Post-predicate that-clauses controlled by verbs in native and non-native academic writing: A corpus-based study

SAMET TAŞÇI a

YUSUF ÖZTÜRK b

a Nevşehir Hacı Bektaş Veli University, Turkey
samettasci@nevsehir.edu.tr

b Muş Alparslan University, Turkey
y.ozturk@alparslan.edu.tr

Abstract

Post-predicate that clauses are one of the indicators of evaluative language which shows the attitude, position, and personal evaluation of the writers in academic discourse. Therefore, this study aims to investigate the distribution and frequently used controlling verbs of post-predicate that clauses controlled by verb (VTHAT) in academic writings of Turkish learners of English and to compare them with those of native English students and expert writers. For this purpose, a specialized research corpus with three main sub-corpora was compiled: MA and PhD theses written by L2 students and native English students and published research articles written by native experienced writers. The results of the study showed that both L2 and native English students significantly overused VTHAT structures compared to the native experienced writers; however, Turkish students used VTHAT structures more frequently than their native peers in academic discourse. The analysis of the controlling verbs in VTHAT constructions showed that eight out of the 10 top frequent verbs in the Turkish students’ theses were common across either three or two sub-corpora. However, there were various significant differences in the frequencies of these controlling verbs between the sub-corpora, which can imply that Turkish students use such constructions to have a low authoritative presence in terms of their authorial stance in academic writing. The findings were discussed with reference to the previous findings, and pedagogical implications were offered.

Keywords: that-clauses, evaluative that, evaluation, academic writing, corpus-based

Introduction

Academic genre is expected to be purely objective and informative, supported by proof; however, it may not be as objective as it is thought. Academic writing is evaluative and interpersonal rather than...
simply informational and objective (Hyland & Tse, 2005a). In other words, in academic writing, writers not only express the objective facts but also evaluate the results, comment on the findings (Hyland and Tse, 2005a) and try to persuade the readers “by claiming solidarity with readers, evaluating their material and acknowledging alternative views” (Hyland, 2004, p. 13). Thus, the writers express evaluation by proposing an idea, stating “a personal opinion” (Hunston, 2010, p.12), adopting “an attitude,” “engagement,” and “graduation” (Martin & White, 2005, p.35), and building and maintaining relations with the readers in their writing (Thompson & Hunston, 2000, p.8). By using evaluative language, the writers define their position in their writings. Du Bois (2007) noted that evaluation is one of the acts of stance-taking. Evaluation in academic discourse is subjective and shows the attitude of the writer (Hunston 1994). In the literature, evaluative language has been studied under different titles such as evaluation (Thompson & Hunston, 2000), stance (Biber & Finegan, 1989), appraise (Martin & White, 2005), metadiscourse (Biber 2006; Hyland, 1998; Hyland & Tse, 2004; Crismore, 1989), sentiment (Taboada & Grieve, 2004), and attitudinal or affective language (Ochs, 1996). Although different terminologies have been used to refer to evaluative language, they are equivalent on the common ground (Hunston, 2010).

Certain lexical and grammatical items can be used to express personal evaluation of the writer in academic discourse. The lexical items that are used to differing extent to indicate evaluation include the lexical choices (e.g., adjectives, verbs, nouns) of the writer. The grammatical items are the grammatical devices such as adverbials, complement clauses, modals, and prepositional phrases (see Biber et. al. 1999). A type of complement clauses, post-predicate that-clause controlled by verbs (VTHAT) is a grammatical device in which a verb in the matrix clause is the controlling element of the following that complementizer and shows the writers’ viewpoint, opinion and attitude about the subordinate clause (Hyland & Tse, 2005b). Ishikawa (2015) noted that post-predicate that-clauses controlled by verbs are widely used to report utterances, thoughts, feelings, and stances of human subjects. Similarly, Hyland and Tse (2005a) stated that this structure is an important means of providing author commitment and evaluation, and therefore, helps writers to manage their discourse in various ways and to signal a clear stance towards the information they present. In spite of its importance, VTHAT clauses are one of the least noticed interpersonal features of academic discourse (Hyland & Tse, 2005b).

As evaluation is the key aspect of academic writing (Hyland & Tse, 2005b), and VTHAT constructions are among the most frequent structures of academic writing (Biber et. al. 1999), investigating the distribution, and frequently used controlling verbs of VTHAT in academic discourse may provide insight on the position of researchers. Moreover, previous corpus-based studies analyzing VTHAT showed that the frequency and frequently used controlling verbs of VTHAT varied according to L1 background (i.e., Chinese, French, Spanish) (Biber & Reppen, 1998; Ishikawa, 2015), disciplines (Hyland & Tse, 2005a, 2005b; Lou, 2014), and proficiency levels (Ishikawa, 2015) of the writers. In addition, the use of evaluation in academic discourse composes difficulty to L2 learners (Hyland & Tse, 2005a; Lou, 2014). In this regard, the current study aims to investigate the distribution and frequently used controlling verbs of VTHAT in academic writings of L2 Turkish learners of English to reveal the attitude and position of Turkish students in academic discourse in comparison to native English student and expert writers, and to reveal to what extent Turkish students differentiate from native speakers in terms of evaluation.

**Literature Review**

**What is evaluation?**

Evaluative language has an important role in text, in that it shows the signals of comparison,
subjectivity and social value in a text (Thompson and Hunston, 2000 p.13). Therefore, the use of evaluative language in a text does not only reflect the writer’s attitude and viewpoint but also a value system of the community which the writer belongs to (Ikeo, 2007). Evaluation has been one of the topics of interest in discourse studies. Lexis, grammar and text were analyzed in order to find out evaluative language because evaluation can be realized by textual aspects such as lexis, grammar and textual organization (Ikeo, 2007). Lexical means involves the use of lexical choices to convey evaluative meaning e.g., hate, nice; grammatical means involves the use of some grammatical devices e.g., adverbials and complements, to express evaluation related to a proposition (Lou, 2014). Textual means involves analyzing throughout the text rather than looking to one particular part of it e.g., the position of a paragraph in a text.

