A Corpus-based Study on Turkish Spoken Productions of Bilingual Adults

Reyhan Ağçam[1,*], Adem Bulut[2]

[1]Faculty of Education, Kahramanmaras Sutcu Imam University, Turkey
[2]Faculty of Arts and Sciences, Cukurova University, Turkey

Abstract   The current study investigated whether monolingual adult speakers of Turkish and bilingual adult speakers of Arabic and Turkish significantly differ regarding their spoken productions in Turkish. Accordingly, two groups of undergraduate students studying Turkish Language and Literature at a state university in Turkey were presented two videos on a computer screen, and asked to narrate each film in Turkish as completely as possible, which was videotaped by the researchers. Subsequently, two sets of corpora were compiled from the video transcriptions and analysed through a computer programme. The findings showed that the bilingual group does not significantly differ from the monolingual group in their choice of word order and voice, and that they overused the present progressive tense, and conversational fillers while narrating video-films as opposed to the monolingual group. It might be concluded that both languages are active in their mind, and that their lexical access to their L1 is slightly stronger than that to their L2. The study concludes with a few suggestions for further directions.

Keywords   Bilingualism, Arabic, Turkish, Interference

1. Introduction

Widely known as competence in more than one language, bilingualism has taken considerable attention in the field of applied linguistics. It was initially defined by Bloomfield [1] as 'native-like control of two languages'. About four decades later, Titone [2] redefined it as 'the individual's capacity to speak a second language while following the concepts and structures of that language rather than paraphrasing his or her mother tongue'. In the Webster's Dictionary [3], on the other hand, bilingual is defined as 'anyone who uses two languages especially as spoken with the fluency characteristic of a native speaker; a person using two languages especially habitually and with control like that of a native speaker'. However, McNamara [4] describes bilinguals as people who possess 'a minimal competence in only one of the four language skills, listening comprehension, speaking, reading and writing, in a language other than his mother tongue'. Although scholars have not negotiated on the meaning of the term 'bilingual', the general consensus is that age factor plays a significant role in the acquisition of a second or third language; namely, it is widely believed that the younger a person is exposed to a second or third language, the more successfully s/he acquires it. Baker [5] and Reich [6], for instance, advocate that children speak two languages as fluently as native monolinguals if they receive enough exposure to both languages by age four. Woolfolk et al. [7], on the other hand, argue that if they learn two languages as toddlers, there is a period between ages two and three when they progress more slowly because they have not yet figured out they are learning two languages and they may mix up the grammar of those languages. Yet, the research on bilingualism has indicated that they are likely to achieve native-like proficiency in both languages. Cummins [8] calls our attention to a different point stating that the more proficient the speaker is in the first language, the faster she or he will master a second language. The last century has experienced various approaches towards bilingualism, which has taken considerable attention in the field of applied linguistics. Cook [9] attributes it to the fact the majority of the world population utilizes two or more languages for communication. Likewise, Grosjean [10] notifies that more than half of the world's population speaks two or more languages. The Republic of Turkey, which currently welcomes more than thirty languages in addition to Turkish, has been no exception in this sense since its foundation. The population of the people who fall into this category is not precisely known; however, it is estimated that monolinguals constitute the minority in the whole country whereby Kurmanji (Kurdish), Arabic, and Azerbaijani could be listed as the most widely spoken minority languages that will be covered in detail in the subsequent section.

1.1. Theoretical Background

Kornfilt Kornfilt [11] suggests that Turkish, which is the native language of over 90% of the population, is the official and dominant language in Turkey. As for minority

languages she describes Kurdish speakers as the largest linguistic minority in the country, and speakers of Arabic, some Caucasian languages and speakers of Gagauz as small minority language communities. In more detail, Lewis et al. [12] report that the number of individual languages listed for Turkey is 36 (35 living and 1 extinct) and that among the living languages, fifteen are developing, three are institutional, six are vigorous, ten are in trouble, and one is dying. Table 1 is intended to show the distribution of the languages in concern.

