The Effect of Latinization on Reading Time and Understanding: Greeklish in Communication and Social Media

Evangelos Kehris¹, George Karavasilis¹, Vasiliki Vrana¹, Dimitrios Kydros²

¹Department of Business Administration, International Hellenic University, Serres, Greece
²Department of Economics, International Hellenic University, Serres, Greece

Email: kehris@ihu.gr, gkaravas@ee.duth.gr, vrana@ihu.gr, dkydros@ihu.gr

Abstract

The increasing prevalence of technology in society has an impact on young people’s language use and development. Greeklish is the writing of Greek texts using the Latin instead of the Greek alphabet, a practice known as Latinization, also employed for many non-Latin alphabet languages. The primary aim of this research is to evaluate the effect of Greeklish on reading time. A sample of 732 young Greeks were asked about their habits when communicating through e-mail and social media with their friends and they then participated in an experiment in which they were asked to read and understand two short texts, one written in Greek and the other in Greeklish. The findings of the research show that nearly one third of the participants use Greeklish. The results of the experiment conducted reveal that understanding is not affected by the alphabet used but reading Greeklish is significantly more time consuming than reading Greek independently of the sex and the familiarity of the participants with Greeklish. The findings suggest that amending social and communication media with software utilities related to Latinization such as language identifiers and converters may reduce reading time and thus facilitate written communication among the users.

Keywords

Social Media, Digital Communication, Latinization, Reading Speed, Text Understanding, Greeklish

1. Introduction

Latinization, the transliteration of native non-Latin scripts into Latin, is a digital practice adopted by many users initially as a response to the limitations and re-
strictions of early Information and Communication Technologies to address local linguistic realities and conventions [1] [2] [3]. The practice gained popularity and it is still in use, despite the fact that the technological limitations and restrictions of the past have been addressed, mainly because users find it fast and/or convenient.

Latinization in computer-mediated communication activities (e.g. emails, forums, online chats, game interaction, instant messaging, social media) and associated areas like SMS messaging (texting) has been acknowledged in many languages, such as Arabic [1] [3], Chinese [4] [5], Cypriot-Greek [6], Greek [7] [8] [9], Japanese [10], Korean [11] [12], Punjabi [13], Sebian [14], Singapore Creole [15] and with the interference of languages and word loans [16] new linguistic phenomena, referred as 3arabizi, Chinglish, Greeklish, Japlish and Singlish appeared [17] [18] [19]. These linguistic phenomena have impact on cultural mind-set of the populace and hold nowadays great sociolinguistic influence over these nations [16]. The phenomenon is particularly popular among the younger generations and as Harada [20] comments on the usage of Singlish: “Singlish is the preferred language of interaction among students”.

The phenomenon linguistically identified as synchronic digraphia [21] is present in a variety of languages that do not adopt the Latin script [19] and refers to the use of two writing systems that coexist for the same language [21]. Androutsopoulos [22] coined the term “computer-mediated digraphia” for the simultaneous use of both the native Greek and the Latin script in computer-mediated interaction and Crystal [17] made reference to a new phenomenon the “Internet linguistics”.

Throughout the 1980s latinization was the only option available to the Greek-speaking internet users [22] [23], as the basic ASCII code, employed at that time for the communication and composition of digital texts, did not support the use of non-Latin alphabetical characters [24]. Transliterating Greek with Latin characters was also employed in mobile communication as mobile phones of the time supported English by default and the Greek language was not always available as an option [9] [25]. Thus, “Greeklish”—the written representation of Greek using Latin characters in computer-mediated environments—emerged. According to Blanas [26], Greeklish was “the de facto standard for electronic communication”. Especially Internet users under 35 that considered Greeklish a necessary evil when using the Internet [7]. During 1990s the development of Unicode, which supported the Greek writing system, provided the technical capability to overcome the constraints imposed by the ASCI code and increased the use of Greek script in the computer-mediated communication environments [8] [22] [23]. However, even nowadays, despite the widespread and predominant use of Unicode and technological developments that have relaxed the conditions necessitated Latinization, Greeklish is still widely used [22] and is a choice of many users, fulfilling their specific communication needs [27].

In Greece, the number of internet and social media users is steadily rising,
Greece’s internet penetration rate stood at 82.2% of the total population [28] at the start of 2022 from 69.1% at the start of 2016 [29]. At the beginning of 2022, the number of social media users in Greece was 71.5 percent of the total population [28]. Koutsogiannis [30], who studied Greeklish in the context of a globalized environment, claimed that the Latin alphabet is used to a greater extent than the Greek, especially by children, who are familiar with digital environments and the English language.

The adoption of Greeklish in digital environments has attracted some attention as Koutsogiannis & Mitsikopoulou ([8], p.17) claimed “attitudes towards the use of Greeklish are deeply embedded in the Greek sociocultural context”. However, research related to Greeklish is rather limited [30] and mainly focuses on software converters, transliteration practices and sociolinguistic issues [31] while research on the use and understanding of Greeklish is also quite limited. Androutsopoulos [32] referring to the users who employ Greeklish wondered “…but how many are they and exactly how they write? Nobody knows”. To our knowledge phenomena like Greeklish, up to now have been approached from a linguistic point of view or from a philosophical perspective. The paper adopts a different approach and aims to address a number of questions related to their use by young people.

The aims of this research are to:

• to measure the extent to which Greeklish is used by young people when communicating through internet and social media;
• to evaluate the effect that Greeklish have on reading time of young Greek speakers;
• to explore if understanding is affected by the alphabet (Greek vs Greeklish).

