Coptic SCRIPTORIUM – Sentence Segmentation Guidelines

Version 1.0.1_2024-02-17

Amir Zeldes
Georgetown University

Preamble

Sentence segmentation is an important type of analysis in any language, and is particularly crucial for ancient texts where there is an interpretative dimension and potential importance for versification decisions vis-a-vis other analysis types, such as textual reuse research, comparisons of translated works and more. These guidelines outline Coptic Scriptorium's recommendations for segmenting Coptic texts which do not have established versification into sentences, primarily based on grammatical criteria. The examples are drawn from Sahidic Coptic, but it is expected that the guidelines can largely be applied to other dialects in a straightforward fashion.

One particular application of sentence splitting lies in its role as the input to syntactic analysis, i.e. parsing. Compatibility with parsing, which expects grammatical units to be self contained, is crucial not just for linguistic analysis of Coptic syntax, but also for downstream applications of analyzed data: for example, entity recognition is typically applied sentence-wise, and constrained by the limitation that no entity mention may cross sentence boundaries. Thus, just as a parser would expect a relative clause to be in the same sentence as the noun it describes, so would an entity span referring to a person described by a relative clause require us to keep that clause in the same sentence as the noun it expands on.

The purpose of these guidelines is therefore to create more consistency in manual sentence splitting, which in turn feeds into training data for automatic tools which are part of the toolchain creating syntactic and semantic analyses of the data.

Sentences and other units of the text

Verses and sentences

Although many established versifications of texts correspond to sentences, it should not be assumed that such verses and sentences correspond to each other one-to-one. Consider the following examples as a case in point:

1. ⲉϥⲧⲱⲛ ⲥⲟⲫⲟⲥ ⲉϥⲧⲱⲛ ⲅⲣⲁⲙⲙⲁⲧⲉⲩⲥ ⲉϥⲧⲱⲛ ⲥⲩⲛⲍⲏⲧⲏⲧⲏⲥ ⲛⲧⲉⲡⲉⲓⲁⲓⲱⲛ ⲙⲏ ⲙⲡⲉⲡⲛⲟⲩⲧⲉ ⲉⲓⲣⲉ ⲛⲧⲥⲟⲫⲓⲁ ⲙⲡⲕⲟⲥⲙⲟⲥ ⲛⲥⲟϭ. (1 Cor. 1:20) - *Where is the wise? Where is the scribe? Where is the lawyer of this world? Hasn't God made foolish the wisdom of this world?*
2. ⲁϥϯⲥⲃⲱ ⲇⲉ ⲛⲁⲩ ⲉⲙⲁⲧⲉ ϩⲛϩⲉⲛⲡⲁⲣⲁⲃⲟⲗⲏ ⲁⲩⲱ ⲛⲉϥϫⲱ ⲙⲙⲟⲥ ⲛⲁⲩ ϩⲛⲧⲉϥⲥⲃⲱ (Mark 4:2) ϫⲉⲥⲱⲧⲙ ⲉⲓⲥ ϩⲏⲏⲧⲉ ⲁϥⲉⲓ ⲉⲃⲟⲗ ⲛϭⲓⲡⲉⲧϫⲟ ⲉⲧϫⲟ. (Mark 4:3) - *He taught them many things in parables, and told them in his teaching, || Listen! Behold, the farmer went out to sow,...*

In the first example, a Biblical verse contains multiple grammatically independent sentences, as shown in the translation which uses multiple question marks, one for each question. In the second example, the content of a speech verb from one verse is expressed in the next.

Although these situations can create difficulties for applications such as parsing and entity recognition, we have opted so far in Scriptorium data **to respect established versifications** and assume sentence boundaries identical to such verses, in the interest of keeping our data model simple and easy to read. Thus verses are the units which are numbered and displayed with interlinear translation for each work, entities displayed in each verse may not overflow such verse boundaries, etc. This has the consequence that entity annotations and syntactic analyses never cross verse boundaries in Scriptorium data.

That said, cases such as 1. and 2. above are fairly rare exceptions, with established verses usually corresponding to syntactic sentences. For texts without a pre-existing binding versification scheme (most texts outside of Biblical data), we therefore strongly recommend using the linguistic guidelines laid out in the following sections to establish any new versifications, which maximize consistency and correspond to the vast majority case in versified works.