As seen in Table 1, Arabic and Kurdish are the most spoken minority languages in Turkey; namely, about 15 million people speak Kurdish and a half million speak Arabic as a first language. It is estimated that the great majority of them acquire the languages in concern and the official language of the country, Turkish simultaneously. The present study is intended to scrutinize whether Arabic-Turkish bilingual speakers who were born and raised in Turkey significantly differ from Turkish monolinguals regarding their spoken productions in Turkish.

Table 1.   Minority languages spoken in Turkey (Adapted from Lewis et al.)

Status of Language	Language	Location	Speakers (N)
Developing	Abaza	Central; Eskişehir, Samsun, Yozgat, Adana, and Kayseri provinces	10,000 (1995)
	Adyghe	Central and western Anatolia, Kayseri, Tokat, Kahramanmaraş, and many other provinces	278,000 (2000)
	Azerbaijani	South: Kars and Igdır provinces	530,000
	Crimean Tatar	Ankara Province, Polatlı district, Karakuyu, several villages	2,000
	Kabardian	Uzunyayla, east of Kayseri; Samsun area; Amasya; Çorum	1,000,000 (2005)
	Tatar	Istanbul, perhaps elsewhere	Not known
	Uyghur	Kayseri and İstanbul provinces	500 (1981)
	Uzbek, Southern	Hatay, Gaziantep, and Şanlıurfa provinces	1,980 (1982)
	Zazaki, Southern	East-central, Diyarbakir, Elazığ and Bingöl provinces	1,500,000 (1998)
Dispersed	Bulgarian	Scattered in Edirne and other western provinces	300,000 (2001)
	Greek	Istanbul, some in Izmir Province	4,000 (1993)
	Kazakh	Manisa, Istanbul; Kayseri Provinces	600 (1982)
	Kyrgyz	Van and Kars provinces	1,140 (1982)
	Turkmen	Tokat Province	920 (1982)
Dormant	Syriac	Southeast, Şanlıurfa Province	Not known
Educational	Zazaki, Northern	Sivas, Tunceli Bingöl, Erzurum, Erzincan, Elazığ, Malatya Provinces	140,000
Shifting	Balkan Gagauz Turkish	Edirne Province	327,000 (1993)
	Ladino	Mainly Istanbul; some in Izmir Province	10,000 (2007)
Threatened	Abkhaz	Northeast, Artvin Province; also Çoruh, Bolu, and Sakarya subprovince	4,000 (1980)
	Albanian, Tosk	Edirne, Istanbul, Kırklareli, and Tekirdağ provinces: center is Arnavut; otherwise scattered throughout western Turkey	15,000 (1980)
	Armenian	Many in Istanbul; and in east Turkey, Kars Province, scattered elsewhere	40,000 (1980)
	Georgian	North and northwest Anatolia, Artvin, Ordu, Sakarya, and other provinces	40,000 (1980)
	Kumyk	Gümüşhane Province	Not known
	Laz	Northeast, Rize, Kemer, Atin, Artasen, Vitse, Arkab, Hopa, and Sarp; Artvin, Sakarya, Kocaeli, and Bolu provinces	20,000 (2007)
	Serbian	Widespread in the west	20,000 (1980)
	Turoyo	Southeast, Şırnak and Mardin provinces	3,000 (1994)
Vigorous	Arabic, Mesopotamian Spoken	Şanlıurfa, Diyarbakir, Mardin, and Siirt provinces; very small area in Gaziantep province	100,000
	Arabic, North Mesopotamian Spoken	Mardin, Şırnak, Batman, Siirt, and Şanlıurfa provinces	400,000 (1992)
	Domari	Mainly west; widespread	28,500 (1985)
	Hértevin	Southeast, most likely Mardin Province; otherwise scattered	1,000 (1999)
	Pontic	Northeast, Trabzon Province, near southeast Black Sea coast	300,000 (2009)
	Romani, Balkan	West; widespread	25,000
Wider Comm.	Kurdish, Northern	Widespread, especially east and southeast	15,000,000 (2009)