This work studies the reading speed of computer users when reading Greek texts written in Greek and in Greeklish with the aim of understanding them. Text understanding is strongly related to reading speed [33]. The evaluation of the effect of Greeklish on reading time has both theoretical and practical implications: On the theoretical level, such an evaluation may 1) provide evidence that texts in Greeklish impose an extra conceptual load on the reader compared to texts in Greek and 2) quantify that conceptual load by estimating the increase in reading time; on the practical level, the results of the evaluation may suggest that software aids, such as language identifiers that identify Greeklish and language converters that automatically convert Greeklish to Greek, may be helpful for readers and their integration in social media may be beneficial for the users.

Although this research focuses on Greeklish, both its findings and approach may have a larger impact since Latinization (or Romanization) is not restricted to the Greek language. It is a practice used for many languages which use logograms or non-latin characters for reasons similar to those that gave birth to greeklish. As a result, Latinization is used for Chinese, Japanese, Russian, Arabic, Thai and many more languages. The common issues shared in all these cases include the existence of multiple systems for Latinization (despite the existence of an international transliteration standard), the development of transliteration
software and the societal concerns about the danger that such a practice represent for the traditional alphabets/logograms.

The paper is structured as follows: Section 2 discusses the emergence and evolution of Greeklish along with transliterating practices and related software. In addition, studies on using, reading and understanding Greeklish are also presented. In Section 3, the methodology and the experiment that was conducted are presented, followed by the analysis of the results and discussion. Finally, conclusions, implications, limitations and possible future threads of research are discussed.

2. Greeklish

2.1. The Evolution of Greeklish

“Greeklish” or Latin-alphabet Greek is the “representation of the Greek language with the Latin script” [32]. Valassakis [34] and Moustaka et al. [35] argue that Greeklish is a hybrid language of written Greek electronic communication, while Androutsopoulos [36] describes the Greeklish phenomenon as Sociolect, in the vein that it is a way of speaking with lexical, factual and structural features that is used under certain communication conditions and is part of the linguistic consciousness of a community.

Sporadic evidence suggests that Latin-alphabet Greek were used in folk poetry and catechism during the early modern era in areas of the Aegean that were under the Venetian rule. The comedy “Fortounatos” by Mark Anthony Foscolo written in 1655, one of the three prototypes of the tragedy “Erofili” by Georgios Hortatis written in Crete in 1595 and the book “I Mera tou Hristianou” stored at the Venetian Museum of Naxos, are written in this way [22] [37] [38]. The Latin alphabet was also used by Greek traders in Chios Island in the 18th century, in letters and telegrams as well as in religious texts. The Levantine traders of Smyrna, Asia Minor, who spoke Greek but had difficulties with the Greek spelling, used to write in Latin characters too. Thus, the terms “Francochiotica” and “Francolevantinika” were coined [32] [35].

In 1930s Greek intellectuals of the time put forward a reform issue proposing the writing of the Greek language using Latin characters. The suggested reform raised opposing reactions from other intellectuals who suggested that the adoption of the Latin alphabet would have negative effects.

The Greeklish phenomenon remained a secondary issue in the post-war decades, with use of Greeklish being limited to telegrams to and from abroad, cash receipts and claims and the first weather reports from the Hellenic National Meteorological Service [22] [39]. Scientific lists and University messages were also written in Greeklish, even when technological solutions were given [32].

The explosion of Greeklish in computer-mediated communication environments [7] [8] [9] [22] [26] [32] [34] [35] [36] [40] arose a lot of discussion and concerns about the future of Greek writing. The Academy of Athens released a declaration on January 6th of 2001, expressing its strong concern and intention...
to resist the replacement of the Greek by Latin characters considering it as "an unholy, but also senseless, attempt to replace the Greek script in its own birthplace" and "a full-fledged attack against the classical Greek thinking" [41]. Professor of Linguistics Georgios Babiniotis, pointed out the risk of the “alienation” of Greeks from the image of the Greek words, due to the increasing use of Greeklish, and mentioned that “Greeklish is the best way of alienating from the image of the word. This is what young people will pay for it. We have Greek fonts and we can, using the Internet and the electronic media, use the Greek fonts that have the advantage of giving the image of the word, the visual idol, and to reconcile us with the spelling of the word and its meaning” [42]. On the other hand, Hatzisavvidis [43] as mentioned in [18] argues that “the Greek language is not going to be lost, because a language is lost when people who use it are lost”.

In a recent study Mouresioti & Terkourafi [44] found that Greek native speakers generally are disapproving attitude about Greeklish. They suggested that the negative attitudes toward Greeklish that are currently being expressed by Greek native speakers are caused by technological, demographic, as well as ideological concerns. Negative attitude towards usage of Greeklish of teachers and parents of primary school students was recorded by Xidopoulos, Tzortzatou &Archakis [45]. They declared that they do not write in Greeklish and appears to project the identity of the “custodian” of the national spelling. On the other hand, the majority of students demonstrate that he is in tune with new writing practices in digital communication by stating that he uses Greeklish.

Koutsoutassiou [46] who investigated the graphemic policy of 106 Greek discussion forums in relation to Greeklish use, discovered, the vast majority of the forums have strict graphemic policies that forbid the use of Greek and implement stringent “preventive” measures to guarantee that all users will follow the rules without exception (use of a convertor that converts Latin characters to Greek ones, reprimanding and imposition of punishments, modification or deletion of messages in Greeklish without warning).

