

# Use of Stance Lexical Bundles by Turkish and Japanese EFL Learners and Native English Speakers in Academic Writing<sup>1</sup>

## Tutum Sözcük Örüntülerinin İngilizceyi Yabancı Dil Olarak Öğrenen Türk ve Japon Öğrenciler ve İngilizce Anadil Konuşucuları Tarafından Akademik Yazımda Kullanımı

Meltem MUŞLU\*

### Abstract

Linguists have been interested in the sequences of words that tend to occur for a long time. The present study examines a particular type of recurrent chunks called lexical bundles (LB). LBs are multi-word expressions and an important component of the fluent linguistic production depends on the control of them (Hyland, 2008a). The purpose of this study was to find out the most common stance LBs used in argumentative essays written by native English speakers and Turkish and Japanese EFL learners. It also aimed at finding the structural and functional characteristics of these bundles, and to what extent these structures used by the Turkish and Japanese EFL learners are similar to the ones used by the native speakers. To answer these questions, Louvain Corpus of Native English Essays (LOCNESS), Japanese International Corpus of Learner English (JPICLE) and Turkish International Corpus of Learner English (TICLE) were used. The structures of the stance LBs were determined by following Biber's (2006) classification and the functions were determined by adapting their classification. The concordancing program WordSmith was used to find and determine the 3-4 word stance LBs. In the statistical analysis, Type-token ratio and Log Likelihood were used. The results showed that native speakers use lexical bundles least; whereas, Japanese EFL learners use them most frequently. The functions and the structures of LBs vary in each group. Suggestions regarding how to teach these devices in foreign language education were also given.

**Keywords:** Lexical bundles, Corpus Linguistics, stance, Contrastive Interlanguage Analysis, academic writing, Turkish and Japanese EFL learners

### Öz

Dilbilimciler uzun zamandan beri bir arada olma eğiliminde olan ardışık kelimeleri incelemektedirler. Bu çalışmada sözcük örüntüsü denilen ve sürekli tekrar eden belirli bir kelime grubu incelenmektedir. Sözcük örüntüsü, 'küme' ya da 'yığın' da denilen birkaç kelimenin yan yana gelmesi ile oluşan kelime gruplarıdır ve dili akıcı bir şekilde kullanmanın önemli bir bölümü bu yapıları hakim olmaktan geçer (Hyland, 2008a). Bu çalışma, anadili İngilizce olan ve İngilizce'yi yabancı dil olarak öğrenen Türk ve Japon öğrenciler tarafından tartışma türünde yazılmış kompozisyonlarda en çok kullanılan 'tutum sözcük örüntüleri'ni (TSÖ) bulmayı amaçlamıştır. Çalışma aynı zamanda, bu kelime grubunun yapısal ve işlevsel olarak hangi özellikleri gösterdiğini ve bu özelliklerin anadili İngilizce olan ve İngilizce'yi yabancı dil olarak öğrenen öğrenciler arasında nasıl bir farklılık gösterdiğini bulmayı hedeflemiştir. Bu amaçlara ulaşmak için veri toplama aracı olarak Louvain Corpus of Native English Essays (LOCNESS), Japanese International Corpus of Learner English (JPICLE) ve Turkish International Corpus of Learner English (TICLE) kullanılmıştır. Tutum sözcük örüntüleri yapısal ve işlevsel olarak analiz edilirken Biber'in (2006) sınıflandırması takip edilmiş ancak bu sınıflandırma çıkan sonuçlara göre adapte edilmiştir. 3-4 kelimeden oluşan sözcük örüntülerini bulmak için WordSmith programı kullanılmıştır. İstatistiksel analizde, Type/token ratio ve Log Likelihood Analizi kullanılmıştır. Çıkan sonuçlara göre anadili İngilizce olan katılımcılar TSÖ'yü en az kullanırken en fazla kullanan grup Japon öğrenciler olmuştur. TSÖ'lerin yapısal ve işlevsel kullanımları her grupta farklılık göstermiştir. Bu yapıların yabancı dil öğretiminde kullanımına yönelik önerilere de çalışmada yer verilmiştir.

**Anahtar Kelimeler:** Sözcük Örüntüleri, Derlem Dilbilim, tutum, karşılaştırmalı aradil çalışması, akademik yazım, İngilizceyi yabancı dil olarak öğrenen Türk ve Japon öğrenciler

### Introduction

Lexical bundles (LBs hereafter) are the most frequently occurring sequences of words which are usually not idiomatic in meaning; the meanings are transparent from the individual words, and are usually not complete grammatical structures (Biber, 2006). They are combinations of three or more words, which are identified empirically in a corpus of natural

<sup>1</sup> This study is adapted from the writer's PhD thesis titled Stance Lexical Bundles In Academic L2 Writing: A Contrastive Interlanguage Analysis.

\* Dr.Öğrt. Üyesi, Gaziantep Üniversitesi, Fen Edebiyat Fakültesi, Batı Dilleri Bölümü, [meltemmuslu@gmail.com](mailto:meltemmuslu@gmail.com)

language; for instance, *take a look at*, *I don't know*, and *as a result of* (Cortes, 2006). They are extended collocations that appear more frequently than expected by chance, helping to shape meanings in specific contexts and contributing to our sense of coherence in a text (Hyland, 2008b). Since gaining sensitivity to expert users' preferences for certain sequences of words is required in gaining the control of a new language (Hyland, 2008a), LBs have a crucial role in both academic prose and spoken discourse. Identifying fixed phrases provides benefits for both language teachers and learners since fluent linguistic production depends on the control of the LBs. Identifying them may help learners acquire the specific rhetoric practices of the texts they are asked to write (Hyland, 2008b). As Hyland (2008a, p. 60) states 'the study of clusters offers insights into a crucial, and often overlooked dimension of language use, providing a better understanding of the ways writers employ the resources of English in different contexts, and with the potential to inform advanced academic literacy instruction'.

Considering their significance in language learning and teaching, researchers have been investigating the use of LBs in classroom teaching and textbooks, and in professional writers' and students' writings. First studies on LBs focused on the identification of these units and a description of their patterns of occurrence in different written and spoken registers in English (Biber et al., 2004; Biber, 2006; Biber & Barbieri, 2007; Conrad & Biber, 2004; Cortes, 2002; Hernandez, 2013; Herbel-Eisenmann et al, 2010; Hyland, 2008a; Hyland, 2008b; Jablonkai, 2010; Nesi & Baştürkmen, 2006; Strunkyte & Jurkunaite, 2008). In addition to these studies, some research focused on the pedagogical aspect of LBs (Cortes, 2006; Neely & Cortes, 2009), some on comparing different languages (Bal, 2010; Cortes, 2008; Nekrasova, 2009; Tenuta et al., 2012), and some on the comparison of their use by native and nonnative speakers of English (Chen & Baker, 2010; Cortes, 2004; Jalali et al., 2008; Juknevičienė, 2011; Taki and Jafarpour, 2011; Wei & Lei, 2011).

However, none of these studies compared the use of LBs by learners from different native language backgrounds. As Hyland (2008b) points out, more work with different disciplines, genres, and first language groups should be done, since as Granger (2002) points out, an improved understanding of learner output illuminates all aspects of pedagogy from tasks to curriculum. Considering these, this study first aims at exploring the structures and functions of stance lexical bundles used by native speakers of English, and by Turkish and Japanese learners studying English as a Foreign Language. The second aim is to find out whether there is a difference in the use of LBs between the native speakers of English and the Turkish and Japanese EFL learners.