According to Lewis et al. [12], Arabic spoken especially in south-eastern provinces of Turkey is called North Mesopotamian Arabic which is a variety of Arabic mainly spoken in Iraq (7,570,000 speakers in 2014), Syria (300,000 in 1992), Turkey (621,000 users in 2014), and Jordan (500,000). The language in concern is spoken by a total number of 8,691,000 users all over the world. Also known as Mesopotamian Qeltu Arabic, Moslawi, Syro-Mesopotamian Vernacular Arabic, Mesopotamian Qeltu Arabic, and Moslawi, the language is classified as Afro-Asiatic, Semitic, Central, South, and Arabic, and has dialects such as Abdul-Massih, Jesrawi, Mardilli, Mardini, and Mardini Aramaic. It is understood from what Shibatani and Bynon [13] remark on Semitic languages that SVO is followed as the default word order in North Mesopotamian Arabic. In this regard, it seems to differ from Turkish as the latter follows SOV pattern. In return, the languages in question are identical in that both require the possessor to precede the possessed, and the adjective to precede the noun in sentences. Nonetheless, the two languages exhibit morphologically distinct behaviours; namely, North Mesopotamian Arabic allows both suffix and prefix conjugations whereas Turkish is restricted to the former. Another similarity between the two languages is that aspect and tense are indicated by verb conjugation in both. Likewise, the agent of a sentence written in both languages could be indicated in this operation. The following are intended to illustrate the similarities in concern.

أكْتُب or aktub(u)
| Yaz | -(a)r | -(I)m. |
| Write | Present | 1 SG |

I write.

كَتَبْتُ or katabtu
| Yaz | - dı | - m. |
| Write | Past | 1 SG |

I wrote.

يُكْتَب or *yuktab(u)*
| Yaz | -(I)l | -(I)r. |
| Write | Pass. | Present |

It is written.

كُتِبَ or kutiba
| Yaz | -(I)l | -dı. |
| Write | Pass. | Past |

It was written.

كَتَبْتُ or *katabtu*
| Yaz | -dı | -m |
| Write | Past | 1 SG |

I wrote.

كَتَبَ or kataba
| Yaz | -dı | |
| Write | Past, 3 SG | |

He wrote

Taking all these into consideration, it might be expected that the spoken productions of Arabic-Turkish bilingual speakers differ from those of Turkish monolingual speakers in word order; namely, they might be expected to use inverted sentences more frequently than the monolingual group.

The study also investigated the use of conversational fillers which are more likely to be employed by the above-mentioned bilingual group and monolingual group while speaking Turkish. Also known as fillers, gap fillers, linguistic fillers and discourse markers, conversational fillers are defined by Lee [14] as sounds, words, phrases, or even clauses that are 'relatively syntax-independent, do not have a particular grammatical function, do not change the meaning of the utterances and have a somewhat empty meaning" themselves'. In a similar vein, they are described by Bies et al. [15] as conventions in using the term to describe a broad set of vocalized space-fillers which includes filled pauses. Keevallik [16] suggests that one of the main functions of these fillers is to inform recipients that the speaker is going to continue. According to Frățilă [17], they might serve such possible functions as discourse markers contributing to the development of conversation in a particular way, and as interactional signals conveying the speakers' attitudes as well as particular emotions. The most common conversational fillers in Turkish could be listed as *yani* ("meaning..."), *şey* ("thing"), *işte* ("that is"), and *falan* ("as such", "so on"), and while ي *ya* ("means") and شو *wallāh(i)* ("by God") fulfil this position in Arabic.

Kroll et al. [18] propose that although bilinguals rarely make random errors of language when they speak, research on spoken production provides compelling evidence to suggest that both languages are active when only one language is spoken (e.g., Poulisse, 1999).

