About this Guide

This annotation Style Guide was created by Chiara Palladino and David J. Wright for the Ugarit Project. The languages involved are Ancient Greek as source language and Latin as translation language.

The Guide was created as follows:
- The authors wrote a preliminary draft with fewer rules, addressing predictable issues and sources of variation, based on previous experience at aligning similar texts in Ugarit.
- The authors manually aligned a sample set of fragments, testing the rules for consistency.
- For each new or unpredicted situation, there was a brief discussion and a preferred annotation style or improvement to the Guide was agreed upon. Each change was incorporated in the final version.

Disclaimer

This Guide specifically targets the alignment of texts in Ancient Greek from the *Fragmenta Historicorum Graecorum,* a collection of fragments from lost ancient historians realized by Karl Müller in five volumes, and printed in Paris between 1841 and 1873 by the publisher Ambroise Firmin-Didot. The texts aligned during this project were provided by the *Digital Fragmenta Historicorum Graecorum* produced by Monica Berti (https://www.dfhg-project.org/). Therefore, while most of the issues here addressed can be generalized to philological Latin as a language of translation, some of them will be peculiar to the texts of this collection.

These guidelines can be used to perform alignments between various texts in Ancient Greek and English and to create gold standard corpora. However, they are not designed to answer specific research questions. If you have a particular research topic and you need specific results for the purpose of language research, pedagogy, or data mining, you should create your own guidelines (you can reuse any part of these guidelines).

General Principles

These guidelines follow the definition of Lambert et al.: "the correspondence between two lexical units (single words or expressions) should involve on both sides as few words as possible, but as many words as necessary, with the requirement that the linked words or groups bear the same meaning [...]. If two single words match, they should be linked together, but if there is no correspondence at a single word level, groups of words should be linked together. Groups of words linked together should be as small as possible." In addition, "the only valid elements in an alignment are single words and indivisible groups of words" (p.275): groups of words are linked together when the meaning of the group is distinct from that of the sequence of each word's

meaning, and single words cannot be separated from the rest of the group without changing their meaning (= indivisible lexical unit).

Links allowed in Ugarit are one-to-one (1-1), many-to-many (N-N), one-to-many (1-N) and many-to-one (N-1). Currently it is not possible to distinguish between ambiguous and unambiguous, complete or partial. Links in Ugarit do not consider lack of alignment: words that do not correspond are simply left unaligned.

Guidelines

1. Punctuation

Punctuation is never aligned.

Exception: the apostrophe in the Saxon genitive, when aligned with a corresponding word/group of words in Ancient Greek.

2. Words omitted or incorrectly translated

If a group of words only appears in one language whose meaning has no correspondence in the other, it is not aligned.

If a group of words is an incorrect translation of a group of words in the source language, it is not aligned.
This includes pronouns or determiners translated explicitly with the names they refer to:
Example: Numerius Fabius Pictor 29: ὅτε τοῦ ἔθνους τούτου κατέστησαν κύριοι - quum Sabinorum gentem sibi subjecissent, 'τούτου' doesn't align with 'Sabinorum'.

If a word that is referred to two or more words, for example a preposition referring to multiple nouns, appears only once in the source or the translation, it is considered as forming a group with only the first word of the sequence and aligned accordingly. Example: Heraclides Ponticus 4: περὶ ἐσθῆτα ἢ κουράν - neque veste neque tonsura > 'περὶ ἐσθῆτα' - 'veste', 'κουράν' - 'tonsura'.

3. Repetition

When a word is repeated in a language but not in the other, only the first instance in each language is aligned.

4. Phrasal Construction and Idioms

These are expressions that constitute an indivisible lexical unit, where each word singularly does not bear the same meaning, but the group of words together is a semantic equivalent in both languages. In these cases, the group is aligned with a N-N link. Example: Hellanici fragmenta 123, p. 61: 'κατὰ Μίδαν' - 'Midae temporibus'.

5. Enclitics

When Greek words include a conjunction that is rendered in Latin with an enclitic, e.g. -que, they are aligned with an N-1 link including the conjunction and the word. <u>Example</u>: Apollodorus Atheniensis 2.5 'καὶ κτείνας' - 'occisoque'

6. Words quoted in the original language

When a word is quoted in the translation and left in the original Ancient Greek, it is not aligned. <u>Example</u>: Neanthes Cyzicenus Lysimachus 29: 'βωμοὺς' - ø.