## Methodology

### Data collection tools and participants

Three existing corpora were used in the present study: Louvain Corpus of Native English Essay (LOCNESS) and two learner corpora from ICLEv2, namely TICLE (Turkish International Corpus of Learner English) and JPICLE (Japanese International Corpus of Learner English). Each corpus consists of argumentative essays written by university students.

In the study, LOCNESS was used since the participants' age was similar to the ones in the learner corpora and the data consists of argumentative essays. The number of words in each corpus is also similar: LOCNESS, the native speaker corpus including the essays written by American university students, contains 149,501 words. The learner corpus TICLE contains 199,532 words and JPICLE contains 198,131 words.

## Data Analysis

### Operationalization

A word combination has to meet a set of defining criteria to be counted as a bundle. The first criterion is the cut-off frequency, which determines the number of LBs to be included in the analysis. The frequency cut-off point of the bundles varies from study to study. Biber (2006) concluded that to be a LB, a four-word expression had to recur ten times per million words and appear in more than five texts. However, Biber et al. (2004) set the criteria as forty times in a one-million word corpus and Cortes (2004) set it as twenty times in one million words. These higher cut-off points were chosen to make the frequency analysis to be more conservative and to ensure that the expressions analyzed in these studies were used in extremely high frequencies. In addition to frequency, lexical bundles must be used in at least 3-5 different texts (10%) to prevent focusing on idiosyncratic uses by the authors of the texts in the corpus used. Since high cut-off points are chosen to be more conservative, in this study, the cut-off point is determined as 3. Also, the bundles which are seen in at least 3 different texts are chosen to make the analysis more accurate. The last criterion concerns the length of word combinations, usually 2-6 word units. Three and four-word lexical bundles were analyzed in this study. Although 'many four-word bundles hold three-word bundles in their structures' (Cortes, 2004, p. 401) and they are more common and present a wider range of structures and functions (Hyland, 2008b), three-word lexical bundles are also included since some of the 3-word bundles found in the study have an important role in the analysis. For instance, the bundles *the fact that* and *I think that* were used very frequently in all corpora. To determine the LBs, the concordance program WordSmith v.5, created by Scott (2008), was used.

### Statistical Measures

Lexical variety and frequency distribution of the corpora were analyzed. To find out the lexical variety, type-token ratio (T/t) was calculated. T/t ratio of LBs represents the percentage of each bundle within all words in a corpus, that is, the number of bundles that falls into per 100 words.

The frequency distribution of the corpora was calculated through Log Likelihood (LL) analysis. LL test considers word frequencies weighted over two different corpora. It measures higher or lower frequencies than expected. In the analysis, LOCNESS frequency results were compared with TICLE and JPICLE (L1 vs. L2 comparison). In order to see the condition among learner groups, overall lexical bundle frequency of TICLE and JPICLE were also compared (L2 vs. L2 comparison).

### Exclusion criteria

The focus of this study was to determine stance lexical bundles; therefore, while doing the analysis, the 1<sup>st</sup> person singular and plural pronouns were included and considered as personal stance. Yet, 2<sup>nd</sup> and 3<sup>rd</sup> person singular and plural pronouns were excluded since as Biber (2006, p. 91) points out, 'Although such structures express some kind of attitude or evaluation, they do not necessarily reflect the personal stance of the speaker/writer'. In this study, the 2<sup>nd</sup> person singular / plural was only included when it was an attitudinal/modality stance; such as *if you want to*.

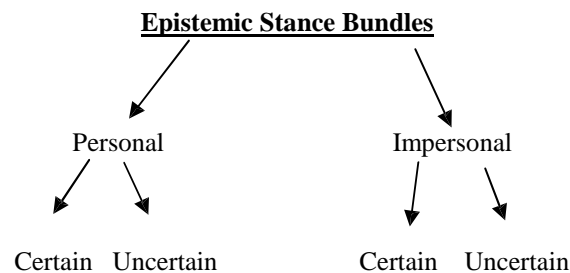
Also, when analyzing the data, bundles including abbreviation are combined with the ones with no abbreviation so that they do not stand as different word groups. For instance, *I do not want* and *I don't want* were combined and counted as one bundle. Similarly, bundles

that are structurally and functionally similar were combined and counted as one. For instance, *I think the*, *I think it*, and *I think we* are all SP+VP structure and they are all impersonal and uncertain. Therefore, these bundles were combined and counted as one impersonal/uncertain and SP+VP bundle. Similarly, the verbs ‘is’ and ‘are’ were combined, such as *is/are very important*. In the analysis, the frequencies of each bundle were summed up.

### Classification of Lexical Bundles

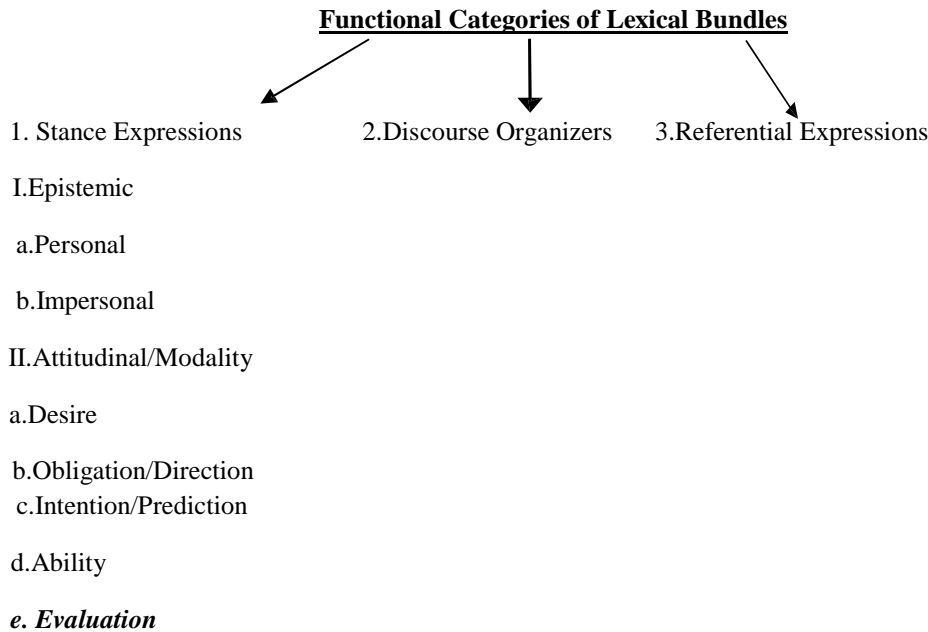
The structural and functional classification of LBs by Biber (2006) has been widely relied on in the studies on LBs in the field. Therefore, when finding the structural and functional features of the stance lexical bundles used by native English speakers, and Turkish and Japanese EFL learners, his classification was used. However, while determining the epistemic stance bundles, a revised version of Biber’s classification was used (see Figure 1 and 2 below).

