

Differences in linguistic and discourse features of narrative writing performance

Dr. Bilal Genç¹
Dr. Kağan Büyükkaracı²
Ali Göksu³

Abstract

The research presented in this paper aimed to investigate the linguistic and discourse characteristics of narratives produced by student-teachers in an ELT department. Thirty-four students from the ELT Department of Inonu University participated in the study. Each was asked to write two stories about an experience in which they were made angry (in English) and an experience in which they made someone angry (in Turkish). A total of 68 narrative texts were collected. Using concordance software the number of types and tokens and the type-token ratio were calculated; and each sentence in the text was rated as simple, coordinate, complex and coordinate+complex by the researcher to reveal the syntactic richness. The analyses suggest that although the participants' Turkish texts are richer than the English ones in terms of lexical richness, there is high parallelism between the texts in terms of syntactic richness. We believe that relatively poor vocabulary in the target language is the main reason for participants to prefer writing in their native language.

Key Words: emotional language, native language, foreign language, vocabulary

Introduction

Although narratives are among the most frequently taught types of written in general foreign language courses, compared with the studies on narratives in oral language, there is not much research on written narratives (Kormos, 2011). Interest in the use of vocabulary of language learners in their written productions has gained momentum during the last ten years.

In their study, Dewaele and Pavlenko (2003) analyzed retellings of movies in English and Russian produced by monolinguals and two types of bilingual to investigate cross-linguistic and cross-cultural effects (as well as possible gender effects) on productivity and lexical diversity. They found that neither the type of material, i.e the movie the participants were asked to watch, nor gender was linked with productivity and lexical variety. In addition, gender was not found as a key variable that determines the number and range of emotion words in speech, either (Dewaele & Pavlenko, 2002).

Again, regarding emotion vocabulary in the target language, Dewaele (2005) argues that since emotions are a crucial aspect of human mental and social life, they should not be absent in foreign language teaching material. That language teaching materials do not include enough emotion vocabulary accounts for the infrequent emotion words in L2 users' interlanguage.

Investigating lexical diversity, Laufer (2003) carried out a study with immigrants in Israel who maintained their first language, Russian, by using it on a daily basis. The study showed that L1 lexical

¹ University of Inonu, bilal.genc@inonu.edu.tr

² Suleyman Demirel University, kaganbuyukkaraci@sdu.edu.tr

³ Gazi University, aligoksu46@gmail.com

diversity declined as the contact time with L2 (Hebrew) increased. The percentage of non-frequent vocabulary and the total number of words produced in free expression significantly decreased as time passed.

In order to find out the effects of topic on the use of modal verbs, Hinkel (2009) conducted a study in which speakers of English, Chinese, Korean, and Japanese were given five topics about which they were to write essays. She found that the median frequency rates of modal verbs in L2 essays are significantly affected by the writing topic; the frequency rates of possibility and ability modals appear to be less topic dependent than obligation and necessity modals in the L2 writing of Chinese, Japanese, and Korean speakers. Again, the same researcher investigated tense and voice differences between written productions of NNS and NS. She especially focused on the usage of three English tenses (the present, the past and the future), two aspects (the progressive and the perfect), and passive verb structures. She found that even after many years of L2 learning and use, advanced NNS students may have difficulty with the conventionalized uses of tenses, aspects and the passive voice in written academic discourse (Hinkel, 2004).

Ho (2009) conducted a study with the participation of 33 students from two universities in Hong Kong. In the study the narrative length, lexical and syntactic richness, and the use of metaphorical expressions in the Chinese and English texts were compared and the students' perceptions of bilinguality and emotionality were also elicited. The analyses and comments of the students led the researcher to argue that to a very large extent language competence affects how emotionally expressive bilinguals can be.