### 2.2. Transliterating Practices in Greeklish and Related Software

In linguistics, transliteration means “the system of conveying as nearly as possible by means of one set of letters or characters the pronunciation of the words in languages written and printed in a totally different script” ([47], p. 1069). For transliterating Greek to Latin characters and Latin to Greek the ELOT-743 and ISO-843 transliteration standard is provided by the Greek Standardization Body. Despite this standard, “there are as many different types of ‘Greeklish’ as the Greek-speaking computer users are” ([48], p.275). This exaggeration describes vividly the inconsistency and the variety in transliterating Greek using the Latin alphabet [49].

One of the main characteristic of Greeklish, from the linguist point of view, is spelling [50] since any Greek letters (vowels, consonants) and combinations of
letters (twin, diphthongs) may be represented in Greeklish in alternative ways [50]. As a result, it is not uncommon for Greek users to transliterate in their own idiosyncratic way [7] depending on personal preferences taking into account the similarity between Greek and Latin letters’ appearance, sound or keyboard layout [49]. Therefore, the use of Greeklish has adopted three main methods of transliteration:

1) Vocal: Based on sound resemblance, aims to represent phonetically the Greek text as accurately as possible while simplifying the historic Greek spelling, i.e., the Greek letter /θ/ yields /th/ and the diphthong /αι/ yields /e/.

2) Visual: Based on similarities between Greek and Latin letter shapes, using visually equivalent Latin characters or, in case of absence, numbers that optically resemble Greek letters. In contrast to the vocal trasliteration, it reproduces as much as possible the Greeklish spelling although it leads to some unorthodox solutions e.g. /8/ for the letter /θ/.

3) Locational: Based on the keyboard layout; it represents some Greek letters by the Latin letters that are placed on the same location on a qwerty keyboard.; it is similar to the visual transliteration but differs from it only in some letters e.g. the Latin letter /u/ is used for the Greek letter / θ/ and /c/ for the letter /ψ/ [7] [8] [48] [49] [51] [52].

Taking into consideration all possible different types of Greeklish along with all their likely combinations, common Greek words can be transliterated into many alternative representations. As an example, Androutsopoulos [22] mentioned that the Greek word “δευθυνση” (“address” in English) can be transliterated in twenty-three different Latin-alphabet versions.

Transliteration systems have been developed in an effort to automate the transliteration of Greek to/from Greeklish: deGreeklish [53], E-Chaos [54], Greeklish Converter v1.0 [55], Greek to Greeklish by Innoetics [56] are some of the transliteration systems developed. The majority of them are based on specific sets of rules that map directly each Greek character to a corresponding symbol of the Latin alphabet or use databases of Greek-Greeklish word pairs [19] [26] [47] [57] [58] [59] [60].

In 2012, Google started the Google Transliteration, an online service which converts Latin characters to phonetically equivalent characters to Greek and many other languages including Arabic, Hindi, Urdu, Punjabi, Persian, and other languages spoken in East Asia [57].

2.3. Using, Reading and Understanding Greeklish

Androutsopoulos [32] investigated the usage and attitudes towards Greeklish in e-mail messages, using a questionnaire electronically distributed to Greek mailing lists in Greece and abroad. Findings indicate that despite the technological advances, the use of Greeklish is quite widespread, as 80% of Greeks who live abroad write more than half of their messages in Greeklish while for the Greek residents the corresponding percentage is 69%. The majority of users do accept Greeklish as a means of electronic communication, although half of the sample
considers it “ugly” but not a “threat” to the Greek language. Greek e-mail users seem to have a high degree of metalinguistic awareness regarding different transliteration conventions. Use of Greeklish poses more problems when used for reading rather than writing. Almost 50% of the responders agree that “reading Greeklish is a hard and tiring task”. Laghos et al. [25] analyzed 1000 email messages, from 48 email accounts, exchanged between Greek/Cypriot academic and administrative staff at the public universities of Cyprus in 2011. They found that 1 out of 3 emails were written in Greeklish. The Greek language with Greek fonts is used in formal documents while Greeklish is highly used in informal e-discussions, mainly due to the speed and flexibility of using it. They concluded that using Greeklish appears to be a conscious decision suggested by convenience rather than technological constraints.

Moustaka et al. [35] investigated the effect of Greeklish use on students’ spelling. As a part of their study they recorded student habits regarding Greeklish use. A significant overall percentage of 77.4% use Greeklish. This percentage rises from 67.8% among middle school to 88.5% for high school students. It seems that students use Greeklish by habit and they feel they save time while a significant percentage mentioned that using Greeklish is a way to avoid spelling mistakes. Mouratoglou [18] investigated the Greeklish phenomenon from Greek language teachers’ point of view. Increasing use of Greeklish was noticed as 49% of the responders reported that they use it while almost 34% did not answer the question. Extensive use of Greeklish was recorded in social media and mobile phones with the main reasons for using it being the speed of writing a message (68.60%) and time saving (52.89%). The majority of the sample (86.25%) considers Greeklish as a form of writing, mainly in informal forms of communication and recognizes its functionality as a tool for electronic communication (77.92%). A significant percentage of them (63.7%) stated that they use Greeklish on a daily basis.

Karagouni [27] studied both students’ and Greek language teachers’ views for Greeklish. The majority of students uses Greeklish primarily in social media as it provides speed and facilitates communication. Students do not believe that Greeklish has a negative impact on their writing. The use of Greeklish is reduced among the final-year Lyceum students, who noticed a negative effect of Greeklish in their writing and decided to discontinue their use. This finding seems to contradict the findings of Moustaka et al. [35] who reported that the use of Greeklish is reduced only among the students of the last grade of the general high schools and not among the students of the other types of schools. Regarding Greek language teachers who participated in the survey, the overwhelming majority do not use Greeklish; only a small minority admit Greeklish use, especially in social media environments. Greek language teachers believe that Greeklish, along with other factors, contribute negatively to students’ correct writing.