As Biber (1999) states, epistemic bundles comment on the knowledge status of the information (certain, uncertain, probable/possible), such as *I don’t know if*, and *I don’t think so*. There are two sub-categories of epistemic stance bundles: Personal and impersonal. Personal epistemic bundles can express certainty and uncertainty, but most of them express only uncertainty. *I don’t know what*, *I don’t think so*, and *I think it was* are examples of personal epistemic bundles. Impersonal epistemic bundles express certainty; *are more likely to*, and *the fact that the* are considered in this group. However, a slight change has been applied to this model by adding two more categories since while analyzing the data, it was found that there were many personal certain and impersonal uncertain LBs. Therefore, epistemic stance bundles were categorized as:



**Figure 1.** Adapted Categories of Epistemic Stance Bundles

When analyzing the data, it was also found that some chunks expressing ‘evaluation’ were fairly used. These chunks indicate what a writer thinks about a specific subject. They are formed by using an adjective and they make an assessment on a particular topic. Therefore, they give a clear indication of the writers’ thoughts on that specific topic and show personal stance. For this reason, they were counted as a new category of SLBs. Biber (2006, p.92) categorizes common lexico-grammatical features used for the stance analyses and he lists some stance adjectives used in evaluation. Some of the adjectives he mentions in this category are *appropriate*, *good/bad*, *better/worse*, *essential*, *important*, *interesting* and *necessary*. While determining the evaluation SLBs, this list is used as a basis. Some of the LBs determined in the evaluation category were *it is difficult to*, *is the best*, *is very easy to*, and *is very interesting*. This new category is added to the Attitudinal/Modality category. Hence, the new attitudinal/modality category consists of five groups. The functional categories used in this study can be seen in the Figure 2 below.



**Figure 2.** Functional Categories of LBs identified in this study

The results of the study and the discussion of the findings can be found in the next part.

## Results and Discussion

### Overall Results of the Study

Table 1 below presents the overall descriptive results of the stance lexical bundles used in all corpora.

**Table 1:** Frequencies and percentages of SLBs Corpora

	Total number of words	Total number of SLBs	% of SLBs
<b>TICLE</b>	199,532	288	0,14
<b>LOCNESS</b>	149,501	184	0,12
<b>JPICLE</b>	198,131	357	0,18
<b>Total</b>	547,164	829	0,15

The total number of the words analyzed in this study is 547,164 and 0,15% of these words consist of SLBs. The results show that native speakers (184) use SLBs least, whereas Japanese EFL learners (357) use them most frequently. The use of SLBs by the Turkish EFL learners (288) is in between these two groups. The findings are consistent with Wei and Lei (2011) and Jalali et al. (2008) that nonnative speakers use much more lexical bundles than native speakers. These results do not coincide with Chen and Baker (2010) for in their study, they found that the widest range of LBs was used by native speakers and the smallest range was used by nonnative learners.

The number of bundles found in this study is higher than the numbers found in previous studies. For instance, Hyland (2008a) found 629 bundles in a corpus of nearly 3.5 million corpus, Wei and Lei (2011) found 241 bundles in a 2.5 million corpus. In these studies, they analyzed the 4-word lexical bundles, since as Wei and Lei state, 3-word bundles are extremely common. Therefore, one of the reasons of the high number of lexical bundles found in this study is both three and four-word lexical bundles are included. Also, as mentioned earlier, a new category (evaluation) is added to Biber's classification and the highest number of lexical bundles constitutes this category in the Japanese corpus.

### Log Likelihood Analysis Results

In order to test the frequency distribution of the corpora, Log Likelihood (LL) analysis was done. LOCNESS frequency results were compared with orderly TICLE and JPICLE (L1 vs. L2 comparison). In order to see the condition among learner groups, overall lexical bundle frequency of TICLE and JPICLE were also compared (L2 vs. L2 comparison). Table 2 shows the LL ratio of the bundles in comparison with TICLE and LOCNESS.

**Table 2:** LL Ratio of TICLE and LOCNESS

	O1 TICLE	%1	O2 LOCNESS	%2	LL ratio
Lexical Bundles	288	0,14	184	0,12	+2,88

O1 is observed frequency in Corpus 1  
 O2 is observed frequency in Corpus 2  
 %1 and %2 values show relative frequencies in the texts.  
 + indicates overuse in O1 relative to O2,  
 - indicates underuse in O1 relative to O2

In the table, O1 and O2 refer to overall frequency of lexical bundles observed in TICLE and LOCNESS. %1 value includes the relative frequency of LBs in the texts, i.e., 0,14 relative frequency means there are approximately 0,14 lexical bundles falling into every 100 words in TICLE. According to the result, LL ratio measurement indicates an overuse in TICLE with an +2,88 LL value. This means that there is a significant difference between the two corpora in terms of LB frequency ( $p < 0.05$ ) and the overuse in TICLE relative to LOCNESS has been approved by the LL calculation.

**Table 3:** LL Ratio of JPICLE and LOCNESS

	O1 JPICLE	%1	O2 LOCNESS	%2	LL
Lexical Bundles	357	0,18	184	0,12	+18,28

As expected from the frequency difference between JPICLE and LOCNESS, LL value between the two corpora revealed an overuse of SLBs. The LL ratio between these two corpora is + 18,28. In the LL analysis, the higher the LL value means the more significant overuse; therefore, it can be concluded that the overuse of lexical items in JPICLE is more than the use of TICLE.

**Table 4:** LL Ratio of JPICLE and TICLE

	O1 JPICLE	%1	O2 TICLE	%2	LL
Lexical Bundles	357	0,18	288	0,14	+7,89

The comparison between JPICLE and TICLE showed that Japanese EFL learners overused SLBs with a +7,89 LL ratio. The results of LL shows that compared to native speakers, both EFL learner groups overused LBs; however, since the language item under consideration is an English language structure, the usage of SLBs in LOCNESS is accepted as the accurate usage.

## Type-token Ratio Results

T/t ratio shows the lexical variety within a text. T/t ratio of SLBs represents the percentage of each bundle within all words in the corpora, that is, the number of bundles that falls into per 100 words.

**Table 5.** Overall Comparative Frequency Distribution of SLBs in general

	<b>Tokens</b>	<b>Types</b>	<b>T/t Ratio</b>
LOCNESS (L1)	1056	184	17
TICLE (L2)	2519	288	11
JPICLE (L2)	8394	357	4

T/t ratio= Percentage of number of SLBs (types) in total of words (tokens) in each corpus

As Table 5 shows, the lowest number of LBs in total belongs to LOCNESS (184). Yet, when the T/t ratio of LBs is compared, it was seen that the T/t ratio of LBs used in LOCNESS is 17%; whereas, it is 11% in TICLE, and 4 % in JPICLE. This means that native speakers used more varied SLBs; 17 of their every 100 words were a SLB. Turkish learners used 11 of their every 100 words as a SLB and Japanese learners used 4. The percentage of T/t ratios is similar in LOCNESS and TICLE corpora, but it is lower than these two groups in JPICLE. The distribution of epistemic and attitudinal/modality SLBs were also compared since the evaluation SLBs are highly used by the Japanese EFLs.

