UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Polysemous Verbs and Modality in Native and Non-Native Argumentative Writing: A Corpus-Based Study

DANICA SALAZAR & ISABEL VERDAGUER [1]
UNIVERSITY OF BARCELONA

ABSTRACT
The present study is a corpus-based analysis of a selection of polysemous lexical verbs used to express modality in student argumentative writing. Twenty-three lexical verbs were searched for in three 100,000-word corpora of argumentative essays written in English by American, Filipino and Spanish university students. Concordance lines were examined to determine their use in the three corpora. After presenting the overall results for all verbs studied, more in-depth linguistic analysis was performed on the polysemous verb *feel.* These analyses revealed that the non-native writers, unlike their native counterparts, had a limited grasp of the full range of meanings of lexical verbs such as *feel.* It also showed that all student writers under study employed only a restricted range of lexical verbs to convey modal meanings in their writing.

KEYWORDS: Polysemous verbs, lexical verbs, modality, writer stance, academic writing, argumentative writing, foreign language writing, second language writing, learner corpora

RESUMEN
En este artículo presentamos un análisis de una selección de verbos polisémicos, utilizados para expresar modalidad, en tres corpus de textos argumentativos escritos en inglés por estudiantes universitarios americanos, filipinos y españoles. Después de exponer los resultados generales, se presenta un análisis más exhaustivo del verbo polisémico *feel*, que revela que los estudiantes no nativos, a diferencia de los nativos, tienen un conocimiento limitado de su diversidad de sentidos. También muestra que todos los estudiantes analizados usaron un repertorio restringido de verbos léxicos que expresan modalidad.

PALABRAS CLAVE: Verbos polisémicos, verbos léxicos, modalidad, textos académicos, textos argumentativos, segunda lengua, lengua extranjera, corpus de aprendices.

[1] ***Address for correspondence***: Danica Salazar & Isabel Verdaguer. *University of Barcelona.* Department of English and German. Gran Via de les Corts Catalanes 585, 08007 Barcelona, Spain. Telephone: 934035683. Email: dlorenzosalazar@ub.edu; i.verdaguer@ub.edu

I. INTRODUCTION

Research has shown that writers' use of modal expressions to communicate their stance towards their statements and their audience constitutes a rhetorical feature crucial to academic writing (Aijmer, 2002; Biber, 2006; Hyland & Milton, 1997). Lexical verbs provide the most precise and versatile means to express modality, but to take full advantage of them, adequate control of their varying senses is necessary. The use of lexical verbs involves "critical lexical, tense and voice choices" (Hyland & Milton, 1997: 191) that most non-native writers are not able to make. This study aims to increase our understanding of learners' difficulties with lexical verbs so that appropriate tools and resources can be designed to help them make more productive use of these linguistic devices in their writing.

In line with the current learner-corpus trend, Salazar (2008) compiled her own corpus of argumentative writing in English by Filipino students and used it with comparable American and Spanish corpora to investigate modality in native and non-native writing. In her analysis, the author determined the overall range and frequency of some modal expressions and examined their grammatical distribution, their function as hedges and boosters, and their use in personalized, impersonalized and depersonalized constructions and in modal combinations. She found that while the Filipino student writers tended towards hyperclarity, the Spanish foreign-language writers transferred Spanish rhetorical strategies to English. Both non-native groups overused personalized forms and had difficulty establishing appropriate author distance through impersonalized constructions. Spoken-language features and inappropriate degrees of author commitment were detected in all texts, suggesting that problems with the expression of modality in academic writing may not only be due to language-learning factors but to novice-writer factors as well.

This previous study also yielded very interesting results in relation to the grammatical distribution of the modal devices used in the three corpora. The data showed that all three student groups relied heavily on modal verbs when expressing modal meanings, to the detriment of other grammatical classes such as adjectives and nouns. A striking difference was also found between the native and non-native groups: while the American student writers slightly favored lexical verbs, the Filipino and Spanish student writers preferred to use adverbs.

Authors of similar studies on modality have also noted the difficulties faced by non-native writers in providing evidential justification for their claims (Carlson, 1988; Hyland & Milton, 1997; McCann, 1989). Lexical verbs effectively fulfill this modal function in academic writing, but non-native writers' excessive dependence on adverbs and incomplete mastery of the whole meaning range of polysemous verbs point to their inability to exploit the full potential of lexical verbs in their writing. This obvious problem led us to take a closer look at lexical verbs with multiple meanings in our native and non-native corpora, in an attempt to shed more light on this issue.