In all cases, **the concept of verses and sentences is meant to overlap perfectly in Scriptorium data.**

Translations and sentences

As shown above, sentences in English translations do not necessarily overlap with established verses, nor do they necessarily overlap with original Coptic sentences. **We hold English translations to be largely irrelevant** to deciding Coptic sentence boundaries. Although we may consider translations as a factor in rare cases where multiple segmentation points are possible purely based on the Coptic text, grammatical criteria should outweigh issues related to translational equivalents, since they, and not translation sentence boundaries, are what constrains the distribution of information (presence of entities, argument structures) in Coptic text.

Chapters and sentences

Chapters in Scriptorium data **must** at a minimum neatly nest verses, and therefore also sentences. It is possible (but uncommon) for a chapter to consist of a single verse (and therefore sentence), but sentences may never cross chapter boundaries. In works with established chapters, sentence boundaries must be chosen to respect this constraint. In works for which chapters are being newly established, these guidelines suggest first ensuring sentence boundaries at any point in which a chapter transition is being proposed.

Grammatical criteria for Coptic sentence segmentation

Whole propositions

At their core, sentences form independent units, which are usually complete propositions, and typically consist of at least one main predication, usually accompanied by a subject phrase. However some independent units are not predicates, such as fragments and interjections, and even sequences of foreign words appearing in a Coptic text can form a ‘sentence’ if they are either isolated or surrounded by otherwise complete sentences.

For example, consider the following spans of text which can be full sentences in context, but all lack a main verb:

- Headings:
    - ⲡⲃⲓⲟⲥ ⲁⲩⲱ ⲧⲡⲟⲗⲓⲧⲉⲓⲁ ⲙⲡⲉⲛⲡⲉⲧⲟⲩⲁⲁⲃ ⲛⲉⲓⲱⲧ ⲉⲧⲧⲁⲓⲏⲩ ⲕⲁⲧⲁⲥⲙⲟⲧ ⲛⲓⲙ ⲁⲡⲁ ⲟⲛⲛⲟⲫⲣⲓⲟⲥ ⲡⲁⲛⲁⲭⲱⲣⲓⲧⲏⲥ ⲛⲧⲁϥϫⲱⲕ ⲉⲃⲟⲗ ⲙⲡⲉϥⲃⲓⲟⲥ ⲛⲥⲟⲩⲙⲛⲧⲧⲁⲥⲉ ⲙⲡⲉⲃⲟⲧ ⲡⲁⲱⲛⲉ ϩⲛⲟⲩⲉⲓⲣⲏⲛⲏ ⲛⲧⲉⲡⲛⲟⲩⲧⲉ - “The life and conversation of our holy father, who was glorious in every way, Apa Onnophrios the Anchorite, who ended his life on the sixteenth day of the month Paone in the peace of God!”
    - ⲥⲓⲛⲟⲩⲑⲓⲟⲩ ⲉⲡⲓⲥⲟⲧⲗⲏ “Epistle of Shenoute”
- Free standing interjections (i.e. not part of a larger sentence):
    - ϥⲑ “Amen!”
    - ϩⲁⲓⲟ “yea!”
- Fragments (e.g. in papyri, ostraca, or due to lacunae in codices)
    - … ϭⲟⲗ… “lie”

Tiling principle

We assume two underlying axioms:

1. All words in a text are assumed to be contained in some sentence.
2. Sentences may never overlap.

As a result of these axioms, we can surmise that splitting a text into sentences will result in a **tiling analysis -** just as tiles cover a wall without gaps, the division into sentences covers an entire text.

A consequence of this principle is that, if we have two sentences with a complete propositional structure as outlined above, but we also have some intervening material between them, then what is left over after segmenting those two sentences must necessarily be at least one more sentence (or multiple ones, if there are reasons for subdividing that span of tokens). Intervening tokens between two well-formed sentences will therefore always be allowed to be a sentence, even if the resulting span is not a well-formed sentence of grammatical phrase.

Punctuation

In texts with good and consistent orthography, punctuation marks may indicate sentence boundaries in similar ways to English sentences. By default, if a possible boundary is ambiguous between a coordination (A and B) or a sequence of sentences (A. B.), then in the presence of punctuation between the two units, we prefer to assume a sentence boundary.

Conversely, if there is no punctuation between two propositions and the text does regularly use punctuation to separate sentences, we assume a single sentence with internal parataxis, just as we might have two propositions in an English sentence:

- Sometimes you want it, sometimes you don't.