**Table 6.** Overall Comparative Frequency Distribution of Epistemic & Attitudinal/ Modality LBs

	<b>Epistemic SLBs</b>		<b>Attitudinal/Modality SLBs</b>			
	<b>Tokens</b>	<b>Types</b>	<b>Types</b>	<b>T/t Ratio</b>		
LOCNESS (L1)	772	118	15	284	66	23
TICLE (L2)	1425	184	13	1094	104	9
JPICLE (L2)	1496	170	11	6898	187	3

**Tokens    Types    T/t Ratio**

T/t ratio= Percentage of number of SLBs (types) in total of words (tokens) in each corpus

As Table 6 presents, the percentage of the use of epistemic bundles is more than the attitudinal/modality bundles in L2 learner groups. Turkish EFL learners used 13 of their every 100 words as epistemic SLB and Japanese EFL learners used 11 of their every 100 words as epistemic SLB. On the contrary, native speakers used attitudinal/modality SLBs more than epistemic SLBs. The difference between the use of epistemic and attitudinal/modality LBs is the highest in Japanese EFL learners group. They used epistemic bundles 11%, but attitudinal/modality bundles 3%. The reason for this could be that Japanese EFL learners used evaluation LBs excessively compared to other bundles. When the most frequent 40 lexical bundles are compared (Table 7), it is seen that Japanese learners used some of the bundles far more frequently than the EFL learners. Therefore, although the number of LBs is more in JPICLE, the lexical variety is lower. Conversely, native speakers used the fewest number of lexical bundles; however, lexical variety is the highest (17%) in this group. That is, native speakers use less but various bundles; whereas EFL learners use more bundles, but they are not various; similar or the same bundles are used in many instances. The results are consistent with Chen (2012) that she also found that nonnative speakers use restricted range of epistemic devices. In order to see the differences between the three corpora, the most common 40 bundles were compared.

## Forty most common stance lexical bundles

Table 7 presents the most common 40 stance lexical bundles used in LOCNESS, TICLE and JPICLE. The bundles that are used in all of the three corpora are written *Italic* and **bold**.

**Table 7.** Most common 40 Stance Lexical Bundles

NATIVE		TURKISH		JAPANESE	
Stance Bundles	Fre.	Stance Bundles	Fre.	Stance Bundles	Fre.
<i>THE FACT THAT</i>	71	<b><i>I THINK+N</i></b> ( <i>THE/IT/THIS/etc.</i> )	88	<b><i>I THINK+N</i></b>	309
<b>BE ABLE TO</b>	50	IN MY OPINION	70	I THINK THAT	279
<b>HAVE TO BE</b>	23	I WANT TO	59	I WANT TO	185
<b>ARE/IS GOING TO</b>	22	BE ABLE TO	53	WE HAVE TO	108
<b>I FEEL THAT</b>	20	<b><i>I BELIEVE THAT</i></b>	43	WE NEED TO	94
<b>DO NOT WANT</b>	18	IT CAN BE	39	<b><i>IS/ARE VERY IMPORTANT</i></b>	68
<b>BE ALLOWED TO</b>	15	DO NOT WANT	36	I DO NOT THINK	67
<b>NEED TO BE</b>	15	IT WILL BE	32	IT IS IMPORTANT	48
<b>WILL CONTINUE TO</b>	14	<b><i>IS/ARE VERY IMPORTANT</i></b>	31	<b><i>DON'T HAVE TO</i></b>	43
<b>THE OPPONENTS OF</b>	13	<b><i>THE FACT THAT</i></b>	31	WE WANT TO	40
<i>I BELIEVE THAT</i>	12	THERE WILL BE	31	<b><i>I BELIEVE THAT</i></b>	38
<b>THERE WOULD BE</b>	12	WE CAN SEE	30	I WOULD LIKE TO	38
<b>WOULD HAVE TO</b>	12	<b><i>I THINK THAT</i></b>	29	IT WILL BE	38
<b>IT IS IMPORTANT</b>	11	DO NOT KNOW	27	I DON'T WANT	34
<b>WOULD LIKE TO</b>	11	DO NOT THINK	30	ARE ABLE TO	32
<b>IS/ARE FORCED TO</b>	10	THERE SHOULD BE	24	IT IS GOOD	32
<b>WOULD HAVE BEEN</b>	10	WE CAN SAY	23	MOST IMPORTANT THING	31
<b>ARGUE THAT THE</b>	9	WE SEE THAT	23	WILL BE ABLE TO	31
<i>DO NOT HAVE TO</i>	9	WE HAVE TO	22	HAVE TO BE	30
<i>I THINK THAT</i>	9	IS THE MOST IMPORTANT	21	IS VERY DIFFICULT	30
<i>I THINK+N (IT/THE)</i>	9	IS/ARE GOING TO	19	IT IS DIFFICULT	30
<b>IN MY OPINION</b>	9	IT MEANS THAT	19	WANT TO SAY	30
<b>IT IS OBVIOUS</b>	9	<b><i>DO NOT HAVE TO</i></b>	18	I DON'T KNOW	28
<b>THERE WILL BE</b>	9	CAN BE DONE	17	IS VERY USEFUL	28
<b>CAN BE SEEN</b>	8	I CAN SAY	17	IT IS SAID	28
<b>I HAVE SEEN</b>	8	IS/ARE NOT IMPORTANT	17	I NEED TO	25
<b>PEOPLE FEEL THAT</b>	8	THERE IS NO NEED	17	IT IS NATURAL	25
<b>SHOULD BE ALLOWED TO</b>	8	PEOPLE THINK THAT	16	IT IS SAID THAT	25
<b>WILL BE ABLE TO</b>	8	WE KNOW THAT	16	IS GOOD FOR	24
<b>WILL HAVE TO</b>	8	IN FACT THERE/THIS/THE	15	IS THE BEST	23
<b>WOULD BE ABLE TO</b>	8	IT IS NECESSARY	15	IT IS NECESSARY	23
<b>ARE WILLING TO</b>	7	SEEMS TO BE	15	IT MAY BE	23
<b>CONSIDERED TO BE</b>	7	THE IMPORTANT THING IS	15	<b><i>THE FACT THAT</i></b>	23
<b>I WOULD LIKE TO</b>	7	WE CAN SAY THAT	15	IT IS VERY IMPORTANT	22
<b>IS/ARE ALLOWED TO</b>	7	ACCORDING TO ME	14	I AM GOING TO	21
<i>IS VERY IMPORTANT</i>	7	DO NOT UNDERSTAND	13	IT IS TRUE	21
<b>IT IS POSSIBLE</b>	7	IT IS BETTER	13	IT MIGHT BE	21
<b>MAY HAVE BEEN</b>	7	IT IS TRUE	13	WILL NOT BE	21
<b>ONE CAN SEE</b>	7	SEEN AS A/AN	13	DON'T NEED TO	20
<b>SHOULD BE ABLE TO</b>	7	THIS MEANS THAT	13	I WANT TO SAY	20

As Table 7 shows; *the fact that*, *do not have to*, *I think that*, and *is/are very important* are the most common LBs used by both the native speakers and EFL learner groups.



When the frequency of these bundles was compared, it was found that *I think+n* is by far the most frequently used SLB by the Japanese and Turkish EFL learners, 309 times and 88 times respectively. Whereas *the fact that* (71) was the most frequently used SLB by the native speakers. *The fact that* is considered an impersonal certain epistemic SLB and it is used 31 times by the Turkish EFL learners and 23 times by the Japanese learners. Unlike native speakers, EFL learners used personal uncertain epistemic stance bundles the most, which are *in my opinion*, *I think+n* and *I think that*. *In my opinion* is used 70 times by the Turkish EFL learners; on the other hand, it was used 9 times by the native speakers and 20 times by the Japanese EFL learners. *I think that* is used excessively by the Japanese EFL learners; they used it 279 times. However, native speakers used it only 9 times and Turkish EFL learners used it 29 times. Another striking finding is Japanese EFL learners used the verb *think* immensely; *I think+n* 309 times and *I think that* 279 times. Turkish EFL learners used *I think+n* 88 times and *I think that* 29 times. Unlike Japanese and Turkish EFL learners, native speakers used *I think that* and *I think+n* only 9 times each.

