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://copticscriptorium.org) and a discussion during the common work in the project *Kellia* (http://copticscriptorium.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://copticscriptorium.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 (ⲡ- ⲧ- ⲛ-; ⲟⲩ- ⳉⲉⲛ-)
All articles combine with the following word.

ⲡⲣⲱⲙⲉ	"The man"
ⲧⲉⲥⳉⲓⲙⲉ	"The woman"
ⲛ̄ϣⲁϫⲉ	"The words"

2. Demonstrative and possessive infixes (-ⲉⲓ- -ⲓ- -ⲕⲉ-; e.g., -ⲁ- -ⲉⲕ- -ⲉⲥ-)
In keeping with the prior rule, these infixes conjoin with the following word.

ⲡⲁⲣⲱⲙⲉ	"My man"
ⲧⲉⲓⲥⳉⲓⲙⲉ	"This woman"
ⲛⲉⲧⲛ̄ϣⲁϫⲉ	"Your (pl.) words"

3. The possessive pronoun and composite nouns (ⲡⲁ ⲧⲁ ⲛⲁ)

Because the placement of the definite article before the subsequent noun distinguishes this construction from homographic first-person possessive forms (ⲡⲁⲣⲱⲙⲉ *versus* ⲡⲁⲡⲣⲱⲙⲉ), 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.

ⲡⲁⲡⲣⲱⲙⲉ	"The one (masc.) of the man"
ⲧⲁⲧⲥϩⲓⲙⲉ	"The one (fem.) of the woman"
ⲛⲁⲛ̄ϣⲁϫⲉ	"The ones of the words"
ⲡⲁⲧϣⲉⲗⲉⲉⲧ	"bridegroom"

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

4. Enclitic ⲡⲉ ⲧⲉ ⲛⲉ

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

ⲛ̄ⲧⲟⲕ **ⲡⲉ** ⲡⲁϣⲏⲣⲉ	"You are my son."
ⲉϥϣ̄ⲡϩⲓⲥⲉ ⲛ̄ϥ̄ⲛⲟϭ̄ⲥ ⲁⲛ **ⲡⲉ**	"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: ⲛ̄ϭⲓ- ϫⲉ- -ϭⲉ.

ⲁⲥⲟⲩϫⲁⲓ **ⲛ̄ϭⲓ** ⲧⲉⲥϩⲓⲙⲉ	"Saved was the woman."
ⲡⲉϫⲁⲩ **ϫⲉ** ⲁϩⲉ ⲡϫⲟⲉⲓⲥ	"They said, 'Yes, Lord.'"
ⲧⲉⲛⲟⲩ **ϭⲉ** ϯϫⲱ ⲙ̄ⲙⲟⲥ ⲛⲏⲧⲛ̄	"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.

ⲁⲥⲛⲁⲩ	"She saw."
ⲛⲉϥϩⲟⲟⲩ	"He was evil."
ϯⲛⲁⲥⲱⲧⲙ̄	"I will hear."

2. Conjugation prefixes — nominal

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

ⲛ̄ⲧⲉⲣⲉⲓⲏ̄ⲥ ⲟⲩⲱ ⲉϥⲟⲩⲉϩ ⲥⲁϩⲛⲉ	"When Jesus finished teaching …"
ⲙⲁⲣⲉⲛ̄ϩⲉⲑⲛⲟⲥ ⲧⲏⲣⲟⲩ ⲉⲓⲙⲉ	"Let all the gentiles know …"
ϣⲁⲣⲉⲧⲉⲅⲣⲁⲫⲏ ϫⲟⲟⲥ	"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.

