Linguistic Features of English Textese and Digitalk of Iranian EFL Students

Azizeh Chalak
Assistant Professor, Islamic Azad University, Isfahan (Khorasgan) Branch, English Department; azichalak@gmail.com

Abstract
This study aimed at investigating the English textese of Iranian EFL learners by scrutinizing the linguistic features through a qualitative design. In doing so, 700 messages were collected from 43 MA Iranian EFL learners of both genders. The features were categorized and analyzed calculating the frequency and percentage. The findings of the study showed that Iranian EFL students used different types of linguistic features with different density in their textese. They had consistency in textism and their text messaging language was mostly standard. The findings of the study could be used to raise students' awareness on the linguistic features in choosing the right kind of language when communicating with others.
Keywords: Iranian EFL Students, Linguistic Features, Text Messages, Textese

1. Introduction

The development of technology has introduced new ways and modes of communication. It has also made significant variation in written languages. Text messaging has become very popular among the users. Electronic discourse has been taking new dimensions, mainly in the way students write. Digitalk as the new way of e-discourse has its own characteristics. It has a unique format and specific features leading to new language forms and functions. Baron (2008) asserts that the digital venues are blended with elements of both written and spoken discourse. Accordingly, the kind of language that youngsters use for their digital communication is called digitalk (Turner, Abrams, Katíc, & Donovan, 2014) and textese is a sound-based/phonological form of spelling that reduces the time and cost of texting.

There have been many studies conducted on textism, netspeak, the features of textese, and its possible negative effect on students' literacy skills. While some scholars have focused on the sociolinguistics of text messages among users (Elvis, 2009; Thurlow & Brown, 2003), others have investigated the impact of texting on literacy of the younger generations (Plester, Wood & Joshi, 2009). A number of researchers have investigated sociocultural aspects of text messaging (e. g. Ling, 2008), while others have focused on gender differences in texting (Heidari & Alibabaee, 2013, Tafida & Okudishu, 2016). Crystal (2008) conducted a comprehensive study on the linguistic features of online communication and presented how they differ from real life speech.

A thorough study of the related literature showed that many researchers have investigated textese in different languages. To name a few, Thurlow and Brown (2003) in Britain, Ling (2008) in Norway, Elvis (2009) in Nigeria, and Akbari (2013) in Iran studied textese features and characteristics. Some of these studies have investigated linguistic features of text messaging through the analysis of different languages while focusing on abbreviations, communicative functions, and grammatical ellipsis; while others have studied message length, gender differences, integration, and comparison with other technologies. Various methods of data collection including coding frameworks, questionnaires, and interviews have been employed in these stuides.

Different classification schemes have been proposed for codifying linguistic forms in the language of textese (Crystal, 2008; Lyddy, Farina, Hanney, Farrell & O'Neill, 2014; Plester, et al., 2009). In Iran also some studies have focused on issues such as gender differences, or Romanized Persian SMS messages (Akbari, 2013; Heidari and Alibabaee (2013), but the number of the studies on linguistic features of textese is not sufficient. This study explored the linguistic features of textese by Iranian EFL students in their digitalk and posed the following research questions:

1. What are the linguistic features of textese sent by Iranian EFL students?
2. What are the most and least frequently used linguistic features in Iranian EFL students' textese?

2. Method

2.1. Design and Corpus

To conduct the present study, the qualitative research paradigm was employed. The study was conducted in Islamic Azad University (IAU), Isfahan Branch, in spring of 2016. The corpus was collected from Iranian MA EFL students majoring English. The participants were asked to send their Telegram text messages to the researcher and 700 messages were collected through volunteer participation. The reason for choosing Telegram App was because of its popularity and number of users among Iranians (Azali, 2016).

2.2. Participants

The target population was students studying at IAU, Isfahan Branch and the availability sampling procedure was employed. The participants were 43 (25 females and 18 males). Their mother tongue was Persian and their target language was English. They were all the researcher's MA students with the average age of 24. The female students were more than male students due to availability; usually the number of the female students exceed males in English classes.