With a narrower perspective, Crossley and McNamara (2009) analyzed the use of cohesive vocabulary in the written productions of first language (L1) writers of English and second language (L2) writers of English. They found that L2 writer differ in their use of cohesive devices as compared to L1 writers and that L2 writers are less likely to create a coherent text that is as readable and thus as comprehensible as the text of an L1 writer. Since L2 texts are also less abstract and less ambiguous, they are more likely to be context dependent. Allman (2005) compared the vocabulary size of English monolinguals, Spanish monolinguals, and English/Spanish bilinguals and found that, regarding the size of vocabulary, the English monolinguals scored the highest, followed by the English dominant bilinguals, followed by the balanced bilingual group, with the Spanish dominant bilinguals scoring the lowest. Allman argues that when learning or acquiring English, whether the learner is monolingual or bilingual, there is an advantage to being in an English dominant environment. However, the impact of being in an English dominant environment is even more negative on the Spanish language development of Spanish monolinguals' and English/Spanish bilinguals' vocabulary size.

In order to study lexical and syntactic richness in Turkish and English texts produced by ELT department students, we asked those pre-service teachers to write about the two topics given to them. With this we aimed to get an overall picture of the participants' lexical and syntactic skills.

Research Questions

The study reported in this paper intends to investigate the structure of narratives in L1 and L2 writing. Given the fact that two narrative tasks make different cognitive complexity demands on the stages of writing, in this study we have tried to find answers to the following research questions:

- To what extent do Turkish and English written productions of Turkish speakers of English exhibit differences in terms of vocabulary variation?
- To what extent do the English written productions of Turkish speakers exhibit lexical variety?
- To what extent do the writings of Turkish speakers exhibit differences in terms of sentence complexity?

Data Collection

This study was carried out with the participation of Turkish students majoring in English at the ELT Department, University of Inonu. In order to elicit written narratives, the participants were first asked

to write two stories about an experience in which they were made angry and an experience in which they made someone angry, the former in English and the latter in Turkish. We also asked the participants to provide written comments on their feelings and language preferences in writing life stories.

The participants were also asked in which language they prefer to produce written productions and provide reasons for their choice. While quantitative analyses were conducted regarding lexical richness, variety and syntactic richness; the comments of the participants were the subject of qualitative analysis.

Participants

The participants in this study were thirty-four junior (3rd year) students from the ELT Department of Inonu University. Given the fact that, the study was conducted in May of the 2010-2011 academic year, it becomes clear that they participants had already been studying English for four years (including one year preparatory education). At the end of the following year they would be qualified as Teachers of English.

Data Analysis

The writing performances of the participants were grouped as separate entities; therefore, two units of writing tasks were investigated in this study. All the units were transformed into soft copies by the researcher using standard orthography for analysis. Contracted forms of auxiliaries were counted as a single word.

As the title of the software suggests, Concordance 3.3 is capable of making indexes and word lists, counting word frequencies, comparing different usage of a word, analysing key words, and finding phrases and idioms*.

Lexical Diversity

As Dewaele and Pavlenko (2003) suggest lexical diversity is measured through a type-token ratio (TTR), which compares the number of different words (types) with the number of total words (tokens) and TTR is mostly considered to be the main parameter of lexical diversity. In this study we also calculated lexical richness through TTR and we analyzed lexical variety in the target language by comparing the token and types with that of Concordance base word lists.

Sentence Types in English

Peters (2004) classifies three sentence types: simple, compound and complex. In this study syntactic analysis of the texts was based on this classification and one more type was added: those sentences formed by the combination of compound and complex sentences. A simple sentence consists of a single clause; the coordinates are usually joined by conjunctions such as *and*, *but*, or *or*, though a semicolon or occasionally a comma can also serve to coordinate; and in complex sentences the clauses are linked so as to give one of them superior status. The main clause (principal clause) occupies superior status and the other clause is labelled as the subordinate (or dependent) clause (Peters 2004).