7. Determiners

Definite articles are used in Ancient Greek, but not in Latin. When a noun appears with the article in AG, the article is not aligned: only the noun is aligned to the corresponding Latin translation.

Sometimes, an article in AG may be rendered in Latin through a demonstrative pronoun. In this case, the two are aligned, if deemed appropriate by the annotator. <u>Example</u>: Hecataeus 329: τὰς περιηγήσεις - illarum descriptionum > 'τὰς' - 'illarum', 'περιηγήσεις' - descriptionum.

Exceptions: substantive participles, for which see Rule on Participles.

7.1 Pronominal uses of articles in Ancient Greek

Sometimes Ancient Greek uses articles with particles like δὲ as pronouns. Latin responds differently depending on context.

8. Pronominal equivalents: the whole group in Ancient Greek is aligned with the corresponding Latin if they have equivalent meaning. <u>Example</u>: Theopompus 5: 'οἱ μέν' - 'alii'.
9. Adverbial constructs: the whole group in Ancient Greek is aligned with the corresponding Latin if they have equivalent meaning. <u>Example</u>: Phlegon Trallianus 45: 'τοὺς νῦν' - 'nostri'.

10. Verbs

8.1 Verbs and Subjects

When the subject is a separate noun or pronoun in both languages, it is aligned with a 1-1 link.
1. When the subject is not explicit in both languages, only the verb is aligned.
2. When the subject is explicit in Latin, but implicit in Ancient Greek:

a. If the Latin subject is a pronoun, it is aligned together with the verb, with a 1-N link. Example: 'λέγω' - 'ego dico'.
b. If the Latin subject is a noun or proper noun that is not expressed in Ancient Greek, it is not aligned and only the verb is aligned.

8.2 Verbs and prepositions

Verbs in Ancient Greek translated into Latin may include a subordinating particle that is structural to the verb construction in Latin, but is not necessary in Ancient Greek. In these cases, the original is aligned to the translated verb + particle as long as the translation is deemed sufficiently accurate by the annotator. Example: Apollodorus Atheniensis 2.2, ἠνάγκαζε … ἀνέχεσθαι - cogebat ut … tenerent > 'ἠνάγκαζε' - 'cogebat ut', 'ἀνέχεσθαι' - 'tenerent'.

9. Participles in Ancient Greek

9.1 Genitive Absolute

When the genitive absolute in Greek is translated as an ablative absolute, the corresponding words should be aligned 1-1. Example: Apollodorus Atheniensis 6.3: γενομένων … τούτων - quibus factis > 'γενομένων' - 'factis', 'τούτων' - 'quibus'.

9.2 Attributive participles

Attributive participles in Greek are aligned with preference for a 1-1 link. The definite article that generally appears in Ancient Greek but not in Latin is not aligned. Example: Heraclides Ponticus 4: 'τοῖς'-'ø' 'τελευτῶσιν'-'defunctos' 'ἀνδράσι'-'viri'.

9.3 Substantive participles

1. Substantive participles translated as relative clauses et sim.: if the article is present in Ancient Greek, we privilege a 1-1 link, aligning the article with the pronoun or equivalent, and the participle with the corresponding verbal form. Example: Hermippus Callimachius 27: τοὺς δὲ κληθέντας (εἶναι) - qui vero invitati (erant) > 'τοὺς'-'qui', 'δὲ'-'vero', 'κληθέντας'-'invitati'.
2. Other translations: sometimes a substantive participle will be translated with a noun. In these cases, the noun is aligned to the corresponding participle, without including the article. Example: Hecataeus 329: ὁ γεγραφὼς - auctor > 'ὁ' - ø, 'γεγραφώς' - 'auctor'.

9.4 Participle clauses

1. Participles translated as a dependent clause, e.g. Lat. *cum* + subjunctive: the full form is aligned as a 1-N or N-N group. <u>Example</u>: Aristoxenus Tarentinus 27: 'εἰπὼν' - 'quum…dixisset'.
2. Participles translated explicitly in other ways, such as prepositional phrases or relative clauses, are also aligned as a 1-N group, provided that the translation is deemed sufficiently accurate by the annotator. <u>Examples</u>: Heraclides Ponticus 1: 'ληστεύοντες' - 'ex latrociniis'; Demetrius Callatianus 1: 'λεγόμενον' - 'qui dicuntur'.