2.3. Coding Framework

To examine the linguistic features of textese, the recent frameworks (such as AbuSa'aleek, 2015; Crystal, 2008; Lyddy et al., 2014; Plester et al, 2009) were

studied and then, the features common among them were identified. Having a pilot study to examine samples of textese used by Iranian EFL learners, the researcher decided to employ Lyddy et al.'s (2008) coding framework with slight modifications. The categories under the investigation included (a) Accent Stylization, (b) Contractions, (c) Emoticons, (d) Clippings, (e) Initialisms, (f) Letter/Number Homophones, (g) Missed Capitalization, (h) Missed Punctuation, (i) Misspellings, (j) Onomatopoeic/Exclamatory, (k) Phonetic/Nonconventional Spellings (l) Semantically Unrecoverable, (m) Shortenings, and (n) Symbols. This coding framework was used to analyze the linguistic features of Telegram textese.

2.3. Data Collection and Analysis Procedures

To collect the necessary data on textese, first, a request was sent to 80 students. They were asked to forward their messages via Telegram App in private messages. They were asked to forward only natural data, not Persian or Romanized scripts. Volunteer participants took part in the study and they were assured that their information and data will be kept confidential. For ethical concern, names and personal identifications were deleted. To contribute to the issue of reliability, the stability (intra-rater reliability) and the reproducibility (inter-rater reliability) of the coding schemes were established by analyzing and categorizing the data by two raters. The raters were the researcher herself, and one of her experienced colleagues who rated the whole corpus twice. It is worth mentioning that the raters did not work together to come to a consensus about what category they would give to the textese. Then, the inter-rater and intra-rater reliability coefficients were computed using Cronbach's alpha. The inter-rater reliability was found to be high enough ($\alpha = 0.93$) because it was statistically significant at $p \leq 0.000$. Moreover, the intra-rater reliability was quite satisfactory ($\alpha = 0.97$) which was statistically significant at 0.000 level. To analyze the data, after codifying the textese, the obtained findings were tabulated by reporting the frequency and percentage of the occurrence of each linguistic feature.

3. Results

In order to find out the linguistic features used in their textese, the messages were analyzed in terms of descriptive statistics. The data were basically nominal; therefore, frequencies and percentages were calculated. Based on the framework, 700 messages were classified and analyzed. The data showed that the participants used different types of linguistic features with different density. To present the data, the features were classified under five main domains and Table 1 refers to Domain One.

Table 3.1 *Results of Domain One in Iranian EFL Students' Textese*

Linguistic Features	Frequency	Percentage
Clippings	75	4.42
Contractions	38	2.24

Initialisms	19	1.12
Shortenings	57	3.35
Symbols	10	0.60
Total of Domain One	199	11.37

As shown in Table 3.1, less than 12 percent of the total corpus of textese falls into Domain One. Clippings such as (goin, hav, I'l) was the most frequently used feature, followed by Shortenings such as (pic, bro, mon). Contractions such as (bday, abt, txt) were the next category, followed by Initialisms such as (ASAP, LOL). The least frequently used linguistic feature was Symbols such as (@, #, $). It was one of the least frequently used features among all the linguistic features. The results showed that the students did not use Symbols in their text messaging a lot. The findings of the study showed that compressing features were not favoured a lot by the students. This might show that the students preferred to spell out the words in English conventionally.

Table 3.2 *Results of Domain Two in Iranian EFL Students' Textese*

Linguistic Features	Frequency	Percentage
Missed Capitalizations	197	11.61
Missed Punctuations	488	28.75
Misspellings	67	3.94
Phonetic/nonconventional Spellings	35	2.06
Total of Domain Two	787	46.36

Table 3.2 refers to Domain Two, the orthographic features of the textese. The most frequently used feature in this domain was Missed Punctuation followed by Missed Capitalization. The least frequently used feature was Nonconventional Non-Standard Spellings such as (shud for should, no, for know) coined as personal spellings. About four percent of the textese had Misspellings, the typos found in the textese.

Table 3.3 *Results of Domain Three in Iranian EFL Students' Textese*

Linguistic Features	Frequency	Percentage
Emoticons	499	29.30
Letter/number Homophones	167	9.84
Total of Domain Three	666	39.14

In Domain Three, the most frequently used feature, that is, Emoticons could be traced. Emoticons in the corpus referred to Emojis and Stickers used by students in their text messages. Sometimes, there were used as substitutes for words or messages. Two of the frequently used emoticons were (❤ and ☺). The findings indicated that the students used many Emoticons in their textese. Another feature in this category was Letter/Number Homophones such as (U, R, b4) which received almost 10 percent of the whole features.

