Creating an Orthography for an Endangered Language: The Graphisation of Cypriot Arabic

Spyros Armostis & Marilena Karyolemou

Abstract. Graphisation, i.e., the development of a signary and associated orthographic conventions for a previously unwritten language variety, constitutes an important component of language corpus planning, particularly in the context of endangered-language revitalisation. This paper presents the graphisation of Cypriot Arabic, a severely endangered Arabic variety spoken by Maronite Cypriots of Kormakitis. Following its recognition as a minority language of Cyprus in 2008 and the establishment of revitalisation initiatives, the need arose for a codified writing system to support language documentation, teaching, and community use. A proposed signary and orthography were developed on the basis of linguistic analysis and subsequently evaluated through a study involving native speakers. The study investigated preferences regarding script choice, orthographic depth, and the representation of selected morphophonological phenomena, as well as attitudes towards the proposed system in terms of acceptability, usability, and learnability. Participants generally considered the proposal acceptable and logical. However, they did not uniformly favour either phonetic transparency or morphophonological depth. Instead, preferences varied according to the linguistic phenomenon under investigation, with surface-oriented spellings preferred for lexical processes and deeper spellings for post-lexical ones. The findings suggest that successful graphisation should combine linguistic analysis with speaker feedback and that orthographic depth may vary across phenomena rather than characterising a writing system as a whole.

1. Introduction

Linguists are often called upon to propose a writing system for language varieties that have hitherto existed primarily as oral traditions with no established written norm. In such graphisation efforts, especially in cases of endangered varieties, multiple parameters must be taken into account in order for the implementation of a new system to have the potential to succeed. Beyond the linguistic soundness of the proposed or-

Spyros Armostis 0000-0003-2927-2096
University of Cyprus, E-mail: armosti.spyros@ucy.ac.cy

Marilena Karyolemou 0000-0003-0914-5794
University of Cyprus, E-mail: makar@ucy.ac.cy

Y. Haralambous (Ed.), *Grapholinguistics in the 21st Century 2024. Proceedings*
Grapholinguistics and Its Applications (ISSN: 2681-8566, e-ISSN: 2534-5192), Vol. 11.
Fluxus Editions, Brest, 2026, pp. 181–207. https://doi.org/10.36824/2024-graf-armo
ISBN: 978-2-487055-10-0, e-ISBN: 978-2-487055-11-7

thography, issues of acceptability and learnability among native speakers of the variety become central. The present study reports the results of an investigation that tested a proposed orthography for Cypriot Arabic, a severely endangered language in the mediterranean island of Cyprus.

1.1. Cypriot Arabic

Cypriot Arabic (endonym: *sanna* 'our language', also referred to as *arapika* 'Arabic' within the community) is spoken by members of the Maronite community of Cyprus, one of the constitutionally recognised ethnoreligious groups of the Republic of Cyprus. In the scientific literature the variety is known as *Cypriot Arabic*, while the designation *Cypriot Maronite Arabic* has appeared to a certain degree in official contexts; the informal label *marōnitika* 'Maronite language' is also attested but imprecise (Armostis, 2025). Maronite Cypriots belong to the Maronite Church, an Eastern Catholic Church based in Lebanon with Saint Maron as its patron saint. Maronite populations arrived on the island in successive waves from the region of Greater Syria between approximately the 7th and 12th centuries, with medieval sources reporting the presence of numerous Maronite villages especially to the north of the island (Frangeskou and Hadjilyra, 2012, p. 8). Today only four Maronite villages remain, all located in the northern part of the island. Cypriot Arabic is spoken exclusively by Maronites originating from the Kormakitis village; Maronites from other villages neither speak the variety nor have any memories of the language ever having been used in their villages in the past.

From a typological and historical perspective, Cypriot Arabic belongs to the group of peripheral Arabic varieties, i.e. varieties that developed outside the core Arabic-speaking area and in sustained contact with non-Semitic languages (Borg, 1985; 1994). In this case, prolonged contact with Cypriot Greek has resulted in extensive lexicogrammatical influence (Armostis and Karyolemou, 2023). Following the turbulent events of 1974 and the displacement of the Kormakitis population, community networks were disrupted and intergenerational transmission virtually arrested. As a result, Cypriot Arabic is today spoken as a first language by fewer than one thousand speakers. Therefore, Cypriot Arabic is classified as severely endangered in UNESCO's *Atlas of the World's Languages in Danger* and is characterised by very low ethnolinguistic vitality (Karyolemou, 2019).

The official recognition of Cypriot Arabic as a minority language of the Republic of Cyprus in 2008, following joint efforts by academics and members of the Maronite community, constituted a crucial step in formal language planning, and more specifically in status planning: this recognition established Cypriot Arabic as a language in its own right rather than as a mere dialectal form of Arabic, thereby raising its insti-

tutional profile and enabling access to funding and coordinated protection measures (Armostis, 2025). In consultation with community representatives, a team of academics subsequently designed a structured programme for documentation and maintenance/revitalisation, while the government appointed an Experts Committee composed of academics and community members to advise on revitalisation strategy. These developments culminated in 2013 in the launch of the 'Research Programme for the Documentation, Preservation and Revitalisation of Cypriot Maronite Arabic' under the Cyprus Ministry of Education (the now-called the Ministry of Education, Sports and Youth), a programme in which the present authors participate (ibid.).

Given that Cypriot Arabic had for centuries functioned almost exclusively as an orally transmitted variety, with very limited written material and no comprehensive grammatical or lexicographic description, the programme initially prioritised systematic documentation as a foundation for corpus planning (Armostis, 2025; Armostis and Karyolemou, 2023). Documentation, however, was conceived not as an end in itself but as a prerequisite for maintenance and revitalisation. From the outset, the development of teaching and learning materials emerged as a central objective. This, in turn, made explicit the absence of a codified orthography for Cypriot Arabic and necessitated its practical development, i.e., its *graphisation*.

Graphisation is part of language corpus planning (Fishman, 1974; Haarmann, 2006; Liddicoat, 2005) and can be defined as the development of a signary and associated orthographic conventions for a given language variety (Fenyvesi and Bakrá-Nagy, 2025; Haarmann, 2006). The signary is 'the complete inventory of the basic signs of a given writing system' (Coulmas, 2003, p. 36), which includes not only signs corresponding to language units (distinctive features, phones, phonemes, syllables, moras, or morphemes/words), but also other symbols, such as typographic blank spaces, punctuation marks, and numerals. Signary development, as a component of graphisation, is called *alphabetisation* in the case of alphabetical writing systems (Haarmann, 2006); for graphisation to be complete, the creation of an associated orthography is required. Orthography is an optional component of a writing system and is understood as regulation of the system by means of prescriptive rules specifying how one may write correctly in the given writing system (Meletis and Dürscheid, 2022, p. 171–172).

1.2. Writing in Cypriot Arabic

In discussing the development of a signary and orthography for Cypriot Arabic, it is necessary to consider the literacy backgrounds of speakers and their familiarity with existing scripts and writing systems. Mem-

bers of the Cypriot-Arabic-speaking community have native competency not only in Cypriot Arabic, but also, in recent years, in Cypriot Greek and are literate in Standard Modern Greek; they therefore have full command of the Standard Modern Greek writing system and the Greek script in general. In addition, through formal education, many speakers have learnt English, French, and/or Italian as foreign languages and are consequently familiar with their writing systems and with the Roman script more broadly. By contrast, there is practically no literacy in Modern Standard Arabic within the community, with the exception of a couple of individuals. As a result, the Arabic script is largely unfamiliar and not functionally available for community-wide literacy practices. Syriac abjad is used in liturgical contexts, but it is not directly accessible to most Cypriot Arabic speakers, since liturgical texts are typically mediated through Greek transliteration rather than read in the original script.

