

A Greek-Chinese Interlinear of the New Testament Gospels

John Lee¹, Simon S. M. Wong², Pui Ki Tang¹ and Jonathan Webster¹

¹The Halliday Centre for Intelligent Applications of Language Studies, City University of Hong Kong

²United Bible Societies

E-mail: {jsylee, pktang, ctjjw}@cityu.edu.hk, ssmwonghk@yahoo.com

Abstract

This paper describes an interlinear text consisting of the original Greek text of the four gospels in the New Testament, and its Chinese gloss. In addition to the gloss, the Greek text is linked to two linguistic resources. First, it has been word-aligned to the *Revised Chinese Union Version*, the most recent Chinese translation of the New Testament; second, it has been annotated with word dependencies, adapted from dependency trees in the PROIEL project. Through a browser-based interface, one can perform bilingual string-based search on this interlinear text, possibly combined with dependency constraints. We have evaluated this interlinear with respect to the accuracy, consistency, and precision of its gloss, as well as its effectiveness as pedagogical material.

1. Introduction

A bilingual corpus typically consists of a source text and its translation in a foreign language. In addition to these two parallel texts, an interlinear text (or, simply, “an interlinear”) provides linguistic information for each word in the source text, most commonly a morphological analysis and a gloss, i.e., a translation of the word in the foreign language. The typical user of an interlinear is primarily interested in the source text, and needs linguistic support to read it.

Interlinears can serve readers at every level of proficiency of the source language. Advanced students can enjoy a quicker and smoother pace of reading, since the glosses reduce the interruptions caused by dictionary look-ups. For beginners, even dictionaries can be difficult to use, since the conversion from the inflected form of a word to its root form, or dictionary form (e.g., in Greek, from the surface form *andra* to the root form *anēr*), is no trivial task; choosing the appropriate meaning from the various senses provided in the dictionary entry (e.g., *anēr* can mean ‘man’ or ‘husband’) can also be challenging. Morphological analyses and glosses remove both obstacles.

For readers who do not aim to learn the source language, interlinears can still reveal important linguistic features of the source text that are obscured in the foreign translation. For example, Koine Greek explicitly places the word *heis* ‘one’ in front of a noun to emphasize the uniqueness of the entity in question (Bauer et al., 2001); unfortunately, the equivalent Chinese word, *yī* ‘one’, cannot convey this emphasis since it is routinely used with all nouns to indicate indefiniteness. A Chinese reader, if made aware of the presence of the word *heis*, would better appreciate the idea of ‘one and only’ that is strongly underlined by the Greek author.

This paper describes a Greek-Chinese interlinear of the four gospels --- Matthew, Mark, Luke and John --- which

constitute the first four books of the New Testament in the Bible. The rest of the paper is organized as follows. Section 2 reviews previous work. Section 3 outlines the design principles of our interlinear. Section 4 discusses the implementation of the interlinear. Section 5 presents evaluations in section 5, followed by conclusions.

2. Previous Work

While interlinears have a long history in opening up Classical works to modern readers (Ovid, 1828; Xenophon, 1896), they are also valued for religious texts, whose readers attach special reverence for sacred texts in their original languages. For example, both the Qur’an (Eckmann, 1976; Dukes and Habash, 2010; etc.) and the Bible have been interlinearized with various languages (Green, 1984; Wang, 1985; etc.)

For the New Testament, most interlinear texts have been in English (Green, 1984; Marshall, 1993; Brown and Comfort, 1993); so far, only one Greek-Chinese version has been published (Wang, 1985). This pioneer work is, however, almost completely outdated from the point of view of both the Chinese and Greek texts.

For the Chinese text, Wang (1985) used the *Chinese Union Version*¹ (1919). Since the glosses were also selected mainly from this 90-year-old translation, their vocabulary, expressions and word senses diverge considerably from contemporary Chinese usage. For the Greek text, Wang (1985) used the 1952 edition of the *Novum Testamentum Graece* (Nestle and Aland, 1952), a version that is no longer acceptable to the scholarly community.

Our interlinear brings both texts up-to-date with the latest Greek edition (Nestle and Aland, 1994) and the recent *Revised Chinese Union Version* (2010). Henceforth, we refer to these as the “Greek text” and the “Chinese text”.