Shook et al. [19] investigated lexical access in spoken sentence comprehension of English-German bilinguals, German-English bilinguals and English monolinguals. They recorded eye movements of the participants while they were listening to target words in spoken English, and found that bilinguals have a weaker lexical access to spoken words relative to monolinguals.

Studying examined syntactic interference in the spoken productions of Chinese-English bilingual children in a study, Wang [20] reported that the bilingual children used the Chinese filler *Tlifw'* frequently whereas the Chinese monolingual children never used it, which the researcher believes indicates that the bilingual children encounter some expression problems and are thinking about the next words while speaking.

Conducting a study with the participation of first year Hispanic college students in the USA, Méndez-Newman [21] observed that the students used Spanish phatic conversational fillers (e.g., *pues* and *es que*...) while speaking English.

1.2. Research Questions

In accordance with the aim of the current study, a broad research question and its sub-questions were formulated.

Q1. Do monolingual and bilingual speakers of Turkish differ in choice of

a) tense?

b) word order?

c) voice?

d) the use of conversational fillers?

2. Methods

20 Students studying Turkish Language and Literature at a state university in Turkey were the participants of the present study. They were all born and raised in Turkey and reportedly, half of them are Arabic-Turkish bilinguals who acquired Turkish at the age of four or five. At the time of the study their mean age was 20,8. They were divided into two groups as Turkish monolinguals (TMs, hereafter), who spoke Turkish as the only language, and Arabic/ Turkish bilinguals (ATBs), who acquired Turkish as an additional language to Arabic. TMs comprised of five male and five female students while ATBs consisted of 4 male and 6 female students.

In order to see whether they significantly differ in their Turkish spoken productions, the groups in concern were simultaneously shown two animated films in a lecture hall. Prior to this, they were informed that they would not be graded based on their task performance with the purpose of providing them with an anxiety-free atmosphere whereby more reliable data could be elicited. Each film was played twice, and the participants were allowed to take notes while watching them. In a subsequent session, they were individually invited to a room where they were requested to tell the videos in Turkish, and during the session, they were tape-recorded. The recordings in concern were transcribed to construct two sets of corpora: TMC (Turkish Monolingual Corpus), and ATC (Arabic-Turkish Corpus). Finally, the corpora were analysed in order to see whether they significantly differ in terms of tense, word order, voice, and conversational fillers they used during the task. The following section is intended to describe and outline general findings and related discussion on them.

3. Results and Discussion

The study has revealed that sentences produced by TMs outnumbered those produced by ATBs. Table 2 illustrates general results obtained from the analysis of spoken productions of the groups.

Even though the monolingual group seems to slightly differ from the bilingual group in the number of words they used, the two groups displayed a great similarity in the number of different words they produced during the task, confirming Kroll et al. [18], and contradicting the finding

previously reported by Shook et al. [19]. That is, it might be concluded that the bilingual group has as strong lexical access to Turkish as the monolingual group, and that both languages are active in their mind. On the other hand, the number of sentences employed by the former was measured considerably higher than the latter. In return, the average length of sentences produced by the monolingual group was counted slightly lower than those by the bilingual group, which was also approved by the finding that the conjunctions were used more frequently by the bilingual group. Namely, this group of words appeared 114 and 101 times in ATC and TMC, respectively. Not surprisingly, *ve* (meaning and) was the mostly frequented conjunction across both corpora appearing at similar frequencies in each corpus (52 times in TMC; 56 times in ATC). The groups in question significantly differ in the use of adversative conjunctions meaning but (e.g. *ama, ancak* and *fakat*). Indeed, they were used 29 and 26 times by the monolingual group and bilingual group, respectively; however, it was revealed that the bilingual group employed the Arabic originated conjunction *ama* (أمّا ʾammā) approximately as twice often as the monolingual group whereas they never used the conjunction *ancak* which is originally Turkish during the task. In return for this, the monolingual group used the conjunctions in concern approximately with the same frequency. This particular finding is considered a clear indicator of L1 interference.