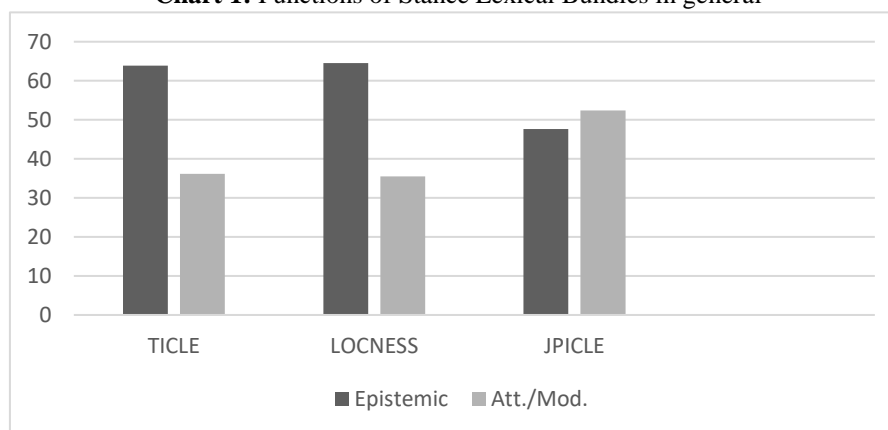
It is also seen that Japanese EFL learners tend to use first person singular and plural pronouns more often than the other groups. 'I' was used 11 times and 'we' was used 3 times by the Japanese EFL learners in the most common 40 lexical bundles. Native speakers used 'I' five times and they did not use 'we'. The use of 'I' and 'we' by the Turkish EFL learners remains in between the native speakers and Japanese EFL learners: 'I' three times, 'we' five times. Since Japanese is a pro-drop language, surprisingly, Japanese speakers used the expletive 'it' more frequently than the native speakers. It is also surprising that although they used the dummy subject it more frequently than the other groups, they did not use the expletive 'there'. Turkish is also a pro-drop language, but Turkish native speakers did not use the expletive 'it' as often as Japanese native speakers. English native speakers used expletive pronouns 'it' and 'there' twice. They used the first person singular six times, but they did not use the first person plural pronoun. The use of first person subject pronouns and expletives by the Turkish EFL learners remains in between the native speakers and Japanese EFL learners.

It can be concluded from these results that the use of epistemic devices in general by the Turkish EFL learners is closer to native speaker use. Hyland (2005) suggests that contrary to what is often thought, academic writing is not an impersonal monologue. Use of self-mention is an important stance marker, but as seen in the results; native speakers used these expressions less frequently than EFL learners. This finding is inconsistent with Taki and Jafarpour (2011) that native speakers use self-mention more than the learners, namely the Persian group. Another important finding the table shows is that Japanese EFL learners used uncertainty verbs *think* and *believe* immensely. Conversely, the verb choice of native speakers was varied: *believe*, *argue*, *seem*, *think*, and *feel* were the verbs used most frequently. In addition, the most frequently used epistemic bundle by the native speakers was *the fact that*, which shows certainty. It can be concluded that Japanese learners used uncertainty bundles more often compared to native speakers and Turkish EFL learners. The reason for this could be the difference between the cultures. As Kamimura and Oi (1998) states, Japanese argumentative style is more subjective rather than factual or objective, and there is more weight on evoking empathy in the audience, rather than more emphasis on persuading the audience. In order to interpret these findings more accurately, a deeper analysis, both functionally and structurally, should be done.

## Functions of stance lexical bundles

A general comparison of the use of epistemic and attitudinal/modality SLBs can be seen in the chart below.

**Chart 1:** Functions of Stance Lexical Bundles in general



As Chart 1 shows, epistemic bundles were used more frequently in TICLE (63,89%) and LOCNESS (64,14%); whereas attitudinal/modality bundles were used more frequently in JPICLE (52,37%). When the use of epistemic bundles and attitudinal/modality bundles are compared in all groups, it is found that epistemic bundles were used most frequently by the native speakers (64,14%). The results are somewhat consistent with Chen (2012) that in her study she found that there is a great similarity in the total number of epistemic devices used by native and non-native (Chinese) speakers. In the current study, the

### FUNCTIONS OF STANCE LEXICAL BUNDLES

Corpora		TICLE		LOCNESS		JPICLE	
Frequency & Percentages		Freq.	%	Freq.	%	Freq.	%
EPISTEMIC BUNDLES	P/C	41	14,24	14	7,6	37	10,36
	P/UC	18	6,25	7	3,8	47	13,17
	IM/C	65	22,57	47	25,54	48	13,45
	IM/UC	60	20,83	50	27,2	38	10,64
ATTITUDINAL/MODALITY	Desire	11	3,82	7	3,8	21	5,88
	Obligation/Direction	29	10,07	29	15,76	47	13,17
	Intention/Prediction	11	3,82	9	4,89	15	4,20
	Ability	3	1,04	2	1,09	7	1,96
	Evaluation	50	17,36	19	10,32	97	27,17
<b>Total</b>		<b>288</b>	<b>100</b>	<b>184</b>	<b>100</b>	<b>357</b>	<b>100</b>

percentage of the use of epistemic devices by the native speakers and Turkish EFL learners is very close to each other; however, the percentage of Japanese EFL learners is lower. Native speakers used epistemic bundles slightly more than the Turkish EFL learners did, whereas they used them relatively more than the Japanese EFL learners. The reason why Japanese EFL learners used attitudinal/modality LBs more is that they used evaluation bundles

excessively. Looking at each functional category would give more insights on the use of stance lexical bundles. Table 8 below summarizes the functions of SLBs used in all corpora more in depth.

**Table 8:** Functions of Stance Lexical Bundles in detail

As can be seen, epistemic bundles were frequently used in all groups. Both native speakers and Turkish and Japanese EFL learners used impersonal SLBs pervasively. Compared to personal bundles, impersonal stance bundles were used more often. The total number of impersonal stance bundles used by the native speakers is 52,74%. Similarly, both Japanese (24,09%) and Turkish (43,4%) EFL learners used impersonal stance, such as *are more likely to*, more often than the personal stance, such as *I think that*. Among all lexical devices used by native speakers, 64,14% of them were epistemic devices, most of which (52,74%) shows impersonality. Similarly, 43,4% out of 63,89% of the Turkish data consists of impersonal epistemic devices, and 24,09% out of 47,62% of Japanese data consists of impersonal epistemic devices.

When the use of certainty and uncertainty devices is compared, it is seen that native and Turkish EFL learners used certainty devices more than uncertainty, whereas the number of certainty and uncertainty devices was the same in Japanese native speakers (23,81%). Turkish EFL learners used 27,06% uncertainty devices and 36,81% certainty devices. Native speakers used 31% of uncertainty and 33,14% of certainty devices. This is somewhat in line with Ağçam (2014) that in her study, nonnatives (Turkish and Spanish) used more certainty devices than native speakers did. The use of certainty/uncertainty devices by the native speakers is slightly more than each other (2,14%), whereas the use of certainty/uncertainty devices by the Turkish EFL learners is a little more visible (9,75%). In order to make better inferences about the data, personal certainty/uncertainty and impersonal certainty/uncertainty should also be compared.

