

## I. INTRODUCTION

The present document proposes shared guidelines for the transcription of Coptic texts into a TEI XML database. The principal goal is to produce transcriptions which are consistent, accessible and generally reflect standard scholarly practice. Because the majority of projects currently underway are focusing on Sahidic texts, the forms and examples in the present discussion appear in this dialect. In particular, the present discussion has reiterated the standards formerly proposed by Walter Till, surveying his main points and examples. According to an agreement with Caroline Schroeder and Amir Zeldes of *Coptic Scriptorium* (<http://copticcriptorium.org>) and a discussion during the common work in the project *Kellia* (<http://copticcriptorium.org/kellia>) these guidelines represent a compromise between the projects aiming at a critical edition of Coptic biblical texts and *Coptic Scriptorium* aiming mainly at a linguistic annotation and analysis of Coptic texts. This compromise enables an automatic segmentation of Coptic texts through the Natural Language Processing (NLP) service in the ANNIS database (<http://copticcriptorium.org/tools>). The Transcription Guidelines mediate actually between the rules of word segmentation for Coptic text publications as published by Walter Till and the word segmentation of Coptic texts based on a bound group concept developed by Bentley Layton in his *Coptic Grammar*, which was adopted by *Coptic Scriptorium*.

W. C. Till, 'Zur Worttrennung im Koptischen,' *ZÄS* 77 (1941): 48–52.

———, 'La séparation des mots en copte,' *BIFAO* 60 (1960): 151–170.

B. Layton, *A Coptic Grammar*. Third Edition, Revised and Expanded (Wiesbaden : Harrassowitz 2011), 19-27.

## II. WORD DIVISION

The issue of word division begins with the creation of a base text, and arises again with individual transcriptions of manuscripts and the eventual creation of a critical edition. Before all, transcriptions should distinguish for the reader between the lexicographic units (*i.e.* adverbs, adjectives, conjunctions, nouns, prepositions and verbs).

### A. Articles, infixes and particles

#### 1. Definite and indefinite articles (π- τ- η-; οΥ- ρεν-)

All articles combine with the following word.

πρωμε	“The man”
τῆςριμε	“The woman”
ἡωα.δε	“The words”

#### 2. Demonstrative and possessive infixes (-ει- -ι- -κε-; *e.g.*, -α- -εκ- -εε-)

In keeping with the prior rule, these infixes conjoin with the following word.

παρωμε	“My man”
τῆςριμε	“This woman”
ἡετῆρωα.δε	“Your (pl.) words”

### 3. The possessive pronoun and composite nouns (πα τα να)

Because the placement of the definite article before the subsequent noun distinguishes this construction from homographic first-person possessive forms (παρωμε *versus* παπρωμε), no separation is necessary. Likewise, the possessive pronoun will appear without division in the case of compound words with distinct entries in Crum’s lexicon.

ΠΑΠΡΩΜΕ	“The one (masc.) of the man”
ΠΑΤΡΩΜΕ	“The one (fem.) of the woman”
ΝΑΝΩΑΔΕ	“The ones of the words”
ΠΑΤΩΕΛΕΤ	“bridegroom”

❖ ANNIS Compliance: All possessive pronouns appear without separation from the possessed noun.

### 4. Enclitic πε τε νε

For clarity, the enclitic particles, which most commonly mark a copulative sentence but also appear almost randomly, should be transcribed separately.

ΝΤΟΚ ΠΕ ΠΑΩΗΡΕ	“You are my son.”
ΕΩΠΡΙΣΕ ΝΩΝΟΒ̄C ΔΗ ΠΕ	“He was tormented, but was not angry.”

### 5. Other common affixes

For clarity’s sake, the following frequently occurring affixes should be transcribed separately whenever possible: ν̄σι- δε- -βε.

ΑΟΥΧΑΙ Ν̄ΣΙ ΤΕΣΙΜΕ	“Saved was the woman.”
ΠΕΧΑΥ ΔΕ ΑΡΕ ΠΧΘΕΙC	“They said, ‘Yes, Lord.’”
ΤΕΝΟΥ ΒΕ †ΧΩ ΗΜΟC ΝΗΤ̄Ν	“Now, therefore, I say to you.”

## B. Verbal entities and Adverbial sentence

### 1. Conjugation prefixes — pronominal

When a pronominal infix marks the subject, the entire subject-verb phrase stands as a single unit.

ΑCΝΑΥ	“She saw.”
ΝΕΦ̄ΡΟΥ	“He was evil.”
†ΝΑCΩΤ̄Ν	“I will hear.”

### 2. Conjugation prefixes — nominal

The Conjugation prefixes are also written together with their nominal subject.

