidates for the Word 'Pose' in Academic Journals*

| | **Collocation Candidate** | | **Frequency** | **All** | **%** | **MI** |
|---|---|---|---|---|---|---|
| 1 | threat | (-) | 350 | 12710 | 2.75 | 6.94 |
| 2 | challenge | (N) | 231 | 26082 | 0.89 | 5.31 |
| 3 | question | (N) | 214 | 60616 | 0.35 | 3.98 |
| 4 | problem | (-) | 213 | 68889 | 0.31 | 3.79 |
| 5 | risk | (-) | 186 | 30040 | 0.62 | 4.79 |
| 6 | serious | (N) | 98 | 10668 | 0.92 | 5.36 |
| 7 | danger | (-) | 72 | 5368 | 1.34 | 5.91 |
| 8 | obstacle | (-) | 32 | 3410 | 0.94 | 5.39 |
| 9 | hazard | (-) | 31 | 2337 | 1.33 | 5.89 |
| 10 | difficulty | (-) | 25 | 13041 | 0.19 | 3.1 |
| 11 | greatest | (+) | 22 | 5818 | 0.38 | 4.08 |
| 12 | dilemma | (-) | 21 | 3409 | 0.62 | 4.78 |
| 13 | strike | (N) | 21 | 7070 | 0.3 | 3.73 |
| 14 | barrier | (-) | 20 | 6434 | 0.31 | 3.8 |
| 15 | interesting | (+) | 16 | 7614 | 0.21 | 3.23 |
| 16 | formidable | (-) | 14 | 842 | 1.66 | 6.22 |
| 17 | severe | (-) | 12 | 6086 | 0.2 | 3.14 |
| 18 | gesture | (N) | 11 | 2308 | 0.48 | 4.41 |
| 19 | grave | (-) | 8 | 2297 | 0.35 | 3.96 |
| 20 | harm | (-) | 8 | 4246 | 0.19 | 3.07 |

*(-) = Negative; (+) = positive; N= neutral*

In Table 3, collocate candidates are given in the second column along with their frequency values in the third column. As can be inferred from these values, they are in frequency order, from the most frequent to the least. This means that among the words that occur frequently with *pose*, the word *threat* occurs most frequently (350 times). The forth column (titled all) exhibits the total occurrence of the word threat in the corpus (12710 times). The percentage column signifies the per cent that the collocating word co-occurs only with the target word. The last column, shows the MI score the meaning of which was mentioned before. Co-occurrence frequencies lower than eight or nine aren't taken into consideration because even with MI scores higher than 3.0, their importance in terms of collocation are quite questionable. For example, at the bottom of the above list which is not displayed in the table, the word *occlusion* (item no. 861) appears to have an MI score of 4.05, which is obviously higher than the critical meaningful level 3.0. However, its frequency, which is only 1, is too low to mention as a colocation candidate. The signs next to each lexical item (+,-,and N) denotes its somewhat subjective category of being positive, negative, or neutral. Table 4 summarizes the results.

**Table 4.** *SP categorizations of the collocation candidates of the target word pose in academic contexts*

| | **Number of words** | **%** |
|---|---|---|
| Potentially positive | 2 | 10 |
| Potentially negative | 13 | 65 |
| Potentially neutral | 5 | 25 |
| Total | 20 | |

In the table above, SP categorizations of the target word in academic contexts are given. Considering that there are 20 items in the list (see Table 3), only two of them appear to have positive associations. Whereas, 13 of these words happen to have negative meanings, and the rest five have neither negative nor positive making them neutral in this sense. As was mentioned before, this part is one of the subjective points of the study, albeit partly, as this categorization doesn't stem from any statistical data.

After the target word was processed in the academic context, it was processed in the other contexts available in the COCA database (see Table 2). The same parameters were calculated for these contexts. The results are given in Table 5.