On the basis of the above considerations, any graphisation effort would be more successful if it was based on the Greek and/or Roman scripts rather than on the Arabic one. This is supported by the limited writing practices in Cypriot Arabic that existed prior to the graphisation efforts of the above-mentioned research programme. For instance, Cypriot Arabic was rendered in writing with the use of the Greek script in the Frangiskou (2000) dictionary and in some other unpublished material. Even though the use of the Greek script (with the addition, for some users, of two letters from the Roman script, namely <b> and <d>) was not conceived necessarily as a full alphabet proposal in those sources, it can be summarised as follows:

(1) <Αα Ββ Γγ Γ̲γ̲ Δδ Εε Ζζ Ζ̲ζ̲ Ηη Θθ Ιι Κκ Λλ Μμ Νν Ξξ Ξ̲ξ̲ Οο Ππ Π̲π̲/Bb Ρρ Σσς Σ̲σ̲ς̲ Ττ Τ̲τ̲/Dd Υυ Φφ Χχ Ψψ Ψ̲ψ̲>

An alphabet based primarily on the Roman script (with the addition of letters from the Greek script) was proposed by linguist Alexander Borg and was officially adopted by the community in a ceremony in 2007 (Armostis, 2020; Karyolemou and Armostis, 2020). This alphabet was subsequently used in community contexts, including the 'Maronite Press' newspaper. Borg, being a speaker of Maltese, a peripheral variety whose alphabet uses the Roman script, modelled his Cypriot Arabic alphabet on the Maltese one, hence its primarily Roman character. Borg's alphabet, as presented in Borg (2008), is replicated below:

(2) <Aa, Bb, Cc, Ċċ, Dd, Δδ, Ee, Ff, Gġ [or Gg], Xx, Ii, Jj, Kk, Ll, Mm, Nn, Oo, Pp, Rr, Ss, Şş, Tt, Θθ, Uu, Vv, Ww, Yy, Zz>[1]

1. Arguably, 'Gġ [or Gg]' found in Borg (2008) was a typo; one would have expected 'Ġġ [or Gg]'.

However, the practical application of this alphabet displayed considerable graphematic variation (Armostis, 2020), a pattern also observed in the orthographic transcription of the oral corpus compiled within the revitalisation programme (Armostis, 2025). The main reason for this inconsistency was that only the alphabetisation component of the graphisation process was provided, i.e., the creation and adoption of the alphabet, whereas the second essential component, namely a coherent set of orthographic rules governing its use, was largely absent, beyond basic graphophonemic correspondences. For instance, in the leaflet (Borg, 2008) in circulation within the community about his proposed alphabet, it was specified that the letters <ċ> and <ġ> represent the affricate consonants [t͡ʃ] and [d͡ʒ] respectively, but only in loanwords. However, it was not explicitly stated how the same affricates would be spelled in inherited words. In the sample text provided in the leaflet for how the alphabet could be applied, it could be found that, in inherited words, these affricates were spelled with the digraphs <tş> and <tj> respectively. Additional limiting factors in the application of Borg's proposal included the scarcity and fragmentation of publicly available documentation about the alphabet, the absence of established letter names, and the fact that speakers had not been systematically consulted regarding their preferences.

More generally, the creation of a signary and orthography for previously unwritten or underwritten varieties is often pursued as a largely top-down process with limited speaker participation. However, as speakers preferences often take precedence over a linguistically ideal orthography (Cahill, 2014; Cahill and Karan, 2008), the research programme opted for a full graphisation process that would involve community members in reaching a signary and orthography proposal, which would be put to test with native speakers. Hence, one of the primary objectives of the Cypriot Arabic documentation and revitalisation programme was the development not only of a revised and fully specified signary, building on and reassessing Borg's earlier alphabetic proposal, but also of a comprehensive orthographic framework (Armostis, 2025). With respect to the signary, this involved evaluating whether the existing alphabet adequately represented all phonemic contrasts of the language and making decisions concerning alphabetical order, letter names, and the inventory of punctuation and auxiliary symbols. With respect to orthography, the aim was to formulate explicit rules covering graphophonemic correspondences and their balance with morphographic considerations, orthographic depth, the marking of suprasegmental features such as stress and consonant gemination, capitalisation, orthographic word division and hyphenation, contraction, loanword spelling, and punctuation practices (ibid.; Coulmas, 1996, p. 379–380). Establishing such conventions was necessary both for ensuring consistency in corpus transcription and for supporting the production of

teaching materials for younger and adult learners within the broader framework of acquisition planning.

1.3. Research Motivation

In designing a writing system from a linguistically informed perspective, a central question concerns the appropriate level of representation: whether priority should be given to surface phonetic forms or to underlying phonemic and morphophonological structure, and the extent to which morphological or etymological information should be encoded. Although linguists typically aim at maximising structural adequacy and internal consistency, linguistic soundness alone does not guarantee the success of a newly proposed writing system. Even a theoretically optimal system cannot be considered 'ideal' if it is not acceptable to its intended users, usable in everyday practice, and teachable within realistic educational settings. Research on orthography development has repeatedly shown that acceptability, usability, and learnability, alongside educational, sociopolitical, and practical production factors, are decisive for successful implementation (cf. Cahill, 2014; Cahill and Karan, 2008). These factors include, among others, the transfer of graphetic and graphematic conventions from familiar writing systems (Malone, 2004; Smalley, 1964), as well as technological support such as Unicode encoding, keyboard and font availability, and spellchecking tools. For this reason, proposals developed exclusively on theoretical linguistic grounds must be empirically tested with native speakers, and user preferences and attitudes must be taken into account before a system is finalised (Armostis, 2025; Karan, 2013).

1.3.1. *Morphophonological Processes Under Investigation*

A writing system is analysed as an ordered language-script pair (Justeson, 1976; Weingarten, 2011) consisting of a signary and associated graphematic correspondences between the language and its written form (Daniels and Bright, 1996, p. xlv), part of which may be codified in an orthographic module. A central design issue is representational depth, namely whether spellings should be phonetic or morphophonological.

As a Semitic language, Cypriot Arabic exhibits templatic morphology characterised by the use of consonantal roots, i.e., abstract discontinuous consonant sequences which, through the insertion of transfixes (vowels and additional consonants within the root), yield inflected and derived word forms. Depending on the morphophonological environment, the consonants of a root may undergo systematic (morpho)phonological processes. For example, the root |ktp| 'write' appears unchanged in (3a),

partially modified in (3b), and doubly modified in (3c), due to a Cypriot Arabic rule whereby plosive clusters undergo dissimilation with respect to continuancy (Armostis, 2020):

(3) The root |ktp|

a. /'**kitẹp**/ → ['**citẹp**] 'he wrote'

b. /'pɐ**ktọp**/ → ['pɐ**xtọp**] 'I write'

c. /**ktup**t/ → [**xtuf**t] 'I wrote'

The question that arises is whether the forms in (3b) and (3c) should be written morphophonologically as <paktop> and <ktupt> or phonetically as <paχtop> and <χtuft>. The first option gives the writing system greater *orthographic depth/opacity* (Coulmas, 2003, p. 213–215; Rogers, 2005, p. 275), in that different phonetic allomorphs of the same morpheme are written in the same way, something that makes their connection apparent in writing. The second option increases *orthographic transparency*, in that it more faithfully represents the surface pronunciation.

A second phenomenon of interest concerns the spelling of the definite article. The underlying form of the definite article in the nominative, accusative, and dative is the consonant /l/, which phonetically assimilates fully to the consonant of the following word when that word begins with a singleton consonant (rather than a geminate or a consonant cluster), with the exception of pharyngeal [ʕ], which never occurs as a geminate in Cypriot Arabic (Christodoulou, Armostis, and Karyolemou, 2021):

(4) The definite article /l/

a. /l#'ʕɐfɐ/ → ['lʕɐfɐ] 'the snake'

b. /l#xɐɾɾ/ → [xːɐrː] 'the heat'

c. /l#kɐ'fẹ/ → [kʰːɐ'fẹ] 'the coffee'[2]

d. /l#'pɐpẹ/ → ['bːɐpẹ] 'the door'

In examples (4b)-(4d), the phonetic realisation differs from the underlying morphophonological form, and the question arises whether the

2. In Cypriot Arabic, the geminate counterparts of the voiceless occlusives [p], [t], [k], [c], [ʧ] are enhanced by aspiration, i.e., [pʰː], [tʰː], [kʰː], [cʰː], [ʧʰː]. This phenomenon is an influence from Cypriot Greek, in which these occlusives are also enhanced by aspiration when geminate (Armostis and Karyolemou, 2023). It should be noted that Cypriot Arabic additionally exhibits two geminate voiced plosives, namely [bː] and [dː] (which do not exist in Cypriot Greek), with no singleton counterparts (singleton [b], [d], and [g] do appear, but only as allophones of /p/, /t/, and /k/ respectively when immediately preceded by a voiced consonant: e.g., /pi'ntuɐ/ → [pi'nduɐ] 'little girl'; cf. [pnɐt] 'daughters').

