Arabizi: An Analysis of the Romanization of the Arabic Script from a Sociolinguistic Perspective

Wid H. Allehaiby
King Saud bin Abdulaziz University for Health Sciences
Riyadh, Saudi Arabia

Abstract:
The purpose of this paper is to present a sociolinguistic analysis of the Romanization of the Arabic Script phenomenon better known as Arabizi. Arabizi is a widely used alternative for the Arabic script in computer-mediated contexts (CMC) and social networking sites, which has emerged as a result of the Arabic script being unsupported in technological tools and Internet resources. Arabic speakers have taken different stances to using Arabizi. In this paper, the following issues are discussed: the historical emergence of Arabizi, a thorough explanation of the script, its characters and its features, the social contexts in which users utilize it, and the different attitudes of Arabic speakers towards its use.

Keywords: Arabic Script, Arabizi, Romanization, sociolinguistics, speakers’ attitudes
Globalization has struck the world like a storm as the result of the spread of technology. During the 90s, most Arab countries, if not all, witnessed an increasing importance of the English language, which rose to dominate many technological devices and realms, such as online chats, short message service (SMS), and mobile phones (Warschauer, Said, & Zohry, 2002). During that time, non-Latin scripts were not supported on many popular devices and the only means to communicate effectively with others was through the use of the English language and/or Latin script. Arabic was not unaffected by this situation. In fact, Arabic speakers were forced to develop a means of communicating in Arabic using the Latin script. The Romanization of Arabic has become widely spread among Arab youth across many Arab countries and this use of the Latin alphabet, as opposed to the Arabic, is known as Arabizi (Bianchi, 2012; Sperrazza, 2011; Yaghan, 2008), Latinized Arabic (Aboelezz, 2009), ASCII-ization (Palfreyman & Khalil, 2003), and Arabic Chat Alphabet (ACA) (Elmahdy, Gruhn, Abdennadher, & Minker, 2011; Mostafa, 2012).

According to Yaghan (2008) and Attwa (2012), the word Arabizi originated from blending the words “Arabic” and “Inglizee” (which is the Arabic name for English). As mentioned above, this phenomenon is believed to have been developed in response to the prevalence of western technology, namely Internet Relay Chat (IRC), text messaging (SMS) and emails, all of which initially required the use of the Latin alphabet. Originally, the majority of these devices or software applications did not support Arabic script; thus, users and consumers of such products had to find a way to communicate in Arabic through the use of Latin script (the only supported keyboard layout available in the Arab world at the time) (Elmahdy et al., 2011). With that, the Arabizi form of writing was invented. Although technical devices and the Internet have since been updated to support non-Latin script, including Arabic, Arabic speakers continue to use Arabizi (Attwa, 2012) where English is seen as the dominant language of cyber space (Aboelezz, 2009). Al-Khatib and Sabbah (2008) make the following observation:

It is worth mentioning that this form of Romanized Arabic did not exist before the advent of the Internet, and it can be noticed that there is a heavy use of this new form of written communication among students to such a degree that the traditional way of writing Arabic is counted out. (p. 46)

The Arabizi Phenomenon

Arabizi is defined as an encoding system that uses the Latin script and Arabic numbers instead of Arabic letters. Each English letter represents an Arabic phoneme that matches it in pronunciation, whereas the Arabic numerals compensate for Arabic phonemes that are non-existent in the English language, but resemble Arabic letters and their shapes. In some cases, accent marks can also be used to refer to an Arabic phoneme along with the Arabic numerals. Bianchi (2012) adds to this definition the fact that Arabizi is mostly used in Computer-mediated Communication (CMC) contexts and refers to the Arabic numerals that are used as “arithmographemes,” while Yaghan (2008) labels this phenomenon as “slang.” The term also refers to the process of code-switching between Arabic and English during conversations or while speaking (Attwa, 2012; Sperrazza, 2011). To further explain this phenomenon, the author has taken a screen capture of a text message conversation between two Native Arabic speakers (see Figure 1).
As can be seen in Figure 1, the interlocutors use Arabizi to carry out a normal daily life conversation (talking about plans for the weekend). Both Arabic and English words are used, but all are in Latin script. The use of Arabic numerals is also evident. For example, one interlocutor uses the number 7 in place of an Arabic phoneme ح /ħ/ (devoiced pharyngeal fricative) in the word “Al7amdellah,” “الحمد لله” ‘Thank God’. As is apparent, the Arabic letter resembles the corresponding Arabic number as far as its appearance is concerned. In Arabizi, ten Arabic letters are replaced with Arabic numerals, and in some cases, an apostrophe is added (see Table 1). Because Arabizi is a non-standardized form of writing, Arabic speakers inconsistently use numbers instead of letters, or vice versa, and this varies according to familiarity and/or personal preference (Sperrazza, 2011).