¹ This translation in turn reflects a 19th-century edition of the Greek New Testament (Palmer, 1881).

3. Interlinear Design

Our interlinear improves upon (Wang, 1985) in two ways. In terms of *new content*, we indicate correspondences to Greek morphology in the Chinese glosses (section 3.1), and provide word alignments between the Greek text and the Chinese text (section 3.2); in terms of *methodology*, we address issues in gloss accuracy, precision and consistency (sections 3.3 to 3.5).

3.1 Correspondences to Greek morphology in Chinese glosses

To indicate case, gender, number, tense or mood, Greek employs an extensive system of suffixes that are attached to the lemmas of nouns, adjectives and verbs. Chinese expresses the equivalent information not with suffixes, but with separate characters. It is critical to convey these concepts to the Chinese speaker; to this end, our Chinese glosses clearly distinguish between characters corresponding to the meaning of the Greek lemma, and those corresponding to the suffix(es).

Greek	<i>boēthēi</i>	<i>kyrie</i>
	‘help!’	‘O Lord’
Chinese gloss in our interlinear	求你#幫助	主#啊
	<i>qiu ni#bangzhu</i>	<i>zhu#a</i>
	‘beg you#help’	‘Lord#O’
Chinese gloss in (Wang, 1985)	幫助	主啊
	<i>bangzhu</i>	<i>zhu a</i>
	‘help’	‘O Lord’

Table 1. Two example glosses illustrating how our Chinese glosses reflect Greek morphology. The pound sign in each gloss separates the meaning of the lemma from that of the suffix. On the left, the characters *qiu ni* ‘beg you’ before the pound sign expresses the imperative mood of the verb *boēthēi*. In the example on the right, the character *a* expresses the vocative case of the noun *kyrie*. In contrast, the glosses in (Wang, 1985) do not distinguish these meanings.

For instance, as shown in Table 1, the Chinese gloss for the Greek word *boēthēi* ‘help!’, in the imperative mood, contains a pound sign that separates the characters *bangzhu* ‘help’ (meaning of the Greek lemma) from *qiu ni* ‘beg you’ (meaning of the imperative mood). Similarly, the gloss for *kyrie* ‘O Lord’ has two clearly demarcated components, *zhu* ‘Lord’, and *a* ‘O’, the latter of which expresses the vocative case. In contrast, the glosses in (Wang, 1985) mix the meaning of the suffix with that of the lemma (e.g., *zhu a* for *kyrie*), and sometimes even omit the former (e.g., simply *bangzhu* for *boēthēi*).

3.2 Word Alignment

Our interlinear is digitally searchable, allowing readers to retrieve Greek word(s) equivalent to a Chinese word, and vice versa. In the two verses in Figure 1, the word *patēr* ‘father’ in the Greek text is shown to correspond to the word 父 *fu* in the Chinese text, and both *ēgapēsen* ‘loved’ and *philei* ‘loves’ correspond to 愛 *ai* ‘love’.

This search function critically depends on word alignments between the two texts. It is inadequate to rely on the Chinese gloss: the search would fail whenever the gloss deviates from the exact wording in the Chinese text. We therefore specify direct word alignments between the Greek and Chinese texts.

Another benefit of these alignments is to connect the Chinese text to the gloss. The connection is sometimes not evident in non-literal translations; for example, at first glance, the word 迎娶 *yingqu* ‘marry’ in the Chinese text does not obviously correspond to 在一起 *zaiyiqi* ‘be together’, the gloss for *synelthein* ‘to be together’. Given the alignment between *yingqu* and *synelthein*, the reader can see that the word for ‘marry’ in the Chinese text in fact translates a word that literally means ‘to be together’ in the original Greek text.

3.3 Gloss Accuracy

In addition to the two types of new content described above, our interlinear also makes advances in gloss accuracy, precision and consistency. An accurate gloss is one that expresses the exact meaning of the word in question, without mixing with other words in the sentence.