Sentence Types in Turkish

As in English, in Turkish there are three basic sentence types: simple, coordinate and complex. As in English simple sentences consist of a single clause and complex sentences are linked so as to give one

* For more information the reader is referred to <http://www.concordancesoftware.co.uk/>

of them superior status. Regarding the identification of coordinate sentences we took into consideration the definition made by Kornfilt (1997). According to her, there are essentially three ways to coordinate sentences:

- a) by simply stringing the coordinated sentences one after another, without using any coordination marker;
- b) by using the unbound conjunction marker *ve* ‘and’, borrowed from Arabic, between two conjoined sentences (or between the last two, if there are more than two conjuncts); or
- c) by attaching the coordination postclitic DA to the first constituent of the second conjunct, if only two sentences are conjoined, and to the first constituent of the last conjunct, if more than two sentences are conjoined (Kornfilt, 1997).

Results and Discussion

Research Question 1: To what extent do Turkish and English writings of Turkish speakers of English exhibit differences in terms of vocabulary variation?

To see the vocabulary variation in Turkish Corpus and English Corpus, we first analyzed the corpora through *Concordance*; then we provided token and type numbers and type-token ratio in Table 1 below.

Table 1: Lexical Diversity in Turkish Corpus (TC) and English Corpus (EC)

	TC	EC
Tokens	3596	4353
Types	1589	838
Type/Token Ratio*	44,18	19,25

*Type-Token Ratio = (number of types/number of tokens) * 100

As seen in Table 1, the participants told their experiences in which they made someone angry (i.e TC) using a total of 3596 words. When we look at the number of types we see that there are 1589 words which means the participants wrote a 3596 token corpus using 1589 words. Type-Token Ratio (TTR) value of the TC (44.18%) also indicates the extent of lexical richness in the corpus. In order to better understand the meaning of percentages in the table, the explanation given by Laufer (2003) is helpful:

For example, in a 300-word composition, 240 words belong to the ‘first 2000 most frequent’ vocabulary. The 60 remaining words therefore are infrequent. The percentage of these words, the lexical richness of the composition, is 20%. The second component of lexical diversity is sometimes termed ‘lexical variation’. For example, in a composition of 300 words, the writer uses 150 *different* words. The lexical variation of the composition is $150/300 \times 100\% = 50\%$. High lexical variation shows that a person is not repetitive in his choice of words (p.26-27).

On the other hand, in English Corpus (EC) the participants told their stories in 4353 tokens using 838 types. The TTR value (19.25%) indicates a relatively poor lexical variety in EC. This means that while in TC almost half of the vocabulary are different from each other in EC only one fifth of the vocabulary are different from each other.

The fact that there 757 fewer tokens in TC does not mean that the texts written in Turkish were shorter than those written in the target language. English and Turkish are both labelled as agglutinating languages, i.e., in both languages words typically contain a linear sequence of morphs – and thus contrast with isolating and inflectional languages. However, while some languages display agglutination to a minor extent, some languages such as Turkish and Japanese display agglutination to a major extent (Crystal, 2008).

In Turkish, nouns display the feature of grammatical case. Grammatical case of a noun or pronoun is an inflectional form that indicates its grammatical function in a phrase, clause, or sentence. There are six cases in Turkish: 1) absolute case: the simplest form of a noun with no suffixes, 2) accusative case: denoting direct object of a verb or the object of certain prepositions, 3) genitive case: denoting possession, 4) dative case: denoting indirect object, 5) locative case: denoting place and 6) ablative case: denoting point of departure (Lewis, 1967). Thus in Turkish certain tasks performed by various prepositions in English could be achieved through suffixes. In addition to grammatical case in nouns, Turkish language marks tense, aspect and person on verbs through conjugation. So in Turkish one can construct subjectless sentences and there are no separate words to mark tense and aspect.

There is also another striking difference between English and Turkish: Turkish language lacks an overt definite article. Thus the above mentioned three differences between the two languages account for how two texts with almost the same length display a great difference in terms of token numbers. Here we would like to quote a famous example given in several context to reveal the complexity of the Turkish language: The question sentence “Çekoslavakyalılaştıramadıklarımızdanmısınız?” is a one word sentence which means “Are you one of those that we couldn't turn into a Chekoslovakian” in English.

Research Question 2: To what extent do the writings of Turkish speakers of English exhibit lexical variety in the L2 writing performances?

