

# A Corpus-Based Study on Epistemic Adjectives in Academic English

Reyhan Ağçam\*

School of Foreign Languages, Kahramanmaraş Sütçü İmam University, Kahramanmaraş, Turkey

\*Corresponding author: reghanagcam@gmail.com

Received November 28, 2014; Revised December 16, 2014; Accepted December 21, 2014

**Abstract** ‘Stance is the way academics annotate their texts to comment on the possible accuracy or creditability of a claim, the extent they want to commit themselves to it, or the attitude they want to convey to an entity, a proposition or the reader’ (Hyland, 2005). Recent studies have revealed that researchers, particularly those who study soft sciences, tend to adopt a certain stance while reporting on their research even though they are supposed to use an objective voice in their academic writing. The present study aims to investigate epistemic adjectives used in conveying author stance in Academic English through a corpus-based approach. It reports the results of the Contrastive Interlanguage Analysis (Granger, 1996) administered to a total number of 136 doctoral dissertations written by native and non-native academic authors of English. Frequencies of the epistemic adjectives were separately calculated for each corpus and a log-likelihood test was utilized to see whether native and non-native academic authors of English significantly differ with regard to these items. Findings of the study have shown that certainty adjectives were mostly used by non-native academic authors of English whilst likelihood adjectives were mostly employed by native academic authors of English, which leads us to the conclusion that the non-native groups are inclined to use cautious expressions less frequently than the native groups in their academic writing. The study concludes with possible reasons for this particular result, implications of the results to academic writing and a few suggestions for further research.

**Keywords:** academic writing, contrastive interlanguage analysis, epistemic adjective, stance

**Cite This Article:** Reyhan Ağçam, “A Corpus-Based Study on Epistemic Adjectives in Academic English.” *American Journal of Educational Research*, vol. 2, no. 12 (2014): 1230-1236. doi: 10.12691/education-2-12-16.

## 1. Introduction

Stance is defined by Hyland (2005: 178) as “the way academics annotate their texts to comment on the possible accuracy or creditability of a claim, the extent they want to commit themselves to it, or the attitude they want to convey to an entity, a proposition or the reader”. Biber and Finegan (1989: 124) propose it “refers to the lexical and grammatical expression of attitudes, feelings, judgments, or commitment concerning the propositional content of a message”. Stubbs (1986) argues that academic texts inevitably contain the author’s attitude. Salager-Meyer (1997) asserts that research from a variety of disciplines (e.g., sociology of science) has revealed ways in which academic discourse is both socially situated and structured to accomplish rhetorical objectives, which are realized as hedges. According to her, these devices are used “to reflect not only fundamental characteristics of modern science (scepticism, uncertainty and doubt), but also the true state of the writer’s understanding and state of knowledge”. Conrad and Biber (2000) classify stance into three semantic categories as epistemic, attitudinal and style. MacIntyre (1999) suggests epistemic stance, in general, represents how certain or reliable the writer’s proposition is, attitudinal stance reports personal attitudes

or feelings and style stance describes how the information is being presented. Table 1 displays epistemic stance devices (ESDs) proposed by Biber (2006).

**Table 1. Epistemic Stance Devices (Adapted from Biber, 2006: 94)**

ESDs	E.g.
Adjective + <i>that</i> -clause	It is <i>evident</i> [that by 1925 Gandhi had fully worked out his approach to, and explanation of the text].
Adjective + <i>to</i> -clause	These “control failures” are <i>certain</i> [to happen occasionally].
Adverb	Now these schools might <i>possibly</i> be able to afford this special fund.
Noun + <i>that</i> -clause	There’s some <i>indication</i> [that prenatal development has an influence on lifespan development].
Verb + <i>that</i> -clause	Therefore there is no reason to <i>suspect</i> [that the two means are different].
Verb + <i>to</i> -clause	Our best evidence <i>seems</i> [to suggest that if you control for size, the more international you are, the less risk you have].

The present study concentrated on the use of epistemic adjectives found in the doctoral dissertations produced by native, Turkish-speaking and Spanish-speaking academic authors of English.