II. CORPORA AND METHODOLOGY

II.1. Corpora

A fairly recent trend in corpus research that holds enormous potential for many fields of linguistic inquiry is the compilation and analysis of learner corpora. Learner-corpus researchers compile learner data following strict design criteria that control a wide range of learner and task variables. They then use various methods of analysis to quantify and examine these data in order to highlight significant patterns in learner language. One of these methods is contrastive interlanguage analysis (CIA) (Granger, 1996), which involves comparing learner corpora to native-speaker reference corpora or to different varieties of learner language.

A groundbreaking collection of learner corpora is the International Corpus of Learner English (ICLE), which contains over three million words of argumentative-essay writing by advanced learners of English as a foreign language from 21 different mother-tongue backgrounds. This large-scale, international project has already proven to be of enormous value in the study of foreign-language writing.

One of the three corpora used in the present study is a sample of around 100,000 words from the Spanish sub-corpus of ICLE (SPICLE), which is composed of around 100,000 words of essay writing by advanced Spanish-speaking learners of English. Another 100,000-word sample was taken from the Louvain Corpus of Native English Essays (LOCNESS), the native-English corpus compiled for ICLE that contains argumentative essays written by native-speaker American students. Following the same design as the other two corpora is a corpus of argumentative essays written in English by Filipino students that Salazar collected for her study on modality (2008). These corpora represent three contexts of English-language use: as a first language (American corpus), as a second language (Filipino corpus) and as a foreign language (Spanish corpus). The exact number of words in each corpus is shown in Table 1.

Name of corpus	**Number of words**
LOCNESS (American - AME)	98,816
Filipino Corpus (Filipino - FIL)	97,695
SPICLE (Spanish - SPA)	98,945

Table 1. Corpora used in the study

II.2. Methodology

A pre-established list of lexical verbs frequently used by native speakers to express modality was used in this study. This list of 23 lexical verbs was taken from the inventory of modal devices used in Salazar (2008), which was based on Hyland and Milton (1997) and Hyland (2005).

appear	demonstrate	guess	realize	suspect
argue	doubt	indicate	seem	tend to
assume	establish	know	show	think
believe	feel	propose	suggest	
claim	find	prove	suppose	

Table 2. List of lexical verbs

Concordancing software was used to search for the target items, including misspellings and morphological variants (singular, plural and past-tense forms). Every occurrence of the target items was carefully examined in its sentential context to determine its use.

To control for the novice-writer effect, some of the frequency results obtained from the American, Filipino and Spanish corpora were compared to a native-speaker standard provided by the Corpus of Contemporary American English. Samples of 300 concordance lines containing the verb *feel* were extracted from the entire corpus and from its fiction and newspaper-editorial sections. The usage patterns of *feel* in these concordance lines were later analyzed and compared to those found in the three novice-writer corpora.

III. RESULTS

III.1. Use of lexical verbs

Table 3 ranks the 23 lexical verbs according to their overall frequencies in the three corpora.

AME			**FIL**			**SPA**		
Rank	**Item**	**No.**	**Rank**	**Item**	**No.**	**Rank**	**Item**	**No.**
1	feel	137	*1*	think	148	*1*	think	269
2	think	115	*2*	believe	94	*2*	know	126
3	believe	98	*3*	find	88	*3*	find	102
4	find	83	*4*	know	86	*4*	seem	67
5	know	83	*5*	seem	58	*5*	believe	54
6	show	76	*6*	realize	30	*6*	feel	51
7	argue	73	*7*	feel	28	*7*	show	51
8	seem	72	*8*	prove	28	*8*	appear	44
9	claim	44	*9*	tend to	28	*9*	suppose	43
10	realize	44	*10*	show	26	*10*	realize	36

11	prove	30	*11*	argue	21	*11*	establish	29
12	establish	24	*12*	claim	19	*12*	prove	25
13	appear	21	*13*	suppose	16	*13*	demonstrate	19
14	suppose	19	*14*	appear	14	*14*	tend to	13
15	suggest	16	*15*	establish	11	*15*	argue	9
16	tend to	14	*16*	suggest	8	*16*	suggest	9
17	assume	13	*17*	assume	6	*17*	claim	6
18	demonstrate	11	*18*	guess	4	*18*	assume	4
19	guess	6	*19*	indicate	2	*19*	guess	2
20	propose	4	*20*	propose	2	*20*	doubt	1
21	indicate	2	*21*	suspect	2	*21*	indicate	1
22	suspect	1	*22*	demonstrate	1	*22*	suspect	1
23	doubt	0	*23*	doubt	1	*23*	propose	0

Table 3. Lexical verbs in order of frequency

The figures show that none of the three student groups has an extensive repertoire of lexical verbs at their disposal, as they all appear to depend on a handful of common speech-like verbs such as *feel*, *seem*, *think* and *know*. There is an especially marked overuse of the verb *think* in the Spanish corpus.