In the second stage of our analyses we compared lexical variety in EC by comparing the tokens in EC against those of *Concordance* base word lists. Concordance has three baseword lists, each of which has 1,000 words.

Table 2: Type-token Variety in EC

	Token		Type		Family	
	n	%	n	%	n	%
Concordance Baseword1	3934	90.37	619	73.87	426	64,94
Concordance Baseword2	221	5.08	115	13.72	98	14,94
Concordance Baseword3	99	2.27	37	4.42	33	5,03
Not in the List	99	2.27	67	8.00	99	15,10
Total	4353	100	838	100	656	100

Looking at Table 2, we see that almost 98% of the vocabularies in EC are among core vocabulary. 26 of the words that do not appear in Concordance Baseword lists are Turkish words. While four of those words are proper names belonging to places and humans, twenty two Turkish words appear as embedded an English narratives.

In addition to relatively poor vocabulary, it seems that the participants' vocabulary use does not reveal much variety. When we refer to Allman (2005), we could understand the reasons for poor vocabulary: Allman postulates that whether the learner is monolingual or bilingual, in learning English there is an advantage to being in an English dominant environment.

Research Question 3: To what extent do the writings of Turkish speakers of English exhibit differences in terms of sentence complexity?

In the third stage of our analyses we measured syntactic richness in both corpora. As mentioned above, the sentences were categorized into four groups as simple, coordinate, complex and coordinate+complex sentences.

Table 3: Syntactic Richness in TC and EC

	TC		EC	
	n	%	n	%
Simple Sentence	187	45,17	154	39,70
Coordinate Sentence	97	23,68	129	33,49
Complex Sentence	106	25,60	78	20,10
Coordinate+Complex	24	5,80	27	6,96
Total	414	100	388	100

As seen in Table 3, in TC there are a total of 414 sentences consisting of 187 (45,17%) simple sentences, 97 (23,68%) coordinate sentences, 106 (25,60%) complex sentences and 24 (5,80%) coordinate+complex sentences. In EC, on the other hand, there are a total of 388 sentences consisting of 154 (39,70%) simple sentences, 129 (33,49%) coordinate sentences, 78 (20,10%) complex sentences and 27 (6,96%) coordinate+complex sentences. Looking at these figures, it is seen that there is much difference between TC and EC in terms of syntactic richness. In Ho's (2009) study the participants generally produced longer texts and developed story lines more fully in Chinese and there was a greater richness of emotion vocabulary, syntactic structures and metaphorical expressions in Chinese than in English. In our study, however, while the participants Turkish texts excelled in terms of lexical richness, the two corpora have similar syntactic richness.

When we compare the number of tokens and sentences in both corpora we see that while there are more tokens in EC than in TC, there are more sentences in TC than in EC. The underlying reason for differentiation becomes clear when we remember that in Turkish sometimes sentences can be made up single words because verbs in Turkish could be added displaying person, tense and aspect of the sentence.

Given the fact that, the participants are third year ELT department students, their proficiency in writing in terms of syntactic richness seems quite promising. Regarding lexical richness and variety, however, the participants' target language lexical richness is poorer than that of their native language and only a few of the vocabulary in English corpus were beyond the 3000 base word list of Concordance. What is worse 90% of the vocabulary are confined with the 1000 base word list. Decarrico (2001) argues that in the past teaching vocabulary was often neglected "because it was thought that vocabulary could simply be left to take care of itself" (p.285). As this study has shown 'vocabulary could not take care of itself'.

Relatively poor vocabulary also accounts for why the participants said they prefer to write in their native language

When we look at the responses about their choice of language in writing, we see that all of the 34 participants stated they would prefer to write in Turkish. Below are some verbatim examples from the participants' remarks:

- *I can write in both languages. However, writing in English may be sometimes difficult for me. I can't find some words immediately. Hence, writing in Turkish is much easier than English.*
- *I prefer to use my native language when I write about my life. Because, I express my feelings best with my native language.*
- *When I write about my life experience, I usually prefer to use Turkish because it's my native language. So I feel more relaxed while writing in Turkish. I can explain my thoughts, feelings, opinions better in Turkish.*

The difficulty the participants face when they are asked to write about an emotional experience is quite clear. As can be inferred from the above quotations, the participants have difficulty in ‘finding some words immediately’ to express ‘their thoughts, feelings and opinions’.