10. Prepositional phrases

A prepositional phrase in Greek that is translated with an equivalent preposition + noun in Latin should be aligned with a 1-1 link. <u>Example</u>: Apollodorus Atheniensis 2.5: μετὰ τῶν παίδων, cum filiis > 'μετὰ' - 'cum', 'τῶν' - ø, 'παίδων' - 'filiis'.

11. Case uses

1. An inflected noun in Greek is often translated as a word group in Latin, e.g. a prepositional phrase: the noun in Greek is aligned with the whole phrase in Latin (1-N link), if the translation is deemed sufficiently accurate by the annotator. Such cases include partitive genitives, accusatives of motion, etc.

   > <u>Examples</u>: Apollodorus Atheniensis 2.2 Κορίνθιον ἰσθμόν - in Corinthio isthmo > 'Κορίνθιον' - 'Corinthio', 'ἰσθμόν' - 'in isthmo'; Apollodorus 6: 'Ἀθήνησιν' - 'in urbe Athenarum'.

   This rule also applies to adverbial constructions. <u>Example</u>: Heraclides Ponticus 4, 'ἐνιαυτός' - 'per annum'.
2. The reverse could also happen, where a prepositional phrase in Greek is rendered with an inflected word in Latin. The same rule applies as well (N-1 link), as long as the translation is deemed sufficiently accurate by the annotator. <u>Example</u>: Heraclides Ponticus 4, 'περὶ ἐσθῆτα' - 'veste'.

Exceptions: The fixed expression in Greek ἐν + book number, e.g. Theopompus 176, ἐν τῇ εἰκοστῇ πέμπτη, is aligned privileging a 1-1 link wherever possible. The Latin translation may contain the preposition 'in', which is aligned to 'ἐν', otherwise the preposition is not aligned. Moreover, if the Latin translation supplies the word 'libro', rule 2 on Words Omitted applies and the word is not aligned. <u>Example</u>: Theopompus 176; ἐν τῇ εἰκοστῇ πέμπτη - libro vigesimo quinto > 'ἐν' - ø, 'τῇ' - ø, 'εἰκοστῇ' - vigesimo, 'πέμπτη' - quinto, ø - 'libro'.

12. Functional words and particles

As a general principle, the functional words which are required for grammatical reasons are aligned when they have an equivalent counterpart in the other language. They are aligned with a 1-1 link if there is an equivalent Latin translation.

a. Particles in combination can be aligned as a group to the corresponding translation (N-N) if they cannot be omitted without changing the meaning of the expression, and if the translation bears an equivalent meaning to their function in combination. E.g. 'ἔτι καὶ νῦν' - 'etiam nunc', Numerius Fabius Pictor 29, 'τότε πρῶτον' - 'tunc demum'; 'οὐ μόνον' - 'non modo'.
b. Functional words and particles: δὲ, μὲν, δή…
    -   Typically, μέν and δέ should be aligned separately to the corresponding translation, if available. Sometimes, δέ at the start of a sentence is translated as *vero* or similar, in which case it should be aligned to the corresponding adverb, provided that there is sufficient semantic similarity: Example: Theopompus 221.a 'δέ' - 'vero'.
    -   When a particle has an intensive function and it is translated in Latin with a semantic equivalent (even if not a particle), the two may be aligned with a 1-1 link. Example: Androtio 39: λέγω δή - ego dico > 'λέγω' - 'dico', 'δή' - 'ego'.

13. Abbreviations

Words that are abbreviated in one of the two languages may be aligned against the corresponding word, including punctuation (in exception to Rule 1: Punctuation). Example: Aristotle 178: 'πολιτείᾳ' - 'rep.'.

Literature

Patrik Lambert, Adrià de Gispert, Rafael Banchs, and José B. Mariño. 2005. Guidelines for word alignment evaluation and manual alignment. *Language Resources and Evaluation.* 39:267-285.

João Graça, Joana Paulo Pardal, Luísa Coheur and Diamantino Caseiro. 2008. Building a golden collection of parallel Multi-Language Word Alignments. *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08).*

I. Dan Melamed, Annotation Style Guide for the Blinker Project. 1998. *IRCS Technical Reports Series.* 53.

C. Palladino and F. Shamsian, Ancient Greek to English Alignment Guidelines. 2022. https://github.com/UgaritAlignment/Alignment-Gold-Standards/blob/main/grc-eng/GRC-ENG%20 0Guidelines.pdf