Table 1. Arabizi Characters

<table>
<thead>
<tr>
<th>Phonetic Description</th>
<th>Arabic Letter</th>
<th>Romanization Possibilities</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced glottal stop /ʔ/</td>
<td>ء</td>
<td>2</td>
<td>Bala2 (disaster)</td>
</tr>
<tr>
<td>Devoiced pharyngeal fricative /ħ/</td>
<td>ح</td>
<td>7</td>
<td>7alwa/ halwa</td>
</tr>
</tbody>
</table>
Another general feature of Arabizi is apparent in the use of vowels. As indicated by Yaghen (2008), vowels are seen as optional and in some cases can be omitted according to the speakers’ background knowledge, contextual lucidity vocabulary items, and the number of characters permitted in one SMS text message. He also noted that Arabizi is highly dependent on context and, given the fact that Arabizi is mostly used in informal situations, is very similar to CMC language, which allows for the use of abbreviations. Lastly, Yaghan notes that, given that there are numerous Arabic dialects, the use of Arabizi not only varies from one Arab country to another, but will also differ from one group to another within the same country (Muhammed, Farrag, Elshamly, & Abdel-Ghaffar, 2011).

A Call for the Romanization of the Arabic Script

During the 1800s, a very extreme call for reforming the Arabic writing system through the Romanization of Arabic script was led by two directors of the Egyptian National Liberty: Wilhelm Spitta in 1880 and Karl Vollars in 1890. However, their proposal was heavily criticized and subsequently rejected (Sakkal, 2008). Then, later in the 1900s, another strong proponent for the Romanization of the Arabic script, Salama Mosa, rekindled the argument that the Latin script should be used, arguing that it is more efficient and effective than the Arabic script (as cited in Széll, 2011). The fact that he was a Coptic Christian led him to believe that Arabic was strongly associated with Islam, and this made him feel a detachment from the Arabic language. He strongly advocated for a movement that sought to convince native Arabic speakers that the Arabic language was incapable of being used in scientific progress as well as the English language, although he completely discounted the significant scientific achievements of Arabic speaking scholars and scientists that had been made during the Islamic Golden Age.

The Arabic language contains 28 phonemes with six vowels sounds (three short and three long vowels) (/a/, /i/, /u/, /a:/, /i:/, /u:/). Apart from the obvious fact that the direction of writing and reading Arabic is from right to left, one of the real complications proposed by pro-Romanization supporters of the Arabic script lies in the fact that the typology of the letters changes according to their position in the word (alone, initial, middle, final). For example, the letter (س) (seen) can be written as (سکر) (sugar) which is in the initial position, (مسمک) (musk) in
the middle position, and (ناس) (people) in the final position. Furthermore, Arabic letters have a vast number of varieties, which range from 600 to 900 forms (Muhammed et al., 2011).

Another significant issue concerns the absence of Arabic short vowels, which are known as diacritics, from most handwritten and printed Arabic texts. This can become troublesome for readers who lack the appropriate grammatical knowledge of the Arabic language. In addition, the similarities in the shapes of the letters and the need to distinguish between them using dots (e.g., د/ث) are seen as another major concern of critics of the Arabic alphabet.

