Differences in Linguistic Features of Spontaneous Speech: The Case of Obama's Speeches at the London, Toronto and Cannes G-20 Summit Press Conferences

Bilal GENÇ[A], Ali GÖKSU[B], Kağan BÜYÜKKARCI[C*]

[a]Inonu University, Faculty of Education, Malatya/Turkey
[b]Gazi University, Faculty of Education, Ankara/Turkey
[c]Suleyman Demirel Unıversity, Faculty of Education, Isparta/Turkey

CrossMark

Article Info

Abstract

The research presented in this paper aimed to investigate the linguistic and discourse characteristics of Mr. Obama's press conferences held after the G-20 summits in London in 2009, in Toronto in 2010 and in Cannes in 2011. The President's speeches were divided into two separate parts: (1) the first part during which he spoke about his plans to deepen international connections and the second during which he answered the questions of journalists from different nationalities– question/answer session – considered as spontaneous speech text and labeled a TC, LC and CC. Using concordance software the number of types and tokens and the type-token ratio were calculated. The results show that although the number of tokens decreases chronically, in terms of the qualitative analysis all speeches display similar characteristics regarding vocabulary variation.

Introduction

Narratives are always among the most taught written text types in foreign language courses They are used for higher level college learners as well as beginning levels. In addition, narratives are of importance in authentic written communication contexts, and they can be embedded in other text types, such as supporting evidence in argumentative writing (Kormos, 2011). Oral language including various narratives is generally characterized as informal, interpersonal, whereas written language also including other narratives is considered formal, planned and expository-like (Horowitz & Samuels, 1987). Although narratives are among the most frequently taught types of writing in general foreign language (FL) courses, compared with the studies on narratives in oral language, there have not been enough studies on written narratives.

Regarding this issue, Pu (2006) compared oral and written narratives in Chinese and English, and demonstrated that a discourse of same genre, same context, and same goal orientation exhibited significant patterns of differences at various levels of structure between the modalities across languages. The results of the study indicated that the differences arose mainly from distinct mental processes and mechanisms between speaking and writing in general, and also from varying degrees of inherent dependency of both on interpersonal involvement in story-telling, while the similarities were attributed to the storytelling genre and narrative nature of the particular discourse. In another study, Bardovi-Harlig (1992) analyzed oral and written narratives produced by sixteen L2 learners of English with different first languages and 24 native speakers of English. She found that L2 learners of English

[*]**Corresponding author:** kaganbuyukkarci@sdu.edu.tr

diverged from L1 speakers in that they made use of past tense forms to mark foreground and non-past forms to mark background.

Furthermore, first language (L1) and second/foreign language narratives in written text have been investigated by many researchers (Crossley and Mcnamara, 2009; Dewaele & Pavlenko, 2002; Hinkel, 2004, 2009; Kormos, 2011; Yee Ho, 2009). For example, Kormos (2011) examined the linguistic and discourse characteristics of narratives produced by upper-intermediate foreign language learners in a bilingual secondary school. This study was aimed at how specific linguistic and discourse characteristics of narrative tasks were influenced by aspects of task-complexity in L1 and FL writing. Hinkel (2004) identified the patterns and median frequency rates of L1 and L2 uses of three English tenses (the present, the past and the future), two aspects (the progressive and the perfect), and passive verb structures encountered in a native student (NS) and non-native student (NNS) corpus of L1 and L2 academic student academic texts She analyzed specific written discourse production in which NNSs' usage of English tenses and voice appears to be dramatically indifferent from that of NSs. The study revealed that most of advanced NNSs simply choose to avoid using such complex verb phrase constructions as passive voice, the perfect aspect, or predictive/hypothetical would. Crossley and Mcnamara, (2009) in their study analyzed how lexical differences related to cohesion and connectionist models can distinguish first language (L1) writers of English from second language (L2) writers of English. Results of the study showed that L1 and L2 written texts varied in several dimensions related to the writer's use of lexical choices. The dimensions correlated to lexical depth of knowledge, variation, and sophistication.

Beside the studies covering the written narratives in the field, the interest in the use of vocabulary of language learners in their written productions has also gained momentum during the last ten years. Laufer and Nation (1995) emphasize that the effective use of vocabulary in the written production provides a well-written composition. In addition, a well-used rich vocabulary has a positive effect on the readers. Moreover, the interest focusing on the measures of the lexical richness has tried to quantify the degree to which a writer is using a varied and large vocabulary. Their study was aimed at examining the relationship between direct measures of learners' vocabulary size and richness of vocabulary in their language production. They also researched whether Lexical Frequency Profile (LFP) bore a relationship to vocabulary size and was an indicator of quality of vocabulary use. The results of the study indicated that LFP was a valid and reliable measure of lexical use in writing. The purpose of this study was to research whether the oral productions of Mr. Obama display the differences in terms of vocabulary variation and to what extent the oral productions of Mr. Obama exhibit items from core vocabulary.