López (2002) noted the same limited use of lexical verbs in her corpus of Spanish student essays, which exhibited overuse of the verb *decir* relative to her corpus of expert academic writing in Spanish. The expert writers showed a wider vocabulary that included less generic verbs such as *sostener*, *afirmar* and *demostrar*. This suggests that overuse of general verbs such as *decir* in Spanish and *think* in English is characteristic of Spanish novice writing, whether in the former language or in the latter.

Further analysis of lexical-verb use revealed the non-native groups' imperfect grasp of the full range of meanings of these verbs. The American essays contain a number of polysemous verbs used not with their concrete, prototypical meanings but with abstract ones. For instance, *feel,* the most frequent lexical verb in the native essays, usually appear with the figurative meaning, "to have a particular way of thinking about something, especially one that depends on your emotions rather than on facts or evidence" (Macmillan English dictionary for advanced learners*,* MED2, Rundell, 2007). In the non-native scripts, on the other hand, *feel* is not among the top ten most frequent lexical verbs, and most of its occurrences carried the literal, non-epistemic meaning of "to experience a particular emotion or physical feeling" (MED2). The same trend is observed in similarly polysemous verbs: in

the Filipino and Spanish corpora, *believe* is generally used to mean "to think that a fact is true" instead of "to have an opinion about what is true or what might happen"; *show*, "to let someone see something" instead of "to prove that something exists or is true"; *claim*, "to say that something is yours, especially as a right" instead of "to say that something is true"; *find*, "to discover something, or to see where it is by searching for it" instead of "to discover a fact or piece of information"; and *appear*, "to begin to be seen" instead of "to seem" (MED2).

In fact, when the raw frequencies are narrowed down to only those occurrences where the lexical verbs are used to express modality, the list of the 10 most frequent lexical verbs changes considerably (Table 4).

AME			**FIL**			**SPA**		
Rank	**Item**	**No.**	**Rank**	**Item**	**No.**	**Rank**	**Item**	**No.**
1	feel	91	1	believe	72	1	think	132
2	believe	86	2	seem	60	2	seem	71
3	seem	69	3	think	46	3	know	35
4	argue	57	4	tend to	28	4	believe	22
5	think	49	5	know	26	5	prove	16
6	show	35	6	prove	23	6	show	15
7	prove	25	7	realize	18	7	tend to	13
8	claim	24	8	claim	17	8	realize	13
9	know	18	9	argue	16	9	demonstrate	10
10	realize	15	10	show	14	10	suppose	8

Table 4. Ten most frequent verbs expressing modality

As can be seen in Table 4, while there are verbs that show a balanced frequency of occurrence in the three corpora (e.g., *seem*, *realize*), others exhibit greater degrees of divergence, which are all the more striking in the comparison between the American and Spanish corpora. Verbs such as *feel, argue* and *claim*, which occur on the American list of ten most frequent lexical verbs, do not appear on the Spanish list. The most remarkable difference is with *feel*, which is the most frequently occurring lexical verb in the native corpus, overall and as a modal device. In the non-native corpora, it is also one of the ten most frequent verbs overall, but it disappears from the top ten when only its use as a modal device is considered.

III.2. Use of lexical verbs in personalized and depersonalized expressions

It is also interesting to note the different ways in which the native and non-native groups use lexical verbs to express modality. The American writers studied often use lexical verbs to attribute their arguments to other sources. Most instances of three of the most frequently occurring lexical verbs in the American compositions, for example, follow the third-person subject + verb in third person + *that* pattern: *believe* (61% of occurrences), *feel* (48% of occurrences) and *argue* (75% of occurrences).