## 2. Background to the Study

Skelton (1988) contends that epistemic comments are widely used in the arts and sciences stressing that they are

found in one third to one half of the sentences in the texts. In a study where Finnish-speaking learners of English and English native speakers were compared in terms of epistemic stance devices, Ventola and Maureanen (1990) revealed that the non-native group used relatively a small number of safe expressions of epistemic possibility than the native group. Investigating scientific research papers with respect to modality markers, Gosden (1990) found that over 7% of grammatical subjects consist of these markers that convey writer's perception of uncertainty. In another study, Gabrielatos and McNery (2005) analysed MA dissertations written by students at the Department of Linguistics and English Language, Lancaster University and concluded that native speakers of English employed epistemic adjectives, epistemic adverbs and epistemic modals more frequently than the non-native speakers. In his study focusing on stance modals, Kafes (2009) compared three sets corpora including research articles produced by Turkish-speaking learners of English, Spanish-speaking learners of English and native speakers of English with regard to stance modals. He found that non-native speakers of English do not significantly differ in the use of stance modals except *may* and that there is a statistically significant difference between the native and non-native groups concerning the use of *can*, *could*, *may*, *might*, *would* and *should*. Kanté (2010) investigated head nouns across legal and academic corpora and revealed that head nouns controlling that-clauses such as fact, statement and assumption might be used for different communicative functions and effects concerning text genres.

Biber et al. (1999: 671) propose that the adjectives that control a that complement clause convey stance and that they fall into three major semantic domains: degrees of certainty (e.g. certain, confident, evident); affective phonological states (e.g. annoyed, glad, sad); and evaluation of situations, events, etc. (e.g. appropriate, odd, good, important, advisable). According to them, certainty adjectives indicate the degree of certitude of the proposition presented in the *that*-clause or they can reflect the writer's assessment of a position held by some third person. The following statement was taken from Longman Grammar of Spoken and Written English (LGSWE) to exemplify the latter (Biber et al., 1999: 672).

e.g. They are undoubtedly right that it has now become clear that the government will not pay for the expansion it desires (NEWS).

They point out that adjectival predicates controlling extraposed that-clauses signals a stance or attitude towards the proposition conveyed in the *that*-clause and that they mostly represent the attitude of the speaker/ writer even though they are not overtly attributed to him/ her. Their study revealed that certainty adjectives constituted the majority of adjectival predicates controlling these clauses.

### 3. Research Design

#### 3.1. Research Questions

This study is intended to provide answers to the following research questions:

1. Do Turkish-speaking academic authors of English and native academic authors of English significantly differ with respect to the use of epistemic adjectives?

2. Do Spanish-speaking academic authors of English and native academic authors of English significantly differ with respect to the use of epistemic adjectives?

#### 3.2. Methodology

The present study is corpus-based in design. It primarily investigated author stance across doctoral dissertations of native and non-native academic authors of English. It includes the analysis of the epistemic adjectives that are frequently reported to occur in the written academic registers in Biber's study (2006), as illustrated in Table 2.

Table 2. Epistemic Adjectives (Adapted from Biber, 2006: 92)

Epistemic Adjectives	
Certainty	apparent, apt, certain, due, clear, confident, convinced, correct, evident, false, guaranteed, impossible, inevitable, obvious, positive, prone, right, sure, true, well-known
Likelihood	doubtful, liable, likely, possible, probable, unlikely

#### 3.3. Data

Three sets of corpora were built with the collection of 133 doctoral dissertations produced by Turkish-speaking Academic Authors of English (TAEs, henceforth), Spanish-Speaking Academic Authors of English (SAEs, hereafter) and Native Academic Authors of English (NAEs, henceforth) within the field of English Language Teaching, English Language and Literature, Applied Linguistics and Modern Languages [TACE (Turkish Academic Corpus of English): 48; SACE (Spanish Academic Corpus of English): 43; NACE (Native Academic Corpus of English): 45]. Most of the dissertations in NACE were submitted to four different universities in the UK while 11 were submitted to five different universities in the USA, one to the one in Netherlands and one to a university in Australia. The majority of the dissertations in TACE were submitted to seven different universities in Turkey and one to a university in Australia. As for SACE, thirty dissertations were submitted to 14 different universities in Spain, eight to six different universities in the USA and six to five different universities in the UK. All were written between 2005 and 2012 and for practical reasons, the corpora were constructed with the collection of dissertations available online. Assuming that those who supervised the dissertations might have influenced the writing process of the sections Abstract, Introduction, Review of Literature, Methodology and References, they were excluded from the corpora. Instead, data analysis was realised through the sections Findings, Discussion, Conclusion, Pedagogical Implications (Implications to English Language Teaching) and Suggestions for Further Research. Finally, all figures, paraphrases, quotations, tables and titles were excluded from the above-mentioned sections of the dissertations in the corpora. Table 3 shows the size of corpora across which epistemic adjectives were investigated in this study.