The concept of evaluation has been studied from different aspects in different contexts and in different registers (Biber & Finegan, 1989; Biber, 1999; Biber & Reppen, 1998; Ishikawa, 2015; Hyland & Tse, 2005a, 2005b; Liu & Thompson, 2009; Lou, 2014). With respect to evaluation in general, the foci of these studies mainly involved speech styles marked by stance (Biber & Finegan, 1989), interactive role of evaluation and evaluative judgements in academic research articles (Thetela, 1997), stance in spoken and written academic language (Biber, 2006), and attitude in native and non-native writing (Liu & Thompson, 2009).

One of the preliminary studies on evaluation (Biber & Finegan, 1989) focused on the lexical and grammatical expression of attitude, feeling, judgment or commitment concerning the propositional content of a message in 500 texts extracted from London-Lund spoken and written corpora. The researchers examined adjectives, verbs and modal markers giving 12 categories of stance markers based on semantic and grammatical criteria. On the other hand, Thetela (1997) analyzed linguistic and contextual clues to find out the ways in which evaluation works in academic discourse to convey the overall purpose of the writer of the text, concluding that different disciplines present different evaluation systems (research oriented or topic oriented) distributed differently in academic register. On a later study, Biber (2006) examined the academic genre in spoken and written register focusing on the lexico-grammatical features used for the expression of stance describing register variations and found that stance is expressed to strikingly different extents, and for different purposes, in different registers. In a case study, Liu and Thompson (2009) focused on students’ argumentative writing in both English and Chinese to find out the use of evaluative language drawing on appraisal theory. The study showed that the use of judgment, appreciation and effect slightly differed in English and Chinese argumentative essays of the students. The abovementioned studies were conducted on the different aspects of evaluative language in different contexts, using different methodology in different genres and all these studies showed that writer evaluation is a part of language in different registers. These studies also concluded that authorial stance and evaluation are prevalent features of academic language, and not only the content but also the writer’s angle is of significance for readers to interpret the content from the intended perspective. In this sense, evaluative language is of great importance for academic writers. Moreover, the use of evaluative features may differ in native and non-native speaker discourse, which suggests that this aspect of academic writing should be focused in corpus-based research attempts to offer pedagogical implications for academic writers and graduate students in EFL contexts.

Post-predicate that-clauses controlled by verbs

One of the powerful ways for expressing evaluative meaning in academic discourse is that constructions or in other words, post-predicate that-clauses, a grammatical pattern in which a that-complement clause is contained in a higher super-ordinate clause to complete its construction and which together project the writer’s attitudes or ideas about the message being delivered (Hyland and
Post-predicate *that*-clauses can be controlled by an adjective or a noun but usually a communication verb such as *suggest* or *prove*, a cognitive verb such as *think* or *believe*, or a speech act verb, like *say* or *state* (Biber et al., 1999, p.661) as is represented in Table 1.

**Table 1 Examples of post-predicate that-clauses controlled by adjective, noun and verb**

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Evaluated entity</th>
<th>Controlled by</th>
</tr>
</thead>
<tbody>
<tr>
<td>She was <strong>glad</strong> <strong>that</strong> today was her day off.</td>
<td>An adjective</td>
<td></td>
</tr>
<tr>
<td>It is the <strong>fact</strong> <strong>that</strong> the cold war is now behind us.</td>
<td>A noun</td>
<td></td>
</tr>
<tr>
<td>We <strong>suggest</strong> <strong>that</strong> more work needs to be done.</td>
<td>A verb</td>
<td></td>
</tr>
</tbody>
</table>

*That*-constructions project the writer’s evaluation such as attitudes or ideas and have attracted the attention of researchers who have examined this phenomenon in terms of frequency and function in native and non-native English contexts. A brief summary of recent studies on the use of *that*-constructions is shown in Table 2. Some of these studies compared the use of evaluative *that*-constructions by non-native speakers and native speakers of English (Biber & Reppen, 1998; Lou, 2014; Ishikawa, 2015) while others focused on the disciplinal variations of the usage (Hyland & Tse, 2005a, 2005b). Both approaches are of great significance for teaching academic writing and English for academic purposes.

The preliminary study of Biber and Reppen (1998) on the use of evaluative *that*-construction by native and non-native speakers in written and spoken registers showed that the use of *that*-construction changes according to register and L1 background. Compared to academic prose, *that*-construction is mostly used in conversation both by native and non-native speakers of English. Spanish learners of English use *that*-constructions more frequently than other learners. While *think*, *say* and *know* are the most common controlling verbs in conversation, none of these verbs are particularly common with *that*-clauses in academic prose.

**Table 2 A brief summary of recent studies on that-construction**

<table>
<thead>
<tr>
<th>Study</th>
<th>Register</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyland &amp; Tse (2005a)</td>
<td>Research abstracts in different disciplines</td>
<td>Academic scholars</td>
</tr>
<tr>
<td>Hyland &amp; Tse (2005b)</td>
<td>Research, master and dissertation abstracts</td>
<td>Academic scholars, and L2 graduate students in Hong Kong</td>
</tr>
<tr>
<td>Wang &amp; Chen (2012)</td>
<td>Abstracts of MA theses in different disciplines</td>
<td>Chinese MA students, and academic scholars</td>
</tr>
<tr>
<td>Kilimci (2014)</td>
<td>Argumentative essays</td>
<td>Turkish and American university students</td>
</tr>
<tr>
<td>Lou (2014)</td>
<td>Dissertation abstracts and article abstracts in different disciplines</td>
<td>Chinese doctoral students, and academic scholars</td>
</tr>
<tr>
<td>Ishikawa (2015; 2016)</td>
<td>Written and spoken</td>
<td>Asian EFL &amp; ESL learners with six different L1 backgrounds, and native speakers</td>
</tr>
<tr>
<td>Kim &amp; Crosthwaite (2019)</td>
<td>Research articles in business and medicine</td>
<td>Academic scholars</td>
</tr>
</tbody>
</table>