The treatment of expressions involving conjunctions and prepositions leaves much room for improvement in (Wang, 1985). First, the glosses for verbs often contain superfluous conjunctions, such as 就 *jiu*, 而 *er*, and 便 *bian* ‘then’. For instance, *anexōrēsen* ‘departed’ was glossed as 就退 *jiu tui* ‘then departed’ (Matthew 2:22), *egeneto* ‘became’ as 而成 *er xing* ‘then became’ (John 1:3, 1:10, 1:17), and *eipen* ‘said’ as 便告訴 *bian gaosu* ‘then said’ (Matthew 3:7). These conjunctions suggest meanings that are not at all present in the Greek, and they risk confusing the reader. Although they might improve the reading fluency, especially when the glosses were to be read as a sentence, the sacrifice in accuracy is hardly justifiable.

A second problematic construction is the Greek prepositional phrase (PP), which is often expressed by two Chinese words that are non-contiguous. Consider the PP *epi gēs* ‘on earth’, which means 在地上 *zai di shang* ‘at earth top-of’. The preposition *epi* ‘on’ corresponds not only to *zai* ‘at’, but rather to *zai ... shang* ‘at ... top-of’. The tendency in (Wang, 1985), unfortunately, is to follow the Chinese word order, often at the expense of accuracy. The PP above is a typical case: *epi* was inaccurately glossed as 在 *zai* ‘at’; while *gēs* ‘earth’, which means only *di* ‘earth’, was assigned 地上 *di shang* ‘earth top-of’, the remainder of the equivalent Chinese PP.

3.4 Gloss Precision

Besides being accurate, the gloss should also be precise. Precision means distinguishing the finer shades of meaning of a word, including its implied meaning, where appropriate.

15. New Testament (Novum Testamentum Graece), John, Chapter 15, Verse 9

καθώς ἠγάπησέν με ὁ πατήρ κἀγὼ ὑμᾶς ἠγάπησα

μείνατε ἐν τῇ ἀγάπῃ τῇ ἐμῇ

Show Word Dependency, POS & Gloss

Show POS

Show Gloss

Search By Example

Hide Aligned word

16. New Testament (Novum Testamentum Graece), John, Chapter 16, Verse 27

αὐτὸς γὰρ ὁ πατήρ φιλεῖ ὑμᾶς ὅτι ὑμεῖς ἐμὲ

πεφιλήκατε καὶ πεπιστεύκατε ὅτι ἐγὼ παρὰ [τοῦ]

θεοῦ ἐξῆλθον

Show Word Dependency, POS & Gloss

Show POS

Show Gloss

Search By Example

Hide Aligned word

聖經新約（和合本修訂版），約翰福音，15 章，9 節
我愛你們，正如父愛我一樣；你們要常在我的愛裏。

聖經新約（和合本修訂版），約翰福音，16 章，27 節
父自己愛你們，因為你們已經愛我，又信我是從上帝而來的。

Figure 1. Two of the verses retrieved from a search of the Chinese word 愛 *ai* ‘love’ aligned to a Greek verb with a noun subject, based on dependency annotations from PROIEL. In the top verse, the Greek verb is *ēgapēsen* ‘loved’, in the bottom one, it is *philei* ‘loves’; the subject for both is *patēr* ‘father’. These results demonstrate that the Greek language distinguishes between two kinds of love that are not reflected in the Chinese, and also list the entities that can be the agents of these kinds of love.

The most frequently occurring word in the Greek text, *kai*, generally means ‘and’ or ‘then’, but also has a wide semantic range including ‘but’, ‘also’, ‘even’, and sometimes bears no particular meaning (Wallace, 1997). In (Wang, 1985), *kai* was indiscriminately given the gloss 就 *jiu* ‘then’ or 而[且] *er[qie]* ‘and’. Our glosses, in contrast, operate at a finer level of granularity, further distinguishing instances where the context calls for 但 *dan* ‘but’, 卻 *que* ‘but’, 又 *you* ‘also’, 甚至 *shenzhi* ‘even’, or <empty>.