Table 3. Corpus Size

Corpus	Dissertation (N)	Word (N)
TACE	48	675.072
SACE	43	668.256
NACE	45	671.475
Total	136	2.014.873

### 3.4. Data Analysis

Since our study aimed to reveal whether TAEs, SAEs and NAEs significantly differ from each other in terms of epistemic adjectives, the Contrastive Interlanguage Analysis (CIA) (Granger, 1996) was used as method of data analysis procedure. The following are the six stages of the process in concern.

1. The Analysis of TACE: The corpus including dissertations produced by TAEs was examined with respect to epistemic adjectives.
2. The Analysis of NACE: The corpus including dissertations produced by NAEs was searched with respect to epistemic adjectives.
3. The Analysis of SACE: The corpus including dissertations produced by SAEs was examined with respect to epistemic adjectives.
4. TACE vs NACE: Findings obtained from the analyses of TACE and NACE were compared to see whether they significantly differ in the use of epistemic adjectives.
5. SACE vs NACE: Results of the analysis administered to SACE and NACE were compared to reveal whether they significantly differ in the use of epistemic adjectives.
6. TACE and SACE: Findings obtained from the analyses of TACE and SACE were compared to find out whether they significantly differ concerning the use of epistemic adjectives.

For the first three stages of data analysis, Wordsmith Tools 6.0, a computer software package developed by Scott (2012), “for analysing the lexis of texts and corpora in order to produce frequency lists, to run concordance searches and calculate collocations for particular words, and to find keywords in a text and examine their distribution” (Baker et al. 2006: 169), was used to identify epistemic adjectives in three corpora. For the last three stages, Log Likelihood Statistics, which was previously used in significant studies conducted in corpus linguistics (e.g. Granger and Rayson, 1998; Scott, 2011), was utilized to compare the frequencies of epistemic adjectives measured in each corpus.

## 4. Results and Discussions

Table 4. Epistemic Adjectives not Found in Three Corpora

Epistemic Adjectives	
Correct	Is it <i>correct</i> that personnel cannot be charged to an award until their COI Training is completed?
Doubtful	With her injuries it's <i>doubtful</i> that she'll ever walk again.
Due	Rose is <i>due</i> to start school in January.
False	Is it true or <i>false</i> that the Robinson projection shows the size of the continents accurately and the water areas are expanded to fill the extra space?
Inevitable	It was <i>inevitable</i> that there would be job losses.
Possible	It's just <i>possible</i> that I gave them the wrong directions.
Right	Is it <i>right</i> that mostly college drop-out students run the biggest businesses in the world?

This section presents findings drawn from the analysis of three corpora including doctoral dissertations produced by TAEs, SAEs and NAEs. Following Biber (2006), a total number of 26 English epistemic adjectives (under two categories as certainty adjectives and likelihood adjectives) were identified over three corpora and their frequencies were computed via WS Tools 6.0.

Subsequently, three corpora were analysed into categorical and individual epistemic adjectives to see whether native and non-native academic authors of English significantly differ with respect to the use of these adjectives. Yet, it might be useful to begin with presenting epistemic adjectives not found in three corpora. Table 4 illustrated the items at stake.

### 4.1. Epistemic Adjectives in Three Corpora

Three corpora were compared in terms of the epistemic adjectives. Figure 1 shows the distribution of epistemic adjective types found across three corpora.