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 demands in terms of cognitive complexity at different stages of writing, in this study we have tried to find answers to the following research questions:
   • To what extent do the oral productions of Mr. Obama display differences in terms of vocabulary variation?
   • To what extent do the oral productions of Mr. Obama exhibit items from core vocabulary?

Method

Data Collection

President Obama delivered news conferences at Exvel Center in London at the conclusion of the G20 summit on April 2, 2009; at Intercontinental Hotel in Toronto on June 27, 2010 and at Claude Debussy Theater in Cannes on November 04, 2011. The President's speeches were divided into two separate parts: (1) the first part during which he spoke about his plans to deepen international connections and

the second during which he answered the questions of journalists from different nationalities– question/answer session –considered as spontaneous speech text and labeled a TC, LC and CC. Since the study aims to find out the characteristics of spontaneous speech, the first parts of the press conferences during which the president read his notes and thus labeled as read speech were excluded from the analyses.

Data Analysis

The G20 News Conference speeches of the US President were analyzed in a two step procedure. First, we divided the conferences into two sections; the first section was labeled as read speech and the second section as spontaneous. The second sections of press conferences held in London, Toronto, and Cannes in consecutive years were analyzed using the software Concordance 3.3. 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, analyzing 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 calculated lexical richness through TTR and also analyzed lexical variety in the target language by comparing the token and types with that of Concordance base word lists.

Results and Discussion

Research Question 1:
To what extent do the oral productions of Mr. Obama display differences in terms of vocabulary variation?

In the first step of the analysis, we summarized the counts for LC, TC, and CC corpora in Table 1. Using Concordance 3.3, we found 4370 tokens and 1082 types in LC, 3326 tokens and 858 types in TC, and 2058 tokens and 613 types in CC in Mr. Obama's oral productions. 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 150/300×100%= 50%. High lexical variation shows that a person is not repetitive in his choice of words (p.26-27).

Table 1.
Lexical Diversity in London Conference (LC), Toronto Conference (TC) and Cannes Conference (CC).

	LC	TC	CC
Tokens	4370	3326	2058
Types	1082	858	613
Type/Token Ratio*	24.76	25.80	29.79

The software also provided us the type/token ratios in LC, TC, and CC. As can be inferred from Table 1, the richest lexical variation in Mr. Obama' speeches was found to be in CC (x=29.79), the second highest lexical variation in TC (x=25.80), and the least lexical diversity in LC (x=24.76). Thus it becomes clear while in terms of number of tokens the president's speeches employs less vocabulary chronically,

in terms of lexical richness a quite different picture emerges in that although the later speeches have less token they display more lexical variety within themselves.

Research Question 2:
To what extent do the oral productions of Mr. Obama exhibit items from core vocabulary?

In the second step of analysis, we analyzed Mr. Obama's speeches in LC, TC, and CC in terms of core vocabulary productions. In Table 2, over 85% of the tokens (n=3782) in LC has been found to be in concordance baseword1. When we look at the types (n=654), we see that 60% of the vocabulary can be found in concordance baseword1, and nearly 54% of the types belong to same word family. Considering type ratio in concordance baseword1, we could say that Mr. Obama's speech, if used in language classes, LC may be easily understood by most language learners as it consists of basic vocabulary in English.

Table 2.

Type-token Variety in LC.

	Token		Type		Family	
	n	%	n	%	n	%
Concordance Baseword1	3782	86.54	654	60.67	449	53.71
Concordance Baseword2	274	6.27	188	17.44	161	19.26
Concordance Baseword3	77	1.76	59	5.47	49	5.86
Not in the List	237	5.42	177	16.42	177	21.17
Total	4370	100	1078	100	836	100

Of the tokens (n=274) in Obama' speech in concordance baseword2, we see that the types (n=188) form 17.44% of the total lexical variation. The words in concordance baseword3 (n=77) and the ones not in the list (n=237) composes almost 7% of the total tokens, and the types in this group (n=59) with the ones not in the list (n=177) forms nearly 22% of the whole lexical variation in LC.

The results of the analyses of Obama's core vocabulary uses in TC are presented in Table 3. When we look at Table 3, we can see that, similar to LC, more than 86% of the total tokens (n=2881) are in concordance baseword1. However, when we compare the types in TC with the ones in LC, it can be seen that a higher percentage (x= 63.87%) of the tokens have been found to be in concordance baseword1. On the other hand, the types found in concordance baseword2 (x=13.17%) and concordance baseword3 (x=3.50%) are lower than LC. In other words, we could say that Mr. Obama's speech can be considered to be easier to understand for language learners when compared with LC.

