

PART I

PHONEMES AND
PHONES

1. COPTIC

Any study of ancient Egyptian phonology must be based on Coptic, because that is phonologically the most transparent stage of the language. Coptic is written in an alphabet derived from the Greek, with additional signs from Demotic primarily for sounds not present or not represented in Greek. It appears fully formed in the third century AD but has written antecedents at least six centuries earlier.¹ Coptic had six major dialects: Akhmimic (A), Bohairic (B), Fayumic (F), Lycopolitan (L, formerly Subakhmimic A₂, also known as Lyco-Diospolitan), Oxyrhynchite (or Mesokemic, M), and Saidic (S). These vary from one another grammatically in some respects, but mainly phonologically.

GRAPHEMES

The graphemes found in texts from the six major Coptic dialects are the following, in the order of the Greek alphabet:

1 Most recently, Quack 2017. The antecedents are often termed “Old Coptic” (OC).

COPTIC	VAR/ALT	GREEK	COPTIC	VAR/ALT	GREEK
ⲗ	ϵ	Α	ⲣ	ⲗ	Ρ
Ⲗ	ⲟϥ, ϥ, ⲡ	Β	Ⲙ	Ⲍ, ϣ	Σ
Ⲓ	κ	Γ	Ⲛ	Ⲕ, †, Ⲟ	Τ
Ⲕ	τ	Δ	ϥ	ϵ, ⲓ, ⲛ	Υ
ϵ	ⲗ, –	Ε	ϥ	ⲡⲌ; ⲡ; ⲡ (B)	Φ
Ⲍ	Ϙ	Ζ	Ⲙ	ⲕⲌ; ⲕ; ⲕ (B)	Χ
ⲛ	ⲓ, ϵ, ϥ, ⲗ	Η	ϥ	ⲡϘ	Ψ
Ⲟ	ⲚⲌ; Ⲛ; Ⲛ (B)	Θ	ⲟ	ⲟ	Ω
ⲓ	ϵ	Ι	ϣ	Ϙ	
ⲕ	Ⲓ, Ⲕ, Ⲙ	Κ	ϥ	Ⲗ, ⲟϥ	
ⲗ	ρ	Λ	Ⲍ	Ⲍ, ϣ	
ⲙ	ⲛ	Μ	Ⲍ (A)	Ⲍ, ϣ	
ⲛ	ⲙ	Ν	ϣ (B)	Ⲍ, ϣ	
ⲛ	ⲕϘ	Ξ	Ⲙ	Ⲛϣ, Ⲕ	
ⲟ	ⲟϥ	Ο	Ⲕ	Ⲙ; Ⲙ (B)	
ⲡ	Ⲗ, ϥ	Π	†	Ⲛⲓ	

The graphemes Ⲓ, Ⲕ, and Ⲍ are used mainly in Greek loan-words, but Ⲓ and Ⲍ also occur as variants of ⲕ and Ϙ, respectively: e.g., ⲗⲛⲕ/ⲗⲛⲒ “I,” ⲗⲛⲌⲛⲖⲉ/ⲗⲛⲘⲖⲉ “school-room.” The graphemes ⲛ, ϥ, and † are monograms in all dialects, for ⲕϘ, ⲡϘ, and Ⲛⲓ, respectively.

The graphemes Ⲟ, ϥ, and Ⲙ are monographic for ⲚⲌ, ⲡⲌ, and ⲕⲌ, respectively, in all dialects except Bohairic, where they replace Ⲛ, ⲡ, and ⲕ, respectively, in certain words and phonetic environments: for example, B ϥⲛⲟϥⲓ vs. AM ⲡⲛⲟϥⲉ, F ⲡⲛⲟϥⲓ,

LS ΠΗΥΕ “heaven.” Bohairic also has a similar alternation between its **ϝ** and the **x** of other dialects: e.g., B **ϝΝΟΥ** vs. AFLS **xΝΟΥ** “ask.”

The graphemes **ϑ** and **ϕ** exist in Akhmimic and Bohairic, respectively; they are replaced by **ϗ** or **Ϙ** in other dialects: e.g., A **ϗΕ**, B **ϕΕ**, F **ϑΙ**, LMS **ϗΕ** “manner” and A **ϗΩΠΕ**, BF **ϘΩΠΙ**, LS **ϘΩΠΕ**, M **ϘΩΠΕ** “become.”