article should be written as pronounced in each case or whether it should always be written unchanged in its underlying form. For instance, (4b) may be represented morphophonologically as <l-χarr> or phonetically as <χ-χarr> (or <χχarr>, if the hyphen is not used). If (4c) is written more phonetically as <kkafé> (without the hyphen), the phrase 'the coffee' becomes homographic with <kkafé> 'cafeteria', whereas a morphophonological spelling <lkafé> (and even better: <l-kafé>) keeps the two forms distinct and encodes additional morphographic information (cf. Rogers, 2005, p. 275). The situation is even more complex in (4d), where assimilation of /l/ to /p/ obligatorily yields [bː] rather than [pʰː]; thus the phonetic spelling <bbape> / <b-bape> departs even further from the underlying form (and from the form without the article, i.e. <pape>) than in the previous examples, whereas the morphophonological spelling <lpape> / <l-pape> may facilitate visual word recognition (Armostis, 2020). Also, even though the assimilation of /l/ to the coronal /t/ typically yields a voiceless geminate [tʰː] (e.g., /l#tik/ → [tʰːik] 'the rooster'), in a couple of lexical items the result of this assimilation is rather a voiced [dː]: for instance, /l#tist/ → [dːist] 'the cauldron'. The reasons for both the regular /l#p/ → [bː] assimilation and the exceptional cases, such as /l#tist/ → [dːist], can be found in the diachrony of the language (Borg, 1985). These irregular cases also pose the same problems as the /l#p/ → [bː] assimilation for the written representation of such phrases: should it be written phonologically as <l-tist> (cf. the form without the article: <tist>) or phonetically as <d-dist>?

Even in case (4a), where phonological and phonetic form do not differ, the issue still arises of how the article should be written: (a) attached to the following word, i.e., <lcafa>; (b) as a separate word, i.e., <l cafa>; (c) with an apostrophe, i.e., <l'cafa> (a practice attested in usage); or (d) joined with a hyphen, i.e., <l-cafa>. Writing the definite article /l/ as a separate word results in a word spelled without a vowel letter (i.e., <l cafa>), which may appear unusual, since phonological words are expected to contain a syllabic nucleus. Writing it attached (i.e., <lcafa>) conforms to phonotactic structure (the cluster /lʕ/ is allowed word-initially: e.g., [lʕɐft] 'I played'), but is not morphologically transparent, as it does not indicate that the article is a separate morpheme. The apostrophe is not optimal (i.e., <l'cafa> or <l' cafa>), since it conventionally signals vowel elision, which is not the case here. Hyphenation indicates that the sequence forms one phonological word but two morphological units. In our proposal, hyphenation was adopted (as also done by Borg), with /l/ written invariantly as <l-> in all contexts: e.g., <l-cafa>, <l-χar>, <l-kafé>, <l-pape>, <l-tist>. These various ways of spelling the definite article were included in the test design of the present study.

A third phenomenon investigated in the present study concerns a small set of voiced 'fricative + plosive' clusters, namely [ðb], [γb], and [γd]. These clusters present a particularly interesting challenge for or-

Table 1. Examples of Cypriot Arabic words containing voiced 'fricative + plosive' clusters (adjusted from Borg, 1985, p. 15)

surface form	Borg's underlying representation	reconstructed historical form	etymology
[ɣbiɾ] 'big (m.)'	← /x̟p̟iɾ/	< *ɡbiːɾ	< (Old Arabic) kabiːɾ
[ˈɾɐɣbe̟] 'neck'	← /ˈɾɐx̟p̟e̟/	< *ɾɐɡbe̟	< (Old Arabic) ɾaqaba
[ˈɾe̟ɣbe̟] 'riding (f.)'	← /ˈɾe̟x̟p̟e̟/	< *ɾe̟ɡbe̟	< (Old Arabic) ɾaːkiba
[ðbɐxt] 'I slaughtered'	← /t̟p̟ɐxt/	< *dbɐxt	< (Old Arabic) ðabaḥt
[pɐˈðbun] 'river'	← /pɐˈθ̟p̟un/	< *pɐˈtmun	< (Cypriot Greek) po̟tɐˈmo̟n
[ɣdiʃ] 'horse'	← /x̟t̟iʃ/	< *ɡdiːʃ	< (Old Arabic) kadiːʃ
[ˈmo̟ɣde̟] 'stove'	← /ˈmo̟x̟t̟e̟/	< *mogde	< (Old Arabic) mawqida

thographic representation, because their synchronic phonological status is far less straightforward than that of the processes discussed above: unlike the alternations involving continuancy dissimilation or full assimilation of the definite article /l/, which can be analysed in terms of productive morphophonological processes, the voiced clusters appear to be the result of historical sound change and have a highly restricted lexical distribution. Such items are presented in Table 1, adjusted from Borg (ibid., p. 15).

Borg (ibid., pp. 15–16) postulated that, synchronically, Cypriot Arabic exhibits a voiced plosive in the underlying representation of such clusters, as, historically, it was the second member of the cluster that was phonemically voiced: e.g., [ɣbiɾ] 'big (m.)' ultimately goes back to Old Arabic /kabiːɾ/. However, postulating voiced plosives in the phonemic form is problematic, as Cypriot Arabic does not have an underlying voicing contrast for singleton plosives (see footnote 2): it exhibits only a series of voiceless plosive phonemes, namely /p/, /t/, /k/, which surface as voiced when immediately following another voiced consonant (ibid., p. 12). The alternative, albeit not historically motivated, would be to assume that, synchronically, the first member of the cluster is underlyingly voiced, to which the second member assimilates. However, this solution is also problematic, as Borg (ibid., p. 36–38) argues that Cypriot Arabic does not synchronically exhibit the Old Arabic /ɣ/ phoneme due to it having merged with /ʕ/ in Cypriot Arabic (e.g., Old Arabic /ɣalaq/ is the origin of Cypriot Arabic [ˈʕilek] 'he shut'). Consequently, it is not al-

ways evident which segment should be regarded as underlyingly voiced from a synchronic perspective, nor is it clear whether the surface clusters should be analysed as the output of an active phonological process converting the voiceless cluster into a voiced one, as such a rule would be active only if historically motivated: e.g. it would apply for /xtiʃ/ → [γdiʃ] 'horse' (cf. Old Arabic /kadiːʃ/), but not for /ktupt/ → [xtuft] (cf. Old Arabic root |tkb|). This conundrum led Borg (1985) to represent the voiced 'fricative + plosive' clusters underlyingly with voiceless phonemes, which however have exceptionally received voicing in certain lexical items due to historical reasons, e.g., /χpiɾ/.

This theoretical opacity of the synchronic underlying form of such clusters has consequences on their written rendition. If a more surface representation is adopted, then spellings such as <γbir>, <paδbún>, <γdiş> etc., are straightforward, as their phonetic form unambiguously contains a voiced fricative followed by a voiced plosive; but, if a deeper representation is adopted, then, apart from the fact that from a theoretical perspective the phonological form is unclear, an additional issue would be how deep the written representation should be with regards to the morphophonology or even etymology of a given item: e.g., would a deeper spelling for [γdiʃ] be <χtiş>, <χdiş>, <gdiş>, or <kdiş>? The phenomenon therefore provides an additional test case for examining speakers' preferences regarding orthographic depth and the extent to which historical, morphophonological, or phonetic information should be represented in the writing system.

1.3.2. *The Proposal*

The proposal of the research programme that was to be tested was the following:

Regarding the signary, it was developed by making decisions about the alphabet, punctuation (e.g., using the Roman question mark <?> rather than the Greek one <;>), numerals (Western Arabic numerals were selected), decimal separator etc. For the purposes of the test of the current study, only the alphabet part of the signary was presented to the participants. The test alphabet was a modification of the one created by Borg:

(5) <Aa Bb Cc Dd Δδ Ee Ff Gg Γγ Ii Jj Kk Ll Mm Nn Oo Pp Θθ Rr Ss Şş Tt Uu Vv Ww Xχ Yy Zz>

This version of the alphabet was not meant as a finalised proposal, but was rather put together to be tested in order to be modified should this investigation yield serious difficulties with certain aspects thereof.