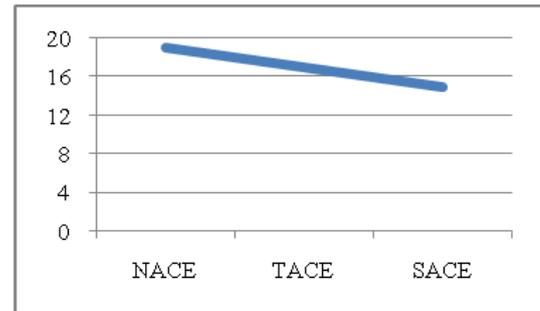


Figure 1. Epistemic adjective types in three corpora

As displayed in Figure 1, epistemic adjectives were found in highest variety in NACE; namely, 19 out of 26 types were used in NACE, 17 types in TACE and 15 types in SACE. In parallel with this, they were mostly used in NACE (586 times) while they occurred 251 and 192 times in TACE and SACE, respectively. Hence, they were underused by TAEs and SAEs in comparison with NAEs, which is illustrated in Table 5.

Table 5. Overall Frequency Distribution of Epistemic Adjectives in Three Corpora

	NACE (L1)	TACE (L2)	SACE (L2)
Corpus Size in words	671.475	675.072	668.256
Epistemic Adj. (n)	586	251	192
n per 10.000	8.7	3.7	2.9
T/t ratio (%)	0.01	0.00	0.00
Eps. Adj. Types (n)	19	17	15

n= raw frequency of epistemic adjectives

T/t ratio= Type/token ratio; percentage of number of epistemic adjectives (types) in total of words (tokens) in each corpus.

As indicated in Table 5, the highest frequency was observed in NACE, which means there is an underuse in the dissertations produced by SAEs and those written by TAEs against NAEs.

Table 6. LL Ratio of Epistemic Adjectives in TACE and NACE

	TACE (N)	NACE (N)	LL Ratio (*p<0.05)
Epistemic Adj.	251	586	-139.71

n=raw frequency of epistemic adjectives in corpus

+ indicates overuse in TACE relative to NACE

- indicates underuse in TACE relative to NACE.

The results showed that epistemic adjectives were underused in TACE with -139.71 LL value (p<0.05), indicating a statistically significant difference between two corpora with respect to epistemic adjectives; thus, the underuse in TACE relative to NACE was approved by LL calculation. Subsequently, SACE and NACE were compared in terms of the overall frequency of the items. Table 7 outlines LL ratio of the frequency in concern.

**Table 7. LL Ratio of Epistemic Adjectives in SACE and NACE**

	SACE (N)	NACE (N)	LL Ratio (*p<0.05)
Epistemic Adj.	192	586	-207.19

n=raw frequency of epistemic adjectives in corpus  
 + indicates overuse in SACE relative to NACE  
 - indicates underuse in SACE relative to NACE.

Table 7 suggests that the underuse in SACE against NACE was confirmed by LL results. Finally, TACE and SACE were compared in order to reveal whether they significantly differ with respect to the use of epistemic adjectives. The test results are demonstrated in Table 8.

**Table 8. LL Ratio of Epistemic Adjectives in TACE and SACE**

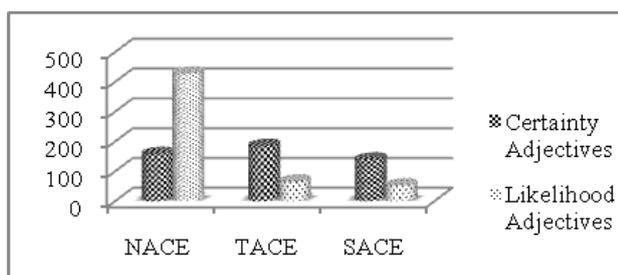
	TACE n	SACE (N)	LL Ratio (*p<0.05)
Epistemic Adj.	251	192	+7.29

n=raw frequency of epistemic adjectives in corpus  
 + indicates overuse in TACE relative to SACE  
 - indicates underuse in TACE relative to SACE.

As indicated in Table 8, epistemic adjectives were relatively more frequented in TACE against SACE, which was approved by a +7.29 LL value. To this point, frequency categories of epistemic adjectives in three corpora as well as the LL results obtained from the analyses of the native and non-native corpora have been provided. However, this does not allow us to claim whether these devices are used to convey the meaning of “certainty” or “likelihood” throughout the dissertations they were found. Therefore, as a further analysis, three corpora were individually compared with each other regarding certainty and likelihood adjectives.