Table 2. Overall results of lexical analysis of the corpora

	TMC	ATC
Total word count (Corpus size)	2720	2883
Number of different words	489	482
Sentence count	361	298
Average sentence length	8.7	10.83

The groups are similar in the most and least preferred aspects while narrating the stories. Figure 1 depicts the distribution of the tense choice used by the two groups during the task.
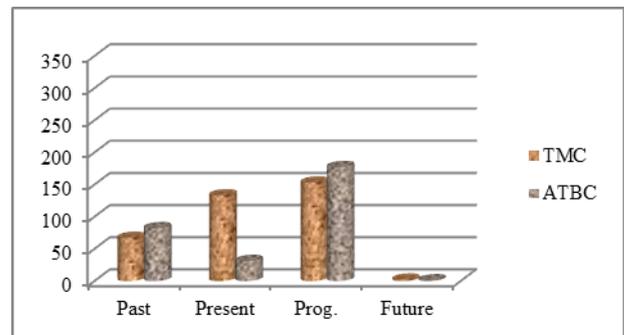
Figure 1. Tense choice across the corpora

As depicted in Figure 1, the two groups used the simple past tense approximately with the same frequency. The progressive tense, which is one of the tenses extensively used in story-telling in Turkish, was used most frequently

by both groups and the future tense was randomly preferred in this regard. The most significant difference between the groups in concern was found in the use of the simple present tense which was another tense commonly employed in Turkish narration; namely, it appeared more than four times as frequently in TMC than ATC. Considering that there is no morphological difference between the conjugation of Arabic verbs in the present progressive tense and the simple present tense (e.g. أكتب /aktubu/ meaning that *I write* and *I am writing*, at the same time), it might be concluded that the bilingual group preferred the former to the latter in the task. The following are the statements taken from the corpora in order to exemplify the use of the three tenses employed by the groups during the task.

Birlikte	kitap	oku	r	lar.
Together	book	read	Present	3PL

They read books together.   (Extracted from TMC)

Sonra	e v	i (n) e	dön	(ü)yor.
Then	house	Poss.Dat.	go back	Prog. 3 SG

Then, he is going back home.   (Extracted from TMC)

Yap	a	ma	dı	lar	o	ev	i.
Build	able	Neg.	Past	3PL	that	house	Acc.

They were unable to build that house. (Ext. from ATC)

Subsequently, the three corpora were examined in terms of word order, and the related distribution was provided in Figure 2.
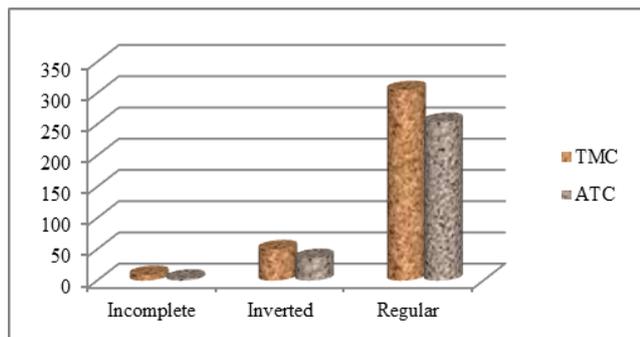
Figure 2. Word order across the corpora

It is seen that the unmarked word order in Turkish language was mostly preferred by both groups. Inverted statements, also allowed in Turkish, were used by the groups with similar frequencies. It should also be noted that incomplete statements were also found in both corpora. Nevertheless, it could not be wise to account this particular finding with the proficiency level of the students in Turkish as they were mostly used by TMs. It might, rather, be attributed to the fact that participants were temporarily distracted or confused while performing the task. If the task in concern were written, for instance, this kind of statements would not probably be found in either corpus. The following extracts were taken from the corpora in question.

