

About this Guide

This annotation style guide is a modification of the general guidelines for aligning Ancient Greek and English in the Ugarit Project

(https://github.com/UgaritAlignment/Alignment-Gold-Standards/blob/main/grc-eng/guidelines_grc-eng.pdf).

This guide modifies the general guidelines to be more useful for the analytical comparison of translations through manual and/or supervised alignment. It still addresses the structural linguistic differences between Ancient Greek and English in the same way as the general guidelines, but it is designed specifically to allow a certain freedom to the annotator, in finding the most appropriate way to align texts to respond to specific research questions. Therefore, it is not functional to the creation of machine-actionable models or gold standards.

Disclaimer

These guidelines are designed with a specific project in mind: using Ugarit as a tool to analytically compare the relationship between a text in Ancient Greek and multiple translations in English. The linguistic aspects are based on the GRC-ENG general Guidelines created for the Ugarit project, but additional aspects are here considered, that pertain to phenomena that a researcher may want to investigate when analytically comparing translations of a literary work.

Rather than prescriptive, these guidelines are conceived as documentation for the alignment of sections of Euripides' *Hippolytus* against four very different English translations. They may be further adapted and modified according to the needs of other types of texts.

General Principles

See the General Guidelines for the foundational principles. In addition, the main principle is that the annotator is to follow more closely the specificities of the text, rather than of the language: e.g. formulaic or fixed expressions, even if only attested in that particular author or text, may be considered as translation units with more tolerance than in the General Guidelines.

The broader guiding principle, when aligning texts to evaluate translations, is whether a particular translation pair constitutes a unit that a researcher may want to extract/analyze further after the alignment is completed. If the answer is yes, this principle governs over the more strict linguistic ones established in the General Guidelines.

These guidelines are intended to be more tolerant in two directions:

1. Redefining the principle of “acceptable translation” in such a way that semantic similarity is not the only governing criteria. If a translation is deemed as not adherent to the

original but the annotator considers it important for retrieval purposes, it may be aligned to the corresponding original text.

2. Introducing more flexibility in grammar and syntax: translations that vary the morphosyntactic structure considerably, or that add elements to the original (e.g. attributes or adverbs added with the purpose of making the context clearer), are usually disregarded in the General Guidelines. Here, they may be something that the annotator wants to retrieve afterwards, as they contain information on the behavior of a translator; therefore, they may be aligned to the original.

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. REVISION: If a translator uses additional words or emphasis in certain expressions, they may be aligned against the corresponding baseline in Ancient Greek, if they represent an intentional feature of that translation.

If a group of words is an incorrect translation of a group of words in the source language, it is not aligned. REVISION: Generally it is desirable to retain this principle, however it could happen that an annotator wishes to retain information about incorrect translations by aligning them, for future queries. These cases should be accurately documented.

3. Expansion, adjectivization, emphasis - Fundamental revision

When a translation shows evidence of expansion on the original, that expansion may be considered as part of the translation pair even though it represents an addition to the literal meaning. A typical example is the addition of adjectives or determiners that are not explicit in the original but are meant to transmit a certain interpretation of the text. In these cases, they should be aligned by the annotator as long as they are considered correct.

Examples:

Hippolytus, 1-20: 'τερμόνων' - 'the very edges'; 'Θησέως' - 'by the seed of Theseus'; 'ἀγνοῦ' - 'that pure man'.

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. They are aligned with a N-N link if there is an English equivalent. Examples: dative of possession: 'ἔστιν αὐτοῖς' - 'they have'; 'Ἔτι καὶ νῦν' - 'even to this day'; 'Ἐκ χειρός' - 'in a hand-to-hand encounter'.

REVISION: In the case of translation analysis, the definition of phrasal construction is broader and more comprehensive. For instance, it will include formulaic expressions typical of one author or text, or recurring phrases that are clearly addressed by the majority of translators as indivisible units.

Examples: 'ἔχει δ' ἀνάγκην' - 'there is a fatal necessity'.

5. Repetition

When a word is repeated in a language but not in the other, only the first instance in each language is aligned (unless context suggests otherwise).

6. Negation

Whenever possible, negation is aligned with a 1-1 link. However, in two cases that won't be possible. In such cases, an N-N strategy is recommended.

1. When the negative English morpheme is abbreviated (-n't):
 - a. When it cannot be separated from the word group, it is aligned N-N. Example: 'ἔστιν οὐ' - 'it isn't'.
 - b. When it only consists of an auxiliary, e.g. 'don't', it may be aligned with the negative particle, and the verb is aligned with the corresponding English word. Example: μὴ ἄλλως ποίει - don't do otherwise: 'μὴ' - 'don't', 'ποίει' - 'do', 'ἄλλως' - 'otherwise'.
2. When the English morpheme is not abbreviated, e.g. 'it is not', 'do not', etc., the auxiliary is not aligned and the verb is aligned with the corresponding English word with a 1-1 link wherever possible.
3. When the structure is changed in English: Example: 'οὐδὲ δυναμένους' - 'were unable'; 'μὴ πείθεσθαι' - 'to refuse obedience'.