ⲉϥϩⲙ̅ⲡⲥⲱⲙⲁ … ⲉⲓⲧⲉ ⲉϥⲙ̅ⲡⲃⲟⲗ	"… he being in the body … or he being outside…"
ⲛⲉϥϩⲓⲡϫⲁⲓⲉ	"He was in the desert."
ⲡⲁⲥⲟⲛ ϩⲙ̅ ⲡⲉϥⲏⲓ	"My brother is in his house."
ϥⲙ̅ⲙⲁⲩ	"He is there"
ⲡⲁⲥⲟⲛ ⲙ̅ⲙⲁⲩ	"My brother is there"
ⲡⲉⲥⲛⲟϥ ⲉⲧϩⲁⲧⲙ̅ⲡⲣⲟ	"The blood which is by the door"

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

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 prenominal state showing reduced stress should be combined with their noun-object. Idiomatic combinations of prenominal verbs with zero-article nouns and Greek infinitives, and sometimes with prepositions (*esp.* ⲉⲓⲣⲉ, ϯ, ϫⲓ, ⲕⲱ, ϥⲓ) will also be connected.

ϯⲥⲱⲧⲡ̅ ⲙ̅ⲙⲟϥ	"I choose him/the man"
ϯⲥⲱⲧⲡ̅ ⲙ̅ⲡⲣⲱⲙⲉ	"I choose him/the man"
ⲁϥⲥⲉⲧⲡ̅ϥ	"He chose him/the man"
ⲁϥⲥⲉⲧⲡ̅ⲡⲣⲱⲙⲉ	"He chose him/the man"
ⲡⲉϫⲉⲡϫⲟⲉⲓⲥ	"The Lord said …"
ⲛⲁⲛⲟⲩⲡⲉⲕⲛⲁ	"Your mercy is good."
ⲟⲩⲛ̅ⲟⲉⲓⲕ	"There is bread."
ⲟⲩⲛ̅ⲧⲉⲡⲉⲓⲱⲧ ⲡⲱⲛϩ̅	"The father has life."
ⲙ̅ⲛ̅ⲧⲟⲩⲗⲟⲓϭⲉ ⲙ̅ⲙⲁⲩ	"They do not have reproach."
ⲙ̅ⲛ̅ⲧⲁⲓⲥⲟⲩ	"I do not have them."
ⲁϩⲣⲱⲧⲛ̅ ⲧⲉⲧⲛ̅ϯϩⲓⲥⲉ ⲛ̅ⲧⲉⲓⲥϩⲓⲙⲉ	"Why do you trouble this woman?
ⲁⲣⲓⲡⲁⲙⲉⲉⲩⲉ	"Remember me"
ⲁϥϯⲥⲃⲱ ⲛⲁⲩ	"He taught them."
ⲁϥϫⲓⲥⲃⲱ	"He learned."
ⲁϥⲣ̅ϩⲁϩ ⲛ̅ⲛⲟⲃⲉ	"He has committed many sins."
ⲁϥⲟⲩⲉϣⲃⲱⲕ	"He wished to go."
ⲁϥⲣ̅ⲁⲓⲧⲉⲓ	"He asked."
ⲁⲩⲣ̅ϩⲁⲣⲱⲓ̈	"They descended on me."

❖ ANNIS Compliance: Prenominal verbal forms combine with subsequent object-nouns (ⲁϥⲥⲉⲧⲡ̅ⲡⲣⲱⲙⲉ, ⲁϥⲣ̅ⲁⲓⲧⲉⲓ).

5. *Relative pronouns*

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

ⲍⲱⲏ ⲉⲧⲉⲧⲁⲓ ⲧⲉ ⲧⲙⲁⲁⲩ ⲛ̄ⲛⲉⲧⲟⲛϩ ⲧⲏⲣⲟⲩ	"… Eve, who is the mother of all living people."
ⲛ̄ⲓⲟⲩⲇⲁⲓ ⲇⲉ ⲛ̄ⲧⲟⲟⲩ ⲉⲧⲉⲙ̄ⲡⲟⲩⲛⲁϩⲧⲉ	"… but the Jews, they that did not believe …"
ϩⲱⲃ ⲛⲓⲙ ⲉⲧϥ̄ⲛⲁⲁⲁⲩ	"… everything which he will do …"
ⲡⲉⲛⲧⲁϥⲧⲥⲁⲃⲟⲓ̈	"… the one who has taught me …"
ⲡⲙⲁ ⲉⲧⲉⲩⲛ̄ⲙⲟⲟⲩ ⲛ̄ϩⲏⲧϥ̄	"The place in which there was water …"
ⲡⲉⲧⲉⲟⲩⲛ̄ⲧϥ̄ ⲙⲁⲁϫⲉ ⲙⲁⲣⲉϥⲥⲱⲧⲙ̄	"The one who has (an) ear, let him hear."
ⲛⲉⲧϩⲛ̄ⲧⲡⲉ	"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 ϩⲛ̄ distinguishes from the variant orthography of the plural indefinite article ϩⲉⲛ-/ϩⲛ̄-. An exception is the suffix pronoun ⲧⲏⲩⲧⲛ̄ which is treated as a fully stressed noun. Therefore, the preceding preposition has the *status nominalis* form.