As it is evident from the research mentioned above, Greeklish is only used for specific communication purposes [61]. In this vein, Laghos et al. [62] investigated the use of Greeklish in YouTube. They collected over 1000 user comments
posted on Greek videos in Youtube, by Greek-speakers living in Greece or Cyprus. The vast majority of the messages (84%) were written in Greeklish, 7% in English and only 9% in Greek. Spilioti [23] investigated the choice of alphabetical encoding in Greek text-messaging. She studied 447 text-messages exchanged among 10 participants aged 15 - 25 and found that writing in Greek characters was the norm in Greek text-messaging. Her findings are in contrast to the findings of Laghos et al. [62] maybe due to the small number of participants. However, the findings give evidence that each participant sticks to their usual choice of alphabetical encoding in text-messaging [61].

Reading and understanding a long text in Greeklish is a demanding task, claimed Chalamandaris et al. [48] partly due to the fact that common Greek words can be transliterated into many alternative representations deriving from a combination of all types of Greeklish and it is rather rare for a user to consistently use only one type of transliteration. Tseliga [63] found that reading and understanding a sentence in Greeklish requires more effort as it is in average more time consuming by over 40% than reading and understanding the same sentence in Greek. Lees, Politis & Koutsogiannis [64] contend that, contrary to popular belief, the use of the Latin alphabet does not pose a threat to the Greek language, but rather contributes to digital communication among youngsters. Their findings also show a definite preference for the Greek script over the Latin script. When using the Latin alphabet, students frequently adopt orthographic transliteration to maintain traditional Greek spelling, notwithstanding gender-related variances. Thoma [65] found that students with a medium to high language profile read Greeklish more readily than students with a low language profile. Another crucial finding was that the use of Greeklish has no bearing on or connection to pupils’ spelling abilities. Students with intermediate to advanced language profiles and students with poor language profiles both utilize Greeklish, and there was no statistically significant difference in the link between code usage and spelling performance. Finally, there was no connection between the use of Greeklish and the amount of free time students spent online on their language profiles.

3. Methodology

The following experiment was conducted to address the aims of this research: Participants (mainly young Greek internet users) were asked to read two short texts of equal length displayed on their computer screens. One of the texts was written in Greek while the other in Latin. After reading each text the participants were required to answer 5 multiple choice questions related to the text they had just read. The reading time of the participants, their responses to the multiple choice questions as well as the participant profile (sex, age, education level, etc) were recorded.

The text was shown to the participants’ screen in black 18 pt Arial characters displayed on white background. The text in Greeklish was produced using the transliteration proposed by Greece’s Standards Organization (ELOT 743).
Both texts, presented to the participants, were short stories related to the theft of Mona Lisa in 1911. The first text presented to the participants (referred to as “the theft”) described how the authorities of the Louvre Museum realized that the famous painting was stolen, while the second text presented to the participants (referred to as “the thief”) described how the theft was carried out and the arrest of the thief. The texts did not require any specific knowledge or terminology and based on standard readability indexes, they are both appropriate for university graduates (see next Table 1).

In order to ensure that text characteristics did not have an effect on reading time, half of the participants were firstly presented the text “the theft” in Greek alphabet and then the text “the thief” in Greeklish while the other half participants were presented firstly the text “the theft” in Greeklish and then the text “the thief” in Greek.

The experiment was implemented in Javascript using jspsych [66], a javascript library dedicated to create and execute psychological experiments. The experiment was hosted in heroku, a cloud application platform [67] and the responses of the participants were saved to a Heroku Postgress database. Statistical analysis of the results was carried out using R [68]. A convenience sample of social media users participated in the experiment. A link to the experiment was posted to Facebook and to a popular blog used exclusively by the students of Aristotle University in Thessaloniki, Greece, along with an informed consent form for the participants. The adolescents that participated in the study were contacted after the explicit consent of their legal guardians.

4. Findings

4.1. Demographics

A total of 732 respondents participated in the experiment. Most of the participants were females (63.1%) whereas 36.9% were males. Table 2 presents the participants’ profile regarding their age and gender. The vast majority of the sample (69.4%) were young people belonging to the age group from 18 to 25 years old, while 23.2% of the sample were middle-aged and only 7.4% were adolescents.

<table>
<thead>
<tr>
<th>Table 1. Readability indexes for the two texts.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Number of words</td>
</tr>
<tr>
<td>Number of syllabuses/100 words</td>
</tr>
<tr>
<td>Number of sentences</td>
</tr>
<tr>
<td>Average sentence length</td>
</tr>
<tr>
<td>Fog (Cunning) Index</td>
</tr>
<tr>
<td>Years of formal education</td>
</tr>
</tbody>
</table>

| Fog (Cunning) classification                  | Very difficult, appropriate for university graduates | Very difficult, appropriate for university graduates |
The participants of the survey were very familiar with internet use since the overwhelming majority uses the internet either daily (71.58%) or at least once a week (25.68%). This finding is rather expected if we take into account the age distribution (Table 2) as well as the education level of the participants (Table 3). The vast majority (97.95%) of the participants were high school graduates (at least) while at the same time 35.25% of them hold a higher education or a postgraduate degree.

Finally, although two out of three participants (66.53%) do not use Greeklish there is a significant percentage (33.47%) that does use Greeklish. Figure 1 presents the percentages of Greeklish use in regard to the gender and education level of the participants.

Table 2. Cross tabulation between the age and gender of the participants.