It is helpful for the gloss to indicate not only the literal meaning of a word, but also its implied meaning. Sometimes it is a matter of restoring grammatical features missing in Greek, such as indefinite articles. To illustrate, while *huios* ‘son’ can be rendered simply as 兒子 *erzi* ‘son’, it is more precise to include the Chinese number and counter word, i.e., (一個)兒子 (*yige*)*erzi* ‘(one) son’. In other cases, the implied meaning can be a narrowing or an expansion of the literal meaning. For instance, *adelphē* generally means ‘sister’, for which 姐妹 *jiemei* ‘sister’ would suffice; but in some contexts, its meaning is clearly limited to ‘Christian sister’, and it should instead be glossed as (信主的)姐妹 (*xinzhude*) *jiemei* ‘(believing) sister’. In the opposite direction, *adelphos* generally means ‘brother’, but when the context calls for the expanded meaning ‘brothers and sisters’, it is more precise to use the gloss 兄弟(姐妹) *xiongdi(jiemei)* ‘brothers (and sisters)’.

3.5 Gloss Consistency

A gloss can be both accurate and precise, but not consistent; i.e., when two synonymous but different glosses are used for two words in the source text that have

the exact same meaning. The treatment in (Wang, 1985) of the phrase *episteusan eis auton* ‘believed in him’ is a case in point. Among its seven occurrences in the Gospel of John, *episteusan* ‘believed’ was variously glossed as 信了 *xinle*, 相信 *xiangxin*, and 就相信了 *jiuxiangxinle*, all near-synonyms meaning ‘believed’; *eis* ‘in’ was glossed in three ways, namely, 深入 *shenru* ‘enter deeply into’, 歸順 *guishun* ‘submit to’, and 歸入 *guiru* ‘return to’. The mix-and-match of these different glosses can easily mislead the reader to the conclusion that the Greek phrase had many different nuanced meanings. To facilitate consistency, we use Paratext, a tool developed at United Bible Societies. This software keeps track of all Chinese glosses that have been used for each Greek word, and lists them in descending frequency. The translator consults this list, and re-uses an existing translation where possible. In our interlinear, *episteusan* is consistently glossed as 相信 *xiangxin* ‘believed’, except when it bears another sense, such as in Luke 16:11, where an alternate gloss 託付 *tuofu* ‘entrust’ is in order.

4. Implementation

The source text of our interlinear consists of about 65,000 Greek words. Their glosses were produced with help from a specialized software (section 4.1). The interlinear has been enhanced with Greek word dependencies, and can be searched on a browser (section 4.2).

4.1 Interlinearization Software

Paratext is a software developed by the United Bible Societies to help translators, exegetical advisors and consultants to produce quality translations from the point of view of both format and content. The translator can first input a first or revised draft of the text, then check

The treebank is based on the 8th edition of the Greek New Testament by Tischendorf (1869 --- 1872), which has a considerable number of differences with our Greek text (Nestle and Aland, 1994). We aligned the two versions, and amended the trees where there are discrepancies.

Search interface. On the browser-based interface, the user can specify up to five terms, and then further constrain their relations. A term can be either a Chinese word, or a Greek word, lemma, or a part-of-speech tag from the PROIEL tagset. A relation can be syntactic, i.e., a head-child relation between two Greek terms. For example, to see what entities can serve as the subject of the verb *philei* ‘loves’, one can search for all occurrences of *philei* that is connected to a child noun with the dependency label “sub”. Alternatively, a relation can be a word alignment between a Chinese word and a Greek term. Figure 1 shows results from a search that uses both languages as well as dependency information from PROIEL.

5. Evaluation

Our interlinear has been evaluated with respect to its gloss quality (section 5.1) and pedagogical effectiveness (section 5.2).

5.1 Gloss Evaluation

After an initial draft, we reviewed the Chinese glosses using a feature called “Biblical Terms” in Paratext. Shown in Figure 3, this tool assists translators and consultants in reviewing key terms of the translated text, based on the source text. About 1000 words long, the list of key terms include attributes, beings, animals, plants, objects, rituals, names and miscellaneous. For each term, we determine its principal senses, consulting if necessary the authoritative dictionary *A Greek-English Lexicon of the New Testament and Other Early Christian Literature* (Bauer et al., 2001). To check collocation units, especially recurring and idiomatic expressions, phrases from (Aland, 1983) have also been selected to complement this list.