In a later study, Hyland and Tse (2005a) explored the disciplinary variations in the frequency, forms and functions of evaluative *that*-constructions in research article abstracts from six disciplines. The result of the study showed that while computer science and business studies have the highest number of evaluative *that*-constructions, electronic engineering made the least use of this construction. Their follow-up study (2005b) explored the effect of genre difference in dissertations and journal articles
from various disciplines with the justification that published writers would have greater experience. They demonstrated that the dissertation abstracts contained a considerably lower number of evaluative that-clauses than the article abstracts in total (4.1 times vs. 6.9 times per 1,000 words); however, the frequencies were quite closer in applied linguistics (5.1 times vs. 6.4 times per 1,000 words) with research articles involving more evaluative that-clauses. In a similar study on graduate academic writing, Wang and Chen (2012) compiled a corpus of MA thesis abstracts written by Chinese students, and research article abstracts published in international reputable journals. According to their findings, the frequencies of evaluative that-clauses were nearly two times less frequent in the MA thesis abstracts than in the article abstracts (3.5 times vs. 7.2 times per 1,000 words), which was in line with the findings of Hyland and Tse (2005b). Lou (2014) also explored the use of evaluative that-constructions in L2 doctoral dissertation abstracts, and academic scholars’ research article abstracts from various disciplines. The study revealed that Chinese graduate students used evaluative that-constructions less frequently than academic scholars (3.6 times vs. 6.9 times per 1,000 words), which might be caused by learners’ misunderstanding of the compositional features of abstracts and lack of consciousness of the evaluative feature of this genre. In a more recent study, Kim and Crosthwaite (2019) examined disciplinary differences in the use of that-complement clauses in research articles from the disciplines of business and medicine. They also reported the existence of a disciplinary variation, with research articles in business involving significantly higher number of that-clauses. Consequently, it can thus be argued that the frequency and usage of that-clauses, including those that are controlled by verbs, vary across different disciplines, and academic scholars make more use of VTHAT in article abstracts than graduate students in dissertation abstracts.

Comparing native speakers and non-native learners, Ishikawa (2015) aimed to examine the use of post-predicate that-clauses controlled by verbs in speech and writing. The distinction of the study from the previous ones is that this study focused on EFL, ESL and native speaker context, and revealed that VTHAT is most often used by native speakers and then by ESL learners and least often by EFL learners. Moreover, the study also revealed that native speakers use VTHAT more often in writing than in speaking, whereas majority of the learners use it mostly in speaking. According to the researcher, this shows that the learners perceive VTHAT as a kind of prefabricated chunk, easing their cognitive load as they struggle to express their thoughts and try to attain fluency. Moreover, the collocating verbs used by native speakers and learners differ to a great extent, with learners having a limited repertoire when it comes to typical combinations of verbs and the complementizer. In a further study, Ishikawa (2016) extended this analysis to how often learners and native speakers omit that as a complementizer, and the type of reporting verbs they use in speech and writing. He reported that native speakers have a nearly balanced use of VTHAT in speech and writing with learners using this structure less often, especially advanced learners in writing. His analysis of omitting showed that learners omit complementizers more often than native speakers, particularly in writing, with advanced learners omitting them even more, while native speakers do the majority of this omitting in speech. The study also reported that although the use of reporting verbs by both groups is similar in variety, with the top frequent verbs being in common, learners are observed to use different sets of verbs for the two registers, inappropriate forms and inflectional forms of basic verbs, contrary to native speakers.

The aforementioned studies either compared the use of evaluative that-construction by non-native speakers and/or native speakers of English or the disciplinal variations regarding this construction. Either of these approaches has importance for teaching academic writing and English for academic purposes since non-native-like usage by graduate students and academic writers and differences in the use of the evaluative that-construction across disciplines can provide insights for academic writing classes as well as the professional needs of non-native scholars as they are expected to
produce texts and publications in English as a lingua franca. Having stated the significance of VTHAT as a research topic and its possible implications, although it has been studied in various contexts with data on learners with varying L1 backgrounds, the research on the use of VTHAT by Turkish EFL learners and academic writers has been extremely limited, with Kilimci (2014) being an exception. He investigated the use of that-clauses controlled by verbs and adjectives in argumentative essays by Turkish and American university students, and reported slightly fewer instances of that-clauses in Turkish students’ essays. However, his study was limited to argumentative essays written on a variety of different topics, and a study concentrating on genres that “represent the key research genres of the academy” (Hyland, 2008, p. 47) such as MA and PhD theses (i.e., not only abstracts) and research articles could possible give a clearer picture of how Turkish EFL learners make use of VTHAT in academic texts.

An analysis of Turkish learners’ use of VTHAT can provide researchers and educators important insights. This is partly due to the characteristic of Turkish as an agglutinative language with a far more complex morphological structure than English. Another challenge for Turkish learners may be due to the sentence structure of Turkish as it is an SOV language, unlike English, with verb at the end of the sentence and all other complementizers and complement phrases accumulated before. Considering the importance of that-constructions in the light of related literature and the lack of research in the Turkish context, this study aims to find out how Turkish speakers of English use VTHAT in their academic writing compared to native speakers of English in terms of frequency and frequently used controlling verbs of VTHAT.

The following questions are formulated for the purpose of the study:

1. How do Turkish speakers of English and native speakers of English use VTHAT in their academic writing in terms of overall frequency?
2. How do Turkish speakers of English and native speakers of English use VTHAT in their academic writing in terms of frequently-used controlling verbs?