In Coptic, this can happen with two predicates not joined by ⲁⲩⲱ:

- ⲡⲗⲏⲛ ⲉⲓⲥⲛϩⲗⲗⲟ **ⲉⲓ** ⲙⲛⲛⲉⲛⲥⲛⲏⲩ **ⲁⲩⲉⲓ** ϣⲁⲣⲱⲧⲛ ⲛⲕⲉⲥⲟⲡ ·

  Furthermore, behold the elders go with their brothers, they came to you another time.

Additionally, if such a paratactic construction is attached jointly to a phrase which modifies the multiple sub-parts, it must also be a single sentence, for example:

- **ⲙⲡⲛⲁⲩ ⲙⲉⲛ ⲉϣⲁⲛⲛⲕⲟⲧⲕ** ⲙⲛⲛⲉⲛⲉⲣⲏⲩ **ϣⲁⲛⲛⲁⲩ** ⲉⲩⲍⲱⲟⲛ ⲛⲧϩⲉ ⲛⲟⲩⲁⲉⲧⲟⲥ ⲉϥϩⲏⲗ **ϣⲁϥⲉⲓ** ⲛϥϩⲱⲥ ⲉⲃⲟⲗ ⲉϫⲙⲡⲉⲛⲙⲁ ⲛⲛⲕⲟⲧⲕ ⲙⲡⲉⲥⲛⲁⲩ

  Moreover, **when** we used to lie together in bed, we **used to see** a creature like unto an a eagle flying in the air, and **he would come** and sing over the bed whereon we two were lying

In this example we must assume the predicates "see" and "come" are coordinate in a single sentence, since they both share the modifier "when" (ⲙⲡⲛⲁⲩ ⲉ...); separating the sentences at ϣⲁϥⲉⲓ would remove the dependency on the joint temporal modifier.

Wackernagel particles

As in Greek, enclitic particles such as ⲇⲉ, ⲅⲁⲣ etc. (sometimes called Wackernagel particles) appear in the second position of Coptic sentences. As a result, their presence is a strong indication of the beginning of a new sentence at the preceding phrase, and this can be especially useful for recognizing sentence boundaries in texts without or with inconsistent punctuation. For example:

- ⲁⲓϭⲟⲙϭⲙ ⲉⲡⲉϥⲥⲱⲙⲁ ⲧⲏⲣϥ ⲁⲓϩⲉ ⲉⲣⲟϥ · ⲉⲁϥⲟⲩⲱ ⲉϥⲙⲟⲩ · ⲉⲣⲉⲡϩⲱⲃ ⲟⲥⲕ || ⲁⲓϭⲱϣⲧ **ⲇⲉ** ⲁⲓⲛⲁⲩ · ⲉⲩⲕⲟⲗⲟⲃⲓⲟⲛ · ⲉϥⲁϣⲉ ⲉϩⲣⲁⲓ ·

  and I felt his body all over, and I found that he was dead, and that the skin had perished. And I looked and I saw a short-sleeved shirt hanging up inside the cave

The position marked with “||” above is a sentence boundary, signaled by the following ⲇⲉ. However the punctuation in the text does not indicate this division (though grammatically we can note that the proposition ⲁⲓϭⲱϣⲧ can stand by itself. The inclusion of ⲁⲓⲛⲁⲩ in the same sentence is more subjective (we could say it too is independent), but the lack of punctuation or any other overt signal of a new sentence suggests that a single sentence containing “ⲁⲓϭⲱϣⲧ ⲇⲉ ⲁⲓⲛⲁⲩ…” is justified.

Exceptions for extremely long sentences

Some Coptic texts have extremely long sequences of ostensibly subordinate clauses, especially ones marked by either circumstantial conversions or conjunctives. If the clauses become very long (well over 100 normalized word units), then it becomes more manageable to separate them, similarly to how we might feel about convoluted legal English. For comparison, consider contracts or declarations in English with very many clauses such as:

“Whereas … (20 words),
Whereas … (40 words),
and whereas (30 words),
…
therefore … (80 words)”

Although this structure might form a single sentence from a purely syntactic point of view, it seems unreasonable and unwieldy for users of the data to leave such blocks as single sentences. We therefore segment each such clause into its own sentence, but revert to normal sentence segmentation as soon as possible.