Taking into consideration all these issues, it is safe to say that the previous claims are extremely exaggerated. In fact, Széll (2011) points out that there has been no evidence that the Arabic language, specifically its alphabet, is more challenging for language learners of Arabic as a foreign language (AFL) than any other language. Similarly, to date, no evidence has been provided that demonstrates the effectiveness of using the Latin alphabet in transcribing Arabic letters. Hence, all efforts suggesting the Romanization of Arabic were relentlessly contested and condemned by Arab nationalists and Muslim enthusiasts (Yaghan, 2008). To them, such proposals were offensive to the Arab identity and language, especially given the fact that the writing system in question has been around for over fourteen centuries.

According to Yaghan (2008), all attempts to replace the Arabic script with the Latin alphabet failed, except in situations where the Arabic or non-Latin script was not supported or unavailable due to technical restrictions, such as on mobile phones; comprehensibility reasons, such as texts meant for non-Arabic speakers; and economical reasons, such as in the case of mobile carriers that charged more for Arabic characters than for Latin ones.

Széll (2011) states that Arabizi is the most, and probably only, successful Romanization system available to date that is widely in use in daily life interactions. However, it has a limited use in specific contexts, namely media and registers. In fact, Abolezz (2012) argues that it is mostly spoken, not written, Arabic that has been Latinized in CMC contexts; and she clearly notes in her 2009 article that “the Arabic we write is not the same as the Arabic we speak” (p.103). Hence, it is thought to be an inadequate form of what Spitta, Vollars and Mosa proposed: a standardized written variety of a Romanized Arabic script that was suitable for all and every context (as cited in Széll, 2011).

The Utilization of Arabizi: The Appeal

As an advocate for Arabizi, Lelania Sperrazza (2009) is eager for her native Arabic Egyptian speaking students to embrace a dual cultural and linguistic identity. She states that pushing her students to become aware of their Arabizi identity has had a positive effect on their learning process, although it is unclear which language has played the most important role in their learning process. In an argument against those who resent the idea of using Arabizi, Sperrazza states the following:

[R]ather than desiring a monocultural identity for an entire lifetime,[…], I want my students to be aware that they are products of a new era, and subsequently, a new identity: one that is increasingly migratory and globalized, and one that is constantly in the process of adapting and reinventing itself. This identity, fueled by the power of an English-based technology, is the Arabizi identity. Therefore, I believe it is necessary for my students to recognize that they are members of a legitimate and valid community, which has emerged from the globalized and interdependent relationship between the East and the West. (p. N/A)

In a similar study, Muhammed et al. (2011) conducted a survey to observe Arab
Facebook users’ perceptions of the use of Arabizi. Out of 70 Native Egyptians, 60% confirmed that using Arabizi had no impact on their Arab identity, although some reported that it might have partially affected those of younger generations, as it appears to have had a negative impact on their use of the Arabic language and resulted in them becoming more westernized. As far as belonging to a speech community, Muhammed et al. report that 21% of the participants who regularly use Arabizi felt that they successfully established a bond with other users of this language system. Other respondents conveyed the urge to become “trendy” and “follow the flow” of other users, a phenomenon that Bassiouney (2009) describes as a sociolinguistic method to accustom oneself to the norms of a speech community.

According to Yaghan (2008), another reason that contributes to the Arabic youth’s use of Arabizi can be attributed to the historical precedence of English over Arabic in both the Internet domain and mobile phone industry. This may have caused Arabic speakers to think that the Arabic language is inadequate for use in the technical world and that it is not technologically friendly (Muhammed et al., 2011). Instead, some researchers see the benefit of Arab user’s ability to use code switching back and forth from Arabic to English as an opportunity for them to practice both languages and, as such, Arabizi is beneficial (Yaghan, 2008).

Several studies have also reported that Arabizi is easier and faster to use; that is, the English letters present on a mobile phone’s keyboard are fewer than the Arabic letters; therefore, it is considered to be less time-consuming to use this system (Al-Khatib & Sabbah, 2008; Attwa, 2012; Muhammed et al., 2011). Thus, as stated by Yaghan (2008), users of Arabizi believe that learning to use one language (English) is much more convenient when applied to both languages as it is far less confusing. This can also have an economic value in the sense that mobile carriers will charge less if SMS text messages contain fewer characters, which is hard for users to achieve when using Arabic letters (Al-Khatib & Sabbah, 2008; Yaghan, 2008).