Regarding the orthographic component, certain spelling conventions were included in the materials to be presented to the participants of this

study. Apart from basic graphophonemic correspondences, the proposal contained rules for marking stress, writing diphthongs, affricates, and voiced plosives. Regarding voiced plosives, the proposed rule presented was that they are written with <b> and <d> only when geminate: e.g., <Rabbi> for ['ɾɐbːi] 'my God'; <şadda> for ['ʃɐdːɐ] 'he tightened it (f. object)'. The rule for phonetically singleton voiced plosives [b], [d], and [g] was that they always occur after other voiced consonants and are represented with the letters for voiceless consonants, i.e., with <p>, <t>, <k> respectively: e.g., <mkass> for [mgɐsː] 'scissors'. The only exception to this rule was the representation of the voiced 'fricative + plosive' clusters: in our proposal, we opted for the unambiguous surface representation, i.e., <δb>, <γb>, <γd>, for example, <δbaχt>, <γbir>, <γdiş>. Regarding the definite article and continuancy assimilation, no explicit proposal was made; the reason for this was for us to be able to observe how native speakers would represent the two phenomena without being influenced by our preferences.

1.3.3. *Research Questions*

Based on the theoretical and practical considerations outlined above, the present study formulated the following research questions concerning the graphisation of Cypriot Arabic:

1. Which script (or scripts) should form the basis of the Cypriot Arabic signary? More specifically, given the literacy backgrounds of the community and their familiarity with particular writing traditions, which graphematic resources are most appropriate for the development of a usable and sustainable writing system?
2. Which level of linguistic representation should a linguistically well-motivated orthography primarily respect? Should it represent:
   (a) surface phonetic forms, resulting in a more transparent orthography; or
   (b) underlying morphophonological representations, resulting in greater orthographic depth (and increased morphography)?
   (c) And should the selected level of representation (surface vs underlying) be uniformly applied to all phenomena of the language?
3. How acceptable, usable, and teachable is the proposed writing system as evaluated by members of the speech community themselves?

Together, these questions link structural linguistic considerations with user-centred criteria and treat graphisation both as a technical design task and as a socially embedded process.

2. Methodology

To answer the above research questions, an empirical study was designed, as presented below.

2.1. Materials

Data were collected with the use of a questionnaire. The questionnaire was administered in printed form, while the instructions were presented on a laptop screen using MS PowerPoint. Each test word or phrase was presented aurally to the participants through headphones.

2.1.1. *The Questionnaire*

The questionnaire consisted of three main parts: (1) a pre-test; (2) a brief teaching session presenting the basic rules of the proposed writing system along with a questionnaire for evaluating the proposal; and (3) a post-test. Its overall design aimed to elicit participants' spontaneous spelling practices, measure their preferences when confronted with alternative spellings, and gauge their evaluation of the proposed alphabet and orthography after minimal instruction.

The pre-test comprised two tasks. In the first task, participants listened to recorded Cypriot Arabic words and short phrases and were asked to write each item down by hand in whatever way they considered appropriate. This task was intended to capture uninstructed spelling strategies and script choices. In the second task, participants again listened to recorded items, but this time a forced-choice format was followed, whereby they were asked to select their preferred spelling among a set of proposed alternatives. These alternatives reflected different graphematic choices (e.g., more phonetic vs more morphophonological spellings, different boundary-marking options) etc.

The second part of the questionnaire consisted of a very brief instructional presentation of the proposed writing system, including the alphabet (signary) and its principal orthographic conventions. Immediately following this introduction, participants completed the second part of the questionnaire evaluating the proposed system. Using five-point Lickert-scaled responses, they were asked to indicate: how much they liked the orthographic rules presented to them; the degree to which they agreed with these rules; to what extent they would personally follow them when writing in Cypriot Arabic; to what extent they believed that all users should follow these rules; how logical they found the rules; and how easy they considered them to be learnt.

TABLE 2. The target determiner phrases used in each part of the test for the investigation of the written rendition of full assimilation of the definite article

Part	target phrase	Gloss
pre-test (a)	/l#ˈvɐɾɐk/ → [ˈvːɐɾɐk]	the books
	/l#tik/ → [tʰːik]	the rooster
pre-test (b)	/l#ˈʒɐ̯ʒʒɐ̯/ → [ˈʒːɐ̯ʒːɐ̯]	the hen
	/l#kɐˈssis/ → [kʰːɐˈsːis]	the priest
post-test (a)	/l#xɐpl/ → [xːɐpl]	the rope
	/l#tist/ → [dːist]	the cauldron
	/l#ˈpɐpɐ̯/ → [ˈbːɐpɐ̯]	the door
post-test (b)	/l#ˈθɐu̯p/ → [ˈθːɐu̯p]	the shirt
	/l#pint/ → [bːint]	the daughter

The post-test also consisted of two tasks, presented in the reverse order compared to the pre-test: participants first completed a forced-choice spelling task based on recorded words, again selecting among proposed spellings; they then performed a final handwriting task, writing down the recorded items. This reversed ordering allowed for comparison of preferences and production before and after exposure to the proposed system and its rules.

2.1.2. *Target Words*

The words and phrases used in the study were embedded as audio files in the PowerPoint file used for the test. To create this material, EZ, a 45-year-old male native speaker of Cypriot Arabic, was recorded in a quiet room. EZ was familiar with the proposed writing system, as he was a member of the research project for the documentation and preservation of Cypriot Arabic. The words and phrases under investigation were presented to the speaker in writing. EZ was instructed to pronounce the words and phrases clearly and to avoid 'list intonation' by pronouncing each one separately.

The present study formed part of a wider investigation, the design of which included a wide range of phenomena to be examined through a total of 129 words and phrases serving as the test material. The current study though was confined to the investigation of three specific phenomena: (1) the complete post-lexical assimilation of the definite article to the following word-initial consonant within a determiner phrase (Table 2); (2) continuancy dissimilation of two-plosive clusters (Table 3); and (3) the voicing of 'fricative + plosive' clusters (Table 4).

TABLE 3. The target words used in each part of the test for the investigation of the written rendition of continuancy dissimilation of two-plosive clusters

Part	target word	Gloss
pre-test (a)	/pkum/ → [fkum]	I get up
	/ptɐˈmin/ → [ftɐˈmin]	rivers
pre-test (b)	/sɾukt/ → [sɾuxt]	I stole
	/tlupt/ → [tluft]	I asked
post-test (a)	/sukt/ → [suxt]	I ploughed
	/ˈtkuɾɐ/ → [ˈθkuɾɐ]	males
post-test (b)	/ktilt/ → [xtilt]	I killed
	/ktupt/ → [xtuft]	I wrote

TABLE 4. The target words used in each part of the test for the investigation of the written rendition of the voiced clusters [ðb], [ɣb], and [ɣd]

Part	target word	Gloss
pre-test (a)	[ðbɐxt]	I slaughtered
	[ɣbiɾ]	big (male-singular)
pre-test (b)	[ðbuˈnɐt]	straw (plural)
	[ˈɾeɣbe]	on horseback (female adjective)
post-test (a)	[ɣdiʃ]	horse
	[pɐˈðbun]	river
post-test (b)	[ˈmo̝ɣde̝]	oven
	[ˈɾɐɣbe̝]	neck

2.2. Participants

Twenty-two participants (10 men and 12 women) took part in the study, ranging in age from 39 to 72 (mean: 52.3; SD: 8.19). Two participants had received only primary education, 12 secondary education, and 8 tertiary education. All participants had Cypriot Arabic as their first language and were bilingual with Cypriot Greek.

2.3. Procedure

Before the main part of the study began, participants filled out a questionnaire about their demographics, as well as about whether and to what extent they use written Cypriot Arabic, and also about whether they know other languages and to what degree.