ϫⲓⲛ ϩⲧⲟⲟⲩⲉ ϣⲁ ⲣⲟⲩϩⲉ	"… from morning until evening …"
ⲉϫ̄ⲛⲟⲩ	"… to ask …"
ⲉϫ̄ⲛ ⲟⲩ	"… on what?"
ϩⲛ̄ ϣⲁϫⲉ ⲛⲓⲙ	"… in every word …"
ϩⲛ̄ϣⲁϫⲉ	"… words …"
ⲉϫ̄ⲛ ⲧⲏⲩⲧⲛ̄	" … 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.

ⲛⲁⲓⲁⲧϥ̄ ⲙ̄ⲡⲣⲱⲙⲉ	"Blessed is the man …"
ⲁⲓⲛⲁⲩ ⲉⲩⲥϩⲓⲙⲉ	"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.

ⲁⲛⲟⲕ ⲟⲩϣⲱⲥ	"I am a shepherd."
ⲁⲛⲅ̄ ⲟⲩϣⲱⲥ	"I am a shepherd."
ⲛ̄ⲧⲉ ⲧϣⲉⲉⲣⲉ ⲛ̄ⲛⲓⲙ	"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 ⲧⲏⲩⲧⲛ attaches to them.

ⲙ̄ⲙⲟⲓ	"… with me …"
ⲛ̄ϩⲏⲧ̄ϥ	"… in him …"
ⲙ̄ⲙⲱⲧ̄ⲛ	"… with you …"
ⲛⲁⲓⲁⲧⲧⲏⲩⲧ̄ⲛ	"Blessed are you …"
ϩⲱⲧⲧⲏⲩⲧ̄ⲛ	"… 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.*, -ⲡ̄ⲙ̄ⲛ̄- -ⲙ̄ⲛ̄ⲧ̄- -ⲙⲁⲓ-) or has abbreviated a less common form to produce a distinct second form, the transcriber will not separate the shorten form from the next entry. The two roots have combined to create a discrete lexicographical entry. This concerns also lexicalized combined forms.

ⲡⲣ̄ⲙ̄ⲛ̄ⲕⲏⲙⲉ	"the Egyptian"
ⲡⲣⲉϥⲛⲟⲃⲉ	"the sinner"
ⲡⲣ̄ⲙ̄ⲛ̄ϩⲏⲧ	"the wise man"
ⲧⲙ̄ⲛ̄ⲧⲣ̄ⲙ̄ⲛ̄ϩⲏⲧ	"The wisdom"
ⲡϩⲟⲩⲙⲓⲥⲉ	"the birthday"
ⲙⲁⲓⲛⲟⲩⲧⲉ	"pious"
ⲁⲧⲙⲟⲩ	"immortal"
ⲡⲉⲕϣⲃ(ⲏ)ⲣ̄ⲙ̄ϩⲁⲗ	"your fellow servant"
ⲡⲙⲟⲩⲛ̄ⲥⲱⲣⲙ	"the torrent"

4. Compound phrases

Stock phrases with unshorten forms should be written separately.

ⲧⲃⲱ ⲛ̄ⲉⲗⲟⲟⲗⲉ	"the olive tree"
ⲡⲙⲁ ⲛ̄ϣⲱⲡⲉ	"the dwelling"
ⲡⲥⲁ ⲛ̄ϩⲟⲩⲛ	"the inside"