Another economical benefit was identified in Aboelezz’s study (2012), which investigated the use of Arabizi in four major published magazines. According to Aboelezz, the study was motivated by the researcher noticing the conspicuous presence of Arabizi in various business and marketing instruments, such as billboards, consumer products, and in printed media (magazines). This marked the transition of Arabizi from CMC online mediums to offline means, thus moving from unregulated to regulated spaces (Sebba, 2002). It is worth mentioning that these magazines started publishing some of their content in Arabizi in the last decade due to the young staff and the fact that the publications were aimed at young readers. This suggests that Arabizi is popular and has achieved general acceptance among members of at least the younger public.

**Arabizi from a Sociolinguistic Dimension**

The use of Arabizi is very common among young Arabic speakers, while it is often rejected by older generations, working seniors, and people of higher class (Palfreyman, & Khalil, 2003; Muhammed et al., 2011). Therefore, within the younger generation, Arabizi is perceived to be up to date, trendy, and “cool”. The youth have established the idea that Arabic script should be confined to the use of classical Arabic, formal and religious contexts and, hence, is inappropriate for the use of slang (Yaghan, 2008; Muhammed et al., 2011).

Arabizi users have also conveyed the fact that this language system allows them to express themselves in a variety of ways that would otherwise be impossible when using Arabic script. According to Yaghan (2008), they report that Arabizi can support both uppercase and
lowercase letters. To them, the benefit of that lies in the fact that they can display a myriad of emotions that are otherwise unmanageable in classical Arabic, such as shouting and calmness. Thus, messages can be labeled as being clearer and easier to understand and convey (Muhammed et al., 2011).

In an attempt to identify the conspicuous sociolinguistics aspects of Arabizi, Al-Khatib and Sabbah (2008) conducted a survey that aimed to investigate the purposes and uses of Arabizi, along with the preferred language choice for SMS text messaging among Jordanian university students. The study included a total of 46 native Arabic-speaking university students of whom 17 were male and 29 female.

The results showed that 61% of the respondents used Arabizi, where 39% used it by word/phrase, meaning that they chose to write words and phrases in Arabizi as opposed to English vocabulary (54%) and Arabic vocabulary in Arabic letters (9%). The authors justified this finding by stating that English is the medium of instruction in the chosen universities and, as such, has become a fundamental means of communication for students in the professional field. He also noted that females (44%) have a stronger inclination than males to switch to English instead of Arabic or Arabizi for purposes such as prestige, euphemism, and as an indication of educational achievements.

Al-Khatib and Sabbah categorized the language choices based on the following social contexts or registers:

1. Socio-cultural and religious contexts: It was found that participants choose to write SMS text messages in Arabizi when discussing cultural or religious matters. In addition, they prefer to use this variety when exchanging wishes related to religious holidays and other social occasions that include engagements, weddings, and graduation. Such language choice is related to the fact that expressions related to these events are so culturally bound that it is not possible to identify an appropriate term or statement in the English language.

2. Greetings: Once again, participants’ responses recorded their use of Arabizi when greeting each other via text messages. When asked about using English vocabulary equivalents, some participants voiced the opinion that the Arabic vocabulary portrays more warmth and is more appropriate. Others explained that this choice showcased their regard for their religion and culture.

3. Quoting: It was observed that participants resorted to Arabizi when quoting someone who used the Arabic script in their text messages. There were also instances where users used Arabizi for direct and indirect quotes, and used either the Arabic script or Arabizi when quoting an Arabic proverb or a well-known maxim. This example was provided by Al-Khatib and Sabbah:

A: CCT الاخت الفاضلة .. الرجاء إحضار كتاب صقر معك غدا [literally (virtuous sister), please bring Saqir’s book with you tomorrow]

B: 😊this z the 1st time someone calls me “AL2O5T ALFADILAH”...😊lol. Anyway, don’t worry, I wont 4get 2 bring the book 2mr. Take care.