Daha sonra	kadın	öl	(ü)yor.
Then	woman	die	Prog., 3 SG

Then, the woman is dead.   (Extracted from TMC)

Ada	da	yalnız	bir adam	var.
Island	Loc.	lonely	a man	There + be / Pres.

There is a lonely man on the island. (Ext. from ATC)

Another expected finding of the study was the frequent use of active voice by the participants in both groups while performing the task. It was not surprising to see that the passive and causative voices hardly appeared in the corpora. The use of the active voice is exemplified in the subsequent statements extracted from the two corpora.

Pelikan	yi	yor	onları
Pelican	eat	Prog., 3 SG	them

The pelican is eating them.   (Extracted from TMC)

Adam	içeri	gir	di.
Man	inside	come	Past, 3 SG.

The man came in.   (Extracted from ATC)

Findings related to the sentence types regarding their predicate have indicated that the groups do not significantly differ in that they extensively came up with verbal sentences rather than nominal ones while narrating the video stories. Nonetheless, it should be noted that verbal sentences occurred more frequently in TMC than the other two bilingual corpora. The subsequent statements were drawn from each corpus to display the use of these sentences by different groups.

İlk	film	beni	çok	etkile	di.
First	film	me	a lot	impress	Past, 3 SG.

The first film impressed me a lot. (Extracted from TMC)

Adam	çok	şaşır	dı.
Man	a lot	get surprised	Past, 3 SG

The man got surprised a lot.   (Extracted from A/TC)

Conversational fillers, defined by Stephen [22] as 'a sound or a word that is spoken in conversation by one participant to signal to others that s/he has paused to think but has not yet finished speaking', were mostly found in A/TC followed by TMC and K/TC, respectively. As Kerslake and Göksel [23] indicated, Turkish speakers use such expressions as *şey* (thing), *valla* (well), *yani* (I mean), *ondan sonra(cığıma)* (after that), and *efendime söyleyeyim* (well/ I mean that..) not to lose their turn in a conversation but has not yet formulated what they are going to say. In the present study, the most common fillers in Turkish were searched across three corpora, and the related distribution is shown in Figure 3.
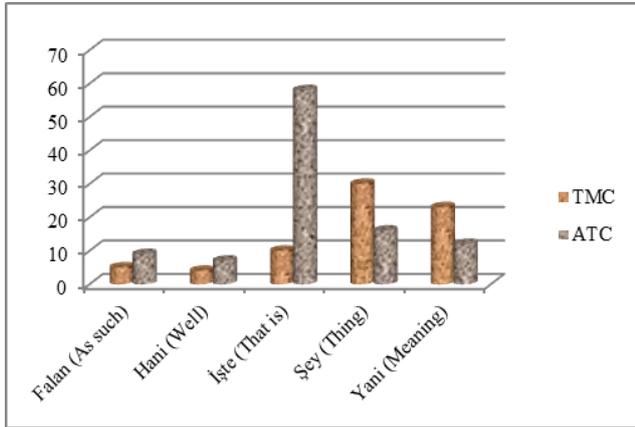
Figure 3.   Conversational fillers across the corpora

The findings have revealed that *falan* (as such) and *hani* (well) were infrequently used by all three groups. It was interesting to see that the Arabic originated *şey* (thing) was the most frequented conversational filler in TMC, and was overused by TMs as opposed to ATBs. Another interesting finding of the study is that *işte* (that is) was mostly found in ATC rather than TMC as it occurred approximately 60 times in the former, and less than ten times in TMC. In order to see whether the groups significantly differ in terms of conversational fillers, a log-likelihood test was administered between ATC and TMC, and the related results are displayed in Table 3.