<table>
<thead>
<tr>
<th>Age</th>
<th>Count</th>
<th>% of Total</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>under 18</td>
<td>54</td>
<td>7.4%</td>
<td>Female 22</td>
</tr>
<tr>
<td>18 - 25</td>
<td>508</td>
<td>89.4%</td>
<td>Female 309</td>
</tr>
<tr>
<td>26 - 33</td>
<td>64</td>
<td>8.7%</td>
<td>Female 47</td>
</tr>
<tr>
<td>34 - 39</td>
<td>47</td>
<td>6.4%</td>
<td>Female 44</td>
</tr>
<tr>
<td>40 - 50</td>
<td>51</td>
<td>7.0%</td>
<td>Female 35</td>
</tr>
<tr>
<td>over 50</td>
<td>8</td>
<td>1.1%</td>
<td>Female 5</td>
</tr>
<tr>
<td>Total</td>
<td>732</td>
<td>100%</td>
<td>Female 462</td>
</tr>
</tbody>
</table>

Figure 1. Bar plots of the participants’ profile regarding the use of Greeklish and their gender or education level.
Table 3. Education level of the sample.

<table>
<thead>
<tr>
<th>Education level</th>
<th>Middle School graduate (MSG)</th>
<th>High School graduate (HSG)</th>
<th>University graduate (UG)</th>
<th>Postgraduate studies graduate (PSG)</th>
<th>Phd degree holder (PDH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage %</td>
<td>2.05</td>
<td>62.70</td>
<td>29.37</td>
<td>5.33</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Thus, the sample consists mainly of young people who use the internet on a daily basis and have a high education level. Furthermore, the sample is comprised by a sufficient number of Greeklish users to conduct reliable inferential statistics. Therefore, the findings of the statistical analysis related to the reading speed of Greeklish may be attributed almost exclusively to the familiarity of the participants with Greeklish since the participants do possess the digital and cognitive competencies at a satisfactory level.

4.2. Reading Speed

In this section, several factors possibly affecting the reading speed of the two texts written in Greeklish and Greek are examined.

Firstly, statistical properties of reading time recorded for both Greek and Greeklish texts are presented. Initially, the users were informed that the aim of the experiment was to measure their reading speed in Greek and Greeklish texts. It was explained to the users that they would have to read two short texts of equal size, one written in Greek and the other in Greeklish and it was pointed that after reading each text they would answer five multiple-choice questions based on the text they just read. To initiate and complete the reading session of each text, the users were prompted to press a designated “start reading” and “end reading” button. The text was not available to the users after they pressed the “end reading” button. The reading time was measured in milliseconds by utilizing jspsych’s automated time recording feature and equaled the time elapsed between the “start reading” and “end reading” buttons press.

There are not any missing values in the sample. Furthermore, the measures of central tendency (mean, median) and dispersion (standard deviation, interquartile range), and especially the low difference between the mean and the median, suggest minor discrepancy from symmetric distributions, which is confirmed by the histograms plotted in Figure 2.

Based on the results shown at Table 4, the mean value of the reading time is much higher for the Greeklish text than for the Greek one, a fact which gives rise to the question whether the mean reading time of the text written in Greeklish statistically differs from the mean reading time of the text written in Greek. The results of the appropriate paired sample t test are presented in Table 5, where it is shown that p value is lower than the common significance level 0.05 and therefore the null hypothesis of the equality of the means is rejected. The mean of the differences (11.47 sec) and the corresponding 95% confidence interval of the difference of the means between the Greeklish and Greek text, which is (10.27, 12.68), indicate that the mean reading speed of the Greek text is signifi-
cantly higher than the Greeklish text, which is rather expected since all the participants are Greek native speakers.

Next, factors that affect reading time were tested. Multi-way ANOVA was conducted on the influence of 6 independent variables (gender, age, education level, frequency of connecting to internet, alphabet of the first text, use of Greeklish) on reading time. The assumptions of the multi-factor ANOVA are satisfied. Multi-way ANOVA identified that gender, use of Greeklish and alphabet of the first text are the only statistically significant factors.

Moreover, the main and interaction effects of: the gender, the use (or nonuse) of Greeklish and the alphabet of the text displayed first, on reading time were explored. A main effect is observed when a factor affects the dependent variable regardless of the effect the other factors have, whereas an interaction effect is detected when two or more factors simultaneously affect the dependent variable. When the combined effect on the dependent variable is observed the impact of one factor depends on specific levels of the other factors. The following research questions are investigated:

- Is there a statistically significant main effect for each one of the above mentioned factors (gender, Greeklish usage, alphabet displayed first) on reading time?

![Figure 2. Histograms of the reading time for the Greeklish and Greek text.](image)

**Table 4.** Statistics for reading time (in seconds) for the Greeklish and Greek text.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Interquartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greeklish</td>
<td>732</td>
<td>65.70</td>
<td>63.55</td>
<td>19.45</td>
<td>25.67</td>
</tr>
<tr>
<td>Greek</td>
<td>732</td>
<td>54.23</td>
<td>52.19</td>
<td>16.84</td>
<td>20.44</td>
</tr>
</tbody>
</table>

**Table 5.** Paired sample t test for the means of reading time (in seconds) between the Greeklish and the Greek text.

<table>
<thead>
<tr>
<th>Differences between reading time (Greeklish-Greek)</th>
<th>Mean of the differences</th>
<th>t statistic</th>
<th>Degrees of freedom</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of the differences</td>
<td>11.47</td>
<td>18.685</td>
<td>731</td>
<td>&lt;2.2e-16</td>
</tr>
</tbody>
</table>
- Is there a statistically significant interaction effect between the above mentioned factors? More specifically, is there a statistically significant impact between the levels of the three factors under investigation?