In some dialects, the grapheme **ι** is also spelled **ει**, as well as **ī** before or after a vowel: e.g., AL **ιΝΕ**, BF **ιΝΙ**, AMS **ειΝΕ** “bring”; AFM **πει**, B **ϕαι**, L **πειει**, S **παι** “this.” The grapheme **Υ** is used primarily in **ΟΥ**, representing [u] and [w], and after vowels: **αΥ/αΟΥ**, **εΥ/εΟΥ**, **ηΥ/ηΟΥ**, **οΟΥ**, and **ωΥ/ωΟΥ**; it occurs by itself either in Greek loan words or as a variant of **ε**, **η**, or **ι**: e.g., F **τεβνη** ~ **τυβνη** “animal.”

A graphemic feature of most Coptic dialects is a supraliteral stroke (e.g., **ḿ**) or, in Bohairic, a dot or acute accent (e.g., **ḿ/ḿ̇**). Both are used in some manuscripts to mark a grapheme that represents a syllabic consonant or a separate syllable: for example, B **ḿΟΟΚ**, S **ḿΤΟΚ** “you.” In some cases, the supraliteral mark varies with **ε** both within and across dialects: e.g., A **ḿḿ**, B **ḿḿ**, F **ḿḿ**, FLMS **ḿḿ** “in.”

PHONES

The phonetic value of Coptic graphemes can be deduced from both the Greek graphemes on which they are based and from language-internal instances of alternation and variation.

For the former, it is clear that Coptic graphemes do not always represent the values they had for Greek speakers in the era when Coptic is first attested, but rather those of the Greek language some six centuries earlier.⁴ The phonetic value of some Greek graphemes changed between the Classical age (fifth and fourth centuries BC) and the Koine period (third century BC to third century AD), and the Coptic values are for the most part those of the older language:⁵

GREEK GRAPHHEME	CLASSICAL VALUE	KOINE VALUE	COPTIC GRAPHEME	COPTIC VALUE
Γ	[g]	[ɣ]	Ⲁ	[k]
Δ	[d]	[ð]	Ⲃ	[t]
Η	[ɛ:]	[ɪ, i]	Ⲅ	[ɛ, e]
Θ	[t ^h]	[θ]	Ⲇ	[t ^h], [t ^h]
Φ	[p ^h]	[φ, f]	Ⲉ	[p ^h], [p ^h]
Χ	[k ^h]	[x]	Ⲋ	[k ^h], [k ^h]

These correspondences agree with the earliest evidence for Egyptian words and texts written in the Greek alphabet during the Ptolemaic Period, and they argue for the preservation of that scribal tradition even as the pronunciation of Greek itself evolved.

⁴ Satzinger 2003.

⁵ Allen 1987, 12–32, 62–79; Horrocks 2010, 117–20. This study uses the symbols of the International Phonetic Alphabet, between square brackets, to indicate pronunciation, with the exception that post-syllabic ' is used to mark stress: e.g., ⲙⲓⲧⲟⲛ [m-tɔn'].

1. COPTIC

7

Greek words that appear in Coptic texts, however, generally reflect contemporary Koine phonology, clearly indicating that the Greek characters used for Coptic sounds in the third century did not derive from contemporary Greek: for example,

GREEK	CLASSICAL VALUE	KOINE VALUE	GREEK EXAMPLE	COPTIC RENDERING		
AI	[ai]	[ε]	δικαιος	ΔΙΚΕΟС	[ti'-ke-ɔs]	“just”
B	[b]	[β, v]	βλάπτειν	ΦΛΑΠΤΕΙ	[flap'-ti]	“hinder”
H	[ε:]	[i, i]	ἐπιστήμη	ΕΠΙCΤΗΜΕΙ	[ε-pis-ti'-mi]	“prudence”
OI	[ɔi]	[i, i]	ἐτοιμος	ΖΕΤΕΙΜΟС	[hε-ti'-mɔs]	“ready”
Y	[y]	[i, i]	πύλη	ΠΙΑΙ	[pi'-li]	“gate”
X	[k ^h]	[x]	χαρακτήρ	ΧΑΡΑΚΤΗΡ	[xa-rak-ter']	“mark”