**4.2. Certainty and Likelihood Adjectives in Three Corpora**

This section is intended to report the results of the analysis administered to three corpora to investigate whether they significantly differ in the use of certainty and likelihood adjectives. Figure 2 illustrates the related results.



**Figure 2.** Certainty and likelihood adjectives in three corpora

As can be seen in Figure 2, certainty adjectives occurred more frequently than likelihood adjectives in the non-native corpora whereas they were much less frequented than likelihood adjectives in the native corpus. Nonetheless, this does not necessarily mean that certainty adjectives were overused in both non-native corpora against the native corpus. Namely, they were mostly used in TACE with a frequency of 185 while they were seen 158 and 139 times in NACE and SACE, respectively. Thus, they were overused by TAEs and underused by SAEs against NAEs.

The most frequented certainty adjective across three corpora is *clear*, which appeared 81 times in TACE, 49

times in NACE and 47 times in SACE. The following statements are extracted from each corpus to exemplify the item in concern.

Example 1.

[From considering their experiences, it is *clear* that the Spanish government needs to radically rethink its plans for the English PAU if it does not want to repeat the mistakes and failures of the past.]

Extracted from <SACE-UPV-2012-DLN>

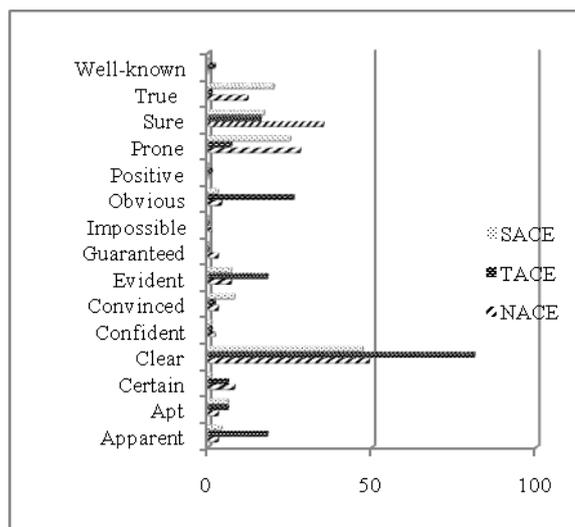
[Table 5.2 makes *clear* that if GENRE effects were to be estimated from all data in this table, i.e. by comparing data in cells A to D with data in cells E to H, we would not be comparing like with like.]

Extracted from <NACE-UE-2011-RJA>

[Think-aloud protocols made it *clear* that gifted learners try to use the exact translation of more complicated expressions they want to use, which makes them fail in activities that are supposed to be quite easy for them.]

Extracted from <TACE-CU-2011-DIA>

It is considered useful to visualize certainty adjective types found across three corpora. Figure 3 displays the frequency distribution of the individual items over three corpora.



**Figure 3.** Certainty adjectives in three corpora

Certainty adjective *sure*, which revealed the second mostly used item following *clear* in the native corpus, was relatively less frequented in the non-native corpora. In response, *true* and *convinced* appeared the second and third mostly used certainty adjective in SACE, respectively while they were found in much lower proportions in NACE and TACE. It is interesting that *obvious*, *apparent* and *evident* appeared among the most frequented items in TACE whereas they were among the least used items across NACE and SACE.

The most striking finding of the study is *probably* that the likelihood adjectives were preferred by NAEs much more frequently than TACE and SACE. Namely, they were found in NACE with a frequency of 428 while they occurred 66 and 53 times in TACE and SACE, respectively, corresponding to the fact that they were employed by NAEs approximately four times as frequently the total frequencies of the items seen in the non-native groups. Figure 4 provides the frequency distribution of these adjectives in three corpora.