We performed a detailed analysis on gloss precision, accuracy and consistency on a small subset of these terms. Bauer et al. (2001) organize the meaning of a word in a hierarchy, and provide, for each meaning, example verses from the New Testament illustrating its use. For each term, we constructed its ‘synsets’ (i.e., its word senses) according to the top-level categories in this dictionary, and retrieved the Chinese gloss assigned to the Greek word in each of the example verses cited. We then measured the following:

- **Consistency:** In principle, each sense should be rendered with one gloss. This metric asks to what extent Chinese glosses vary within a synset (but does not judge the appropriateness of the glosses);
- **Accuracy:** A gloss should reflect the meaning of the corresponding Greek word. This metric asks

whether the Chinese glosses in a synset express an appropriate meaning (even if their wordings differ);

- **Precision:** A gloss should be fine-grained enough to distinguish between major word senses, therefore the same gloss should normally not be used for words in two different synsets. This metric measures how often this occurs.

Nine Greek words were chosen for this analysis based on two criteria: their frequency in the gospels, and amount of word-sense ambiguity. Each of these words straddles at least two, and up to four, synsets. Altogether, 329 instances of their use are cited as examples by Bauer et al. (2001), and are examined in this analysis².

The average gloss accuracy is 98%. Almost all of the inaccurate cases are due to interpretations. For instance, *sēmeion* generally means ‘sign’, ‘token’, or ‘indication’, but in one instance it is glossed as 對象 *duixiang* ‘target’. This is an interpretation beyond the word’s surface meaning given in the dictionary, but is supported by at least one mainstream Chinese translation of the New Testament.

The average consistency is 77%. Often, the presence of multiple Chinese glosses in one synset was caused by a mismatch of the level of granularity with Bauer et al. (2001). To illustrate, three different glosses 預兆 *yuzhao* ‘sign’, 記號 *jihao* ‘token’ and 神蹟 *shenji* ‘miracle’ were found in one synset for *sēmeion*. These nuances all belong to the same top-level category in the dictionary entry of *sēmeion*, but are in fact further differentiated as subcategories.

Word	Accuracy	Consistency	Precision
archē	100%	63%	100%
alētheia	100%	92%	87%
hamartōlos	100%	94%	62%
epitimaō	100%	66%	100%
krinō	100%	45%	81%
logos	98%	95%	96%
peithō	100%	86%	100%
sēmeion	95%	91%	91%
psychē	94%	62%	70%
Average	98%	77%	87%

Table 2. Evaluation results on the Chinese gloss accuracy, consistency and precision for nine common Greek words.

The average precision is 87%. For some terms, subtle differences in shades of meaning were difficult to distinguish, resulting in the application of the same Chinese gloss in multiple synsets. This phenomenon is most frequent for the adjective *hamartōlos*, which, according to Bauer et al. (2001), can be used substantively to mean either ‘a sinner’ or ‘an irreligious person’. The

² The total number of their occurrences in the Gospels is, of course, greater than 329.

difference is often not straightforward to discern in the New Testament context, and the Chinese gloss uses the Chinese word 罪人 *zuiren* ‘sinner’ for both. Similarly, the noun *psychē* can mean ‘soul’ or ‘life’. As acknowledged by Bauer et al. (2001), some instances of *psychē* in the ‘soul’ synset may also serve as a metaphor for ‘life’. In a few of these cases we did use the gloss 生命 *shengming* ‘life’, further contributing to a lower precision.

5.2 Pedagogical Evaluation

We have deployed this interlinear in the course “Elementary Ancient Greek” at our university, with the goal to expose students to authentic Greek texts as early as possible. Twelve students, with no previous knowledge of Greek, were enrolled in the course. At the first two lectures, they were introduced to the Greek alphabet and pronunciation. The third lecture presents the concepts of case and definiteness, for which there are no Chinese equivalents. At the following lecture, students learned the forms of the definite article and adjectives of the first and second declensions; our interlinear was used at this point to teach the following two topics.

Adjectival constructions. In Greek, an attributive adjective can appear after an article but in front of the noun³ (“adjective-first”); alternatively, it can also follow the noun, in which case both the noun and the adjective have an article⁴ (“noun-first”). Rather than directly teaching these constructions, the instructor asked students to induce them from our interlinear, by examining the occurrences of a handful of paradigm adjectives that they had just learned.

With help from the Chinese glosses, the students independently identified the noun modified by the adjective in each instance, and observed the relative positions of the article, adjective and noun. Eleven out of the 12 students were able to formulate the two possible constructions, citing examples such as *ton kalon oinon* (‘the good wine’, in “adjective-first” construction) and *ho poimēn ho kalos* (‘the good shepherd’, in “noun-first” construction).