Table 3.   Log-likelihood results for ATC and TMC

Conv. Fillers	ATC (f)	TMC (f)	LL Ratio (*p< 0.05)
Falan / As such	9	5	-0.10
Hani / Well	7	4	+0.71
İşte / That is	58	10	+33.23
Şey / Thing	16	30	-4.92
Yani / Meaning	12	23	-3.98
Total	102	72	+4.05

n=raw frequency of conversational fillers in corpus

+ indicates overuse in ATC relative to TMC

- indicates underuse in ATC relative to TMC

Table 3 shows that two out of five fillers were underused by ATBs as opposed to the monolingual group. Namely, *işte* (that is) and *hani* (well) were overused by ATBs in comparison to TMs, which was confirmed by log-likelihood results. In return, such fillers as *falan* (as such), *şey* (thing) and *yani* (meaning) were slightly underused by the same group against the monolingual group. It should be noted that the highest difference between the two corpora was measured concerning the use of *işte* (that is), which might be contributed to its intensive use in the speakers' L1. Overall, it seems that the fillers were overused by the bilingual group, which is in line with the finding of the studies previously conducted by Wiese [24] (cited in Khojastehrad [25]), and Wang [20]. It might be suggested, according to Dörnyei and Kormos [26], that they might

have been relatively in more need of stalling, and gaining time in order to keep the communication channel open and maintain discourse at times of difficulty.

5. Conclusions

This study investigated Turkish spoken productions of a Turkish monolingual group and Turkish-Arabic bilingual group in order to see whether they significantly differ with respect to their choice of tense, word order, voice, and conversational fillers. The findings have revealed that they largely differ in the use of the simple present tense while narrating the events. Namely, the bilingual group significantly underused the tense in question as opposed to the monolingual group. Taking into account that they slightly overused the present progressive tense during the task, and that the two tenses have only one counterpart in Arabic, it might be concluded that the group tend to use the present progressive tense which is extensively used in Turkish narration.

The unmarked word order in Turkish (SOV) was predominantly followed by both groups during the task. Besides, inverted sentences, which are also allowed in Turkish, were used slightly more frequently by TMs. Likewise, the active voice proved the mostly preferred voice by both groups. Not surprisingly, verbal sentences were employed more often than nominal sentences by the two groups, and nominal sentences were slightly underused by the bilingual group in comparison to the monolingual one. The study also demonstrated that the groups did not significantly differ regarding word variety in their spoken productions. On the other hand, productions of the monolingual group revealed to be lexically denser and easier to understand when compared to those made by the bilingual group.

Lastly, it has been indicated that the conversational fillers were overused by the bilingual group, indicating that they need more time to think about the upcoming words when compared to the monolingual group since two languages are active in their mind.

As a specific finding of the study, it might be suggested that the Arabic originated conjunction *ama* (meaning 'but') was employed twice as often by the bilingual group while they never preferred its Turkish originated counterpart *ancak*, which might be considered, not surprisingly, as an indicator that they have a stronger access to their first language than the language they acquired at a relatively later age. Nevertheless, in general, the study has shown that the bilingual group does not significantly differ from the monolingual group regarding their choice of word order, and voice, and overused conversational fillers as opposed to the monolingual group.

The present study is limited to the investigation of Turkish spoken productions of a limited number of native monolingual and bilingual undergraduate students attending a state university in Turkey. In order to obtain more generalisable results, further studies could be conducted

with the participation of a larger number of students with different L1 backgrounds. The study might also be furthered to investigate written productions of bilingual speakers from various age groups. Another limitation of the study is that a limited number of videos were presented to the participants to narrate the stories; therefore, more videos might be shown to them, or they might be asked to read a particular short story/ scenario, and to write a report on it in a further study to extend the size of the data to be analysed.

REFERENCES

[1] L. Bloomfield. Language. Allen & Unwin, London, 1935.

[2] R. Titone. Le Bilinguisae Pre'coce Bressels: Dessart. 1972.

[3] Webster's Third New International Dictionary of the English Language. 1961.

[4] J. McNamara. Bilingualism in the Modern World. 23. Journal of Social Issues 1, 1967.

[5] C. Baker. Foundations of bilingual education and bilingualism. Clevedon, England: Multilingual Matters, 1993.