When personal certainty and uncertainty is compared, it is seen that personal certain LBs, such as *I know that, we have seen* are more common than personal uncertain LBs, such as *I think that, it seems to me* in LOCNESS and TICLE. On the contrary, personal uncertain LBs are used more frequently in JPICLE. Native speakers used more personal certain LBs, they preferred using words with certainty; whereas words with uncertainty were preferred in impersonal LBs. Contrary to what Biber et al. (2004) state, personal certain LBs are used more frequently in learner corpora. Turkish EFL learners (14,24%) used personal certain bundles most frequently, but Japanese EFL learners (13,17%) used personal uncertain bundles most frequently. Although Japanese EFL learners used personal uncertain lexical bundles more frequently, their use of these devices is slightly more than personal certain lexical devices (10,36%). Hyland (2005) says that contrary to what is often thought, academic writing is not an impersonal monologue. He (2002 a, p. 1110) also states that "Self-mention constitutes a central pragmatic feature of academic discourse since it contributes not only to the writer's construction of a text, but also of a rhetorical self. The authorial pronoun is a significant means of promoting a competent scholarly identity and gaining acceptance for one's ideas." Since use of self-mention is an important stance marker, it should be wise to look at the use of self-mention words; such as the subject pronouns 'I' and 'we', object pronouns 'me' and 'us', and possessive pronouns 'my' and 'our'. The table below summarizes the self-mention words used in LOCNESS, TICLE and JPICLE.

**Table 9:** Use of Self-mention

Self mention words	TICLE (freq.)	LOCNESS (freq.)	JPICLE (freq.)
'I' (I think that, I've seen that, I cannot say, etc.)	388	107	1383
'WE' (We don't know, we can understand, we must think, etc.)	254	32	416
'MY' (in my opinion, in my view, my opinion is that)	70	9	26
'ME' (according to me, it seems to me, to me that)	14	7	24
Total	726	155	1849

As can be seen, the subject pronouns 'I' and 'we', and the object pronoun 'me' are used most frequently by the Japanese EFL learners. They used the subject pronouns excessively. The possessive pronoun 'my' is used most frequently by the Turkish EFL learners. Although Turkish EFL learners used 'my' most frequently, it was used in only one lexical bundle, which is *in my opinion*. Similarly, they used 'me' in only one lexical bundle: *according to me*. Both native speakers and Japanese EFL learners used *to me that* and *it seems to me*. *To me that* was used 14 times by the Japanese EFL learners and 4 times by the native speakers. *It seems to me* was used 10 times by the Japanese EFL learners and 3 times by the native speakers. It is surprising that Japanese EFL learners used self-mentioning more than the native speakers and Turkish EFL learners since as Ohta (1991) states, first person pronoun is largely unacceptable because collective identity is more valued than individuality in Asian cultures. However, the results are consistent with Ishikawa (2009) that Japanese speakers of English overused the pronoun *I* and the expression *I think*, which shows a subjective point of view. The second group which used self-mention frequently was Turkish EFL learners. Uysal (2012) compared the argument preferences of Turkish people when writing in Turkish (L1) and in English (L2). She compared the indirectness markers and found that there is a similar use of indirectness markers in L1 and L2, specifically point of view of distancing devices, such as *I think* and *I believe that*. When writing in L2, the participants used these devices a little less than they do in their L1 and she concluded that when writing in English, Turkish speakers are influenced by their L1.

When impersonal certainty and uncertainty is compared, it was found that Turkish and Japanese EFL learners used more impersonal certain LBs than impersonal uncertain; however native speakers used more impersonal uncertain LBs. Turkish EFL learners used more certainty words in both personal and impersonal bundles. However, Japanese EFL learners used more certainty words in impersonal lexical bundles, whereas they used more uncertainty words with personal lexical bundles.

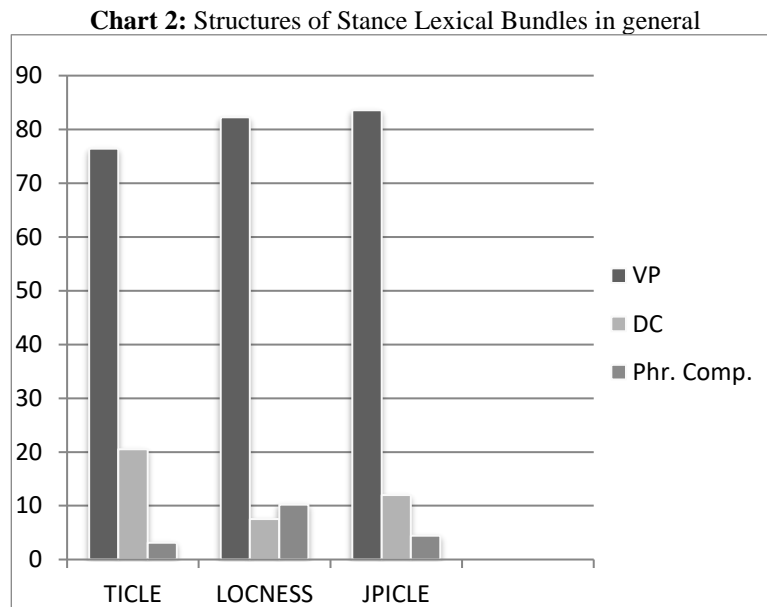
Table 8 also presents the comparison of attitudinal/modality SLBs in the three corpora. Japanese EFL learners used desire, ability and evaluation bundles more frequently than the native speakers and Turkish EFL learners. Compared to other attitudinal/modality LBs, obligation/direction and evaluation LBs were used most frequently by all the participants. Native speakers used obligation/direction and intention/prediction bundles more frequently than the EFL learners. The results are consistent with Tenuta et al. (2012) that in their study they also found that native speakers used modals that denote intention/prediction and obligation/direction more frequently than the nonnative speakers.

As seen in the results, Japanese EFL learners used evaluation bundles, such as *it is not good*, *is easy to*, *it is wonderful*, and *it is so hard* excessively. 27,17% of the bundles they used were evaluative. Kamimura and Oi (1998) compared the argumentative essays written by native English speakers and Japanese EFL learners and found that while more rational appeals are used in Western rhetoric, more affective appeals are used by Japanese speakers, giving more weight on evoking empathy on the reader. They use more softening and

emotional devices, such as *I think that* and *sad*. This may be the reason that Japanese EFL learners used more evaluative bundles; whereas English native speakers used them the least. The use of evaluative bundles by the Turkish EFL learners was in between the native speakers and Japanese EFL learners.

### Structures of stance lexical bundles

The summary of the structures of the SLBs used in all corpora are shown in the Chart 2 below.



As Chart 2 shows, verb-phrase structures were used most frequently in all corpora. 83,47% of the bundles constitute verb-phrase structures in JPICLE, 83,15% in LOCNESS and 76,39% in TICLE. Dependent clauses follow verb-phrase structures in TICLE (20,49%) and JPICLE (12,05%). Phrasal components were used the least in TICLE (3,12%) and JPICLE (4,48%); whereas 9,78% of bundles was phrasal components in the native speaker data. This result is inconsistent with the literature stating that noun phrase structures are immensely used (Strunkyte and Jurkunaite, 2008; Bal, 2010; Wei and Lei 2011). However, it is consistent with Hernandez (2013). A more in-depth analysis of the structures of LBs was also needed in order to compare them more effectively.