Of vowels, **ε** is the most common, as well as the most common Coptic grapheme. Its correspondence with Koine [ε], as in ΔΙΚΕΟС for δικαιος, indicates that it had a similar value in Coptic. Its use as a variant of the signs for a syllabic consonant, however, point to a realization closer to [ə]: e.g., S **Μ̄ΤΟΝ** ~ **ΕΜΤΟΝ** “rest” [m-tɔn'] ~ [əm-tɔn']. Its value may therefore have encompassed, and lain between, mid-central [ε] and [ə], with realization probably conditioned by both dialect and phonological environment. **ε** also occurs as a variant of **α**, both within and across dialects – for example, S **ΧΑCΤϢ**, M **ΧΕCΤϢ** “exalt him” (Matt. 23:12). This suggests a phonetic value for **α** close to that of **ε**, probably back central [a] ~ [æ]. Coptic **η** was likely pronounced [i/i] in Greek loanwords, and this may account for its occasional variance with **ι** in Coptic

words – e.g., S **NHBE** ~ **NIBE** “swim” – but it also varies with **ε** and **α** – e.g., S **PAṬ** ~ **PEṬ** ~ **PHT** “foot” – and was therefore most likely close in value to those vowels in native words, probably ranging between [ɛ] and [e]. The other vowels correspond to their Greek counterparts in loanwords and presumably had similar phonetic realizations: **ι** [i/i], **ο** [ɔ], **οϣ** [u], and **ω** [o].

The consonants represented by Greek letters correspond pretty much to their pre-Hellenic ancestors. **Γ**, **Δ**, and **Ζ** were probably pronounced like **κ**, **τ**, and **ϙ**, respectively, judging from their variance with those graphemes in Coptic words. **Β** alternates with **Π** and varies with **Ϙ** and **οϣ**: A **οϣααβε**, B **οϣαβ**, FLS **οϣααβ**, M **οϣεβ** “pure” and A **οϣαπ**, S **οϣοπ** “become pure”; S **ωβτ** ~ **ωϘτ** “goose”; B **οϣιϙι** ~ **βιϙι**, S **οϣειϙε** ~ **βιϙε** “saw.” The alternation suggests a phonetic realization not only as a stop ([b] → [p]) but also as a bilabial fricative [β], which explains its variance with **οϣ**. Variation with **Ϙ** suggests that the latter may also have been bilabial, distinguished from **β** by voicing. Thus, **β** → [b]/[p]/[β] and **Ϙ** → [ϕ].

The values of the other graphemes derived from Demotic can also be deduced from variances and correspondents: **ϣ** [ʃ] (Arabic أشمون *ašmūn* from **ϣμοϣν** “Hermopolis”), **ϛ** [ħ] (**ϛεβρων** for **ḥbrōn** *hebrōn* “Hebron”), **ϛ** [tʰ] (F **ϛοϣια**, B **τϣοϣιε** “dry” – [tʰ] ~ [tʃ]), **ϛ** [kʰ] (S **ϛαδαρην** from Greek φακιάριον “turban”⁶ – [kia] → [kʰa]).

6 Girgis 1967–1968, 58.

In most dialects, Φ Θ χ are monograms for $\Pi\zeta$ $\tau\zeta$ $\kappa\zeta$, respectively; AFLS $\Pi\zeta\omega\beta$ and M $\Pi\zeta\omega\beta$ “the (Π) thing ($\zeta\omega\beta$ / $\omega\beta$),” for example, can also be spelled $\Phi\omega\beta$ / $\Phi\omega\beta$. In Bohairic, however, they represent, like their Greek ancestors, the aspirated counterparts of Π τ κ , respectively. Aspiration occurs before a stressed vowel and before a sonant (β λ μ ν ρ) or $\omega\gamma$ and $\iota/\epsilon\iota$ preceding a stressed vowel:⁷ e.g., $\Phi\lambda\iota$ [p^hai] “this one” vs. $\Pi\alpha\iota\rho\omega\mu\iota$ [pai-ro'-mi] “this man,” $\chi\beta\omega\psi$ [k^hβoʃ] “you loosen” vs. $\kappa\sigma\omega\phi$ [ksoφ] “you defile.” Similarly, in Bohairic σ is [t^h], the aspirated counterpart of χ [t^j]: e.g., B $\delta\iota\sigma\iota$ “exalt” (ALMS $\chi\iota\sigma\epsilon$, F $\chi\iota\sigma\iota$). Its phonetic value in that dialect can also be gauged from variants such as $\mu\omicron\rho\delta\acute{\nu}\lambda\gamma\zeta \sim \mu\omicron\rho\psi\eta\lambda\zeta$ [mɔr'-t^hɪnawħ] ~ [mɔr'-ʃnəħ] “scapular” ([t^h] lenited to [ʃ]) and $\delta\omega\eta\tau \sim \chi\omega\eta\tau$ [t^hio'-nt] ~ [t^jo'-nt] “try” ([t^h] deaspirated to [t^j]).