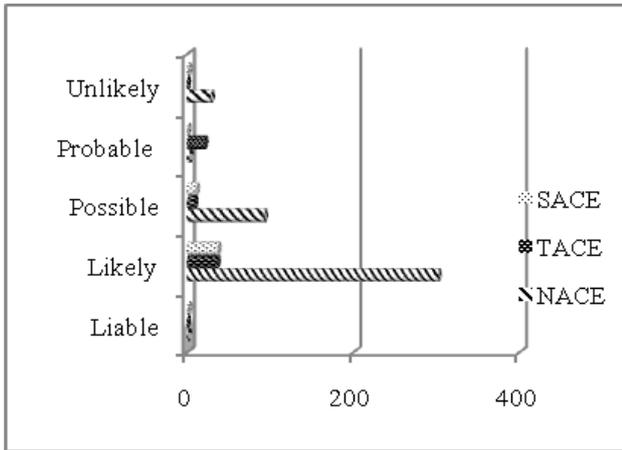


Figure 4. Likelihood adjectives in three corpora

As suggested in Figure 4, one cannot mention a consistency across three corpora concerning the distribution of the likelihood adjectives except *likely*, which is the most frequently used item in all corpora. It constituted approximately 71% of all likelihood adjectives in NACE, 70% in SACE and 55% in TACE. Nevertheless, it was remarkably underused by TAEs and SAEs against NAEs. Example 2 is intended to illustrate the item in question.

Example 2.

[It is therefore *likely* that the repetition task which DEAP effectively became for them was in fact a NWR task as they had no lexical representations available to assist them.]

Extracted from <NACE-NU-2006-CS>

[It is *likely* that the profile of the ELT preparatory students will change in the following years; therefore, a new interest checklist might be applied to review the changing topics of interest.]

Extracted from <TACE-AU-2008-GE>

[Although the exact reasons for its extinction in that area are not known, it is *likely* that it died out in competition with the dingo.]

Extracted from <SACE-UZ-2012-MPBA>

The second mostly used likelihood adjective in NACE revealed *possible*, which is the second mostly used item in SACE and third in TACE. However, as in the case of *likely*, it was underused by both TAEs and SAEs in comparison to NAEs. This time, it was found in NACE (93 times) approximately five times as often they occurred in the non-native corpora (SACE: 11 times; TACE: 9 times). Another underused item in the non-native corpora against the native corpus is *unlikely*, which appeared 29 times in NACE, 2 times in SACE and none in TACE.

*Probable* is not only the mostly preferred likelihood adjective in TACE but it also the only one overused by TAEs against NAEs in this category. For instance, it was seen 21 times in TACE, 3 times in NACE and once in SACE. In return, *liable* revealed the only likelihood adjective that was overused by SAEs against NAEs. Namely, it occurred twice in SACE, once in NACE and none in TACE.

Finally, a log likelihood test was administered to three corpora to see whether the overuse and underuse values in non-native corpora against native corpus are statistically significant and whether TACE and SACE significantly differ with respect to categories of epistemic adjectives.

The first test was applied between TACE and NACE and the results are given in Table 9.

Table 9. LL Ratio of Certainty and Likelihood Adjectives in TACE and NACE

Epistemic Adj.	TACE (N)	NACE (N)	LL Ratio (*p<0.05)
Certainty Adj.	185	158	+1.99
Likelihood Adj.	66	428	-298.31

n=raw frequency of epistemic adjectives in corpus

+ indicates overuse in TACE relative to NACE

- indicates underuse in TACE relative to NACE.

Table 9 demonstrates that certainty adjectives were slightly overused in TACE against NACE, which was confirmed with a +1.99 LL value. Likelihood adjectives, on the contrary, were significantly underused in TACE in comparison to NACE. Namely, they were preferred 185 times by TAEs while they were employed 158 times by NAEs. It was again approved by the test results (-298.31 LL value). The analysis in concern was repeated between SACE and NACE and the results are exhibited in Table 10.

Table 10. LL Ratio of Certainty and Likelihood Adjectives in SACE and NACE

Epistemic Adj.	SACE (N)	NACE (N)	LL Ratio (*p<0.05)
Certainty Adj.	139	158	-1.13
Likelihood Adj.	53	428	-331.28

n=raw frequency of epistemic adjectives in corpus

+ indicates overuse in SACE relative to NACE

- indicates underuse in SACE relative to NACE

As suggested in Table 10, both certainty and likelihood adjectives were underused in SACE against NACE. Yet, the value measured between the two corpora in terms of certainty adjectives is not so high as the one calculated in terms of likelihood adjectives. Namely, certainty adjectives were found 158 times in NACE and 139 times in SACE whereas likelihood adjectives occurred in NACE over 8 times as frequently in SACE. Both findings were confirmed by LL results. As a last step, LL test was conducted to TACE and SACE in order to reveal whether the differences observed between them are statistically significant. The related results can be seen in Table 11.