**Table 10:** Structures of Stance Lexical Bundles in detail

Corpora		TICLE		LOCNESS		JPICLE	
Fre. & %		Freq.	%	Freq.	%	Freq.	%
VP Frag ments	SP+VP	<b>109</b>	37,85	<b>75</b>	40,76	<b>177</b>	<b>49,58</b>
	VP	<b>110</b>	38,19	<b>77</b>	<b>41,85</b>	<b>119</b>	33,33
	Q	<b>1</b>	0,35	<b>1</b>	0,54	<b>2</b>	0,56

Dependent Clause	DC	59	20,49	13	7,07	43	12,05	
	Phrasal Components	NP+Post Mod.	7	2,43	13	7,07	12	3,36
		PP+Emb.	2	0,69	5	2,71	4	1,12
<b>TOTAL</b>			<b>288</b>	100	<b>184</b>	100	<b>357</b>	100

As stated previously, the most commonly used structure was VP fragments in all corpora. Dependent clause structures follow VP fragments in TICLE and JPICLE, whereas phrasal components follow it in LOCNESS. VP fragments consist of a subject pronoun followed by a verb phrase, such as *I'm not going to*, a verb phrase, such as *is very important*, or a question, such as *do we think that*. Native speakers and Japanese EFL learners used more SP+VP structures than VP; whereas Turkish EFL learners used VP structures slightly more than SP+VP. Question forms were very rarely used in all corpora. Dependent clause structures are reasonably used by the EFL learners (Turkish 20,49%, Japanese 12,05%). Native speakers used DCs 7,07% and the highest number of phrasal components was used by native speakers (9,78%). There is an insignificant difference in the use of VPs between Turkish and Japanese EFL learners. These results are consistent with Hernandez (2013) that all three groups used VP fragments most frequently. Both in this study and Hernandez's study, native speakers used more phrasal components than dependent clause structures. In the learner data, it is seen that dependent clause is used more frequently than phrasal components. Conversely, phrasal components are used more frequently in Hernandez (2013). The reason why Turkish and Japanese EFL learners use more DC structures could be related to their mother tongue. As stated in Allen (2009), a similar structure, like *it can be said that*, is used in Japanese. Likewise, similar structures are used in Turkish.

The findings do not coincide with Strunkyte and Jurkunaite (2008) that in their study they found that NP components are used most frequently, followed by VP and DC fragments. Similarly, Bal (2010) found that Turkish scholars use more PP and NP fragments than VP fragments. Wei and Lei (2011) found a similar result that native speakers used NPs more often; whereas Chinese EFL learners preferred VPs.

The structures of 40 most common lexical bundles were also compared in detail. The comparison can be seen in Table 11 below, which gives a broad idea about the structures used in native and nonnative corpora.

**Table 11:** Structures of 40 Most Common LBs in detail

	Structures	TICLE	LOCNESS	JPICLE
NP/PP	NP based	1	2	2
	PP based	3	1	-
VP	Copula be+NP/Adj	3	2	5
	VP with active verb	5	1	1
	Passive verb	2	4	2
	VP+that clause	-	1	-
	Modals+verbs	2	17	4
SP+VP	SP+VP	9	4	15
	Dependent clause	6	3	2
	There is/are	3	2	-
	Anticipatory it+VP/Adj Pr.	6	3	9
<b>TOTAL</b>		<b>40</b>	<b>40</b>	<b>40</b>

When the structures of the 40 most common lexical bundles were compared, the results showed that prepositional phrases were not used in JPICLE; whereas a small number of PP was

used in TICLE and LOCNESS. VP+that clause was used only in LOCNESS: *argue that the*. The pronoun *there* is not used by Japanese EFL learners. Copula be+NP/Adj phrase was mostly used in JPICLE; the reason for this is that evaluative lexical bundles were more frequently used by Japanese EFL learners and the adjectives used in this group were evaluative. Similarly, anticipatory it+VP/Adj phrase was mostly used by Japanese EFL learners. Again, the reason for that is the excessive use of evaluative lexical bundles by the Japanese learners. Native speakers used more passive structures than EFL learners; which coincides with Wei and Lei (2011). These results are not consistent with Hyland (2008a) and Wei & Lei (2011) that in their studies, they found that learners used less anticipatory it and more passive structures. This could be because teaching materials and professors in China suggest learners to use devices of impersonality, such as passives. The conflicting results may due to cultural differences.

### Conclusion

As aforementioned in the findings, the EFL learners used more stance bundles (Turkish EFL learners 288, Japanese EFL learners 357) than native speakers (184). According to the frequency and LL analysis, there is a certain overuse of lexical bundles in the EFL learners' essays when compared to the native speaker group (L1 vs. L2). The Japanese EFL learners used lexical bundles far more frequently than the native speakers-with a +18,28 LL ratio. When the LL ratio of JPICLE and TICLE was compared (L2 vs. L2), an overuse of lexical bundles by the Japanese EFL learners was observed. In general, the Japanese EFL learners used lexical bundles more frequently than the native speakers and the Turkish EFL learners. The overall results of LL ratio showed that lexical bundle usage by the Turkish learners remains somewhere between the native speakers' and the Japanese learners' usage. The general tendency of overuse of these devices by the EFL learners might be due to interlanguage properties, which tend learners to determine certain ways of using stance lexical bundles. However, since SLBs have not been determined in Turkish and Japanese languages yet, it is not possible to make a totally accurate interpretation whether the (over)use of them is related to the learners' native language (L1 transfer), cultural writing styles or cultural writing habits. However, the detailed analysis of the functions, structures and the use of self-mention in the corpora can give us a comprehensive picture of how to express stance.

When the functions of stance lexical bundles were compared, the results revealed that epistemic stance bundles were used more frequently than attitudinal/modality stance bundles by the native speakers and Turkish EFL learners; however the use of epistemic and attitudinal/modality stance markers is just the opposite in Japanese EFL learners. When personal and impersonal bundles were compared, the results showed that impersonal stance bundles were used more frequently than personal bundles in all groups. When the certainty and uncertainty were compared, it was seen that the native and the Turkish EFL learners used certainty devices more than uncertainty; whereas certainty and uncertainty devices were equally used by the Japanese EFL learners. When the use of im/personal un/certainty devices is compared, it was seen that the native speakers used impersonal uncertain and personal certain lexical devices more frequently (than personal uncertain and impersonal certain). The Turkish EFL learners used personal certainty devices more than personal uncertainty devices; however, their use of impersonal certainty devices was more than personal certainty devices. Similarly, Japanese EFL learners used impersonal certainty devices more than impersonal uncertainty devices. Yet, they used personal certainty devices less than personal uncertainty. It can be concluded that the native speakers are more confident in expressing their point of view. This can be related to learners' cultural background.

When the use of self-mention words and most common 40 stance lexical bundles are looked at, it is found that self-mention words (I, we, my, and me) were used most frequently by the Japanese EFL learners. The use of these words is in between in the Turkish EFL learners and the native speakers used them least. The analysis of the 40 most common lexical bundles also showed that EFL learners used devices that are more subjective; on the other hand, native speakers used more objective devices when expressing stance.

When the structures of stance lexical bundles were compared, it was found that verb-phrase structures were used most frequently in all corpora. Dependent clauses follow VP structures in EFL data. In LOCNESS, phrasal components follow VP structures. This could be related to the participants' mother tongue that as Allen (2009) states, similar structures; such as *it can be said that*, are also used in Japanese. Likewise, these structures are also used in Turkish. It is surprising that both Turkish and Japanese EFL learners used anticipatory it+VP/Adj pronouns more frequently than the native speakers since both languages are pro-drop. However, it is seen that the native speakers used more passive structures. This may be due to cultural differences and the more use of evaluative lexical bundles by the EFL learners.