**Methodology**

As mentioned in the previous sections, the main focus of the study is to find out the overall frequency and frequently used controlling verbs of VTHAT. The primary reason for focusing on the VTHAT construction is that they show the evaluative language (proposing an idea, stating personal opinion (Hunston, 2010 p.12), adopting attitude (Martin & White, 2005, p.42), and building and maintaining relations (Thompson & Hunston, 2000, p.8) used by the writer. Although evaluation has been shown to play a central role in text and discourse, its identification in text is not always straightforward (Thetela, 1997) and a single evaluative item is often ambiguous until supported by other items that make the same point (Hunt & Vipond, 1986 p.67). This is why, that-constructions supported by other items such as verbs are examined to find out the writers’ evaluation because VTHAT constructions are a grammatical way of finding the evaluative language used by the writers.

**Research corpus**

In order to examine the frequency and frequently used controlling verbs of VTHAT in the texts of native and non-native speakers of English, a specialized research corpus with three main sub-corpora was compiled. The corpora used in this study consisted of published research articles, MA and PhD theses written in the area of language teaching. Theses and research articles were chosen as the genres for analysis because “they represent the key genres of the academy” (Hyland, 2008 p.47). Moreover, these genres would allow for more occurrences of the structure under examination
compared to shorter text-types such as argumentative essays.

Due attention was paid to ensure a high level of comparability across the sub-corpora. The non-native English academic sub-corpus used in the study consisted of 30 MA theses and 20 PhD theses approved to be published in a 10-year interval (2005-2015) and extracted from the Turkish Higher Education Council’s Thesis Center (https://tez.yok.gov.tr/UlusalTezMerkezi/giris.jsp). For the native speaker reference corpus, a similar process was followed. The native academic sub-corpus consisted of 30 MA theses and 20 PhD theses extracted from ProQuest Dissertation and Theses Center (http://www.proquest.com/products-services/dissertations/). In order to make the sub-corpora comparable, the theses having similar topics were chosen. As the theses are not at the same length, there is a slight difference in terms of their size. Table 3 shows a brief summary of distribution of total number of texts and words in native and non-native academic sub-corpora.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>No. of Texts</th>
<th>No. of Words</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TARC MA</td>
<td>30</td>
<td>615,724</td>
<td>1,358,691</td>
</tr>
<tr>
<td>TARC PhD</td>
<td>20</td>
<td>742,967</td>
<td></td>
</tr>
<tr>
<td>NARC MA</td>
<td>30</td>
<td>462,661</td>
<td>1,257,808</td>
</tr>
<tr>
<td>NARC PhD</td>
<td>20</td>
<td>795,147</td>
<td></td>
</tr>
<tr>
<td>NRAC</td>
<td>50</td>
<td>435,521</td>
<td>435,521</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>150</strong></td>
<td><strong>435,521</strong></td>
<td><strong>3,052,020</strong></td>
</tr>
</tbody>
</table>

TARC: Turkish Academic Register Corpus  
NARC: Native Academic Register Corpus  
NRAC: Native Research Article Corpus

The research corpus also included a sub-corpus research article published by native English speakers. The research articles were taken into the analysis as a reference corpus of experienced writers. They were regarded as experienced writers compared to students as writers of their thesis. Sommers (1980) used “the terms student writers and experienced writers because the principal difference between these two groups is the amount of experience they have had in writing” (p. 380). In this regard, a native research article corpus was compiled as the baseline data while comparing MA and PhD theses of native and non-native writers. Fifty research articles from a respectable journal in the field, namely Language Learning, between were obtained from a 10-year interval.

**Data analysis**

Following the collection of texts for the research corpus, the text files were converted into plain text (i.e., .txt) and all sections but the main texts produced by their authors such as abstracts, references, tables, figures and direct quotations were excluded. This process was followed by a crucial step, tagging the corpus. As the purpose of the study was to examine the post-predicate *that*-constructions, information regarding the parts of speech of the words in the corpus was needed to conduct the analysis. In this respect, the CLAWS (the Constituent Likelihood Automatic Word-tagging System) online tagger software was used (http://ucrel.lancs.ac.uk/claws/) to tag the corpus compiled in the study. This software helps automatically tag parts of speech in a corpus, and it was employed for this purpose to tag the words in the theses and articles to find out grammatical parts of speech such as verbs, nouns, adjectives, adverbs, determiners, and so on. Freeware corpus concordance software Antconc 3.4.3 (Anthony, 2011) was used to identify the word combinations that matched VTHAT in the tagged corpus. All such instances were manually examined. The number of the instances as well
as the controlling verbs were compared between the sub-corpora to see if there were any significant differences in the usage of native and non-native writers in terms of VTHAT. For significance analysis, log-likelihood scores were calculated using an online log likelihood calculator (http://ucrel.lancs.ac.uk/llwizard.html), which calculates the effect size of the results. Raw frequencies as well as normalized frequencies were used while reporting the results since the sub-corpora contained varying number of words although they had the same number of texts.

Results

The first research question of the study relates to the overall frequency of VTHAT structures in native and non-native theses and native research articles. Table 4 below gives the overall frequency of occurrences of the VTHAT structures in native and non-native theses. According to the table, in total, the Turkish students used VTHAT structures more frequently than their native peers in their MA and PhD theses (4304 times vs. 3818 times, respectively). To find out if there were any significant differences in-between, log-likelihood scores were calculated. The analysis revealed no statistically significant difference between the theses of both groups (LL = 3.69, p > .05). Yet, a more detailed analysis is needed to gain an in-depth understanding of their usage of VTHAT.

The table shows that for both native-English speaking students and Turkish students, VTHAT structures were used more frequently in PhD theses than in MA theses, but this was for raw frequencies, and would be expected considering the length of PhD theses. Therefore, it would be more meaningful to compare the normalized frequencies that showed that both natives English and Turkish students made use of VTHAT structures to the same extent in their PhD theses (i.e., 2.99 times per 1,000 words), while in the MA theses the Turkish students had a relatively higher frequency of VTHAT compared to the native students (3.38 times vs. 3.11 times per 1,000 words, respectively).