These constructions are what Holmes (1988) terms "depersonalized expressions". These expressions shift responsibility for statements away from the writer by attributing them to another:

(1) The opposing argument lies in the hands of ***evolutionists***, who ***believe*** that creationism is religion and it does not belong in a science class. (AME)

(2) ***Opponents of the death penalty claim*** that it is not only ineffective, but also immoral and unethical. (AME)

The Filipino and Spanish student writers, in contrast, exhibit a tendency towards more personal expression. Half the occurrences of *believe* and *think*, the most frequent lexical verbs in the Filipino corpus, co-occur with the first-person pronoun *I*. This personalized tone is even more noticeable in the Spanish scripts. *Think*, the most common lexical verb in this corpus, appears 132 times (compared to 49 times in the American and 46 times in the Filipino essays) and accounts for 36% of the total occurrences of lexical verbs. Over 50% of these 132 instances co-occur with the first-person pronoun *I*. This proves that the Spanish writers tend to introduce their opinions with the *I think (that*) construction.

III.3. The lexical verb *feel*

The marked difference in the frequency of the use of *feel* between the native and non-native essays made this particular verb worthy of further investigation. The overall frequency of *feel* in the American corpus is 137, in the Filipino 28 and in the Spanish 51. These figures clearly show that the American student writers use this verb much more often than the non-native writers.

A detailed analysis of the range of meanings in which *feel* is used revealed relevant differences in the three corpora. The distinct meanings found in the three corpora can be summarized as follows:

1. to experience a particular emotion or physical feeling
   Who can say that a person in a coma does not ***feel*** pain? (AME)
2. to be in a particular state because of a certain emotion
   They ***feel*** defeated before they begin. (AME)
3. to have a particular way of thinking about something
   We ***felt*** the right decision was to get married. (AME)

These meanings are connected to different grammatical patterns, which help to disambiguate them. While meaning 1 is used with a noun phrase, meaning 2 usually occurs with an adjective phrase and meaning 3 is followed by a finite clause, usually a *that-, like* or *as if* clause.

	AME	**FIL**	**SPA**
Meaning 1	18	16	13
Meaning 2	26	5	29
Meaning 3	92	7	8

Table 5. Meanings of *feel* and their frequency of occurrence

Table 5 shows that meanings 1 and 2 appear more frequently in the non-native corpora, whereas meaning 3, which is more abstract, is predominant in the native.

Meaning 3 constitutes 67% of the occurrences of *feel* in the American corpus but only 25% and 16% in the Filipino and Spanish corpora respectively. When *feel* was searched for in a sample from the entire Corpus of Contemporary American English, it was found that meanings 1 and 2 were also more frequent in general English. However, the argumentative nature of the essays can explain the prevalence of meaning 3 in the American corpus, given the obvious need in argumentative writing for verbs that introduce opinions.

From an etymological point of view, the first two meanings predate the third one. The evolution of this verb shows the usual trend of verbs of physical perception widening their meanings to include more abstract senses. Abstract meanings usually have their origin in concrete domains, which are extended due to a process of metaphoric transfer (Sweetser, 1990; Verdaguer & Poch, 2005). In Old English, *feel* was used to express perception by touching, which was generalized to physical sensation and further to emotions or mental perception in Middle English. Finally, in early Modern English it began to be used in its modal sense to introduce an opinion.

This historical development, with the more abstract meanings appearing later, correlates with learners' acquisition, as it has also been shown that figurative meanings are acquired or learned later than concrete meanings. As a result, non-natives often find it difficult to learn the figurative senses of polysemous verbs, as illustrated by Laso (2009) in her description of non-native speakers' difficulty with the verbs *draw* and *lead*. This is further demonstrated by the non-native writers in the present study, who overuse the prototypical meanings of *feel* while virtually ignoring its figurative sense.

Feel carrying a modal meaning occurs both in personalized and depersonalized constructions. Since *feel* usually introduces an opinion with a certain degree of emotional

implication, it appears quite often in the native corpus with first person subjects to convey the writer's strong attachment to a proposition:

(3) **I *feel*** very strongly that Exxon owes the people of Alaska. (AME)

However, *feel* is more predominantly used in the native corpus in depersonalized constructions. The American writers show a remarkable tendency to authorize their propositions by means of testimonials, using *feel* to relate their opinions to public opinion (4), attribute their statements to other individuals (5) and present different sides of an argument (6).