7. Verbs

6.1. Auxiliaries, Verbal Groups

Verbal groups are groups formed by more than one verb: this includes auxiliaries, passivization, translations of tense and aspect, etc.

All verbal groups are aligned as a whole regardless of the number of words. The type of link can be N-N, N-1, or 1-N. Example: 'τελέεσθαι' - 'will be brought to pass'.

6.2. Verbs and Subjects

Since Ancient Greek is a highly inflected language, often the subject of a verb is left implicit and only expressed by means of the personal ending. However, in English the subject of a verb is almost always explicit.

1. When the subject is a separate noun or pronoun in both languages, it is aligned with a 1-1 link.
2. When the subject is not explicit in both languages, only the verb is aligned.
3. When the subject is explicit in English, but implicit in Ancient Greek:
 - a. If the English subject is a pronoun, it is aligned together with the verb, with a 1-N link. Example: 'ἔσαν' - 'they were'.
 - b. If the English subject is a noun or proper noun that is not expressed in Ancient Greek, it is not aligned and only the verb is aligned.

These rules also apply to infinitives in infinitive+accusative constructs in Ancient Greek.

6.3. Verbs and Implicit objects (revision)

Implicit direct object that is not in Greek but indicated in English with a pronoun et sim. Is aligned together with the verb, as long as the annotator deems it appropriate.

Example: 'πρεσβεύω' - 'I respect them'.

6.4. Passive verbs translated into active

Verbs are aligned regardless of voice (active, middle, or passive), provided that the English translation bears the same meaning and is correct in the judgment of the annotator. This means that a passive verb in Ancient Greek may be aligned with an active verb in English. Example: 'φάανθεν' - 'shone'.

6.5. Compound verbs with prepositions

Compound verbs (verbs constructed with prepositions in Ancient Greek, e.g. περι-πλέω) are aligned with the corresponding English expression, when they are equivalent in meaning.

Example: 'περιπλέω' - 'I sail around'.

8. Participles in Ancient Greek

7.1. General rules

1. Participles translated with an attributive or implicit form, e.g. -ing form in English: the participle is aligned with the corresponding word/s in English. The noun that agrees with the participle is aligned separately, with the corresponding translation in English. Example: 'φέρων' - 'bearing'.

2. Participles translated as a dependent clause in English: the participle is aligned with the whole expression in the English translation (1-N). The noun that agrees with the participle is aligned separately, with the corresponding translation in English. Example: - 'κιῶν' - 'when he had gone'.
3. Substantive and attributive participles: when translated explicitly (mostly as a relative clause), the participle is aligned with the full explicit translation. Example: τοὺς μὲν ἔχοντας - some people who had: 'τοὺς μὲν' - 'some people'; 'ἔχοντας' - 'who had'.
 - a. When an article is present, it is aligned together with the participle (N-N), unless the translation does not bear equivalent meaning. Example: πάντες δὲ οἱ καλούμενοι - all who are called: 'πάντες' - 'all'; 'οἱ καλούμενοι' - 'who are called'.

7.2. Genitive absolute

The participle forming a genitive absolute is aligned with the corresponding translation in English, following the general rules of participles (1-1 or 1-N link type). The noun that agrees with the participle is aligned separately, with the corresponding translation in English. Example: ἐμεῦ ζῶντος - while I live: 'ἐμεῦ' - 'I', 'ζῶντος' - 'while live'.

9. Prepositional phrases

Every component of a prepositional phrase is aligned separately with the corresponding English translation. Sometimes a preposition will be translated with an imperfect equivalent in English: such cases should be evaluated by the annotator to decide whether the translation is just partial or incorrect, and aligned accordingly. Example: 'ἐν' - 'in', 'τῇ' - 'the', 'νῆσῳ' - 'island'.

10. Case uses

Inflected nouns (including substantivized adjectives, pronouns, etc.) are aligned with the corresponding English grammatical translation, provided that it bears equivalent meaning (1-N link). Example: θυμοῦ - 'of the heart'.

Any modifier of the noun (attributes, determiners, etc.) does not obey this rule, since its inflection is considered as originated from agreement with the noun.

REVISION: there may be cases where the English preposition used is a "free" translation in the judgment of the annotator. In these cases, the general principle prevails: if that particular translation is something worth to be extracted from the aligned data, then a higher tolerance is warranted and it may be matched even though it is an imperfect rendering of the original. Example: 'θυμοῦ' - 'against the heart'.

NOTE that this principle does not apply when the preposition is implied in the use of the verb, and not of the case: in these cases, the preposition should be aligned with the verb.

9.1. Vocatives

Nouns in the Vocative case in Ancient Greek are aligned to the corresponding translation in English. The exclamation (Ancient Greek ὦ or Eng. 'oh') is aligned to the corresponding word when present in both languages. When absent in one of the two, it is not aligned.

11. Pronouns and anaphoric structures

Pronouns are aligned as a 1-1 link with the corresponding word in English when they are equivalent in meaning.

Pronouns or other determiners (e.g. possessives, demonstratives, etc.) that replace a noun in one of the two languages but not in the other are not aligned. Example: 'he' and Ἀγαμέμνων are not aligned (even though it's the same person in the context).