Table 4  The frequency of the VTHAT in the student sub-corpora

<table>
<thead>
<tr>
<th></th>
<th>MA</th>
<th></th>
<th></th>
<th>PhD</th>
<th></th>
<th></th>
<th>TOTAL</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw</td>
<td>Per</td>
<td>Per text</td>
<td>Raw</td>
<td>Per</td>
<td>Per text</td>
<td>Raw</td>
<td>Per</td>
<td>Per text</td>
</tr>
<tr>
<td></td>
<td>freq.</td>
<td>1,000</td>
<td></td>
<td>freq.</td>
<td>1,000</td>
<td></td>
<td>freq.</td>
<td>1,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>words</td>
<td></td>
<td></td>
<td>words</td>
<td></td>
<td></td>
<td>words</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td>2080</td>
<td>3.38</td>
<td>69.3</td>
<td>2224</td>
<td>2.99</td>
<td>111.2</td>
<td>4304</td>
<td>3.16</td>
<td>86</td>
</tr>
<tr>
<td>Native E</td>
<td>1440</td>
<td>3.11</td>
<td>48</td>
<td>2378</td>
<td>2.99</td>
<td>118.9</td>
<td>3818</td>
<td>3.03</td>
<td>76</td>
</tr>
</tbody>
</table>

Comparing the frequency of VTHAT in the MA theses, the statistics showed that the Turkish students used these structures more frequently than their native peers. Log-likelihood analysis revealed that this difference was significant, and the Turkish students overused VTHAT in their MA theses compared to their native English speakers (LL = 5.74; p < .05) whereas there was no such significant difference in VTHAT constructions between Turkish and native English students’ PhD theses (LL = 0.00; p > .05). Log-likelihood scores were also calculated to compare MA and PhD theses within each group. For the Turkish students, VTHAT structures occurred 3.38 times per 1,000 words in their MA theses and 2.99 times per 1,000 words in their PhD theses. This difference was statistically significant (LL = 15.68; p < .0001), indicating that the MA theses written by the Turkish students involved an excessive use of VTHAT combinations compared to their PhD theses. On the other hand, however, for the native students, no statistically significant difference was observed in the occurrences of VTHAT between their MA and PhD theses (LL = 1.42; p > .05).

The relative frequencies of VTHAT structures were compared across the three sub-corpora examined.
in the study that involved native English speakers research articles in addition to the Turkish and native students’ MA and PhD theses. As far as relative frequencies were concerned, the research articles contained the lowest frequency (2.28 times per 1,000 words) when compared to the Turkish students’ (3.16 per 1,000 words) and native students’ theses (3.03 per 1,000 words). The difference between the three sub-corpora is represented in Figure 1.

![Figure 1](image)

**Figure 1** Normalized frequency of VTHAT in theses and native research articles

Log-likelihood statistics showed that VTHAT clauses were significantly underused in the research articles written by native English speakers when compared to the native and Turkish students’ theses ($LL=80.61, p > .05; LL=107.82 p > .05$). Scholars who speak English as their first language, and who can be regarded as experienced writers as they have published in a reputable international journal in their field used VTHAT combinations to a significantly lesser extent in their research articles.

Following the frequency analysis of VTHAT structures in Turkish and native students’ theses and native speakers’ research articles, top ten frequently-used controlling verbs for these structures were also examined in the three sub-corpora. Table 5 shows that some of the controlling verbs namely *find, suggest, indicate, state, show, note, say, argue,* and *mean* were common in the three corpora with striking differences in their frequency.

While the verb *find* was the most frequent controlling verb in native students’ theses with the raw frequency of 336 times, it was the tenth most frequent verb in the Turkish students’ theses with 110 times. Raw as well as the normalized frequencies (in 1,000 words) suggest that the verb *find* was used three times more frequently in native students’ theses than their Turkish peers. This verb was also quite frequent in the research articles with 109 times, and in fact its relative frequencies were close in the research articles and the native students’ theses (.250 times vs. .267 times per 1,000 words, respectively). In a similar vein, there was a statistically significant difference in the frequency of this verb between the native and Turkish students’ theses ($LL=138.16, p < 0.0001$), but almost no difference between the native students’ theses and the native scholars’ research articles ($LL=-0.35, p > 0.05$).
The second most frequent controlling verb in native students’ theses (i.e., 260 times), suggest, was common to all three sub-corpora and the most frequent one in the research articles (i.e., 137 times). As far as the normalized frequencies are concerned, it was used considerably more frequently in native students’ theses and native scholars’ research articles than the Turkish students’ theses. Based on the log-likelihood statistics, the Turkish students significantly underused suggest as a controlling verb in VTHAT combinations when compared to the native students’ theses ($LL=-5.74, p < 0.05$) and the research articles ($LL=-47.12, p < 0.0001$).

The analysis also showed that the verb indicate appeared as the third most frequent controlling verb in native theses (235 times) while it was the second in non-natives (375 times) and forth in native academic articles (61 times). Normalized frequencies and the log-likelihood statistics suggest a significant overuse by the Turkish students when compared to the native students’ theses ($LL=32.31, p < 0.0001$) and the research articles ($LL=28.28, p < 0.0001$).

The most frequent controlling verb in Turkish students’ theses was found to be the verb state occurring 516 times. Although the verb state was the most frequent one in the Turkish student sub-corpus and forth most frequent in the native student sub-corpus (235 times), interestingly it was not in the top ten frequently used controlling verbs in the native research articles, which also reflected in the significance analysis. When compared to the Turkish and native students’ theses, the scholars who produced the texts in the research article sub-corpus significantly underused state as a controlling verb in VTHAT combinations ($LL=201.72, p < 0.0001$; $LL=70.17, p < 0.0001$, respectively).