Ν̄ΤΕΡΕῙΗC ΟΥΩ ΕΦΟΥΕΞ CΑΡΝΕ	“When Jesus finished teaching ...”
ΜΑΡΕ̄Ν̄ΡΕΘΝΟC ΤΗΡΟΥ ΕΙΜΕ	“Let all the gentiles know ...”
ΩΑΡΕΤΕΓΡΑΦΗ ΧΟΟC	“The scripture says (habitually) ...”

### 3. Pronominal and nominal prefixes with prepositions and adverbs (adverbial sentence)

Since pronominal prefixes normally attach to a subsequent verb, in those instances where a pronominal (or a relative pronoun) verbalizes a subsequent preposition, the pronoun and preposition will be connected. Consequently, nominal subjects in the same circumstance will be separated.

εφ̄ξ̄μ̄π̄σ̄ω̄μα ... εῑτε εφ̄μ̄π̄βολ	“... he being in the body ... or he being outside...”
νεφ̄ρῑπ̄χ̄αῑε	“He was in the desert.”
πᾱσον̄ ξ̄μ̄ πεφ̄νη	“My brother is in his house.”
φ̄μ̄μᾱγ	“He is there”
πᾱσον̄ μ̄μᾱγ	“My brother is there”
πε̄σ̄νο̄ῡ ε̄τ̄ξ̄ᾱτ̄μ̄π̄ρο	“The blood which is by the door”

- ❖ ANNIS Compliance: The object of a verbalized preposition must attach to the preposition (εφ̄ξ̄μ̄π̄σ̄ω̄μα).

### 4. Pronominal verbs

In those cases where verbal forms presuppose a following noun, the noun should be written separately if the verb form is in the absolute state or the direct object is connected by a preposition. However, the verbs with only suffixed objects and the subjects of the existential sentence verb forms should be connected in writing. In the same way, Verbs in the pronominal state showing reduced stress should be combined with their noun-object. Idiomatic combinations of pronominal verbs with zero-article nouns and Greek infinitives, and sometimes with prepositions (*esp.* εῑρε, †, χ̄ι, κ̄ω, φ̄ι) will also be connected.

†σ̄ω̄τ̄π̄ῑ μ̄μο̄υ	“I choose him/the man”
†σ̄ω̄τ̄π̄ῑ μ̄π̄ρω̄με	“I choose him/the man”
ᾱφ̄σε̄τ̄π̄ῑ	“He chose him/the man”
ᾱφ̄σε̄τ̄π̄ῑπ̄ρω̄με	“He chose him/the man”
πε̄ξε̄π̄χ̄ο̄εῑς	“The Lord said ...”
νᾱνο̄ῡπε̄κ̄νᾱ	“Your mercy is good.”
ο̄ῡν̄ο̄εῑκ	“There is bread.”
ο̄ῡν̄τε̄πεῑω̄τ̄ π̄ω̄ν̄ξ̄	“The father has life.”
μ̄ν̄τ̄ο̄ῡλο̄ῑσε̄ μ̄μᾱγ	“They do not have reproach.”
μ̄ν̄τ̄αῑσο̄υ	“I do not have them.”
ᾱρ̄ρω̄τ̄η̄ τε̄τ̄η̄†ξ̄ῑσε̄ η̄τεῑς̄ρῑμε	“Why do you trouble this woman?”
ᾱρῑπᾱμ̄ε̄ε̄ῡε	“Remember me”
ᾱφ̄†σ̄β̄ω̄ νᾱγ	“He taught them.”
ᾱφ̄χῑσ̄β̄ω̄	“He learned.”
ᾱφ̄ρ̄ξ̄ᾱξ̄ η̄νω̄βε	“He has committed many sins.”
ᾱφ̄ο̄ῡε̄ω̄β̄ω̄κ	“He wished to go.”
ᾱφ̄ρ̄αῑτεῑ	“He asked.”
ᾱγ̄ρ̄ξ̄ᾱρω̄ῑ	“They descended on me.”

- ❖ ANNIS Compliance: Pronominal verbal forms combine with subsequent object-nouns (ᾱφ̄σε̄τ̄π̄ῑπ̄ρω̄με, ᾱφ̄ρ̄αῑτεῑ).

## 5. Relative pronouns

Relative pronominal forms combine with the following verbal or adverbial component as below. Notably, in some cases, even the pronominal form ετε will elide with a following vowel.

ΖΩΗ <b>ε</b> τεται τε τμααγ̄ ἡνετονε τηρογ̄	“... Eve, who is the mother of all living people.”
ἡιογδαῖ δε ἡτοογ̄ <b>ε</b> τεῖῃπογναρτε	“... but the Jews, they that did not believe ...”
ζωβ νιμ <b>ε</b> τῆῃναααγ̄	“... everything which he will do ...”
πενταγτσαβοῖ	“... the one who has taught me ...”
πμα <b>ε</b> τεγῆῃμοογ̄ ἡρητῆ	“The place in which there was water ...”
πετεογῆῃτῆ μααχε μαρεφωτῆ	“The one who has (an) ear, let him hear.”
νετῆῃππε	“Those who are in heaven.”