Table 3.

Type-token Variety in TC.

	Token		Type		Family	
	n	**%**	**n**	**%**	**n**	**%**
Concordance Baseword1	2881	86.62	548	63.87	397	57.04
Concordance Baseword2	166	4.99	113	13.17	102	14,.66
Concordance Baseword3	38	1.14	30	3.50	30	4.31
Not in the List	241	7.25	167	19.46	167	23.99
Total	3326	100	858	100	696	100

In TC, token types that are not in the list (n=167) has a higher percentage (x=20%) than LC. Type-token variety in CC is presented in Table 4 below. Out of the whole tokens (n=2058), 88.73% in CC is in concordance baseword1, and the types found in concordance baseword1 show the highest percentage (x=70.96%) among these three conferences.

Table 4.

Type-token Variety in CC.

	Token		Type		Family	
	n	**%**	**n**	**%**	**n**	**%**
Concordance Baseword1	1826	88.73	435	70.96	319	65.50
Concordance Baseword2	104	5.05	1887	12.56	71	14.58
Concordance Baseword3	31	1.51	25	4.08	21	4.31
Not in the List	97	4.71	76	12.20	76	15.61
Total	2058	100	613	100	487	100

The types in concordance baseword 2 and 3 (n=1912) consist of almost 17% of the whole lexical variation CC, and the words not in the list (n=76) involves 12.20 percent. When we compare the types in these three conferences, CC can be considered to be most understandable speech among English learners due to having the highest percentage of types being in concordance baseword1.

When we evaluate G-20 summit speeches of the presidents in the light of Krashen's Input Hypothesis, we could argue that the president's speeches prove to be highly beneficial authentic material for language learners around the world. In all the three speeches the percentage of vocabulary found the in the Concordance baseword1 is higher than 86%, which indicates that those speeches could be understood by having an intermediate or higher proficiency level in English. Thus it is clear that the president's G-20 summit press speeches provide language teachers around the world a precious resource through which authentic, comprehensible input is available.

Mr. Obama's G-20 summit press speeches not only provide comprehensible input but also an assessment and evaluation material which has content validity. As teachers of English we might make use of those speeches in assessing our students' listening skills. When we examine the material to see what kind of validity they exhibit, we could argue that they have content validity. Content validity in generally described as the degree to which the content of the test matches a content domain associated with the construct. It is clear that Mr. Obama's speeches mentioned in this study prove to be appropriate material for testing listening due to the vocabulary variation.

Conclusion

The study aimed at finding out whether the oral productions of Mr. Obama display the differences in terms of vocabulary variation and to what extent the oral productions of Mr. Obama exhibit items from core vocabulary. The first finding was that although the number of tokens in his speeches, decreased chronically, in terms of qualitative analysis all speeches show similar characteristics. On the other hand, when we consider these conferences in the sense of lexical variation, among Mr. Obama' speeches the Cannes Conference was found to have the richest, and London Conference to have the least lexical diversity.

Another finding was that although these three conferences differed in terms of lexical variation, the tokens and types of vocabulary in Mr. Obama's speeches can be considered to be mostly in concordance baseword1. That is, his speeches may be easily understood by many people. Moreover, these speeches can be used as authentic material for English language learners, especially for second and/or foreign language learners.

In conclusion, Mr. Obama's G-20 summit press speeches are of pedagogical significance for language teachers for two main reasons: For one thing, they can be embedded in our language teaching activities, for another they can be used in evaluating language students' proficiency levels.

References

Bardovi-Harlig, K. (1992). The telling of a tale: Discourse structure and tense use. learners' narratives. *Pragmatics and Language Learning 3*, 144-161.

Crossley, S.A & Mcnamara, D.S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing 18*, 119-135.

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.

Hinkel, E. (2004). Tense, aspect and the passive voice in L1 and L2 academic texts. *Language Teaching Research 8(1)*, pp. 5-29.

Hinkel, E. (2009). The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics 41*, 667-683.

Horowitz, R. & S.J. Samuels (Eds.). 1987. *Comprehending oral and written language*. San Diego, CA: Academic Press Inc.

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

Laufer, B. & Nation, P. (1995).*Vocabulary size and use: Lexical richness in L2 written production*. Oxford University Press.

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.

Pu, M.M. (2006). Spoken and written narratives: A comparative study. *Journal of Chinese Language and Computing 16(1)*, 37-61.

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