The last verb that was common among the top ten most frequent controlling verbs in the three sub-corpora was show with Turkish students having the highest frequency (.255 times per 1,00 words) followed by research articles (.204 times per 1,00 words) and native students’ theses (.121 times per 1,00 words). Taking into consideration the log-likelihood statistics, the Turkish students overused show as a controlling verb when compared to the native students and scholars, but only the difference between the two student groups was statistically significant ($LL=63.06, p < 0.0001$).

Among the top ten most frequent controlling verbs, there were verbs that were not common across the three sub-corpora including believe, note, feel, report, claim, say, reveal, mean, argue, ensure, demonstrate and conclude, meaning that these verbs did not occur at the top of the frequency list in one or two of the sub-corpora. A closer look at these verbs shows that, in comparison to the native speakers’ theses and native scholars’ research articles, the Turkish students’ theses showed a significant underuse of note ($LL=-69.72, p < 0.0001$; $LL=-22.76, p < 0.0001$) and ensure ($LL=-5.60, p < 0.05$; $LL=17.94, p < 0.0001$). Likewise, the Turkish students were observed to significantly overuse reveal ($LL=62.09, p < 0.0001$; $LL=26.63, p < 0.0001$), claim ($LL=180.75, p < 0.0001$; $LL=87.34, p < 0.0001$) and say ($LL=14.41, p < 0.001$; $LL=56.41, p < 0.0001$) when compared to the

---

**Table 5** Top ten frequently-used controlling verbs of VTHAT in three sub-corpora

<table>
<thead>
<tr>
<th>No</th>
<th>Native Verbs</th>
<th>Raw</th>
<th>Normalized</th>
<th>Turkish Verbs</th>
<th>Raw</th>
<th>Normalized</th>
<th>Research Articles Verbs</th>
<th>Raw</th>
<th>Normalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Find</td>
<td>336</td>
<td>.267</td>
<td>State</td>
<td>516</td>
<td>.379</td>
<td>Suggest</td>
<td>137</td>
<td>.314</td>
</tr>
<tr>
<td>2.</td>
<td>Suggest</td>
<td>260</td>
<td>.206</td>
<td>Indicate</td>
<td>375</td>
<td>.276</td>
<td>Find</td>
<td>109</td>
<td>.250</td>
</tr>
<tr>
<td>3.</td>
<td>Indicate</td>
<td>235</td>
<td>.186</td>
<td>Show</td>
<td>347</td>
<td>.255</td>
<td>Show</td>
<td>89</td>
<td>.204</td>
</tr>
<tr>
<td>4.</td>
<td>State</td>
<td>229</td>
<td>.182</td>
<td>Claim</td>
<td>311</td>
<td>.228</td>
<td>Indicate</td>
<td>61</td>
<td>.140</td>
</tr>
<tr>
<td>5.</td>
<td>Show</td>
<td>153</td>
<td>.121</td>
<td>Suggest</td>
<td>226</td>
<td>.166</td>
<td>Argue</td>
<td>50</td>
<td>.114</td>
</tr>
<tr>
<td>6.</td>
<td>Believe</td>
<td>145</td>
<td>.115</td>
<td>Say</td>
<td>221</td>
<td>.162</td>
<td>Mean</td>
<td>40</td>
<td>.091</td>
</tr>
<tr>
<td>7.</td>
<td>Note</td>
<td>143</td>
<td>.113</td>
<td>Reveal</td>
<td>200</td>
<td>.147</td>
<td>Note</td>
<td>38</td>
<td>.087</td>
</tr>
<tr>
<td>8.</td>
<td>Say</td>
<td>136</td>
<td>.108</td>
<td>Mean</td>
<td>169</td>
<td>.124</td>
<td>Ensure</td>
<td>33</td>
<td>.075</td>
</tr>
<tr>
<td>9.</td>
<td>Feel</td>
<td>124</td>
<td>.098</td>
<td>Argue</td>
<td>129</td>
<td>.094</td>
<td>Demonstrate</td>
<td>27</td>
<td>.061</td>
</tr>
</tbody>
</table>
native students’ theses and the native research articles. With regard to say, there was also a significant difference between the native students’ theses and the research articles (LL=25.84, p < 0.0001), indicating that both the Turkish and the native students overused say compared to experienced native English scholars, with the Turkish students having even a more significant overuse. In addition to the Turkish students’ underuse of ensure, the native speaker students also used it with a significantly lower frequency than the native scholars did (LL=-5.89, p < 0.05). Apart from ensure that was more frequently employed in the research articles than in both groups of students’ theses, the articles contained the verb demonstrate significantly more frequently than the theses (LL=13.75/10.42, p < 0.001) with no such difference being observed between the Turkish and native speaker students.

A relatively subjective verb in the case of academic writing, believe, was significantly underused in the research articles compared to both Turkish and native students’ theses (LL=-39.23, p < 0.0001; LL=-13.00, p < 0.001); moreover, it occurred with a significantly lower frequency in the Turkish students’ theses than in those of the native students (LL=-18.20, p < 0.0001), indicating a clear overuse by the native students. A similar verb, feel, was also significantly overused by the native speakers in contrast to the Turkish students and native scholars (LL=59.92, p < 0.0001; LL=38.12, p < 0.0001) who did not show a significant difference regarding this verb. Likewise, report was also overused by the native speakers as opposed to both groups (LL=42.15, p < 0.0001; LL=23.09, p < 0.0001) who again did not show a significant difference. Lastly, there was an underuse of mean in the native students’ theses when compared to both the Turkish students and the native scholars (LL=-30.28, p < 0.0001; LL=-4.63, p < 0.05) with no such difference between the latter groups.