## C. Prepositions

### 1. Prepositions of more than one character

To distinguish distinct lexicographic entities more clearly, prepositions should be written separately when they precede a noun. Pronominal prepositions are written together with their pronominal suffixes. As in the example below, the separation of the preposition εἰν distinguishes from the variant orthography of the plural indefinite article ζειν-/εἰν-. An exception is the suffix pronoun τηγῆ which is treated as a fully stressed noun. Therefore, the preceding preposition has the *status nominalis* form.

εἰν εἰτοογε ωα ρογρε	“... from morning until evening ...”
εἰῃνογ̄	“... to ask ...”
εἰῃ ογ̄	“... on what?”
εἰῃ ωαχε νιμ	“... in every word ...”
εἰῃωαχε	“... words ...”
εἰῃ τηγῆ	“... about you ...”

### 2. Single-character prepositions

In order to avoid one-character words and recognizing that single character prepositions can elide with initial vowels, transcriptions will combine single-character prepositions with their objects.

ναιατῆ ἡπρωμε	“Blessed is the man ...”
αιναγ̄ εγχειμε	“I saw a woman...”

## D. Nouns and Adjectives

As already mentioned, the goal is to distinguish independent lexicographic entities. Therefore, as a rule, nouns and adjectives are distinguished from other principal grammatical groups (*i.e.* adverbs, conjunctions, prepositions, verbs).

### 1. Abbreviated pronouns

Shortened pronominal forms stand separate from the modified term.

ανοκ ογωωσ	“I am a shepherd.”
ανῆ ογωωσ	“I am a shepherd.”
ἡτε τωερε ἡνιμ	“Whose daughter are you?”

## 2. Pronominal suffixes

As a rule pronominal suffixes attach to the end of verbs, prepositions and adjectives. In case of verbs with only suffixed objects and nouns which can only take a suffix pronoun as possessive modifier (typically body parts) also  $\tau\eta\gamma\tau\eta$  attaches to them.

$\bar{\mu}\mu\text{oi}$	“... with me ...”
$\bar{\eta}\eta\tau\bar{\eta}$	“... in him ...”
$\bar{\mu}\mu\omega\tau\bar{\eta}$	“... with you ...”
$\eta\lambda\alpha\tau\tau\eta\gamma\tau\bar{\eta}$	“Blessed are you ...”
$\eta\omega\tau\tau\eta\gamma\tau\bar{\eta}$	“... yourselves ...”

## 3. Compound nouns

In those cases where the Coptic language has either produced a derivative form (e.g. a noun from a verb, or a noun from another noun) with a standard infix (e.g.,  $-\bar{\rho}\bar{\mu}\bar{\eta}-$   $-\bar{\mu}\bar{\eta}\bar{\tau}-$   $-\mu\lambda\iota-$ ) or has abbreviated a less common form to produce a distinct second form, the transcriber will not separate the shortened form from the next entry. The two roots have combined to create a discrete lexicographical entry. This concerns also lexicalized combined forms.

$\bar{\rho}\bar{\mu}\bar{\eta}\kappa\eta\mu\epsilon$	“the Egyptian”
$\rho\epsilon\sigma\eta\text{no}\beta\epsilon$	“the sinner”
$\bar{\rho}\bar{\mu}\bar{\eta}\eta\eta\tau$	“the wise man”
$\tau\bar{\mu}\bar{\eta}\bar{\tau}\bar{\rho}\bar{\mu}\bar{\eta}\eta\eta\tau$	“The wisdom”
$\rho\eta\sigma\eta\mu\epsilon$	“the birthday”
$\eta\lambda\iota\eta\sigma\eta\tau\epsilon$	“pious”
$\alpha\tau\mu\sigma\eta$	“immortal”
$\rho\epsilon\kappa\omega\beta(\eta)\rho\eta\bar{\mu}\eta\lambda$	“your fellow servant”
$\rho\mu\sigma\eta\bar{\eta}\omega\tau\eta\mu$	“the torrent”

## 4. Compound phrases

Stock phrases with unshortened forms should be written separately.

$\tau\omega\delta\ \bar{\eta}\epsilon\lambda\omega\omega\lambda\epsilon$	“the olive tree”
$\rho\mu\lambda\ \bar{\eta}\omega\omega\pi\epsilon$	“the dwelling”
$\rho\epsilon\alpha\ \bar{\eta}\rho\sigma\eta\eta$	“the inside”