Table 3.4 *Results of Domain Four in Iranian EFL Students' Textese*

Linguistic Features	Frequency	Percentage
Accent Stylization	1	0.06
Onomatopoetic Exclamatory	8	0.47
Semantically Unrecoverable	38	2.24
Total of Domain Four	47	2.77

Table 3.4 presents miscellaneous linguistic features including Accent Stylizationns such as (wantz, gona), Onomatopoetic Exclamatory such as (WOW! Woohoo), and Semantically Unrecoverable features. The findings showed that the least frequently used feature was Accent Stylization. It seemed that they preferred to use more formal words than the words spelled as they are pronounced in casual speech.

Figure 1 summarizes all the linguistic features employed by Iranian EFL students in their textese. The findings of the study showed that Emoticons was the most frequently used feature. The findings presented in the following figure showed that the variety and density of the features used by the students were not the same.

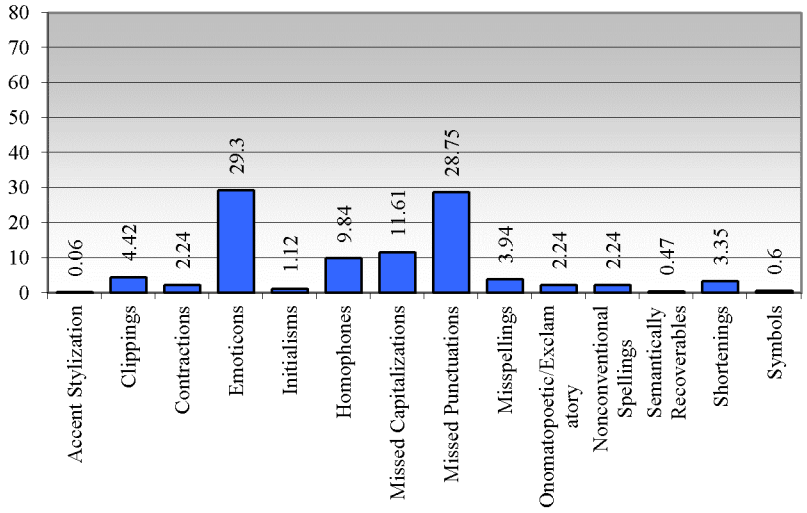

Figure 1 Percentage of linguistic features in Iranian EFL students' textese.

4. Discussion

This study aimed at investigating the linguistic features of Iranian EFL students' textese in order to find out what features were employed by the Iranian English students and to what extent. It also examined which features were the most and least frequently used ones. In doing so, the corpus of the study was collected and analyzed using the adapted framework. The findings of the study showed that Iranian students did not use all the features equally. Some of them were employed

by almost everyone, while others were used only by some students. Some features such as Accent Stylization or Symbols were rarely used. Unlike what the researcher expected, the students did not use Initials, Contractions, and Shortenings a lot while many of them did not follow the conventions of writing such as Capitalization or Punctuations. The findings of this study are not in total agreement with all the other studies. For instance, in Lyddy et al's (2014) study, the most and least frequently used features were Missed Capitalizations and Semantically Unrecoverables, respectively, while in the present study, the most and least frequently used features were Emoticons/Missed Punctuation and Accent Stylization. This might be because of cultural differences, but any generalization needs further investigation.

The study also showed that there was a consistency in textism spelling of the students. In other words, they were using the same features in all their textese. For example, if they were using U (you), 4 (for), or R (are) as Letter/Number Homophones, this feature was present consistently in all of their messages. But, if they were absent in the textese, they were using them rarely. Therefore, it can be concluded that every individual has his/her personal style in writing texts and sometimes every message sender devises his/her own scheme. This finding is in line with Turlow and Brown's (2003) findings.

This study also supports the idea that majority of text messaging language was standard showing that most of the students followed the conventions and norms of writing; this finding is in agreement with the findings of Crystal (2008) and Thurlow and Brown (2003), but further research with different corpus is needed to confirm or contrast the idea. The findings also showed that the students mostly used short sentences or phrases rather than long sentences or paragraphs in each cloud and sometimes these were as short as a word such as Thanks, or letters such as Y (Why), a punctuation mark like (? or !), or even an emoticon or sticker (☺, ♥).