In conclusion, from a communicative perspective we could argue that the participants need to improve their lexical competence, which is considered a sub-component of linguistic competence. Linguistic competence consists of grammar, phonology and vocabulary. This study shows that even advanced learners of English face difficulty in expressing their thoughts due to lack of vocabulary.

Conclusion

Previous research on certain aspects of written productions in target language has focussed on lexical variety and richness in written productions, the use of certain modal verbs and tense-aspect, and how certain topics affect the lexical and syntactic properties of written productions. Our study has a different scope than previous studies but also has a different focus.

In this study we examined the written productions of Turkish speakers of English both in their native tongues and in English. Findings show that the participants generally produce texts with almost the same length and syntactic richness. The lexical variety, however, in each text types is quite different: while the participants wrote a 3596 token corpus using 1589 words with a 44.18% Type-Token ratio in TC, in English Corpus (EC) they told their stories in 4353 tokens using 838 types with a 19.25% Type-Token ratio. Both the linguistic data and the written comments of students regarding in which language they prefer to write suggest that to a very large extent proficiency in vocabulary affects the bilinguals’ preference of language in which to write and their lexical productivity. Therefore, we believe that, as Kormos suggests (2011), ‘even at higher levels of L2 competence, not only do writing classes need to focus on the improvement of text-level composition skills, but should also include a lexical development component’ (p.159).

Since this study investigated the results of an emotional writing task we believe further research might be necessary to gain more insight into how different writing tasks with different task-complexity influence the linguistic quality of FL writing.

References

- Allman, B. (2005). Vocabulary Size and Accuracy of Monolingual and Bilingual Preschool Children. In J. Cohen, K. T. McAlister, K. Rolstad, and J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 58-77). Somerville, MA: Cascadilla Press.
- Crossley, S. A. & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing* 18 (2009), 119–135.
- Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics*. Malden, MA: Blackwell Publishing.
- Decarrico, J. S. (2001). Vocabulary Learning and Teaching. In Marianne Celce-Murcia (Ed.), *Teaching English As a Second or Foreign Language*. (pp.285-299). USA: Heinle & Heinle.
- Dewaele, J-M. (2005). Investigating the psychological and emotional dimensions in instructed language learning: Obstacles and possibilities. *The Modern Language Journal* 89 (3), 367–380.
- Dewaele, J-M., & Pavlenko, A. (2002). Emotion vocabulary in interlanguage. *Language Learning* 52 (2), 263–322.
- Dewaele, J-M., & Pavlenko, A. (2003). Productivity and lexical diversity in native and non-native speech: A study of cross-cultural effects. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 120–141). Clevedon, OH: Multilingual Matters.
- Hinkel, E. (2004). Tense, aspect and the passive voice in L1 and L2 academic texts. *Language Teaching Research* 8 (1), 5–29
- Hinkel, E. (2009). The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics* 41 (2009), 667-683.

Genç, B., Büyükkaracı, K. ve Göksu, A. (2012). Differences in linguistic and discourse features of narrative writing performance. *Dil ve Edebiyat Eğitimi Dergisi*, 1(2), 40-47.

Ho, J.W.Y. (2009). The language of anger in Chinese and English narratives. *International Journal of Bilingualism* 13 (4),481-500.

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing* 20 (2011), 148-161.

Kornfilt, J. (1997). *Turkish*. London:Routledge.

Laufer, B. (2003). The Influence of L2 on L1 Collocational Knowledge and on L1 Lexical Diversity in Free Written Expression. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 19–31).Clevedon, OH: Multilingual Matters.

Lewis, G. L. (1967). *Turkish Grammar*. Oxford: Oxford University Press.

Peters, P. (2004). *The Cambridge Guide to English Usage*. Cambridge: Cambridge University Press.