**Table 5.** *Collocation candidates for the word pose in spoken language, fictions, magazines, and newspapers*

| | Collocation Candidate | | Frequency | All | % | MI |
|---|---|---|---|---|---|---|
| 1 | threat | (-) | 764 | 29778 | 2.57 | 7.42 |
| 2 | risk | (-) | 418 | 53555 | 0.78 | 5.7 |
| 3 | question | (N) | 385 | 181364 | 0.21 | 3.83 |
| 4 | strike | (N) | 325 | 46415 | 0.7 | 5.55 |
| 5 | problem | (-) | 285 | 176834 | 0.16 | 3.43 |
| 6 | danger | (-) | 211 | 20952 | 1.01 | 6.07 |
| 7 | challenge | (N) | 188 | 46421 | 0.4 | 4.76 |
| 8 | health | (N) | 176 | 97014 | 0.18 | 3.6 |
| 9 | serious | (N) | 141 | 46667 | 0.3 | 4.33 |
| 10 | hazard | (-) | 121 | 3832 | 3.16 | 7.72 |
| 11 | picture | (N) | 116 | 74664 | 0.16 | 3.37 |
| 12 | model | (N) | 69 | 47579 | 0.15 | 3.28 |
| 13 | safety | (+) | 61 | 32491 | 0.19 | 3.65 |
| 14 | nude | (N) | 50 | 2822 | 1.77 | 6.89 |
| 15 | significant | (N) | 44 | 23448 | 0.19 | 3.65 |
| 16 | assume | (N) | 41 | 25791 | 0.16 | 3.41 |
| 17 | pose | (N) | 38 | 13978 | 0.27 | 4.18 |
| 18 | classic | (N) | 31 | 19865 | 0.16 | 3.38 |
| 19 | greatest | (+) | 29 | 20018 | 0.14 | 3.27 |
| 20 | immediate | (N) | 23 | 11504 | 0.2 | 3.74 |
| 21 | grave | (-) | 23 | 12047 | 0.19 | 3.67 |
| 22 | yoga | (N) | 22 | 3535 | 0.62 | 5.38 |
| 23 | dramatic | (-) | 21 | 13604 | 0.15 | 3.37 |
| 24 | dilemma | (-) | 20 | 4036 | 0.5 | 5.05 |
| 25 | naked | (N) | 20 | 10665 | 0.19 | 3.65 |
| 26 | portrait | (N) | 20 | 10735 | 0.19 | 3.64 |
| 27 | difficulty | (-) | 20 | 12504 | 0.16 | 3.42 |
| 28 | adopt | (N) | 19 | 14670 | 0.13 | 3.11 |
| 29 | obstacle | (-) | 18 | 5827 | 0.31 | 4.37 |
| 30 | warrior | (N) | 18 | 7888 | 0.23 | 3.93 |
| 31 | heroic | (+) | 14 | 2681 | 0.52 | 5.12 |
| 32 | sexy | (+) | 14 | 7745 | 0.18 | 3.59 |
| 33 | plank | (N) | 13 | 2638 | 0.49 | 5.04 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 34 | corpse | (N) | 13 | 4332 | 0.3 | 4.32 |
| 35 | casual | (N) | 13 | 6658 | 0.2 | 3.7 |
| 36 | playboy | (N) | 12 | 1208 | 0.99 | 6.05 |
| 37 | mimic | (N) | 12 | 2955 | 0.41 | 4.76 |
| 38 | lighting | (N) | 11 | 4840 | 0.23 | 3.92 |
| 39 | existential | (N) | 10 | 600 | 1.67 | 6.8 |
| 40 | imminent | (N) | 10 | 2535 | 0.39 | 4.72 |
| 41 | ethical | (+) | 10 | 4280 | 0.23 | 3.96 |
| 42 | awkward | (-) | 10 | 5348 | 0.19 | 3.64 |
| 43 | stiff | (-) | 10 | 7819 | 0.13 | 3.09 |
| 44 | choking | (-) | 9 | 331 | 2.72 | 7.5 |
| 45 | provocative | (-) | 9 | 2245 | 0.4 | 4.74 |
| 46 | snapshot | (N) | 9 | 2279 | 0.39 | 4.72 |
| 47 | autograph | (N) | 9 | 2421 | 0.37 | 4.63 |
| 48 | relaxed | (+) | 9 | 3467 | 0.26 | 4.12 |
| 49 | languid | (-) | 8 | 460 | 1.74 | 6.86 |

*(-) = Negative; (+) = positive; N= neutral*

Table 5 reveals the results of the calculations that were carried out with the word *pose* in the following contexts taken from the COCA database: spoken language, fiction, magazines, and newspapers. The collocation candidates are again arranged in frequency order, from the most frequent to the least. Total occurrences of the words and their percentages are also displayed in the same manner. Finally, relevant MI scores are exhibited in the last column.

Results that are revealed in Table 3 and 5 are both similar and different in certain ways. First of all, both in Table 3, where the target word is displayed with its collocates in academic contexts, and Table 5, where the same results for other contexts are revealed, appear to have similar collocations particularly with the most frequent ones like *threat*, *problem*, and *danger*.