**Discussion**

The present study explored the use of evaluative language in native and non-native academic corpora aiming to contribute to the existing knowledge on previous corpus-based studies with findings from different L1 backgrounds. This study focused only on one aspect of the native and non-native academic discourse, that is VTHAT structures. The results of the study showed that the Turkish students used VTHAT structures -although not significantly- more frequently than their native peers in academic register, which is in line with the study of Biber and Reppen (1998) suggesting that that-clauses were much more common in all student groups (Spanish, French, Chinese, Japanese) than in any native registers (conversation, fiction, academic, news). Biber and Reppen (1998) presumes that such differences might be due to a transfer effect, which is an issue that needs detailed investigations. In the case of the difference between the Turkish and native graduate students in this study, this presumption may be valid as Turkish is an agglutinative language with a complex morphological structure and also an SOV language. These characteristics lead to a heavy use of complement clauses that come before a main verb and modify it, which may explain more frequent usage of VTHAT by the Turkish students.

The current study also found that both native and Turkish students used VTHAT structures more frequently in their MA theses than their PhD theses. However, a within-group comparison showed that the Turkish students’ MA theses contained significantly more instances of VTHAT than their PhD theses, which was not the case for native students’ MA and PhD theses which involved a similar ratio of VTHAT. The finding that both the Turkish and native MA theses contained more VTHAT constructions than their PhD theses can be attributed to their inexperience, particularly taking into account the significant underuse of VTHAT in native scholars’ research articles. Yet, a further finding showed that the Turkish MA students had a significantly more frequent usage of VTHAT than the native MA students. This implies a repetitive and abundant use of VTHAT combinations in the Turkish MA students’ academic writing. It can thus be argued that as they move towards the PhD
level, Turkish students’ use of VTHAT in academic discourse becomes more native-like and is more approximate to the use of native English speaking and experienced scholars. Another reason behind the Turkish students’ over-reliance on VTHAT can be explained by the argument that learners regard VTHAT as a kind of prefabricated chunk (Ishikawa, 2015) and thus use this construction to fluently express their message in discourse. This may actually be observed in analyses of such chunks in Turkish students’ writing. Turkish students are reported to make an abundant use of lexical bundles (i.e., a type of formulaic language) such as *it can be concluded (that)*, *it was found that*, *it can be said (that)* and *it is/was seen that* in MA/PhD theses and research articles (Güngör & Uysal, 2020; Öztürk & Köse, 2016). From a grammatical perspective, such bundles are VTHAT constructions, and Turkish students seem to regard them as prefabricated chunks as in the case of the current study. On the other hand, the finding that the native MA and PhD students did not significantly differ in their usage of VTHAT indicates that regardless of their level, they had a similar familiarity with the extent to which this structure is used in academic writing.

A noteworthy finding revealed in the present study is that the research articles written by native English-speaking scholars contained a significantly lower frequency of VTHAT in comparison to both the Turkish and the native graduate students. It should be noted that the normalized frequency of VTHAT in the research article corpus in this study (i.e., 2.28 times per 1,000 words) seems considerably lower than the figures reported by Hyland and Tse (2005a; 2005b) who found an evaluative *that* frequency of 6.4 times per 1,000 words in their applied linguistics sub-corpus. Moreover, the frequency-based findings reported here do not overlap with Wang and Chen (2012), Lou (2014) and Kim and Crosthwaite (2019) who revealed significantly higher frequencies of evaluative *that*-clauses in research articles in comparison to graduate dissertations. This is most likely because of the nature of their data (i.e., only research article and dissertation abstracts, not full texts) and their focus being on evaluative *that* in general, meaning that they included not only post-predicate *that*-clauses controlled by a verb, but also those controlled by a noun and an adjective. To our knowledge, there are no other studies that specifically studied VTHAT constructions in research articles (not only the abstract sections) published by native English speakers, and thus the frequency of these constructions can be as low as 2.28 times per 1,000 words as reported in the present study.

The Turkish and native students may have used VTHAT constructions in their theses with significantly higher frequencies compared to the usage in research articles because the genre may require the use of these constructions. Students are almost always asked to have a solid background from the literature in their area of study, which leads them to frequently refer to, report or comment on other scholars’ work. Biber et al. (1999) found that *suggest* and *show* are the most commonly used controlling verbs of *that*-clauses and Conrad (2010) claimed that “the verb + *that*-clause structures in academic prose are used to report previous research, often with non-human entities acting as the subject” (p. 230), for example:

> “Wray (2002) suggests *that* identifying something obviously relies on how you define it.”
> (TARC, MA thesis 6)

As Conrad (2010) states, the most important function of *that*-clauses is “to report thoughts, feelings, and in the case of academic prose, previous research” (p. 230) and Tribble (2002) says that VTHAT is used to make claims in academic register. In fact, Tribble (2002) reported that “of the ten instances (that clauses) in the text (academic), five are either introduced by a verb which comments on claims made by others or introduce a firm claim that the author is making” (p. 141-142). In this respect, both the Turkish and the native students whose texts were included in the research corpus in the present study used VTHAT structures to comment on claims made by others or introduce a firm claim that they as authors have made, which is relatively reflected on their choice of the controlling
verbs in their VTHAT constructions.