They were then given the first printed section of the questionnaire with the instruction to freely write down with a pen the word/phrase recordings they would hear through headphones connected to the laptop displaying the PowerPoint show. It was emphasised to them that this was not a test, as there is no right or wrong answer; on the contrary, it

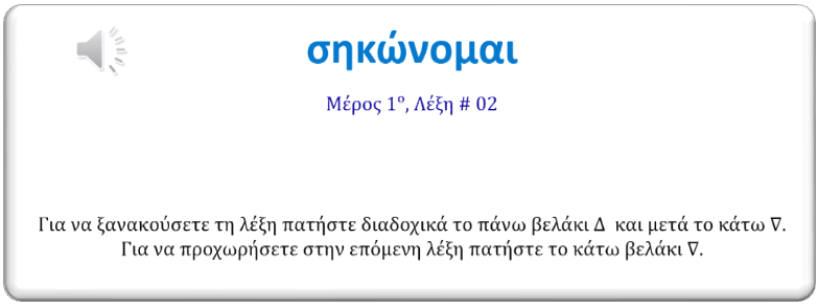

FIG. 1. A slide from the PowerPoint presentation used during the test

was stated that the researchers wanted to learn from the participants regarding their preferences in rendering Cypriot Arabic in writing. Each PowerPoint slide displayed the serial number of the word being played as well as its translation into Standard Modern Greek. In addition, instructions were provided on how to replay a word if they wished (see Figure 1). As each slide appeared, the corresponding audio clip began playing automatically.

2.4. Analysis

In each part of the test, the rates of written representation of the underlying morphophonological level vs the surface phonetic level of the three phenomena under examination (i.e., plosive cluster continuancy dissimilation, full assimilation of the definite article to the following consonant, and the voicing of 'fricative + plosive' clusters) were measured. Regarding the evaluation of the proposed writing system, as responses were given on a five-point Likert scale (1–5), the appropriate measures of central tendency and dispersion for such ordinal scale were calculated, namely, the median (*Mdn*), mode (*Mo*), and range.

3. Results

3.1. Evaluation of the Proposed System

As shown in Figure 2, participants rated the proposed writing system as fairly logical (*Mdn* = 4; *Mo* = 4) and stated that they somewhat liked it and would probably use it (*Mdn* = 3.5; *Mo* = 4 for both). They did not seem to have a strong opinion on whether they agreed with this proposed system (*Mdn* = 3; *Mo* = 3), and there was not much variation in their responses (range = 2–4).

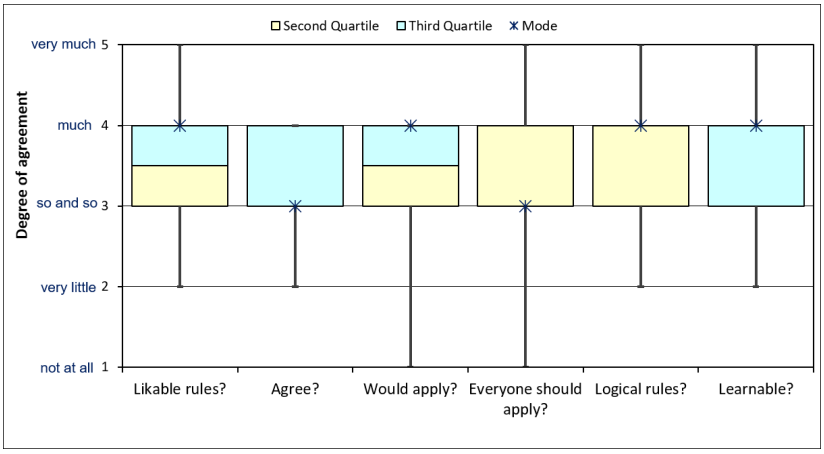

FIG. 2. Boxplots of evaluation of the proposed system

Furthermore, they did not have a strong opinion on how easy it is to learn this system, although the most frequent answer was that it is easy to learn (*Mdn* = 3; *Mo* = 4). Finally, they seem to agree that the proposed system should be implemented by everyone, although the most frequent answer was that they did not have a strong opinion on the matter (*Mdn* = 4; *Mo* = 3), while there was a wide range of responses (range = 1–5).

3.2. Graphematic Choices

Figure 3 shows the graphematic choices of the participants in the parts in which they could write down the target words and phrases by hand, i.e. in the 'pre-test (a)' and the 'post-test (b)'. Regarding script choice, they would initially write in entirely the Greek script at a rate of 34.3% at the very start of the test; this percentage dropped to 14.9% at the very last part of the test. Regarding the use of the proposed alphabet and spelling conventions, their choices followed the proposed system (even before it was introduced to them) at a rate of 10.3% at the beginning of the test (this number counted only segmental graphemes ignoring the use of stress marks); at the end of the test, i.e. after they had been briefly taught the proposed conventions, this use of the proposed system increased to 18.2%. In the vast majority of the cases, they would use mixed conventions, i.e. they would oscillate between the proposed and ad hoc conventions, or they would use completely ad hoc conventions sometimes influenced by the writing systems of foreign languages they had learnt through education, such as English, French or Italian; as expected, no use of the Arabic script emerged. The percentage of such mixed use

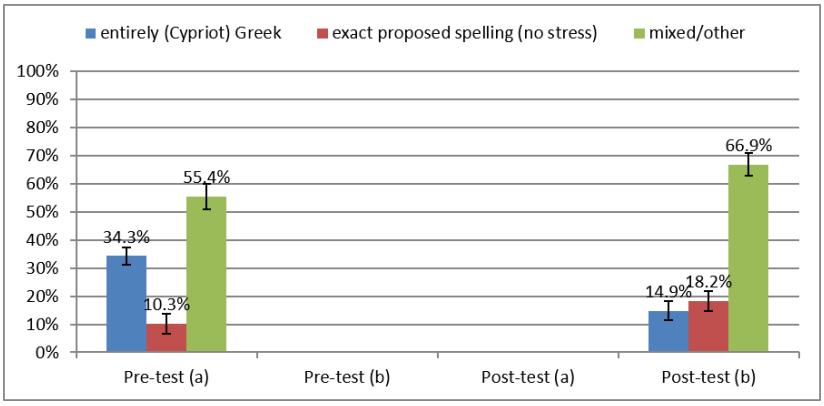

FIG. 3. Percentages of use of the proposed orthography vs. alternatives in the non-forced-choice parts of the test. [Note: In this and the following graphs, error bars represent standard error for the mean.]

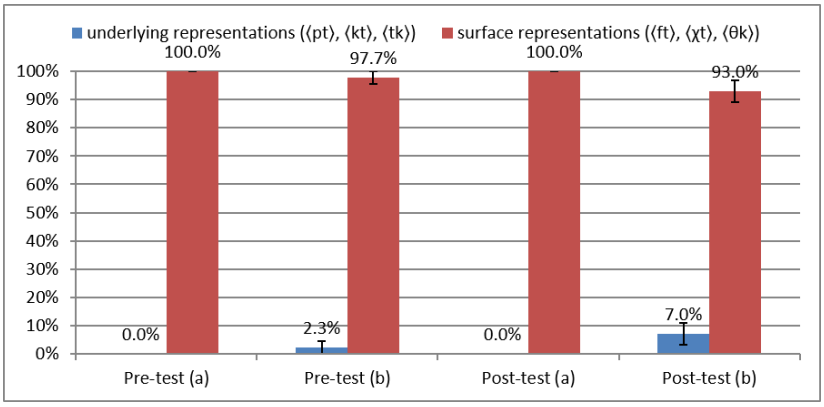

FIG. 4. Percentages of underlying vs surface spelling of two-plosive clusters

was higher at the end of the test compared to what the situation was at the beginning.

With regards to the written representation of two-consonant clusters consisting phonologically of two plosives, a clear preference was shown (M = 98.1%; SD = 13.7%) for representing the surface phonetic result of the continuancy dissimilation process (e.g., [xt] rendered as <χt>) rather than the underlying form (e.g., /kt/ as <kt>). As shown in Figure 4, phonetic spelling was the most prevalent in all parts of the test.