The alternation of $\Phi \leftrightarrow \Pi$ is environmentally conditioned and therefore reflects a single phoneme, but the other alternants are phonemic: B $\theta\omega\rho\iota$ “willow” vs. $\tau\omega\rho\iota$ “handle,” $\chi\rho\omega\mu$ “fire” vs. $\kappa\rho\omega\mu$ “safflower,” $\delta\omega$ “plant” vs. $\chi\omega$ “hunchback.” The phonemic status of the aspirates is reflected in their preservation where environmental aspiration is not required: e.g., $\delta\iota\sigma\iota$ [t^hi'-si] “exalt” and $\delta\epsilon\epsilon\phi\eta\omega\gamma\tau$ [t^hes-p^hnu'-ti] “exalt God.”

7 Shisha-Halevy 1991, 54. In turn, therefore, aspiration was perhaps neutralized in other environments, similar to [t^h] ~ [t] in American English: e.g., *bat* [hæt^h] vs. *batter* [hæt'-ɪ].

Aspiration is not visible in the other dialects: for example, B **ΘΩΡΙ** vs. S **ΤΩΡΕ** “willow,” B **ΧΡΩΜ** vs. F **ΚΛΩΜ** and S **ΚΡΩΜ**, B **ΘΙCΙ** vs. ALMS **ΧΙCΕ** and F **ΧΙCΙ**. Whether this reflects an absence of aspirates in these dialects or merely a graphemic neutrality (i.e., AFLMS **τ** representing both [t] and [t^h]) is not self-evident. The fact that these dialects use graphemes derived from the unaspirated graphemes of (Classical) Greek (κ, π, and τ) might suggest the former. Arabic renderings of Coptic words, however, sometimes show a correspondence between [t] and *t*, on the one hand, and [t^h] and *t*, on the other: e.g., A **ΤΩΒΕ**, B **ΤΩΒΙ**, S **ΤΩΒΕ**/**ΤΩΒΕ** “brick” ≙ Arabic طوبة *tūba*; B **ΘΛϞ**, S **ΤΛϞ** “spit” ≙ تف *taff*.⁸ This may or may not reflect the influence of Bohairic,⁹ but it is also visible in place-names from non-Bohairic areas: e.g., S **CΙOΟΥΤ** “Asyut” ≙ أسيوط *asyūt*.¹⁰

PHONOTACTICS

Coptic words have a single nodal stress around which everything else is reduced as much as is possible phonetically: e.g., S **ΖΟΕΙΝΕ** [hɔi'-ne] “some” + **ΡΩΜΕ** [ro'-me] “man” + **†ΜΕ** [ti'-me] “village” → **ΖΕΝΡΜ†ΜΕ** [hɛn-rm-ti'-me] “villagers.” In

8 Bishai 1964, 46.

9 The prevalence of Bohairic in the north at the time of the Arab conquest has also been called into question: Kahle 1954, 249–52.

10 B **СΙOΟΥΤ**. The association of Arabic *t* with unaspirated [t] is also visible in Greek Πτολεμαῖος “Ptolemy” ≙ Arabic بطليموس *batlaimūs*. Cf. also Bishai 1964, 41: “The velarization of τ is normal owing to its unaspirated nature.”

native words, the vowels **o**, **ω**, and usually **h** carry primary stress; the other vowels can be stressed or not: e.g., SB **ANAO** [a-naʃ'] "oath," SBF **ENEZ** [ε-neħ'] "eternity," BF **INI** [i'-ni] "bring," ABFLMS **OYNOY** [u-nu'] "hour."

A basic distinction in Coptic words is between stressed syllables that end in a vowel (open) and those that end in a consonant (closed). These have an effect on vowel quality for the following pairs:

CLOSED	OPEN	EXAMPLES
α/o	ω	AFL CAN , BS CON "brother" vs. ALS CONE , BF CONI "sister": [san/sɔn] vs. [so'-ne/so'-ni]
ε/α	h	AFLM ZEK , BS ZAK "your (ms) face" vs. ABFLMS ZHTN "your (pl) face": [ħrek/ħrak] vs. [ħre'-tn]
ε/α	i	AFLM XECTOY , S XACTOY "exalt them" vs. ALMS XICE , F XICI "exalt": [tʰes'-tu/tʰas'-tu] vs. [tʰi'-se/tʰi'-si]

These alternants have traditionally been described as "short" (**α ε o**) and "long" (**h i ω**) vowels.¹¹ In Oxyrhynchite, however, the first alternation usually does not occur, while the second and third do: M **CON** "brother" vs. M **CONE** "sister." This indicates a difference in vowel quality rather than length: probably lax (-T) **α ε o** versus tense (+T) **h i ω**. The

11 The classic study is Edgerton 1947 (published before the description of Oxyrhynchite).