The analysis of the controlling verbs in VTHAT constructions showed that eight out of the 10 top frequent verbs in the Turkish students’ theses were common across either three or two sub-corpora, implying that the Turkish students were mostly familiar with the frequent controlling verbs in English as they employed evaluative language in academic discourse. However, there were various significant differences in the frequencies of these controlling verbs between the sub-corpora. Considering that the Turkish students significantly underused *find* but significantly overused *reveal*, they seem to *reveal* results based on their data instead of *finding* them. A quick search on COCA also shows a six times higher frequency for *find* compared to *reveal* in academic prose. Two verbs, *state* and *say*, that are often used to convey other scholars’ arguments were significantly overused by the Turkish students who showed great reliance on these verbs. The Turkish students were also observed to use these two verbs to mitigate the assertiveness of the message to follow after *that*. For example, the Turkish students produced sentences such as “It can be *said* that the scores of the productive task group increased after the treatments.” (TARC, MA thesis 17) although they may be reporting an objective finding (i.e., an increase in scores), which can imply that Turkish students use such VTHAT constructions to have a low authoritative presence in terms of their authorial stance in academic writing. Turkish academics’ tendency to exhibit a low stance in academic writing has also been reported in the literature. Çakır (2016) investigated stance adverbs in abstracts of journal articles written by native English writers and Turkish writers, reporting that Turkish academic writers adopted a less clearly independent stance. Focusing on authorial self-mention words, Kafes (2017) analyzed research articles published by American and Turkish academic writers in applied linguistics, and concluded that Turkish writers significantly underused authorial self-mention words, indicating a downplayed their role in their writing and, similarly, a less clearly independent stance in comparison to American writers. Though examining different aspects of academic writing, the findings of these two studies seem to be in line with the findings of the present study in terms of the student writers’ choices of controlling verbs in VTHAT.

Two verbs, *feel* and *believe*, that can be regarded as relatively subjective verbs in the context of academic prose were significantly overused by the native speaker students in their theses unlike native scholars’ research articles. This indicates that native students had some space for subjectivity in their academic writing, which can be expected at novice levels, while the Turkish students’ performance in this regard was closer to native scholars.

Biber *et al.* (1999) explained the semantic domain of the controlling verbs of VTHAT as: “the verbs that take a that-complement clause in post-predicate position fall into just three major semantic domains: mental verbs, mainly of cognition (e.g., think, know), but including a few with emotive/affective content (e.g., hope and wish); speech act verbs (e.g., say, tell); and other communication verbs that do not necessarily involve speech (e.g., show, prove, suggest)” (p. 661). Therefore, looking at the frequently used controlling verbs in VTHAT constructions, while Turkish writers preferred to use speech act verbs or other verbs, native writers preferred to use mental verbs in their academic writing. According to Biber *et al.* (1999) “mental verbs with that-clauses are an important device used to express stance. For example, verbs such as *think, feel*, and *assume* convey a sense of possibility combined with uncertainty, while verbs such as know, find, and see convey a definite sense of certainty” (p. 665). While the verb *find* was the most frequently used controlling verb found in native students’ theses, it is the tenth frequently used controlling verb in Turkish students’ theses, suggesting that rather than stating a definite sense of certainty, non-natives writers rely on the findings of other researchers by using speech acts verbs.

With respect to the controlling verbs, it should also be noted that the Turkish students relied on a
restricted number of verbs in their VTHAT constructions, and used these verbs abundantly, while the native students made use of more varied verb types. When type-token ratios were calculated for the verb types and the total occurrences of VTHAT constructions, the Turkish students were found to have the lowest ratio (2.88%) followed by the native students (4.24%) and the native scholars (8.63%). It can thus be argued that Turkish EFL learners may need to pay more attention to using a wider range of controlling verbs and use less VTHAT constructions, perhaps by employing other types of constructions that are more apparent in expert academic writing.

Conclusion and Implications

As a result, the present study showed that although the frequency of VTHAT constructions differed between native and Turkish students’ theses, this difference was not significant; however, a detailed examination showed that there were statistically significant differences in terms of the level of the writers and types and frequencies of the controlling verbs. It can be concluded that although the theses written by native and Turkish writers seemed similar in terms of the usage and distribution of VTHAT clauses, a detailed analysis showed the differences suggesting that non-native writers should be careful in choosing the controlling verbs in their theses. The result of the study raised awareness with regard to the controlling verbs used in academic writing and suggests that in academic writing lessons, instructors should mention the frequently used controlling verbs used in academic genre.

Based on the findings of this study, several pedagogical implications can be offered. Firstly, since the Turkish students exhibited an overuse of VTHAT combinations compared to both native students and scholars, a specific emphasis should be placed on the use of VTHAT in academic writing classes at both undergraduate and graduate levels. Thus, Turkish students and academic writers can be encouraged not to make an excessive use of this structure, especially in cases where they can convey their message in a more direct way without a complement phrase.

Secondly, as the Turkish students relied on a limited number of controlling verbs, materials and curricula in academic writing classes can involve activities to expand students’ repertoire of these verbs so that they do not excessively use the same verbs in their academic papers. Therefore, EAP instructors should encourage learners to use a wider range of controlling verbs as experienced writers do.

Thirdly, since Turkish students seem to regard VTHAT constructions as prefabricated chunks (formulaic language), EAP students can be trained to use corpus tools and consult both general and specialized corpora as a direct help in academic writing (Chang, 2014). This could enable academic writers and graduate students to see how and to what extent VTHAT constructions are used in academic writing in general and in the academic texts in their disciplines.

Lastly, considering the finding that the Turkish students may be overusing VTHAT structures as a way of showing a weak authorial presence, awareness-raising activities can be carried out as part of academic writing courses to give students an understanding of the extent to which native English writers show commitment to what they say in their writing. Although authorial stance is a phenomenon that can reflect culture-based variations across academic communities in different countries (Çakır, 2016), students should have an idea about the average level of authorial presence in their disciplines internationally.

Although the current study dealt with the comparison of lexical and structural characteristics of academic writing in native and non-native theses by examining VTHAT structures, it doesn’t allow making any firm statements as to whether the data reflect all the characteristics of academic genre
because the data analyzed here come from the theses written by researchers; therefore, it contains student writers, non-professionals. Another important limitation of the study is that it does not include VTHAT structures in which that are omitted. In this respect, this study may not reflect the whole picture about post predicate that clauses controlled by verbs. In order to understand the complete nature of VTHAT structures, more detailed and cross-linguistics studies are needed. Analyzing the structure in conversation and comparing it with academic genre would provide more insight into the issue.

References