In the case of the definite article, phonetic spelling was also generally preferred (e.g. [tʰːik] rendered as <t-tik>) over the morphophonological one (e.g., /l#tik/ as <l-tik>), but not quite as clearly (M = 69.8%; SD =

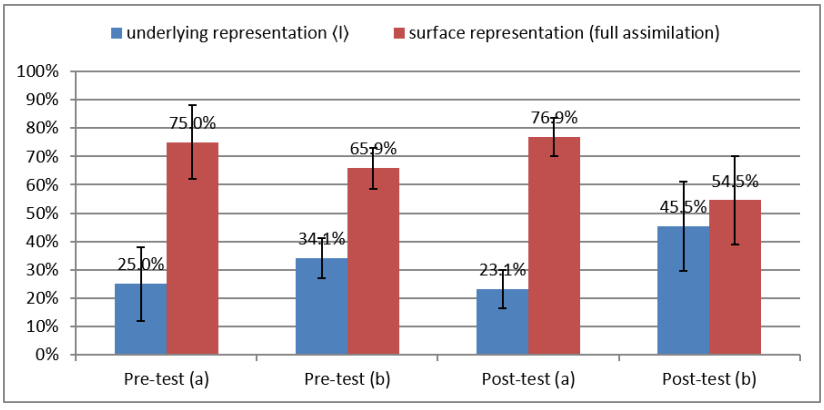

FIG. 5. Percentages of underlying vs surface spelling of the definite article /l/

46.1%). As shown in Figure 5, the morphophonological spelling <l> was used between 23.1% and 45.5%, with the highest percentages recorded in the final part of the test.

As for the rendering of the word boundary between the definite article and the noun it modifies, in the pre-test (a), where participants wrote freely, the main tendency was to write them together as a single word (e.g., <ltik> or <ttik>). In post-test (b), writing with the hyphen (e.g., <l-tik>) became the most prevalent representation. As shown in Figure 6, the other spellings declined in favour of the hyphenated spelling; the only exception was the spelling without the article, which consistently had high percentages in both parts. The use of a simple typographical space between the two words (e.g., <l tik>) was not used in either part. It should be noted that Figure 6 does not include the forced-choice tasks, as in those ones the options given always used the hyphen (e.g., <l-kassís> vs <k-kassís>).

The results regarding the representation of the voiced 'fricative + plosive' clusters [ðb], [ɣb], and [ɣd], are presented in Figure 7. In both the initial handwriting task and the following forced-choice task of the pre-test, the representation of only the fricative as voiced and the plosive as voiceless (e.g., <δpaχt>, <ɣpir>, and <ɣtiş>) was clearly preferred. In the forced-choice task of the post-test immediately after the presentation of the proposed orthographic rules, there was a reversal of preferences, with the proposed surface representations (e.g., <δbaχt>, <ɣbir>, and <ɣdiş>) being favoured. In the final handwriting task though, they reverted to the preference of the 'voiced fricative + voiceless plosive' representation. It should be noted that spellings representing both members of the cluster as voiceless were minimally selected only in the forced-choice tasks.

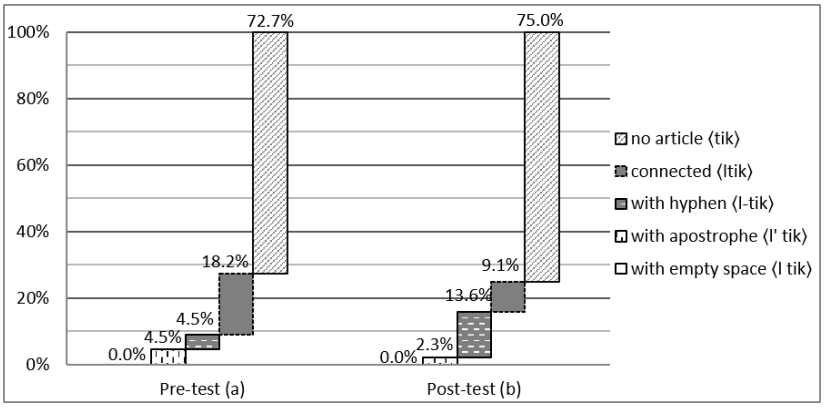

FIG. 6. Percentages of the different ways in which the definite article /l/ was linked in writing to the following noun per the two non-forced-choice parts of the test

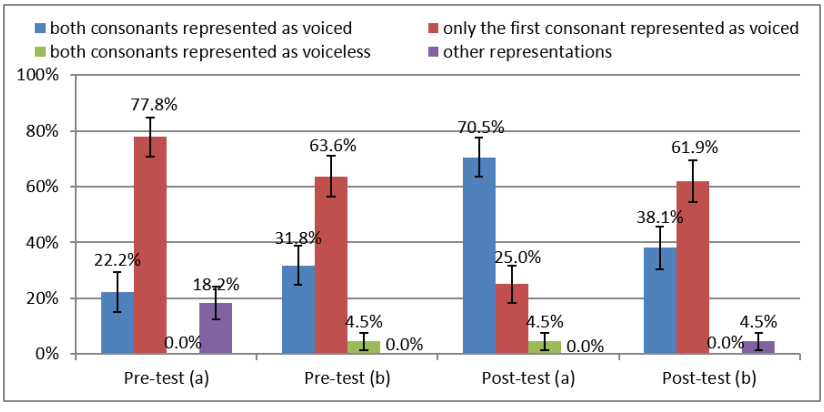

FIG. 7. Percentages of various ways in which the [ðb], [ɣd], [ɣb] clusters were represented in writing

4. Discussion

The present study set out to investigate three interconnected questions regarding the graphisation of Cypriot Arabic: which script should form the basis of the proposed signary and associated orthography, which linguistic principles should govern the orthography, and to what extent native speakers find the proposal logical, learnable, and likeable. The results provide insights into all three issues and demonstrate that graphisation cannot be reduced to solely the pursuit of a linguistically ideal orthography; rather, as argued in the language-planning and

grapholinguistic literature (Cahill, 2014; Cahill and Karan, 2008; Haarmann, 2006), the successful implementation of a proposed orthography depends on balancing structural adequacy with acceptability, usability, and learnability.

With regard to the choice of script, the results confirm the importance of speakers' literacy backgrounds. No participant made use of the Arabic script, despite the fact that Cypriot Arabic is an Arabic variety. This finding is unsurprising given that community members are literate in Standard Modern Greek and familiar with the Roman script through foreign-language education, but generally have no literacy in Modern Standard Arabic (with the exception of a few who learnt it as a foreign language). The spontaneous use of Greek-script spellings at the beginning of the test and the frequent adoption of mixed conventions drawing on both Greek- and Roman-script practices further illustrate the tendency of writers to transfer graphematic conventions from familiar writing systems. At the same time, the increased use of the proposed alphabet following even a brief instructional intervention suggests that script choice is not determined solely by prior literacy experience, but can be influenced through teaching and standardisation efforts.

The most theoretically interesting findings concern the question of orthographic depth and the level of linguistic representation that a writing system should respect. As discussed earlier, a central issue in orthography design is whether spellings should (consistency) privilege phonetic transparency vs. morphophonological deeper representations. The results demonstrate that the answer cannot be reduced to a simple choice between a shallow and a deep orthography. Rather, speakers showed different preferences for different types of morphophonological phenomena.

The strongest preference observed in the study concerned the representation of continuancy dissimilation in dyadic plosive clusters. Participants overwhelmingly preferred spellings reflecting the surface phonetic outcome of the process rather than the underlying morphophonological representation. Thus, forms such as /ktupt/ → [xtuft] were overwhelmingly written as <χtuft> rather than <ktupt>. This finding can be interpreted within the framework of Lexical Phonology (Kiparsky, 1982). Lexical phonological rules are integrated into speakers' lexical representations and are often sensitive to morphological structure. Consequently, the outputs of such rules may acquire psychological reality for speakers and be perceived as the 'actual' form of a word. In Cypriot Arabic, continuancy dissimilation in plosive clusters exhibits precisely such characteristics and is not merely an automatic (i.e., post-lexical) phonological process: more precisely, the dissimilation rule is supressed in lexical items such as ['ʕɐptɐ] '(Santa) Maura', ['piptẹl] 'I change', [ptaɾ] 'seed'; as Borg (1985, p. 18) observed, there is even a minimal pair, namely [ptilt] 'I changed' (root: |ptl|) vs [ftilt] 'I twisted' (root: |ftl|), with the for-

mer suspending the dissimilation rule arguably for remaining distinct in form from the latter. The existence of lexical exceptions, the morphological conditioning of the process, and also the functional load of the process due to the existence of the above-mentioned minimal pair, suggest that speakers store and recognise the resulting forms as part of their lexical knowledge. It is therefore unsurprising that they strongly preferred surface representations such as <χtuft> and <paχtop>.