The results are presented in Table 6. It is important to highlight the fact that there are no interactions effects between the factors examined. The two tables indicate that gender and use of Greeklish are both statistically significant factors affecting reading time for both Greeklish and Greek texts. Based on the results presented in Figure 3, reading time for women is higher than that for men and is also higher for non-Greeklish users compared to Greeklish users. The first important conclusion is that men were faster than women in reading both texts possibly because females read the text more carefully than males [68]. The second important conclusion is that Greeklish users achieve higher reading speed than those who do not use Greeklish for both texts; as it is expected, the participants who were familiar with the Latin alphabet read significantly faster the Greeklish text compared to those who do not use Greeklish; However, the results showed that respondents who use Greeklish read faster the Greek text, also. The specific result may imply that the use of different alphabets helps the readers to improve the reading process and to increase their reading capability.

**Table 6.** Three-way ANOVA; Factors: gender, use of Greeklish, alphabet of the first text; Response variable: Reading time.

<table>
<thead>
<tr>
<th>Sum of squares</th>
<th>df</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greeklish</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First alphabet</td>
<td>64</td>
<td>0.1716</td>
<td>0.67883</td>
</tr>
<tr>
<td>Gender</td>
<td>2243</td>
<td>6.0290</td>
<td>0.01431*</td>
</tr>
<tr>
<td>Greeklish use</td>
<td>1980</td>
<td>5.3210</td>
<td>0.02135*</td>
</tr>
<tr>
<td>First alphabet: Gender</td>
<td>809</td>
<td>2.1757</td>
<td>0.14064</td>
</tr>
<tr>
<td>First alphabet: Greeklish use</td>
<td>18</td>
<td>0.0487</td>
<td>0.82549</td>
</tr>
<tr>
<td>Gender: Greeklish use</td>
<td>704</td>
<td>1.8926</td>
<td>0.16934</td>
</tr>
<tr>
<td>First alphabet: Gender: Greeklish use</td>
<td>1279</td>
<td>3.4380</td>
<td>0.06412</td>
</tr>
<tr>
<td>Residuals</td>
<td>269,342</td>
<td>724</td>
<td></td>
</tr>
<tr>
<td>Greek</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First alphabet</td>
<td>3449</td>
<td>12.5637</td>
<td>0.00042***</td>
</tr>
<tr>
<td>Gender</td>
<td>1336</td>
<td>4.8685</td>
<td>0.02767*</td>
</tr>
<tr>
<td>Greeklish use</td>
<td>2307</td>
<td>8.4040</td>
<td>0.00386**</td>
</tr>
<tr>
<td>First alphabet: Gender</td>
<td>32</td>
<td>0.1180</td>
<td>0.73131</td>
</tr>
<tr>
<td>First alphabet: Greeklish use</td>
<td>301</td>
<td>1.0960</td>
<td>0.29550</td>
</tr>
<tr>
<td>Gender: Greeklish use</td>
<td>21</td>
<td>0.0771</td>
<td>0.78137</td>
</tr>
<tr>
<td>First alphabet: Gender: Greeklish use</td>
<td>702</td>
<td>2.5563</td>
<td>0.11029</td>
</tr>
<tr>
<td>Residuals</td>
<td>198,745</td>
<td>724</td>
<td></td>
</tr>
</tbody>
</table>

Type II sum of squares have been computed; df = degrees of freedom.
A rather unexpected finding was the fact that reading time of the Greek text was affected by the alphabet of the text displayed first. More specifically, the reading time of the Greek text was increased when it was read after the Greeklish text for all but the male participants who use Greeklish. This is shown in Figure 4, where reading time for the Greek text when displayed first (denoted with the letter A) and when displayed second (denoted with the letter B) is shown. A possible explanation for this finding may be that participants used different reading strategies for the Greek and the Greeklish text, with the strategy used for the Greek text being more efficient. When the Greeklish text is read first, the participants (excluding males who use greeklish) delayed to change to the more efficient strategy they used for the Greek text. However, this is an issue that needs to be further investigated.

4.3. Understanding

In this section, the understanding of the two texts by the participants is examined. It is reminded that after reading each text (Greek and Greeklish) the participants had to answer 5 multiple choice questions related to the text. Figure 5 demonstrates the distribution of correct answers of the respondents for Greeklish and Greek texts. It is worthwhile to mention that the frequency distributions of the correct answers have exactly the same shape despite the fact that the reading times are statistically different between the Greeklish and Greek text as it was explained in the previous section.
The first issue regarding the understanding of texts is related to the pairwise independence between the score achieved by the participants and the factors of the experiment (gender, age, education level, use of Greeklish, alphabet of the first text). Chi-square test of independence revealed that only the alphabet of the first text influences significantly the score of correct answers, both for the Greeklish and Greek text. The chi square statistics were 98.467 for the Greeklish and 95.376 for the Greek text with p values less than 2.2e-16 in both cases, indicating highly dependence between the two variables. Figure 6 visualizes the discrepancy between the observed and expected frequencies using the Pearson residuals defined by the quotient (observed-expected)/(square root of expected). In the mosaic plots, introduced by Friendly [69], the area of each tile is proportional to the relative frequency observed when the two variable take the specific values determined by the tile.