Table 11. LL Ratio of Certainty and Likelihood Adjectives in TACE and SACE

Epistemic Adj.	TACE (N)	NACE (N)	LL Ratio (*p<0.05)
Certainty Adj.	185	158	+6.09
Likelihood Adj.	66	428	+1.29

n=raw frequency of epistemic adjectives in corpus

+ indicates overuse in TACE relative to SACE

- indicates underuse in TACE relative to SACE.

Table 10 clearly indicates that both certainty and likelihood adjectives were overused in TACE against SACE. More specifically, certainty adjectives were employed 185 times by TAEs whereas they were used 139 times by SAEs. Similarly, likelihood adjectives appeared 66 times in TACE while they were found 53 times in SACE. Both results were approved by LL results (p<0.05).

## 5. Conclusion

In our study, epistemic adjectives in the doctoral dissertations written by TAEs (Turkish-Speaking

Academic Authors of English), NAEs (Native Academic Authors of English) and SAEs (Academic Authors of English) were investigated using a quantitative analysis method. Overall, it has been revealed that 7 out of 26 types were not found in any corpora. This section offers the evaluation of the research questions, implications to language teaching and a few suggestions for further research.

### 5.1. Evaluation of Research Questions

Our first research question investigated whether TAEs and NAEs significantly differ in the use of epistemic adjectives. The findings have suggested that TAEs used epistemic adjectives in their doctoral dissertations less frequently than NAEs, which was confirmed with the results of statistical analysis ( $p < 0.05$ ). In line with this finding, an overuse was observed in TACE against NACE with respect to certainty adjectives. On the other hand, an underuse was found in TACE against NACE in that of likelihood adjectives, which was also confirmed by statistical results.

The second question was posed to find out whether SAEs and NAEs significantly differ regarding the use of these adjectives. The results have demonstrated that both certainty and likelihood adjectives were employed less frequently in SACE than NACE, which indicates an underuse in SACE against NACE in the categories in concern. Even though the statistical results have approved both cases, it is noteworthy that the difference between the two corpora in terms of certainty adjectives is not so significant as the one counted for likelihood adjectives.

The third question was formed to reveal whether TAEs and SAEs significantly differ concerning the use of the items in question. In accordance with the results, both certainty and likelihood adjectives occurred more frequently in TACE than SACE, which indicates that they were significantly overused by TAEs against SAEs.

### 5.2. Implications to Language Teaching

The study has shown that TAEs and SAEs display a strong tendency to use certainty adjectives while NAEs mostly prefer likelihood ones, which lead us to the conclusion that the non-native academic authors of English are more assertive whilst their native colleagues are more cautious in their academic writing. That is, the native group is more inclined to use tentative statements while conveying their points when compared to the non-native groups. At this point, as put by [1] Adel (2010), instructors who teach academic writing to graduate students should be both descriptive linguists and language policy makers. Namely, they should guide students with recommendations as well as good arguments in the writing sessions. In this case, for example, graduate students, especially those who are studying social sciences, should be provided the principle that likelihood expressions are expected to be preferred in academic writing rather than certainty ones while "taking a position". In other words, they might be instructed to be cautious in presenting their views and interpreting the results of their works. Namely, they might be advised to avoid using expressions like A is the result of B or B is caused by X and to prefer A might have stemmed from B or B might have triggered A, respectively.

Spack (1988) argues that especially non-native graduate students can learn academic writing by observing the process during which academic texts are produced by professional writers or by attending programmes designed with the aim of teaching them how to write a research paper. Likewise, Adel (2010) advocates that corpus-based courses on academic writing are needed for non-native graduate students proposing that non-native speakers lack native-speaker intuitions and they are likely to transfer writing patterns in their L1 to a rhetorically different L2 in an appropriate manner. So, a second suggestion might be that instructors teaching academic writing should utilize corpora including academic works produced by native academic authors of English. At this point, using the British Academic Written English (BAWE) Corpus (Nesi and Gardner, 2012) could be practical as it consists of a large number of academic works produced by native speakers of English. Namely, the instructor might analyse an article in BAWE with students regarding all or some selected principles of academic writing. In our case, it could be analysed in terms of certainty and likelihood adjectives and the students might be asked to write reports on the authorial stance used in the article at stake. Activities of such kind can be employed with individual students as well as small groups of students.