(4) **Many people *feel*** that there is nothing morally wrong with surrogate motherhood. (AME)

(5) **Susan B. Anthony** originally led the cause because **she *felt*** that women should be given equal treatment in society. (AME)

(6) **Many supporters of euthanasia *feel*** that the patients would have wanted to die. **Opposers *feel*** that letting a person die or assisting them in the ending of their life is an immoral act. (AME)

A comparison of the native students' use of *feel* was also carried out with samples from the newspaper-editorial and fiction sections of the Corpus of Contemporary American English. It shows that although the modal use of *feel* appears much more frequently in editorials than in fiction, it is not the predominant meaning in either corpus, accounting for 30% of the occurrences of *feel* in the sample taken from editorials and 5% from fiction. This indicates that native speakers know how to exploit the different meanings of *feel* but tend to overuse its figurative meaning. Published editorial writers have a range of lexical verbs other than *feel* at their disposal to express modality.

In summary, with regard to the use of *feel*, non-native novice writers show incomplete mastery of its full range of meanings since they only use it to express emotions and physical feelings. Native novice writers, on the other hand, have a good command of its different meanings but rely excessively on its modal meaning in their writing.

IV. CONCLUSIONS AND RECOMMENDATIONS

This analysis, which has investigated contrasts between native and non-native performance in argumentative writing, can give a clearer understanding of the difficulties that non-native and non-expert writers meet and thus can be the basis for pedagogically useful conclusions. The recognition of the differences both between native and non-native corpora and between them and an expert corpus may help teachers become aware of the choices writers make and design the resources and tools which are appropriate at different levels of expertise.

From a language teaching perspective the following observations derived from the comparison of the corpora can be useful:

- Teachers need to raise learners' awareness of the range of lexical verbs that can be used to express modality. Spanish learners overuse verbs such as *think, seem* and *know* while

ignoring verbs such as *feel, argue* and *claim*. A comparison between the performance of Filipino and American writers shows more balanced results, except in the case of *feel*.

- Teachers need to make learners aware of the lexico-grammatical patterning of polysemous verbs and focus on their more abstract senses, which are rarely used by learners.
- Learners should be made aware of the close interrelation between the meaning of verbs and their complementation patterns. A change in the type of object may correlate with a change in the meaning of the verb. They can learn the meaning of verbs more effectively if they are explicitly taught the relationship between meaning and syntax.
- Teachers should help non-native writers develop a more indirect style of writing by making greater use of depersonalized constructions in their writing.

ACKNOWLEDGMENTS

The authors acknowledge the support of the Ministerio de Ciencia e Innovación and FEDER (project no. HUM2007-64332/FILO) and the Generalitat de Catalunya (grant no. 2008FI 00044).

REFERENCES

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5, 97-116.

Carlson, S. (1988). Cultural differences in writing and reasoning skills. In A. Purves (Ed.), *Writing across languages and cultures: Issues in contrastive rhetoric*. Newbury Park, CA: Sage, pp. 227-260.

Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg & M. Johansson (Eds.), *Languages in contrast: Text-based cross-linguistic studies*. Lund: Lund University Press, pp. 37–51.

Holmes, J. (1988). Doubt and certainty in ESL textbooks. *Applied Linguistics*, 9, 20-44.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London and New York: Continuum.

Hyland, K. & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6, 183-205.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Laso, N. J. (2009). *A corpus-based study of the phraseological behaviour of abstract nouns in medical English: A needs analysis of a Spanish medical community*. Unpublished Doctoral Dissertation. University of Barcelona, Barcelona, Spain.

Laso, N. J. & Verdaguer, I. (2005). An invitation to explore *conclusion* and its restricted collocations. *Specific: English for Specific Purposes,* 2, 47-54.

López, C. (2002). La comunicación del saber en los géneros académicos: Recursos lingüísticos de modalidad epistémica y de evidencialidad. In F. Luttikhuizen (Ed.), *V Congrés Internacional sobre Llengües per a Finalitats Específiques*. Barcelona:

Publicacions de la Universitat de Barcelona, pp. 164-172.
McCann, T. (1989). Student argumentative writing knowledge and ability at three grade levels. *Research in the Teaching of English*, 23, 62-76.
Rundell, M. (Ed.) (2007). *Macmillan English dictionary for advanced learners* (2nd ed.). Oxford: Macmillan Education.
Salazar, D. (2008). *Modality in student argumentative writing: A corpus-based comparative study of American, Filipino and Spanish novice writers*. Unpublished Thesis. University of Barcelona, Barcelona, Spain.
Sweetser, E. (1990). *From etymology to pragmatics.* Cambridge: Cambridge University Press.
Verdaguer, I. & Poch, A. (2005). Metaphoric mapping and argument structure in semantic change: A case study. In *Moderne Methoden und Technologies in der Germanistikforschung.* Staatsuniversitätsverlag Samara, pp. 50-60.