The shading provides a means to visualize the pattern of dependence when the null hypothesis of independence is rejected. The highlighted tiles correspond to deviations from independence with two significances levels 0.05 (colors light
Figure 6. Mosaic plots relating the variable Experiment which represents the order of the alphabet appearance and the score of correct answers. For Greeklish, the tile with coordinates (1, 6) in the first row (EN-GR) and in the sixth column (score = 5) colored red, represents negative Pearson residuals (less than −4) meaning a very low observed frequency compared to the expected frequency in the case of independence. On the contrary, the tile with coordinates (2, 6) in the second row (GR-EN) and in the sixth column (score = 5) colored blue, represents positive Pearson residuals (more than 4) meaning a very high observed value compared to the expected frequency in the case of independence. The tiles with blue and light red) and 0.01 (colors dark blue and dark red). For Greeklish, the tile with coordinates (1, 6) in the first row (EN-GR) and in the sixth column (score = 5) colored red, represents negative Pearson residuals (less than −4) meaning a very low observed frequency compared to the expected frequency in the case of independence. On the contrary, the tile with coordinates (2, 6) in the second row (GR-EN) and in the sixth column (score = 5) colored blue, represents positive Pearson residuals (more than 4) meaning a very high observed value compared to the expected frequency in the case of independence. The tiles with
coordinates (1, 3), (1, 4), (2, 3) and (2, 4) are also statistically significant, at significance level 0.05, with similar interpretation. The grey color characterizes the non-significant tiles meaning there is no discrepancy from the independence for these tiles. The results are similar for the Greek text. The conclusion in both cases is that the understanding of the second text is significantly better than the first one regardless of the alphabet of the text displayed first. A possible explanation is that the participants after reading the first text and answering the corresponding questions are more prepared about the type and the difficulty of the task when they read the second text and they therefore achieve a better score.

In order to investigate the explanatory variables that affect the “perfect” understanding of the text a multiple logistic regression model is adopted (e.g. [70]). In this model, the binary response variable is one (denoting the event of “success”) when the respondent answers correctly all five multiple choice questions and zero (denoting the event of “failure”) otherwise. The statistically significant explanatory variables (with p values < 0.01 in both cases) are the alphabet of the text shown to the participant first and the score achieved in answering the multiple choice questions of the text written in the other alphabet.

More explicitly, the model for Greeklish is:

$$\ln \left( \frac{p}{1-p} \right) = -3.550 + 1.907 \cdot \text{Experiment} + 0.339 \cdot \text{scoreGR}$$

$$p = \frac{1}{1 + \exp (3.550 - 1.907 \cdot \text{Experiment} - 0.339 \cdot \text{scoreGR})}$$

where:

- $p$ is the probability for a participant to answer correctly all five multiple choice questions related to the Greeklish text;
- Experiment is an indicator variable which shows the alphabet of the first text; It takes the value one if the Greek text is displayed first and zero if the Greeklish text is displayed first;
- ScoreGR (the score for the Greek text) is the number of multiple choice questions of the Greek text answered correctly by the participant.

The model suggests that when the GR-EN procedure is followed, the logarithm of the ratio $p/(1-p)$ increases by 1.907 compared to the EN-GR procedure. Thus, for the mean value of the variable scoreGR which is 3.42, the probability of success $p$ increases from $p = 0.0839$ when the EN-GR is followed, to $p = 0.3816$ if the GR-EN is followed. Likewise, for every unit increase of scoreGR the logarithm of the ratio $p/(1-p)$ increases by 0.339 both for EN-GR and GR-EN procedures.

In a similar manner the model for the Greek text is:

$$\ln \left( \frac{p}{1-p} \right) = -1.977 - 1.913 \cdot \text{Experiment} + 0.449 \cdot \text{scoreEN}$$,

where

ScoreEN is the discrete variable with values the score of the respondents in the
Greeklish text. The coefficient of Experiment in the second model is very closely to the opposite number of the corresponding coefficient in the first model. This means that it has the opposite effect, that is when the GR-EN procedure is followed the logarithm of the odds decreases by 1.913 compared with the EN-GR procedure. These findings are similar with the ones from the chi-square independence tests but the great benefit of the logistic regression models is the quantification of the relationship between the probability of success, the experimental procedure and the score of the respondents in both texts. It should also be noticed that the coefficient of scoreEN is higher than the corresponding coefficient in the first model indicating that the number of correct answers in Greeklish influences the probability of success in Greek text more than the opposite case in the first model.

5. Conclusions

In the era of globalization and social media new linguistic phenomena, like Greeklish, 3arabizi, Chinglish, Japlish, Singlish, Spanglish appeared and important questions about the cognitive processes arise. Up to now, research has focused on a linguistic point of view or from a philosophical perspective. This research goes a step beyond and measures the extent to which Greeklish is used by young people when communicating through internet and social media and proposes a method for evaluation of the effect of these phenomena on understanding and reading time.

A sample of 732 participants read two short texts of equal difficulty, one written in Greek characters and one in Latin. After reading each text the participants answered five questions related to the respective text. Half of the participants started the procedure reading first the text written in Greek characters while the other half started the procedure reading first the text written in Latin characters.

Nearly one-third of the participants stated that they use Greeklish for writing. This is a rather low percentage of young people who use Greeklish when compared to similar reported findings. According to Kappatou [71], for example, 69.4% of young people aged 13 to 25 use Greeklish when chatting while an even higher adoption of Greeklish is reported in Moustaka et al. [35] who found that 77.4% of middle and high school students use Greeklish. The findings of this research are closer to Karagouni [27] who reported that 40% of the high school students use Greeklish and this finding may be an indication that although the adoption of Greeklish by young people is still very high it may be decreasing in time.

Male participants were slightly faster, compared to female participants of the study, in reading both Greek and Greeklish texts. More specifically, male read the Greek (Greeklish) text almost 5% (6%) faster than females. This result seems to contradict the general view that females are generally better at reading activities [72] but is in line with Emam & Youssef [68] who reported that there is a difference in reading speed between males and females. According to them, “fe-
males spend longer time between consecutive segments than males and they spend longer time on each segment than males in the reading process.” They argued that “females might have taken long time to focus and concentrate on understating the topic while males might have read fast without guarantee to comprehend”.