12. Determiners (articles, etc.)

1. When determiners are present in both languages and are equivalent in meaning, they are aligned with a 1-1 link. Examples:
 - a. 'τις ἄνθρωπος' transl. 'a man'/'some man'/'any man': 'τις' - 'a/some/any', 'ἄνθρωπος' - 'man'.
 - b. 'ὁ ἄνθρωπος' transl. 'the man'/'this man'/'that man': 'ὁ' - 'the/this/that', 'ἄνθρωπος' - 'man'.
 - c. Possessive adjectives are included among the options of acceptable translations. Example: τῶν ἵππων - of their horses: 'τῶν' - 'their', 'ἵππων' - 'of horses'.
2. When a determiner is present in both languages, but they do not bear equivalent meaning, they are only aligned if the annotator considers them relevant to their investigation. Example: 'ὁ ἄνθρωπος' transl. 'a man': 'ὁ' - 'a', 'ἄνθρωπος' - 'man'.
3. When the determiner is present in English but not in Ancient Greek: it is aligned together with the noun to the corresponding Ancient Greek word, provided that they are semantically equivalent (N-1 link). Example: 'ἄνθρωπος' - 'the man'/'some man'/'a man'/'that man', etc.
4. When the determiner is present in Ancient Greek but not in English: it is not aligned, unless absolutely necessary, i.e. in nouns that require the article in Ancient Greek but not in English. Examples: 'ὁ Κῦρος' - 'Cyrus', 'τῇ Εὐρώπῃ' - 'Europe'. Incl. substantivizations of adjectives and verbs in certain contexts. Examples: 'τὸν αἰσχρόν' - 'moral wrong', 'τὸ μανθάνειν' - 'learning'.
5. Phrasal translations are aligned to the group of determiner + noun, when they bear equivalent meaning. Example: 'τις ἄνθρωπος' - 'someone'.

13. Pronominal uses of articles in Ancient Greek

Sometimes Ancient Greek uses articles with particles like δὲ as pronouns. English translations of this phenomenon differ quite significantly:

1. Pronominal equivalents: the whole expression in Ancient Greek is aligned with the corresponding English if they are equivalent in meaning. Example: 'Οἱ δὲ' - 'they', 'those', 'those people', 'all the people who', 'other people', 'men', etc.
2. Pronominal uses + prepositional phrases:
 - a. When translated literally, the pronominal expression is aligned with the corresponding English expression, and the prepositional phrase is aligned following the rules of prepositional phrases. Example: Οἱ δὲ ἐν τῇ νήσῳ - Those in the island: 'Οἱ δὲ' - 'those'; 'ἐν' - 'in'; 'τῇ' - 'the'; 'νήσῳ' - 'island'.
 - b. When translated with a close semantic equivalent: both groups are aligned, following the rules of idioms, with a N-N link. Example: 'Οἱ δὲ ἐν τῇ νήσῳ' - 'the islanders'.
 - c. When translated with the addition of words not present in Ancient Greek: the additional words are not aligned, the rest is aligned according to the guidelines. Example: Οἱ δὲ ἐν τῇ νήσῳ - the inhabitants in the island: 'Οἱ δὲ' - ∅; 'ἐν' - 'in'; 'τῇ' - 'the'; 'νήσῳ' - 'island'.

14. Adjectives and Adverbs

Adjectives and adverbs are aligned as individual lexical units, regardless of their function in the context (attributes, predicatives, nouns, etc.).

Comparatives and superlatives, which are usually single words in Ancient Greek, are aligned with the corresponding English translation, with a 1-1 link or, when necessary, a 1-N link.

Example: 'κάκιστος' - 'ugliest' / 'the ugliest' / 'the most ugly' / 'by far the most ugly', etc..

15. 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 English translation.

Exceptions and special cases:

1. ἼAv (when not equivalent of ἑάν):
 - a. + Verbs: if the English translation of this group bears an equivalent meaning, then the group is aligned with a N-N link to the whole English expression. If that is not the case, ἼAv is not aligned. Example: ἼAv δοκῶσι - 'seem likely'.
 - b. In fixed phrasal constructions, e.g. ὅπως ἼAv: ἼAv is aligned within the phrasal construction (N-1 or N-N link). Example: ὅπως ἼAv - 'just as'.
2. τε, μέν, δέ, γάρ, γε, δὴ et sim.: are aligned 1-1 to the corresponding English translation, if it is explicit, clearly identifiable, and isolated from other words.
 - a. τε καὶ group: when the translation is only one word, only καὶ is aligned. Example: 'τε καὶ': 'τε' - ∅; 'καὶ' - 'and'.
 - b. Combined uses (e.g. μέν + δέ, οὐδέ + οὐδέ, etc.): particles in combination can be aligned as a group to the corresponding English translation if they cannot be omitted without changing the meaning of the expression, and if the English

translation bears an equivalent meaning to their function in combination.

Example: 'μέν...ἄέ' - 'on the other hand'/'on the one hand...on the other hand',
'καί...καί' - 'both...and', 'οὔτε...οὔτε' - neither...or.

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.