Verb endings. In a second exercise, the instructor asked students to induce the verb endings for the present indicative active forms, by examining a large number of occurrences of the verb *legō* ‘say’, and separating the stem from the ending of the verb. As in the previous exercise, eleven out of the 12 students successfully hypothesized the endings, and specified the person (first, second or third) and number (singular or plural) to which the each ending corresponds.

In both exercises, the Chinese glosses and word alignments played a critical role in the learning experience. Since their vocabulary was very limited, the

students relied on the glosses to locate the Greek adjectives and verbs, as well as the relevant linguistic context, e.g., articles in the first exercise and noun subjects in the second. The Chinese information was thus indispensable in enabling these beginners a direct encounter with authentic Greek text barely a month after their introduction to the language.

6. Conclusion

We have reported the development and evaluation of a Greek-Chinese interlinear text of the gospels in the Greek New Testament. Based on the most current texts available for both languages, it emphasizes gloss accuracy, precision and consistency, and contains new linguistic information such as word alignments and correspondences to Greek morphology in the Chinese gloss. A Greek dependency treebank has been adapted for the text. A search interface, offering bilingual string-based retrieval with dependency constraints, has been developed and successfully deployed in a Greek language class.

7. References

- Aland, K. (1983). *Vollstaendige Konkordanz zum Griechischen Neuen Testament unter zugrundelegung aller modernen kritischen textausgaben und des Textus Receptus*. Berlin: Walter de Gruyter.
- Bauer, Walter, and Danker, Frederick William (2001). *A Greek-English Lexicon of the New Testament and Other Early Christian Literature*. Chicago: University of Chicago Press.
- Brown, Robert K. and Comfort, Philip W. (1993). *The New Greek-English Interlinear New Testament*. Tyndale House Publishers.
- Chinese Union Version* 和合本 (1919). Hong Kong: Hong Kong Bible Society.
- Dukes, Kais and Habash, Nizar (2010). Morphological Annotation of Quranic Arabic. In *Proc. LREC*.
- Eckmann, J. (1976). *Middle Turkic Glosses of the Rylands Interlinear Koran Translation*. Budapest: Publishing House of the Hungarian Academy of Sciences.
- Green, Jay P. (1984). *The Interlinear Greek-English New Testament, with Strong's Concordance Numbers Above Each Word*. Peabody: Hendrickson Publishers.
- Haug, Dag and Jøhndal, Marius (2008). Creating a Parallel Treebank of the Old Indo-European Bible Translations. In *Proc. LREC Workshop on Language Technology for Cultural Heritage Data*.
- Marshall, Alfred (1993). *The Interlinear NASB-NIV Parallel New Testament in Greek and English*. Grand Rapids: Zondervan.
- Nestle, Erwin and Kurt Aland et al. (1952). *Novum Testamentum Graece*. 21st edition. Stuttgart: Deutsche Bibelgesellschaft
- Nestle, Erwin and Kurt Aland et al. (1994). *Novum Testamentum Graece*. 27th edition. Stuttgart: Deutsche Bibelgesellschaft.
- Ovid (1828). *The first book of Ovid's Metamorphoses*,

³ i.e., in the sequence article-adjective-noun

⁴ i.e., in the sequence article-noun-article-adjective

- with a literal interlinear translation and illustrative notes.* London: University of London.
- Palmer, Edwin (1881). *KAINH ΔΙΑΘΗΚΗ - The Greek Testament with the Readings Adopted by the Revisers of the Authorised Version.* Oxford: Clarendon Press.
- Revised Chinese Union Version* 和合本修訂版 (2010).
Hong Kong: Hong Kong Bible Society.
- Wallace, Daniel B. (1997). *Greek Grammar Beyond the Basics.* Grand Rapids: Zondervan.
- Wang, C. C. (1985). *Chinese-Greek-English Interlinear New Testament.* Taichung: Conservative Baptist Press.
- Xenophon (1896). *The first four books of Xenophon's Anabasis: the original text reduced to the natural English order, with a literal interlinear translation.* Harrisburg: Handy Book Corp.