The average reading speed for Greek texts is calculated to 238.55 words per minute while the average reading speed for Greeklish texts is 193.89. Thus, reading the Greeklish text was more time-consuming than reading the Greek text. Both males and females need nearly 21% more time to read the Greeklish text compared to the Greek text. Similarly, Greeklish users need 23% more time to read the Greeklish text compared to the Greek text while for the non-Greeklish users the corresponding percentage was 20%. This result seems to support the hypothesis that Greeklish is adopted mainly because people find it a convenient writing practice but reading a Greeklish text is “hard and tiring task” [32] even for Greeklish writers. Similar results with even higher differences between the Greek and the Greeklish text have been reported by Tseliga [51] although the conditions of the reported experiment were different since the participants had to read a set of simple sentences and state if each sentence was true or false. Greeklish users are slightly faster readers than non-Greeklish users. More specifically, Greeklish users read the Greek (Greeklish) text 3% (4%) faster that the Greeklish non-users. Finally, as far as the reading speed is concerned, an issue that needs further investigation, is the finding that reading time of the Greek text was affected by the alphabet of the text displayed first.

Regarding the subject of text understanding, the participants achieved the same score when answering the multiple choice questions for the Greek and Greeklish text as shown by the corresponding (nearly identical) score distributions which have a mean of 3.42 and 3.34 respectively. The score is not differentiated by the gender and thus one could not support the hypothesis that the slower reading speed observed by female participants leads to a better performance. On the other hand, it should be noted that the limited size of text as well as the limited number of questions asked, may not provide a sufficient framework to uncover the possibility of a better and deeper understanding due to more careful reading.

The score achieved by the participants of the study in answering the multiple choice questions of the second text was better than the score achieved for the first text regardless of the alphabet of the text displayed first. This difference on the performance of the participants may be attributed to the fact that when they read the second text they were better prepared about the type and difficulty of the questions asked.

The study has some limitations that are mentioned here. Firstly, in terms of the experiment design, participants were asked to read two short texts of equal length displayed on their computer screens. Users reading times were identified based on mouse clicks. While this method provides valuable data, it would be
beneficial to incorporate eye-tracking systems in future research. Previous studies have shown a relationship between reading behavior and eye movements [33] [73] [74], and utilizing eye-tracking technology could enhance the validity of the experiment.

Additionally, the current study focused exclusively on young students. To gain a more comprehensive understanding, future research could involve a wider age range of participants with varying educational backgrounds. This would enhance the generalizability of the findings and allow for a comprehensive understanding of how different demographics engage with the texts.

Latinization in computer-mediated communication activities appeared as a writing practice during the first era of Internet and mobile phones. Latinization, which began as an ad hoc sideway approach for surpassing technological constraints, is nowadays a conscious decision of users. This research focused on Greeklish, the Latinization of the Greek language, however, the results found in this paper may have wider application and the approach proposed may be employed to quantify the effect of Latinization on reading time and understanding. Similar research could be conducted to establish whether similar findings apply to other languages. The findings of this research suggest that Greeklish is still a popular practice among young Greeks despite the fact that it hinders reading time. The findings support the position that amending social and communication media with software utilities that automatically identify and change the keyboard to Greek characters or transform Greeklish into Greek when necessary, would significantly reduce reading time and facilitate written communication among the users.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References


http://users.skynet.be/p.alevantis/gr/gr-001.htm


https://doi.org/10.1163/15699846-02102004


https://doi.org/10.1002/asi.10306


http://tools.wcl.ece.upatras.gr/degreeklish


Greeklish Transcriber Modeled by Real User Data. Artificial Intelligence: Theories,
Models and Applications, 6th Hellenic Conference on AI, SETN 2010, Athens, 4-7
May 2010, 363-368.

a New Transformation Environment for Greeklish Characters. Themata Episthmon

Transliterated Text to Phonetic Strings Case Study: Greeklish. Proceedings Inters-
https://doi.org/10.21437/Interspeech.2007-499

Transcription Effects in Biscriptal Readers: The Case of Greeklish. Psychonomic

American International Journal of Contemporary Research, 2, 1-5.

Greek (i.e. Greeklish) Emails. Internet in Linguistics, Translation and Literary Stu-
dies, 16, 391-401.

[64] Lees, C., Politis, P. and Koutsogiannis, D. (2017) Roman-Alphabeted Greek and
Transliteration in the Digital Language Practices of Greek Secondary School Pupils on
https://doi.org/10.1386/jgmc.3.1.53_1

[65] Thoma, R. (2021) The Use of Greeklish as a Shaping Factor of Students’ Language
Awareness and Ability. Doctoral Dissertation, University of Thessaly, Thessaly. (In
Greek)

in a Web Browser. Behavior Research Methods, 47, 1-12.
https://doi.org/10.3758/s13428-014-0458-y


the American Statistical Association, 20, 190-200.
https://doi.org/10.1080/01621459.1994.10476460

ley, Hoboken. https://doi.org/10.1002/0470114754

and Sociological Analysis of Online Discussions among Young People. Doctoral
Dissertation, University of West Macedonia, Florina.

Research. Learning and Individual Differences, 21, 145-149.
https://doi.org/10.1016/j.lindif.2010.11.021

Chinese Texts with or without Marks for Word Boundaries. 2011 IEEE Internation-
al Conference on Asian Language Processing, Penang, 15-17 November 2011, 291-
294. https://doi.org/10.1109/IALP.2011.35