Another interesting finding was that most of the Iranian EFL students participating in the research mentioned that they use Persian or Romanized scripts in most of their text messages. Although they were all English students, some of them even mentioned that they rarely sent English text messages, or at least they used code switching or mixing while sending text messages. Therefore, collecting the necessary data was not easy job because the researcher had to send so many requestive text messages to collect enough data for the analysis.

5. Conclusion

This study was carried out to investigate the English textese of Iranian learners by scrutinizing the linguistic features through a qualitative design. The analysis of the data showed that the students used different types of linguistic features with different density, they had consistent conventions, and most of their text messaging language was standard. The most and least frequently used features were Emoticons and Accent Stylization, respectively.

The study was carried out with a certain group of participants and corpus. Therefore, similar to other studies, it might have some limitations which could be overcome through further research. For example, a gender-based study with larger corpus considering both genders could be beneficial. Studying other applications such as Viber, Whatsapp or SMS and comparing them with the findings of this study might be an interesting topic.

Understanding how young adults are texting in their everyday communication could have implications for range of issues both at micro level, that is, the daily relationships with others in their lives, and at macro level, for policy makers and decision makers at school and academic settings. To conclude, as Thurlow and Bell (2009) have mentioned, because the technologies of texting are constantly changing, the practices and meanings of texting are also changing. As a result, any research on texting should be constantly updated.

References

AbuSa'aleek, A. O. (2015). Internet linguistics: A linguistic analysis of electronic discourse as a new variety of language International. *Journal of English Linguistics, 5*(1), 135-145.

Akbari, M. (2013). A preliminary linguistic analysis of Romanized Persian SMS messages. *Journal of Novel Applied Sciences, 2*(8), 197-205. Retrieved from https://search.ricest.ac.ir/Ricest/show12.aspx?Doc2310056

Azali, M. (2016). The impact of telegram on Iran. Retrieved from http://techrasa.com/2016/01/13/ impact-telegram-iran/

Baron, N. S. (2008). *Always on: Language in an online and mobile world*. New York, NY: Oxford University Press.

Crystal, D. (2008). *Txting: The gr8 db8*. Oxford: Oxford University Press.

Elvis, F. W. (2009). The sociolinguistics of mobile phone SMS usage in Cameroon and Nigeria. *The Internationa*l *Journal of Language Society and Culture, 28*(28), 25-41.

Heidari, M., & Alibabaee, A. (2013). Linguistic and discoursal features of text message language created by Iranian male and female SMS users. *Sheikhbahaee EFL Journal, 2*(1), 55-72.

Ling, R. (2008). *New tech, new ties: How mobile communication is reshaping social cohesion*. London: MIT Press.

Lyddy, F., Farina, F., Hanney, J., Farrell, L., & O'Neill, N. K. (2014). An analysis of language in university students' text messages. *Journal of Computer-Mediated Communication, 19*(3), 546-561. Retrieved from http://dx.doi.org/10.1111/jcc4.12045

Plester, B., Wood, C., & Joshi, P. (2009). Exploring the relationship between children's knowledge of text message abbreviations and school literacy outcomes. *British Journal of Developmental Psychology, 27*, 145–61.

Tafida, A. G., & Okudishu, C. O. (2016). Gender difference in text messaging in the written English of one hundred level undergraduates of Ibrahim Babangida University, Lapai, Niger State. *Palgo Journal of Education Research, 4*(2), 170-175. Retrieved from http://www.palgojournals.org/PJER/Index.htm

Thurlow, C., & Brown, A. (2003). Generation txt? The sociolinguistics of young people's text messaging. *Discourse Analysis Online, 1*(1), 1-27 *London*. Retrieved from http://www.faculty.washington.edu/thurlow/papers/chapter.pdf.

Thurlow, C., & Bell, K. (2009). Against technologization: Young people's new media discourse as creative cultural practice. *Journal of Computer-Mediated Communication, 14*, 1038–1049.

Turner, K. H., Abrams, S., Katic, E., & Donovan, M. J. (2014). Demystifying digitalk: The what and why of the language teens use in digital writing. *Journal of Literacy Research, 46*(2), 157-193.