[6] R. Reich. Gestural facilitation of Expressive language in Moderately /severely retarted Pre-schoolers. Mental Retardation, Vol.16., No.2, 113-117, 1986.

[7] A. Woolfolk, M. Hughes, & V. Walkup. Psychology in education. Essex: Pearson Education, 2008.

[8] J. Cummins. Bilingualism and special education: Issues in assessment and pedagogy. Clevedon, UK: Multilingual Matters, 1984.

[9] V. Cook. Background to the L2 user. In V. Cook (Ed.), Portraits of the L2 user. Clevedon: Multilingual Matters, 1-28, 2002.

[10] F. Grosjean. Biligualism's Best Kept Secret. Life as a Bilingual, 2010.

[11] J. Kornfilt. Turkish and the Turkic Languages", in The World's Major Languages (2. Edition); B. Comrie (ed.); London & NY: Routledge; 519-544. (Substantial revision of 1987 original.), 2009.

[12] M. P. Lewis, F. S. Gary, & D. F. Charles (eds.). Ethnologue: Languages of the World, Eighteenth edition. Dallas, Texas: SIL International, 2015. Online version: http://www.ethnologue.com.

[13] M. Shibatani, M. & T. Bynon. Approaches to language typology. Oxford: Clarendon, 1995.

[14] H. Lee. Discourse-marker use in native and non-native English speakers. In C.L. Moder and A. Martinovic-Zic (Ed.), Discourse across languages and cultures. Amsterdam; Philadelphia, Pa, 117-127. John Benjamins, 2004.

[15] A. Bies, S. Strassela, H. Leea, K. Maedaa, S. Kulicka. & Y. Liub. The Report on Linguistic Resources for Speech Parsing, 2006

[16] L. Keavallik. The interactional profile of a placeholder: The Estonian demonstrative see. In N. Amiridze, B. H. Davis, & M. Maclagan (Eds.), Fillers, Pauses and Placeholders (pp. 139-171). John Benjamins Publishing Company: Amsterdam, 2010.

[17] L. Frățilă. On fillers and their possible functions. English Language Overseas Perspectives and Enquiries, Vol.7, No.2, 45-56, 2010.

[18] J. F. Kroll, S. C. Bobb, M. Misra, & T. Guo. Language selection in bilingual speech: Evidence for inhibitory processes. Acta Psychol (Amst), Vol.128, No.3, 416–430, 2008.

[19] A. Shook, M. Godrick, C. Engstler, & V. Marian. Bilinguals Show Weaker Lexical Access During Spoken Sentence Comprehension. Journal of Psycholinguistic Research, 2014.

[20] E. Y. Wang, (2005). Syntactic interference in Chinese-English bilingual children. Unpublished master's thesis, National Taiwan University, Taipei, Taiwan. 2005.

[21] B. Méndez-Newman. The discourse of first year writers at border sites: Discerning the transcultural, bilinguistic strategies of English language learners in college. Rhetoric, Professional Communication, and Globalization, Vol.6, 1-21, 2014.

[22] J. Stephen. Why do we say 'um', 'er', or 'ah' when we hesitate in speaking? Retrieved from http://www.theregister.co.uk/2, 2010.

[23] R. Wiese. Language production in foreign and native languages: Same or different? In H. W. Dechert, D. Mohle, & M. Raupach (Eds.), Second language productions, 50–68. Tubingen: Gunter Narr, 1984.

[24] C. Kerslake, & A. Göksel. Turkish: An Essential Grammar. Routledge, 2011.

[25] S. Khojastehrad. Hesitation strategies in oral L2 test among Iranian students shifted from EFL context to EIL. International Journal of English Linguistics, Vol.2, No.3, 10-21, 2012.

[26] Z. Dörnyei, Z. & J. Kormos. Problem-solving mechanisms in l2 communication: A psycholinguistic perspective. SSLA, Vol.20, 349–385. Cambridge University Press, 1998.