A different pattern emerged in the case of the definite article. Although phonetic spellings were generally preferred throughout, the preference was substantially weaker, and morphophonological spellings increased once this option was presented to the participants in pre-test (a) and post-test (b). Importantly, as mentioned above, in the presentation of the proposed rules, no explicit preference for the writing of the definite article was introduced to the participants; this means that the substantially increased percentage of selecting the orthographically deep option in post-test (b) may be attributed to it being a rather intuitive option. This interpretation may be supported by the fact that in Borg's (2008) leaflet for his proposed alphabet, the phonetic level of representation was used in writing the definite article (e.g., <t-tarp> 'the road'), as is the case in Maltese (e.g., <it-tifel> 'the child'); if Cypriot Arabic speakers had already been familiar with Borg's (2008) conventions, then their increased percentages of using a deep representation by the end of the test is quite revealing. This result is likewise compatible with a Lexical Phonology interpretation: assimilation of the definite article /l/ occurs automatically across word boundaries and therefore constitutes a post-lexical process. Unlike lexical rules, post-lexical rules are typically exceptionless,[3] automatic, and less accessible to conscious awareness, because they apply after lexical representations have been assembled. Consequently, speakers may be more willing to preserve the underlying representation of the article in writing, especially when doing so enhances morphological transparency. The increased acceptance of forms such as <l-tik> by the end of the test therefore suggests that speakers recognise the value of maintaining a fixed image for the writing of this grammatical morpheme, despite the fact that its phonetic realisation varies, exactly because the phonetic outcome is completely predictable by the phonological rules of the language.

A somewhat different pattern emerged in the case of the voiced 'fricative + plosive' clusters [ðb], [ɣb], and [ɣd]. Unlike the phenomena discussed above, these clusters do not correspond straightforwardly to an active morphophonological process. As outlined in section 1.3.1, their

3. Even the /l#tist/ → [dːist] case, albeit an exception to the /l#t/ → [tʰː] rule, it is not really a suspension of the full assimilation rule, but rather an application thereof with an unusual assimilation outcome; as such, it remains an automatic process and arguably below the level of conscious awareness.

synchronic status is theoretically opaque under Borg's (1985) analysis: although they are phonetically realised with two voiced obstruents, their historical origins differ from item to item, and it is not clear whether their synchronic written representation should preserve information about etymology, an abstract phonological representation, or the surface phonetic form. Consequently, if we follow Borg's (1985) analysis, the phenomenon occupies an intermediate position between synchronically productive morphophonology and lexically restricted historical residue.

The results indicate that participants were generally reluctant to represent both members of the cluster as voiced. In both tasks of the pretest and in the final handwriting task, they clearly preferred spellings such as <γpir>, <δpaχt>, and <γtiş> over the proposed surface spellings <γbir>, <δbaχt>, and <γdiş>. Only immediately after presentation of the proposed orthographic rules did participants temporarily reverse this preference and favour the fully voiced representations. This suggests that, while speakers found the proposed convention understandable and were willing to adopt it when explicitly prompted, it was not their most intuitive solution when writing spontaneously. An explanation for this preference may lie in the phonological structure of Cypriot Arabic itself. As discussed earlier, the language does not possess an underlying voicing contrast among singleton plosives; instead, voiced plosives are normally predictable allophones of /p/, /t/, and /k/ occurring after voiced consonants (Borg, 1985). Participants therefore appear to have extended this general graphematic principle to the exceptional voiced clusters as well. By writing forms such as <γpir> rather than <γbir>, they preserved the salient surface voicing of the fricative while simultaneously maintaining the conventional representation of the plosive with a grapheme corresponding to the underlying voiceless phoneme. At the same time, the near absence of spellings such as <χpir> demonstrates that speakers did not ignore the phonetic voicing of these clusters altogether. Rather than favouring either a fully phonetic representation or a deeper morphophonological-etymological one, they tended to adopt what appeared to be an intermediate solution that captured part of the surface realisation while remaining consistent with broader orthographic principles of the writing system.

This outcome led us to reconsider our proposed <γb>, <δb>, and <γd> spellings and adopt the apparently more intuitive <γp>, <δp>, and <γt> spellings in our final version of the proposed orthography. Part of the reason for this was that we revisited the phonological status of /γ/: Borg (ibid.) did not recognise phonemic status for this segment, as it was thought to be non-contrastive and also historically deriving from a voiceless consonant, such as Old Arabic /k/ or /q/ in clusters with a voiced consonant; however, in preparing our materials for teaching the language, we observed that /γ/ is regularly used in loanwords from Cypriot Greek or Turkish (usually through Cypriot Greek), such as

/ɣɐˈlliɐ/ ‘France’, /ɐzɣɐˈltin/ ‘almost’, /ɣuɾˈsuz/ ‘a jinx male person’ etc., something that makes its phonemic status plausible. Moreover, Borg (personal communication), has recently found a minimal pair exhibiting a contrast between /ɣ/ and /x/, namely [ˈɣdɐʕɐt] ‘I touched’ vs [ˈxtɐʕɐt] ‘I cut (past tense)’; this independent finding provides further support to the idea that /ɣ/ does have functional load and hence should be considered a distinct phoneme in Cypriot Arabic. This justifies our decision to include a letter for it in our proposed alphabet, i.e. <ɣ>, something that had been missing from Borg’s alphabet. Accepting phonemic status for the voiced velar fricative may have a further implication: one may argue that the native speakers’ preference for the <ɣp>, <ðp>, and <ɣt> spellings mirrored their intuition for the underlying representation of these clusters, as, synchronically, they may be represented phonemically as straightforward /ɣp/, /ðp/, and /ɣt/ clusters. This may mean that, even though historically it was the second member of the cluster that was voiced, the clusters may have been reanalysed as consisting of a voiced velar fricative phoneme /ɣ/ or /ð/ and a voiceless plosive phoneme /p/ or /t/. If this analysis is accepted, then the synchronic underlying representation of the [ɣb], [ðb], and [ɣd] clusters becomes considerably more straightforward than under Borg’s (1985) account. More importantly, if the fricatives in the cluster are phonemically voiced, then the voicing of the plosives need not be sought in their diachrony, but rather in a phonological rule of regressive voice assimilation of the voiceless plosive to the preceding voiced fricative. Such a process may plausibly be analysed as post-lexical, insofar as the voicing of the plosive follows automatically from the presence of the preceding voiced fricative and does not appear to be lexically conditioned, as, e.g., /ɣt/ → *[ɣt] would be not an acceptable outcome for any lexical item. If this is indeed a post-lexical process, then the interpretation that the <ɣp>, <ðp>, and <ɣt> spellings clearly represent their phonemic form gains support, as post-lexical processes tend to be represented in their underlying form, as is the case with the definite article. These findings concerning the phonemic status of the voiced velar fricative, together with the results concerning script preferences discussed earlier, as well as other considerations that go beyond the scope of the present study, led to certain modifications of the proposed signary. The final version of the alphabet adopted for the teaching materials is shown below:

(6) <Aa Bb Cc Γγ Dd Δδ Ee Ff Θθ Ii Jj Kk Ll Mm Nn Oo Pp Rr Ss Şş Tt Uu Vv Ww Χχ Yy Zz>

All in all, the results support the view that orthographic design should not be conceived as a binary choice between shallow and deep orthographic representations; instead, different phenomena may warrant different representational strategies. Thus, the emerging Cypriot Arabic

orthography cannot be characterised as either entirely phonetic or entirely morphophonological; rather, it combines shallow and deep orthographic principles depending on the linguistic phenomenon under consideration. The results regarding script preference also confirmed the importance of literacy transfer from familiar writing systems and hence the relevance of sociolinguistic and educational factors in graphisation. The other main outcome of this study pertains to the significance of involving speakers in graphisation processes. Participants generally evaluated the proposed system positively, considering it logical and expressing moderate willingness to use it. More importantly, their responses provided direct evidence regarding which features of the writing system were intuitively acceptable and which required reconsideration. This supports the argument advanced by Karan (2013) and others that writing systems should not be imposed exclusively through top-down expert decisions. Particularly in the case of endangered languages, where community acceptance is crucial for successful implementation, graphisation should be viewed as a collaborative process involving consultation with speakers (Armostis, 2025). Native-speaker preferences offer valuable information not only about attitudes towards particular orthographic options but also about the cognitive salience of linguistic distinctions and the practical usability of competing representations.