For example:

- ϫⲉⲛⲧⲟⲕ ⲡⲉ ⲛⲧⲁⲕⲉⲓⲛⲉ ⲙⲙⲟϥ ϣⲁⲣⲟⲓ ϩⲙⲡⲉϩⲟⲟⲩ ⲡⲁⲓ ⲉⲧⲉⲥⲟⲩⲙⲛⲧϣⲟⲙⲧⲉ ⲡⲉ ⲛⲁⲑⲱⲣ · || ⲉⲣⲉⲡⲉⲕⲣⲁⲛ ⲛⲁϣⲱⲡⲉ ⲛϩⲟⲧⲉ ϩⲛⲧⲧⲁⲡⲣⲟ ⲛⲟⲩⲟⲛ ⲛⲓⲙ · ⲉⲩⲉⲙⲟⲩⲧⲉ ⲉⲣⲟⲕ ϫⲉⲁⲃⲃⲁⲧⲱⲛ ⲡⲁⲅⲅⲉⲗⲟⲥ ⲙⲙⲟⲩ · || ⲉⲣⲉⲡⲉⲕⲉⲓⲛⲉ ⲙⲛⲧⲉⲕϩⲓⲕⲱⲛ ⲛⲁϣⲱⲡⲉ ϩⲛⲟⲩⲕⲣⲟⲙⲣⲙ ⲙⲛⲟⲩϭⲱⲛⲧ ⲙⲛⲟⲩⲁⲡⲉⲓⲗⲏ ⲉϩⲟⲩⲛ ⲉⲯⲩⲭⲏ ⲛⲓⲙ ϣⲁⲛⲧⲟⲩϯ ⲙⲡⲉⲩⲡⲛⲉⲩⲙⲁ · || ⲉⲣⲉⲛⲉⲕⲃⲁⲗ ⲙⲛⲡⲉⲕϩⲟ ⲛⲁϣⲱⲡⲉ ⲛⲧϩⲉ ⲛⲛⲓⲧⲣⲟⲭⲟⲥ ⲛⲕⲱϩⲧ ⲉⲩϥⲓ ϩⲟⲉⲓⲙ ϩⲟⲉⲓⲙ ϩⲓⲧϩⲏ ⲙⲙⲟⲓ :— || ⲉⲣⲉⲡⲉϩⲣⲟⲟⲩ ⲛϣⲁⲛⲧⲕ ⲛⲁϣⲱⲡⲉ ⲛⲧϩⲉ ⲙⲡⲉϩⲣⲟⲟⲩ ⲛⲧⲗⲓⲙⲛⲏ ⲛⲥⲁⲧⲉ ⲉⲧϫⲉⲣⲟ ϩⲛ ⲟⲩⲕⲱϩⲧ ⲙⲛⲟⲩⲑⲏⲛ · || ⲉⲣⲉⲡⲉϩⲣⲟⲟⲩ ⲙⲡϩⲙϩⲙ · ⲛⲛⲉⲕⲥⲡⲟⲧⲟⲩ ⲛⲁϣⲱⲡⲉ ⲛⲧϩⲉ ⲙⲡⲉϩⲣⲟⲟⲩ ⲛⲧⲥⲁϣϥⲉ ⲛϩⲣⲟⲩⲃⲃⲁⲓ ⲉⲩⲛⲁϣⲁϫⲉ ϩⲣⲁⲓ ϩⲛⲛⲉⲩⲁⲥⲡⲉ ·⁝·— || ⲉⲣⲉⲧⲉⲕⲁⲡⲉ ⲛⲁϣⲱⲡⲉ ⲛⲧϩⲉ ⲛⲛⲉⲓⲛⲟϭ ⲛⲥⲧⲩⲗⲗⲟⲥ ...

  for it was thou who didst bring him to Me on this day, which is the thirteenth for the month Hathor. Thy name shall be a terror in the mouth of every one. They shall call thee Abbaton, the Angel of Death. Thy form and thine image shall be [associated with] complaining, and wrath, and threatening in all souls, until they have yielded up their spirits. Thine eye and thy face shall be like unto a wheel of fire which beareth waves and waves [of fire] before me. The sound of thy nostrils shall be like unto the sound of a lake of fire wherein burn fire and sulphur (or, naphtha).

> The sound of the noises made by thy lips shall be like unto the sounds of the seven thunders which shall speak with their tongues. Thy head shall be like unto these great pillars ...