Lastly, graduate students might be asked to examine the essays they wrote before they were introduced academic writing course or their own exam papers and to write a reflection paper or rewrite the papers in concern considering what they have learned about author stance during the course. Alternatively, they might be asked to compare their papers with the ones written by native and/or other non-native graduate students with respect to certainty and likelihood adjectives used in those papers and to write a report on the similarities and/ or differences between them.

### 5.3. Suggestions for Further Research

Our study is limited to the investigation of doctoral dissertations written by Turkish-speaking, Spanish-speaking and native academic authors of English between 2005 and 2012. It is also limited to the field of English Language Teaching, English Language and Literature, Applied Linguistics and Modern Languages. Hence, its results cannot be generalized to any other settings. Therefore, further research might involve building and investigating much broader corpora which include dissertations authored by Turkish-speaking, Spanish-speaking and native academic authors of English produced in various fields prior to 2005. In addition, author stance could be investigated concerning the use of certainty and likelihood expressions through spoken productions of academics in events such as conferences or symposiums. Finally, the reasons why certainty items are used more frequently than likelihood items by non-native academic authors when compared to native academic authors might be researched.

### Acknowledgement

The author would like to thank academic authors of English who produced dissertations investigated in this particular study.

## References

- (modal verbs) in research articles. PhD Thesis, Anadolu University.
- [1] Ädel, A. (2010). Using corpora to teach academic writing: Challenges for the direct approach. In M. C. Compoy-Cubillo, B. Belles-Fortuno & M. L. Gea-Valor (Eds.), *Corpus based approach to English language teaching* (pp. 39-54).
- [2] Baker, P., Hardie, A., & McEnery, T. (2006). *A glossary of corpus linguistics*. Edinburgh: Edinburgh.
- [3] Biber, D. & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*, 9, 93-124.
- [4] Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.
- [5] Conrad, S. & Biber, D. (2000). Adverbial making of stance in speech and writing. In S. Hunston and G. Tompson (Eds.), *Evaluation in Text*, 56-73. Oxford: Oxford University Press.
- [6] Gabrielatos, C. & McEnery (2005). Epistemic modality in MA dissertations. In Fuertes Olivera, P.A. (Ed.), *Lengua y Sociedad: Investigacion e srecientes en lingüística aplicada* (pp. 311-331). *Lingüística y Filología*, 61. Valladolid: Universidad de Valladolid.
- [7] Gosden, H. (1990). Discourse functions of subject in scientific research articles. *Applied Linguistics*, 14(1), 56-75.
- [8] 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*. Lund: Lund University Press, 37-51.
- [9] Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192.
- [10] Kafes, H. (2009). Authorial stance in academic English: Native and non-native academic speaker writers' use of stance devices
- [11] Kanté, I. (2010). Head nouns as modal stance markers –academic texts vs. legal texts. *Enonciation et rhetorique dans l'écrit scientifique, Lidil* 41, 121-135.
- [12] MacIntyre, R. (2010). *How do they stand? Research into the representation of stance in academic writing and its implications for pedagogy*. Sophia University Faculty of Liberal Arts, 19-32.
- [13] Oxford advanced learner's dictionary (2000). Oxford: Oxford University Press.
- [14] Salager-Meyer, F. (1997). I Think that Perhaps You Should: A Study of Hedges in Written Scientific Discourse. In Miller, T. (Ed.), *Functional Approaches to Written Text: Classroom Applications*. English Language Programs: United States Information Agency, 105-118.
- [15] Scott, M. (2012). *WordSmith Tools version 6*, Liverpool: Lexical Analysis Software.
- [16] Skelton, J. (1988). Comments in academic articles. In P. Grunwell (Ed.), *Applied Linguistics in Society: British Studies in Applied Linguistics* 3. 98-108. London: Centre for International Language.
- [17] Stubbs, M. (1986). A matter of prolonged field work: notes toward a model grammar of English. *Applied Linguistics*, 7(1), 1-25.
- [18] Ventola, E. & Maureanen, A. (1990). *Research and Writing in English*. Helsinki: Yliopistopaino.