In another study by Bianchi (2012), who analyzed 460,220 forum posts on a Jordanian Arabic website to measure the occurrence of Arabizi and the domains in which it occurred, the following results were found as displayed in Table 2.
As can be seen from Table 2, overall Arabizi is the most frequently used variety of language in these posts. However, users seem to prefer to use either the Arabic script or the English language when discussing certain topics. The prevalence of Arabic script is evident in topics that relate to poetry and humor, and local culture. This suggests that these topics are best expressed in the Arabic language given the fact that they are closely related to cultural and religious topics, which are not expressed as effectively in either Arabizi or the English language.

On the contrary, topics dealing with work/ school matters were mostly discussed in the English language. This can possibly be attributed to the fact that English is the medium of instruction in Jordan and the existence of jargons that have no equivalent in the Arabic language. Finally, the dominance of Arabizi can be found in topics relating to family/ friends, gender/ age, hobbies, and general topics. This may suggest that users prefer to utilize the informal nature of Arabizi in order to eliminate social gaps and establish intimacy with one another.

### Table 2. Adapted from Bianchi 2012 Results (some categories omitted)

<table>
<thead>
<tr>
<th>Topics</th>
<th>Arabizi</th>
<th>Arabic Script</th>
<th>English Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Topics</td>
<td>39.6%</td>
<td>22.3%</td>
<td>22.1%</td>
</tr>
<tr>
<td>Humor</td>
<td>19.5%</td>
<td>66.7%</td>
<td>4.3%</td>
</tr>
<tr>
<td>Gender/ Age</td>
<td>48.8%</td>
<td>12.4%</td>
<td>23.5%</td>
</tr>
<tr>
<td>Hobbies</td>
<td>44.7%</td>
<td>8.7%</td>
<td>25.7%</td>
</tr>
<tr>
<td>Local Culture (Nationalistic and Religious)</td>
<td>25.6%</td>
<td>39.5%</td>
<td>16.5%</td>
</tr>
<tr>
<td>Poetry</td>
<td>10.3%</td>
<td>70.6%</td>
<td>7.3%</td>
</tr>
<tr>
<td>Family/ Friends</td>
<td>49.4%</td>
<td>20.4%</td>
<td>15.9%</td>
</tr>
<tr>
<td>Work/ School</td>
<td>28.4%</td>
<td>8.9%</td>
<td>36.4%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>35.5%</strong></td>
<td><strong>32.3%</strong></td>
<td><strong>17.5%</strong></td>
</tr>
</tbody>
</table>

Resentment towards Arabizi: Anti-Arabizi Arguments

The relationship between the Arabic language and native Arabic speakers is very strong and closely tied with the Arab identity. Barakat (1993) gives a vivid description of the sense of nationalism Arab nationalists possess as native speakers of the Arabic language:

It has often been stated that the great majority of Arabs speak Arabic as their mother tongue and thus feel that they belong to the same nation regardless of race, religion, tribe, or region. This explains the tendency to dismiss the existing states as artificial and to call for political unity coinciding with linguistic identity. [...] There is, in fact, unanimous agreement among theoreticians of Arab nationalism on the great significance of language. (pp. 33-34)
Such a close tie has generated a severe backlash against the call for the Romanization of the Arabic language by some philosophers and Arab linguists while many people resent the use of Arabizi among native Arabic speakers. Yaghan (2008) explains that some view the Latin alphabet as representing colonial powers, ones that Arabs long resisted and fought against until the declaration of independence from the grasp of European colonialism. Others view the use of Arabizi as “culture treason” in which discarding the Arabic alphabet would symbolize letting go of the historical and cultural ties that the Arabic language represents.

Similarly, Muhammed et al. (2011) report in their findings that some speakers refuse to use Arabizi out of respect for the Arabic language as the language of Islam (namely Quran). For them, it is the Arabic language that preserves their identity as Arabs. These same respondents also explain their reason not to use Arabizi because they believe it is confusing, since it is not standardized. They believe that computer software and mobile phones now support the Arabic script and, as such, it is no longer necessary to use Arabizi.