Although this sentence could be parsed as a sequence of circumstantial subordinate clauses, it would be extremely long, and the punctuation also suggests the division into multiple sentences, which is indicated again above using “||” (these bars are not present in the source MS, but the punctuation marks are). As a rule of thumb, we adhere to strict syntactic criteria for up to 100 normalized word units, allow flexibility for sentence structures with up to 300 words (left at the editor’s discretion whether to split), and mandate some split to prevent sentences of over 300 words, ideally splitting at a new predication unit.

Such contentious cases, though rare, most often involve long sequences of circumstantial clauses, and the editor’s task is then to choose a clause from which to start a new sentence.

Further examples

The following illustrative examples taken from the UD Coptic Treebank point out the application of the principles laid out above. In general, the Treebank is a good resource to find precedents for similar examples of tricky constructions, with the caveat that, in texts coming from versified sources (i.e. the Bible), the principle of preserving established versification may conflict with the grammatical criteria detailed here.

Double pipes (||) denote the selected sentence splitting in the examples and translations.

1. ⲛⲧⲉⲣⲟⲩⲣϣⲡⲏⲣⲉ ⲇⲉ ⲛϭⲓ ⲛⲇⲓⲕⲁⲓⲟⲥ · ⲁⲩⲱ ⲁⲩϫⲟⲟⲥ ϫⲉⲡϫⲟⲉⲓⲥ ⲛⲧⲁⲛⲣⲛⲁⲓ ⲛⲁⲕ ⲛⲁϣ ⲛⲟⲩⲟⲉⲓϣ · || ϥⲛⲁϫⲟⲟⲥ ⲛⲁⲩ ϫⲉϩⲁⲙⲏⲛ ϯϫⲱ ⲙⲙⲟⲥ ⲛⲏⲧⲛ ϫⲉⲉⲡϩⲟⲥⲟⲛ ⲁⲧⲉⲧⲛⲁⲁⲥ ⲛⲟⲩⲁ ⲛⲉⲓⲕⲟⲩⲓ ⲉⲧⲥⲟⲃⲕ · ⲁⲛⲟⲕ ⲡⲉⲛⲧⲁⲧⲉⲧⲛⲁⲁⲥ ⲛⲁⲓ ·

   Then the righteous, having marvelled (at these words), shall say, 'Lord, at what time did we ever do these things unto Thee?' || And He shall say unto them, ' Amen. I say unto you, inasmuch as ye have done it unto one of these few little ones, it is to Me that ye have done these things.' (On Mercy and Judgment, Chapter 34)

In example 1., note the precursive ⲛⲧⲉⲣⲉ followed by the particle ⲇⲉ, which form strong signals for a new sentence. Also note that the first ⲁⲩⲱ was interpreted as continuing the original sentence, despite the raised dot, since the precursive clause (“after marveling…”) is clearly syntactically subordinate to the verb ϫⲟⲟ⸗ ‘said’. By contrast, the response “he shall say” begins a new, syntactically independent proposition.

2. ⲁⲛⲟⲕ ⲇⲉ ϩⲱ ⲁⲛⲅⲟⲩⲣⲙⲛⲧⲟⲟⲩ ⲛⲧⲉⲕϩⲉ · ⲉⲓϣⲟⲟⲡ ϩⲙⲡⲉⲓϫⲁⲓⲉ · ⲉⲧⲃⲉⲛⲁⲛⲟⲃⲉ · || ⲛⲧⲟϥ ⲇⲉ ⲡⲉϫⲁϥ ⲛⲁⲓ ϫⲉⲛⲧⲕⲟⲩϣⲃⲏⲣ · ϩⲱⲱⲕ ⲟⲛ ⲛⲧⲉⲡⲛⲟⲩⲧⲉ ·|| ⲁⲓϩⲙⲟⲟⲥ ⲟⲛ ⲙⲡⲉϥⲙⲧⲟ ⲉⲃⲟⲗ · ⲁⲩⲱ ⲁⲓⲡⲁⲣⲁⲕⲁⲗⲉⲓ ⲙⲙⲟϥ · ⲉⲧⲣⲉϥϫⲱ ⲉⲣⲟⲓ ⲙⲡⲉϥⲣⲁⲛ ·

> I also am a man of the mountain like unto thyself, and I am living in the desert because of my sins.' || And he said unto me, ' Thou art a friend of God.' || And I sat down before him, and I conjured him to tell me his name. (Life of Onnophrius, Chapter 9)

In example 2., note how the alternation of speakers corresponds to sentence transitions. Once the quoted speech ends, we have new predications "said" for sentence 2 and "sat" for sentence 3. This is mirrored by the presence of clitic particles, ⲇⲉ and ⲟⲛ. The final ⲁⲩⲱ has been interpreted as sentence internal coordination, which is plausible since both verbs have the same subject and there is no boundary marker (clitic or other construction) between them, while 'ⲁⲩⲱ' forms an explicit link.