The final Cypriot Arabic writing system that emerged from this process therefore reflects both linguistic and social considerations. It is grounded in phonological and morphological analysis, but it is also informed by empirical evidence regarding speaker preferences, literacy practices, and usability. In this sense, the study illustrates how graphisation can function as a form of corpus planning that is simultaneously linguistically informed and community-oriented.

5. Conclusions

The present study examined native speakers' responses to a proposed writing system for Cypriot Arabic and explored how different linguistic phenomena should be represented in the process of graphisation. More specifically, it investigated speakers' preferences regarding script choice, orthographic depth, and the representation of selected morphophonological phenomena, while also assessing the acceptability, usability, and learnability of the proposed system.

The findings demonstrate that the development of a writing system cannot be guided exclusively by theoretical linguistic considerations: although linguistic soundness remains an important criterion, the success of a writing system ultimately depends on whether its intended users find it acceptable, usable, and learnable. The relatively positive attitudes expressed by participants towards the proposed orthography con-

firm the importance of incorporating speaker perspectives into graphisation and orthographic planning, particularly in the context of endangered-language revitalisation. With regard to orthographic depth, the results indicate that a writing system need not be uniformly phonetic or uniformly morphophonological. Rather, different linguistic phenomena may call for different representational solutions: participants tended to favour deeper spellings in cases where the phonological process was post-lexical, while more surface-oriented solutions were preferred in cases whereby the phonological process was lexical. The findings therefore suggest that orthographic depth is not an all-or-nothing property of a writing system but may vary according to the linguistic characteristics of individual phenomena.

More broadly, the study demonstrates the importance of involving speakers in graphisation. Native speakers' preferences provide valuable insight into which orthographic conventions feel intuitive and which create difficulties. Consequently, graphisation should be viewed not as a purely top-down linguistic exercise but as a process in which both linguistic analysis and community perspectives are taken into consideration. The case of Cypriot Arabic graphisation demonstrates the value of empirically evaluating orthographic proposals before their wider implementation. Testing orthographies allows researchers to identify potential difficulties, assess competing design choices, and refine orthographic conventions on the basis of actual speaker behaviour rather than theoretical assumptions alone. Such an approach may be particularly beneficial in the graphisation of endangered languages, where the successful adoption of a writing system is closely linked to broader efforts of language maintenance and revitalisation.

References

Armostis, Spyros (2020). "Η γραπτή απόδοση της κυπριακής αραβικής: μια πειραματική προσέγγιση [The written rendition of Cypriot Arabic: an experimental approach]". In: *Έρευνες για την Ελληνική Γλώσσα [Studies on Greek Language]*. Ed. by Elena Ioannidou, Marianne Katsoyannou, and Charalambos Christodoulou. Nicosia: Cyprus Linguistics Society, pp. 1–19.

———— (2025). "Ethical considerations and good practices in linguistic work on endangered languages: the case of a research programme on Cypriot Arabic". In: *Linguistics* 63.2, pp. 429–460.

Armostis, Spyros and Marilena Karyolemou (2023). "Contact-Induced Change in an Endangered Language: The Case of Cypriot Arabic". In: *Languages* 8.1, p. 10.

Borg, Alexander (1985). *Cypriot Arabic: a historical and comparative investigation into the phonology and morphology of the Arabic vernacular spoken by the Maronites of Kormakiti village in the Kyrenia district of north-western Cyprus*. Stuttgart: Komissionsverlag Steiner Wiesbaden.

———— (1994). "Some evolutionary parallels and divergences in Cypriot Arabic and Maltese". In: *Mediterranean Language Review* 8, pp. 41–67.

———— (2008). "An alphabet for Cypriot Maronite Arabic". unpublished manuscript, https://zenodo.org/records/21086024.

Cahill, Michael (2014). "Non-linguistic factors in orthographies". In: *Developing orthographies for unwritten languages*. Dallas: SIL International, pp. 9–25.

Cahill, Michael and Elke Karan (2008). "Factors in designing effective orthographies for unwritten languages". In: *SIL Electronic Working Papers* 2008-001.

Christodoulou, Charalambos, Spyros Armostis, and Marilena Karyolemou (2021). "The Definiteness Marker of Cypriot Arabic: A Phonological Account". In: *Journal of Semitic Studies* 66.2, pp. 503–521.

Coulmas, Florian (1996). *The Blackwell encyclopedia of writing systems*. Oxford (UK) & Cambridge (Mass.): Blackwell.

———— (2003). *Writing Systems: An Introduction to their Linguistic Analysis*. Cambridge: Cambridge University Press.

Daniels, Peter and William Bright, eds. (1996). *The World's Writing Systems*. New York.

Fenyvesi, Anna and Marianne Bakrá-Nagy (2025). "Endangered Languages and Language Revitalization". In: *The Handbook of Plurilingual and Intercultural Language Learning*. Ed. by Christiane Fäcke, Xuesong (Andy) Gao, and Paula Garrett-Rucks. 1st ed. Wiley, pp. 83–95.

Fishman, Joshua A., ed. (1974). *Advances in language planning*. Berlin, Boston: De Gruyter Mouton.

Frangeskou, Marianna and Alexander-Michael Hadjilyra (2012). *Οι Μαρωνίτες της Κύπρου [The Maronites of Cyprus]*. Nicosia: Press and Information Office, Republic of Cyprus.

Frangiskou, Antonis G. (2000). *Το Λεξικό της Αραβικής Διαλέκτου του Κορμακίτη [The Dictionary of the Arabic Dialect of Kormakitis]*. Nicosia: Cultural Foundation of "Kormakitis" Association.

Haarmann, Harald (2006). "Language Planning: Graphization and the Development of Writing Systems". In: *An International Handbook of the Science of Language and Society*. Ed. by Ulrich Ammon et al. Berlin • New York: De Gruyter Mouton, pp. 2402–2420.

Justeson, John S. (1976). "Universals of Language and Universals of Writing". In: *Linguistic studies offered to Joseph Greenberg on the occasion of his 60th birthday*. Ed. by Alphonse Juilland, Andrew M. Devine, and Laurence D. Stephens. Vol. 1. General Linguistics. Saratoga: Amna Libri, pp. 57–94.

Karan, Elke (2013). "The ABD of orthography testing: Integrating formal or informal testing into the orthography development process". In: *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 53.1.

Karyolemou, Marilena (2019). "A story at the periphery: Documenting, standardising and reviving Cypriot Arabic". In: *International Journal of the Sociology of Language* 2019.260, pp. 1–14.

Karyolemou, Marilena and Spyros Armostis (2020). "Langue ancienne, écriture nouvelle : codification et standardisation de l'arabe de Chypre". In: *Cahiers du Centre d'Etudes Chypriotes* 50, pp. 407–429.

Kiparsky, Paul (1982). "From cyclic Phonology to lexical phonology". In: *The structure of phonological representations*. Ed. by Harry Hulst and Norval Smith. Vol. 1. Dordrecht, The Netherlands: Foris, pp. 131–175.

Liddicoat, Anthony J. (2005). "Corpus Planning: Syllabus and Materials Development". In: *Handbook of Research in Second Language Teaching and Learning*. Ed. by Eli Hinkel. New York: Routledge, pp. 993–1012.

Malone, Susan E. (2004). *Manual for developing Literacy and Adult Education Programs in Minority Language Communities*. Bangkok: UNESCO.

Meletis, Dimitrios and Christa Dürscheid (2022). *An Overview of Graphotinguistics*. Berlin, Boston: De Gruyter Mouton.

Rogers, Henry (2005). *Writing systems: a linguistic approach*. Malden, MA: Wiley-Blackwell.

Smalley, William Allen (1964). "How Shall I Write this Language?" In: *Orthography Studies: Articles on New Writing Systems*. Ed. by William Allen Smalley and et al. Helps for Translators. London: United Bible Societies, pp. 31–52.

Weingarten, Rüdiger (2011). "Comparative graphematics". In: *Written Language & Literacy* 14.1, pp. 12–38.