The fact that Arabizi is mostly used in informal settings has also been cited as a strong criticism against its use. Many users argue that the use of Arabizi is inappropriate and highly offensive if used in the place of the Arabic script in some contexts such as formal social encounters and religious practices. Thus, Muhammed et al. (2011) argue that this language standard fails to provide a conclusive solution to the problems associated with the Arabic writing system.

Arabizi may also hinder the learning process of learners of Arabic as a foreign language (AFL), especially when they try to communicate with native Arabic speakers via CMC (Attwa, 2012). These learners have reported that Arabizi contributed to mistakes in their pronunciation, hinders their acquisition of the important language skills of reading and writing, creates confusion because of its inconsistent and unstructured nature, and deprives them of the opportunity to practice the Arabic script with native Arabic speakers.

Conclusion

The globalization of English-based technology has caused the present-day Arab generation to rely on the use of the Latin script as the main writing system when communicating through electronic medium. This indeed has helped with mixing Arabic and English in such a way that has given birth to the Arabizi phenomenon. As Attwa (2012) describes, Arabizi is becoming a trend that cannot be overlooked, especially in the realm of CMC. Echoing the finding of Sperrazza’s study (2011), Arabizi users should not only explore its use, but also encompass and embrace this variety. As such, it could be perceived as an important method through which Arabs can solidify and construct their identity through what Bourdieu (1977) refers to as “legitimate speakers,” without harming their Arab identity, which is potentially inextricable from the Arabic language and its use.

It is clear that the use of Arabizi has become evident in many, if not all, Arab countries and that the youth generation in particular relies upon it (Attwa, 2012). Sperrazza’s (2011) observations suggest that the use of Arabizi serves to endorse the legitimacy of Arabizi speakers, individuals who have obtained a continually expanding, very tangible cultural-linguistic platform in the community. Furthermore, Sperrazza notes that it serves to acknowledge the linguistic and cultural contributions that these individuals have made, since they are the modern creators of a “globalized world that has helped to shape them in the first place.” (p. 102)

However, teachers should take into consideration Attwa’s (2012) caution that the Arabizi phenomenon may cause confusion and even difficulty for AFL learners. Attwa also warns that
the use of such writing system could lead to the potentially undesirable phenomenon of digraphia: the use of more than one writing system for the same language. In addition, Attwa stresses that, while it is important for teachers to introduce Arabizi to AFL learners, it is not necessary for them to teach the actual writing variety itself. Finally, Attwa claims that Arabizi will eventually acquire a standardized or prestigious status among Arabs. This could be potentially detrimental, as teachers and learners of AFL would be faced with a dilemma: a writing code that learners want to learn will be in existence, but nobody will be appropriately equipped or even willing to teach it. Therefore, Széll (2011) and Attwa (2012) suggest the avoidance of using Arabizi altogether by urging language planners and CMC developers to integrate the Arabic script into platforms and systems so as to encourage the use of the Arabic keyboard as opposed to the English one.

Whether Arabizi is looked at as a positive or negative phenomenon, its use has become visible and widespread throughout many Arab countries. Based on Yaghan’s (2008) reflection that Arabizi is the most successful form of Romanization of the Arabic script up to date, in accordance with the current situation in the real world, it can be inferred that Arabizi users will continue to use this variety of Arabic, despite the fact that the Arabic script is now supported on the majority of advance mobile devices and computer software. This strong tendency to use Arabizi ensures that this variety of language will continue to be utilized by native Arabic speakers and might pave the way for its acceptance among wider social groups and ages.

About the Author:
Wid Allehaiby has graduated in 2013 with a Master’s of Arts degree in Teaching English to Speakers of Other Language from the University of Northern Iowa in US. She has worked as an English Teaching Assistant at King Saud bin Abdulaziz University for Health Sciences ( KSAU-HS), Saudi Arabia. She is now a Lecturer at KSAU-HS and is a recipient of King Abdullah Scholarship Program to pursue a Ph.D. degree in Applied Linguistics in the US.

References