3. ⲁⲣⲓⲁⲡⲁⲧⲟⲟⲧⲕ ⲉⲧⲣⲉϥⲥⲙⲟⲩ ⲉⲣⲟⲕ ⲛϣⲟⲙⲛⲧ ⲛⲥⲟⲡ · ϫⲉⲉⲣⲉⲡⲉⲥⲙⲟⲩ ⲙⲡⲁⲅⲅⲉⲗⲟⲥ ⲉⲧⲙⲟⲟϣⲉ ⲛⲙⲙⲁϥ ⲉⲓ ⲉϫⲱⲕ · || ⲉⲧⲃⲉⲧⲡⲓⲥⲧⲓⲥ ϩⲱⲱⲥ ⲛⲧⲕⲁⲑⲟⲗⲓⲕⲏ ⲉⲕⲕⲗⲏⲥⲓⲁ ⲙⲡⲣⲕⲁⲁⲩ ⲛⲧⲕ ⲉⲡⲁϩⲟⲩ ⲛϩⲏⲧⲥ ⲟⲩⲇⲉ ⲉⲃⲟⲗⲕ ⲉⲃⲟⲗ · || ⲧⲛⲡⲓⲥⲧⲉⲩⲉ ⲉⲩⲛⲟⲩⲧⲉ ⲛⲟⲩⲱⲧ ⲡⲉⲓⲱⲧ ⲡⲡⲁⲛⲧⲟⲕⲣⲁⲧⲱⲣ ⲙⲛⲡⲉϥⲙⲟⲛⲟⲅⲉⲛⲏⲥ ⲛϣⲏⲣⲉ ⲓⲏⲥⲟⲩⲥ ⲡⲉⲭⲣⲓⲥⲧⲟⲥ ⲡⲉⲛϫⲟⲉⲓⲥ ⲡⲉⲛⲧⲁⲡⲧⲏⲣϥ ϣⲱⲡⲉ ⲉⲃⲟⲗ ϩⲓⲧⲟⲟⲧϥ · ⲙⲛⲡⲉⲡⲛⲉⲩⲙⲁ ⲉⲧⲟⲩⲁⲁⲃ · ⲉⲧⲉⲧⲁⲓ ⲧⲉ ⲧⲣⲓⲁⲥ ⲉⲧⲥⲙⲁⲙⲁⲁⲧ · ⲡⲁⲓ ⲡⲉ ⲡϫⲱⲕ ⲛⲧⲙⲛⲧⲛⲟⲩⲧⲉ ·

   Do thine utmost to make him bless thee three times, so that the blessing of the angel who walketh with him may come upon thee. || And as regards the Faith itself of the Catholic Church, do not let thyself backslide therein, neither do thou put thyself outside it. || We believe in the One God, the Father the Almighty, and in His only-begotten Son, Jesus the Christ, our Lord, through Whom the Universe came into being, and in the Holy Spirit, that is to say, in the Blessed Trinity, || which is the complete Godhead. (Letter of pseudo-Ephrem, Chapter 8)

In example 3, note first that an imperative can form the independent predicate of a sentence, which in this case takes a subordinate clause with ϫⲉ after the raised dot. The second sentence has several signals: a fronted prepositional object (ⲉⲧⲃⲉⲧⲡⲓⲥⲧⲓⲥ), an enclitic ϩⲱⲱⲥ, and a new negative imperative predicate. After the punctuation, sentence 3 begins with a declarative, i.e. a change in grammatical mood for the new sentence, now with the subject 'we'. Finally note that in the translation, sentence 4 is a relative clause which is part of sentence 3. However in the Coptic original, there is no relative clause, and the final "ⲡⲁⲓ ⲡⲉ ⲡϫⲱⲕ ⲛⲧⲙⲛⲧⲛⲟⲩⲧⲉ" is actually a grammatically independent sentence, which refers back using a demonstrative pronoun, not a relative. A more literal translation might read "This is the complete Godhead", using a new sentence.