The main difference between these two collocation sets is that the collocations taken from the academic contexts exhibit less variety than the ones taken from the other contexts. In academic contexts, *pose* appears to have a very limited but a stable collocational variety signaling a negative SP. It collocates with 20 words more than eight or more times in academic contexts, but it appears to be collocating with about 50 words in the other contexts. From this picture, one can conclude that *pose* has a gender-specific collocational variety. The question here is

whether this variety is also valid in terms of SP. The following table should help us make relevant conclusions.

**Table 6.** *SP categorizations of the collocation candidates of the target word pose in spoken language, fictions, magazines, and newspapers*

| | Number of words | % |
|---|---|---|
| Potentially positive | 6 | 12.24 |
| Potentially negative | 15 | 30.61 |
| Potentially neutral | 28 | 57.14 |
| Total | 49 | |

SP categorizations of the word *pose* are shown in Table 6. As is clear in the table, from the 49 words in total, the target word has six potentially positive, 15 potentially negative, and 28 potentially neutral collocations in spoken language, fictions, magazines, and newspapers. When it is compared with the results of SP categorizations in academic contexts (Table 4), it is quite clear that the target word *pose* has a gender specific collocational variety and it appears to have a negative SP in academic contexts and a neutral SP in other contexts.

**Summary and concluding remarks**

Semantic prosody is relatively a new topic in linguistics with a paradigm to move the concept of collocations a step further by determining prosodic natures of lexical items as negative, positive or neutral. This naturalistic point of view is actually far away from being merely theoretical. Actually, in the course of language production, the distinction between native speakers and the learners of a language lies in such seemingly minor details.

In this study, observations as to the usage of the word *pose* in academic contexts were evaluated through statistical analysis. From these observations, it was seen that the target word (pose), when used in academic contexts, tend to collocate with negative words such as *problem* or *threat*. In other contexts, such as newspapers TV movies, or spoken language, it had a wider variety of usages like *posing for a magazine* or *posing a question*. In order to see whether these observations could be statistically validated, the word was processed in a 464-million-word corpus (COCA). Since a comparison of the word in academic contexts and others (spoken, fiction, magazines, and newspapers) was foreseen, the former contexts were compared with the latter ones as a whole. The outcomes were analyzed based on frequency, percentage and MI (mutual information) scores.

Statistical analyses validated the observation that the target word has a mostly negative SP in academic contexts and a mostly neutral one in other contexts available in the corpus.

In the related literature, as was mentioned before, certain lexical items were found to have positive and negative SPs. Sinclair (1991), for example, found that *break out, happen, set in, bent on* tend to have negative SPs; Stubbs (1995, 1996), in the same manner, claimed that *accost, cause, signs of* have negative SPs while the words *provide* and *career* appear to have positive SPs. Partington (1998, 2004) analyzed the words *commit, peddle/peddler, and rife* and stated that they have negative SPs whereas *impressive* has a positive one. It is clear that none of these studies take into account the potential differences among genres. In other words, as Tribble (2000) puts it, SP could be subject to specific genres. It means that some lexical items might collocate with positive words in one context while collocating with negative ones in another. The results of this study revealed that this insight is actually the case with the word 'pose' as it seems to have a mostly negative SP in academic contexts but somewhat neutral one in other contexts.

Future studies to focus on SP might try to discover features of academic English in terms of SP to be used in teaching academic writing to EFL learners. Furthermore, the effects of SP supported (or informed) vocabulary instruction might be compared with other teaching techniques.

**References**

Aitchison, J. (1987). "Reproductive Furniture and Extinguished Professors". (Eds.: Ross Steele & Terry Threadgold), *Language Topics. Essays in Honour of Michael Halliday*, Vol. 2, Amsterdam: Benjamins, pp. 3–14.

Altenberg, B. (1998). "On the Phraseology of Spoken English: The Evidence of Recurrent Word-combinations". (Ed.: Anthony P. Cowie), *Phraseology,* Oxford: Clarendon Press, pp. 101–122.

Baker, P., Hardie, A. & Mcenery, A. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

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

Conzett, J. (2000). "Integrating Collocation into a Reading and Writing Course". (Ed: Lewis Michael), *Teaching Collocation*, England: LTP, pp. 70–87.

Davies, M. (2008-). The Corpus of Contemporary American English: 450 million words, 1990-present. Available online at http://corpus.byu.edu/coca/.