### Implication and Recommendations

Hyland (2008 a) states that formulaic language is difficult to acquire and that failing to use it correctly identifies learners as outsiders. As Jafarpour et al. (2013) point out; language teachers should attract their learners' attention to collocations, the co-occurrence of words, since grammar and individual words are not sufficient in language proficiency. Different academic discourse relies on different lexical clusters (Hyland, 2008b), therefore lexical bundles specific to different genres should be taught and students should be made aware of these differences and the importance of these expressions. Bundles should be taught by presenting the complete spectrum of their functions in context by analyzing these functions in discourse similar to the one they encounter daily in their academic lives. Language teachers should be made aware of corpus-based teaching and how they can use corpora in their classrooms to make students more aware of these structures and reach native like performance. Finally, textbooks and other teaching materials should be prepared considering these frequently used structures. However, as Flowerdew (2001) suggests, pedagogic materials and teaching methods should be corpus based, not corpus bound. Nonnative data helps researchers to see how mother tongue and cultural differences affect metadiscourse use. Therefore, more nonnative corpora should be collected to make accurate inferences about language learning and teaching by making better contrastive interlanguage analysis. These may help researchers understand human mind better at an abstract level. After understanding how these structures are processed and where the problematic areas are, more experimental studies can be conducted to help learners acquire these structures and become more fluent and proficient language learners. In this study, only stance lexical bundles were analyzed. Studies looking at other functions of lexical bundles should also be conducted to understand these structures completely. Focusing on different genres, academic and nonacademic, spoken and written language would also give more insights on this topic.

### References

- Allen, D. (2009). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS Learner Corpus. *Komaba Journal of English Education*, 1, 105-127.
- Ağçam, R. (2014). *A Corpus-Based Study on Author Stance in Academic English*. Unpublished doctoral thesis. Çukurova University, Adana.
- Bal, B (2010). Analyses of Four-word lexical bundles in published research articles written by Turkish scholars. Unpublished Master's Theses. Georgia State



- University, Georgia. Retrieved 5 July 2011 from: [http://digitalarchive.gsu.edu/alesl\\_theses/2](http://digitalarchive.gsu.edu/alesl_theses/2)
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Biber, D., Conrad, S. & Cortes, V. (2004). If you look at: Lexical Bundles in University teaching and textbooks. *Applied Linguistics*, 25 (3), 371-405.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amstersam, Philadelphia: John Benjamins.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263–286.
- Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology*, 14 (2), 30-49.
- Chen, Z. (2012). Expression of epistemic stance in EFL Chinese university students' writing. *English Language Teaching*; 5 (10), 173-179.
- Cortes, V. (2002). Lexical bundles in freshman composition. In R. Reppen, S. M.Fitzmaurice and D. Biber (Eds), *Using Corpora to Explore Linguistic Variation*, 131-145: Amsterdam, Philadelphia: John Benjamins.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23, 397–423.
- Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17, 391–406.
- Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora*, 3, 43-57.
- Conrad, S. & Biber, B. (2004). The frequency and use of lexical bundles in conversation and academic prose. *LEXICOGRAPHICA*, 20, 56-71.
- Flowerdew, L. (2001). The Exploitation of small learner corpora in EAP materials design. In Ghadessy, M., Henry, A. & Roseberry, R.L. (Eds.) *Small Corpus Studies and ELT Theory and Practice*. Amsterdam: John Benjamins
- Granger, S. (2002). *A Bird's-eye view of learner corpus research in computer learner corpora, second language acquisition and foreign language teaching*. Granger, S. & Hung, J. & Petch-Tyson, S. (Eds), 3-33. John Benjamins, Amsterdam.
- Herbel-Eisemann, B., Wagner, D. & Cortes, W. (2010). Lexical bundle analysis in mathematics classroom discourse: The significance of stance. *Educ Stud Math*, 75, 23-42.
- Hernandez. P.S. (2013). Lexical bundles in three oral corpora of university students. *Nordic Journal of English Studies*, 13(1), 187-209.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7 (2) 173-192.
- Hyland, K. (2008a). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4–21
- Hyland, K. (2008b). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18 (1), 41-62.
- Ishikawa, S. (2009). Phraseology overused and underused by Japanese learners of English: A contrastive interlanguage analysis phraseology. *Corpus Linguistics and Lexicography*, 87-100.
- Jablonkai R.R. (2010). *A Corpus-linguistic investigation into the lexis of written English EU Discourse: An ESP Pedagogic Perspective*. Unpublished doctoral thesis. Eötvös Loránd University, Budapest.

- 
- Jafarpour, A.K., Hashemian, M. & Alipour, S. (2013). A corpus based approach toward teaching collocation of synonyms. *Theory and Practice in Language Studies*, 3 (1), 51-60.
- Jalali, H., Rasekh, A.E., & Rizi, M.T. (2008). Lexical bundles and intradisciplinary variation: The case of applied linguistics. *International Journal of Language Studies*, 2 (4), 447-484.
- Juknevičienė, R. (2011). *Lexical bundles in non-native speaker and native speaker written English*. Unpublished doctoral dissertation. Vilnius University.
- Kamimura, T. & Oi, K. (1998). Argumentative strategies in American and Japanese English. *World Englishes*, 17 (3), 307-323.
- Neely E. & Cortes.V. (2009). A little bit about: Analyzing and teaching lexical bundles in academic lectures. *Language Value*, 1(1), 17-38. Servei de Publicacions de la Universitat Jaume I: Castelló, Spain. Retrieved 9 July 2012, from: <http://www.erevistasuji.es/languagevalue>.
- Nekrasova, T.M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning*, 59 (3), 647-686.
- Nesi, H., & Baştürkmen, H. (2006). Lexical Bundles and Discourse Signaling in Academic Lectures. *International Journal of Corpus Linguistics*, 11, 283-304.
- Ohta, A.S. (1991). Evidentiality and Politeness in Japanese. *Issues in Applied Linguistics*, 2 (2), 211-238.
- Scott, M. (2008). WordSmith Tools version 5, Liverpool: Lexical Analysis Software.
- Strunkyte, G. & Jurkūnaite, E. (2008). *Written academic discourse: lexical bundles in humanities and natural sciences*. Unpublished BA thesis. Retrieved 9 July 2012, from, [http://www.lexically.net/wordsmith/corpus\\_linguistics\\_links/Lexical%20Bundles%20in%20Humanities%20and%20Natural%20Sciences.pdf](http://www.lexically.net/wordsmith/corpus_linguistics_links/Lexical%20Bundles%20in%20Humanities%20and%20Natural%20Sciences.pdf)
- Taki, S. & Jafarpour, F.(2011). Engagement and stance in academic writing: A study of English and Persian research articles. *Mediterranean Journal of Social Sciences*, 3 (1), 157-168.
- Tenuta, A. M., Oliveira, A.L.A.M. & Orfanó, B. M. (2012). How Brazilian learners express modality in their writing: A corpus-based study on lexical bundles. *Revista Intercâmbio*, 5.(26), 1-15. São Paulo: LAEL/PUCSP.
- Uysal, H. H. (2012). Argumentation across L1 and L2 Writing: Exploring Cultural Influences and Transfer Issues. *Vigo International Journal of Applied Linguistics*, 9, 133-159.
- Wei, Y. & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC Journal*, 42 (2), 155-166.
-