Finch, G. (2005). Key concepts in language and linguistics. New York: Palgrave Macmillan.

Firth, J. R. (1935). "The Technique of Semantics". Transactions of the Philological Society, 34, 1, 36-72.

Firth, J. R. (1957). A Synopsis of Linguistic Theory, 1930-1955. (Ed.: Liu Runqing et al. ), *Readings in Linguistics: Seventy-Five Years Since Saussure.* Vol. 1, Beijing: Cehui Press, pp. 328-359.

Guasti, M. T. (2002). *Language Acquisition: The Growth of Grammar*. Massachusetts: The MIT Press:.

Hoey, M. (2000). "A World beyond Collocation: New Perspectives on Vocabulary Teaching". (Ed: Lewis Michael), *Teaching Collocation*, LTP: England, pp.224-243.

Hunston, S., & Francis, G. (2000). *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English.* Philadelphia: John Benjamins.

Leed, R. L. & Nakhimovsky, A. D. (1979). "Lexical Functions and Language Learning". Slavic and East European Journal, 23,1, 104-113.

Lewis, M. (1993). *The Lexical Approach*. England: LTP.

Lewis, M. (1997). *Implementing the Lexical Approach: Putting Theory into Practice.* Hove: LTP.

Louw, B. (1993). Irony in the Text or Insincerity in the Writer? — the Diagnostic Potential of Semantic Prosodies. (Eds.: Mona Baker, Gill Francis & Elena Tognini-Bonelli), *Text and Technology: in Honour of John Sinclair,* Amsterdam: Benjamins, pp. 157–176.

McEnery, T., & Wilson, A. (2001). *Corpus Linguistics: An Introduction.* Edinburgh: Edinburgh University Press.

Meyer, C. (2004). *English Corpus Linguistics: An Introduction*. Cambridge and New York: Cambridge University Press.

Nation, P. (2001). *Learning Vocabulary in another Language.* Cambridge: Cambridge University Press.

Nesselhauf, N. (2003). "The Use of Collocations by Advanced Learners of English and some Implications for Teaching". Applied Linguistics, 24, 2, 223-242.

Palmer, H. E. (1933). Second Interim Report in English Collocations. Tokyo: Kaitakusha.

Partington, A. (1996). *Patterns and Meaning: Using Corpora for English Language Research*. Amsterdam: John Benjamin Publishing Company

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

Partington, A. (2004). ''Utterly Content in Each Other's Company''. Semantic Prosody and Semantic Preference. International Journal of Corpus Linguistics, 9,1, 131–56.

Pose. (2005). In *Longman Dictionary of Contemporary English Writing Assistant Edition CD-ROM.* Essex: Pearson Education Limited.

Prodromou, L. (2004). "Collocation". Language Study Section of the Macmillan Essential Dictionary Webzine,

http://www.macmillandictionaries.com/MED-Magazine/May2004/19-language-study-collocation-uk.htm. (Accessed: 12.15.2012)

Schilk, M. (2011). *Structural Nativization in Indian English Lexicogrammar*. Amsterdam: John Benjamins.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: OUP.

Sinclair, J. (1996). "The Search for Units of Meaning". Textus, 9, 75–106.

Stubbs, M. (1995). "Collocations and Semantic Profiles: On the Cause of the Trouble with Quantitative Methods". Function of Language, 2,1, 1–33.

Stubbs, M. 1996. Text and Corpus Linguistics. Oxford: Blackwell.

Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.

Stubbs, M. (2004). "Language corpora". (Eds.: Alan Davies & Catherine Elder ), *The handbook of applied linguistics,* London: Blackwell Publishing, pp. 106–132.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam and Philadelphia: Benjamins.

Tognini-Bonelli, E. (2010). "Theoretical overview of the evolution of corpus linguistics". (Eds.: Anne O'Keeffe & Michael McCarthy), *The Routledge Handbook of Corpus Linguistics,* New York: Routledge, pp. 14-27.

Tribble, C. (2000). "Genres, Keywords, Teaching: Towards a Pedagogic Account of the Language of Project Proposals". (Eds.: Lou Burnard & Tony McEnery), *Rethinking Language Pedagogy from a Corpus Perspective: Papers from the Third International Conference on Teaching and Language Corpora*, New York: Peter Lang, pp. 74-90.

Wray, A. (2002). *Formulaic Language and the Lexicon.* New York: Oxford University Press.