

**Zeitschrift:** Asiatische Studien : Zeitschrift der Schweizerischen Asiengesellschaft = Études asiatiques : revue de la Société Suisse-Asie

**Herausgeber:** Schweizerische Asiengesellschaft

**Band:** 56 (2002)

**Heft:** 3: Textual scholarship in Chinese studies : papers from the Munich conference 2000

**Artikel:** Applying stemmatology to Chinese textual traditions

**Autor:** Simson, Wojciech Jan

**DOI:** <https://doi.org/10.5169/seals-147580>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 11.07.2025

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

# APPLYING STEMMATOLOGY TO CHINESE TEXTUAL TRADITIONS

Wojciech Jan Simson, University of Zurich

## 1. Introduction

When we read the books that have come down to us from Chinese antiquity, we usually take it for granted, that the editions we use have been provided with the same care as comparable editions of ancient Greek and Latin authors. We have become accustomed to rely on the supposed meticulousness and the abilities of a long-standing tradition of Chinese philologists. There is hardly anyone who would be eager to tackle the painstaking and boring task of comparing different versions of the same text again and again. There is also hardly anyone who could claim to possess an infallible sense to distinguish between right and wrong among the many variant readings, which very often seem to be meaningless differences of style that do not dramatically affect the general meaning of the texts. Therefore we accept, all too willingly, the well-established universe of Chinese classical literature that partly has even been sanctioned by Emperors themselves and engraved into stone to remain unaffected by the ravages of time. The official protection for the canonical literature, together with the long history of printing in China, have helped to promulgate very uniform printed editions of the ancient texts that now obscure the diversity and ambiguity of the manuscript traditions, on which these prints were originally based. However, our seemingly well-grounded universe of classical Chinese literature is being shaken more and more as philology and archeology bring to light older versions of ancient texts.

<sup>1st</sup> Already the Qing philologists became aware of the fact, that in Japan there had been preserved lost Chinese texts and text versions that went back to the Tang (618–907) dynasty when Chinese culture was most intensely adopted overseas. They realized that these texts went back to sources older than the Song (960–1279) and Ming (1368–1644) prints they could rely on for their own editions, and they started to reprint Japanese books or to incorporate their variants into collections of variant readings. They, however, did not dare to emend their texts accordingly, although they some-

times clearly stated, that they believed the Japanese readings to be more original.<sup>1</sup> Therefore the finds and insights of these philologists have remained mostly an unnoticed appendix to the voluminous commentaries and subcommentaries of the texts. The Sino-Japanese conflict of the 20<sup>th</sup> century disrupted this philological tradition in China and the national pride spurred by the conflict has prevented a sober assessment of the material preserved in Japan.

- 2<sup>nd</sup> At the beginning of the 20<sup>th</sup> century an enormous manuscript cache was discovered in a walled-up cave at Dunhuang, an ancient oasis and a monastic center on the Silk Road. It consists not only of tens of thousands of Buddhist sutras dating from the 4<sup>th</sup> to the early 11<sup>th</sup> centuries, but also comprises thousands of pieces of secular literature. These finds were not only unknown to the Qing-philologists who were responsible for most of the currently authoritative editions of ancient literature, but a full edition of these materials has not yet been completed and thousands of hitherto unknown scrolls are going to be published for the first time in the years to come.
- 3<sup>rd</sup> Insights into writing culture and the textual history of single texts, reaching even farther back than the Japanese and the Dunhuang materials, are due to the numerous archeological excavations that have been conducted in China since the Cultural Revolution. From ancient tombs manuscripts written on bamboo and silk were unearthed that date back as far as 2,000 years or even more.

Some of these finds represent text versions that differ from current editions mainly in orthography, style or arrangement, while others give texts that are almost beyond recognition. There is a general agreement that this material is of utmost scientific value. Consequently it is made accessible to the interested public in transcriptions and facsimile editions of growing quality and availability. Now and then we find also synoptic arrangements of different text versions, but the text-critical treatment of these precious testimonies almost never goes beyond the scope of such reproductions and renderings. The copious and often illuminating variant readings that can be found in archeological texts, never lead to the emendation of our standard editions, as would seem mandatory from the

1 Such a treatment of the Japanese variants can be found repeatedly in the collection of variant readings (*Jiaokan Ji* 校勘記) appended to the standard edition of the Thirteen Classics by Ruan Yuan 阮元 (1764–1849). *Shisan Jing Zhushu* 十三經注疏, ed. Ruan Yuan, Beijing <sup>2</sup>1983.

view point of a textual critic trained in the western philological tradition. Therefore, the new insights into the content of ancient Chinese literature and into the structure of the language that could be expected from these archeological materials and that would constitute their scientific value, are still wanting.

I do not think this unsatisfactory state of affairs can be excused only by the relatively short amount of time that has elapsed since the publication of most of the manuscripts in question. It is rather an outcome of the editorial attitude of Chinese scholars, who are used to relying more on authorities than on methods and reasoning. The excavated texts are treated as authorities in their own right and are supplied with some explanatory notes that make difficult passages readable by applying to them traditionally accepted but nonetheless questionable concepts such as *loan characters* (*jiajiezi* 假借字) or *superfluous characters* (*yanzi* 衍字).<sup>2</sup> When we turn to Chinese handbooks of what would seem to be the equivalent of western textual criticism, we find lengthy catalogues of different types of scribal errors with numerous but often far-fetched examples, the sources of which often remain unrevealed. Usually we find only vague explanations of how and why such errors come about and we never get any hint as to the circumstances under which, and the frequency with which, such errors can be expected. Moreover, there are no rules given that would help us to discern right from wrong readings. To put it in a nutshell: There is a manifest lack of method, and this absence of method explains also the lack of text editions that could be as accurate and reliable compared to those provided and used by western classical philology. This lack of method has incited me to explore in the present article the applicability of stemmatology—one of the methods of western textual criticism—to the radically different linguistic and scriptural conditions of ancient Chinese texts.

## 2. Methodology

Stemmatology was originated by Karl Lachmann (1781–1851), a classical philologist, who first applied this method to the text of the New Testament and to Greek and Latin authors. Most of the authoritative editions of European classical literature are a result of the application of this method. The name *stemmatology* is derived from the Greek word *stemma*, which indicates a

2 For a critical discussion of at least the concept of *loan characters* see Bernhard Karlgren: *Loan Characters in Pre-Han Texts*, Stockholm 1963–1967 p. 1–18.



diagram that visualizes the kinship relations of different textual witnesses of the same text. This stemma has usually the shape of a genealogical tree. We assume the root to be the original text. From the root springs the trunk, which branches out into the different textual traditions, which mostly split into further branches and twigs.

The textual critic, of course, does not possess the root. Mostly he will not possess the trunk either, but usually finds himself confronted with some scattered twigs from the treetop. These may be manuscripts or their fragments, early prints, but also quotations in ancient works. They are all treated equally as *textual witnesses* that bear witness to certain stages of the transmission of the text. Stemmatology attempts to understand the succession of these various stages of transmission, i.e. it reconstructs the genealogical tree from the scattered branches. If it succeeds, such a reconstruction allows us to discern which of the branches grew out of the trunk directly and which are offshoots from other branches and do not stem directly from the trunk. The branches that are closer to the trunk can be regarded as more original, whereas the lateral branches can be cut off, i.e. their particular variant readings can be discarded. By performing such an elimination process we move along the branches towards the trunk. When we have eliminated all variant readings that cannot be direct offsprings of the trunk, we have arrived at the trunk itself or have at least grasped the few branches that stem directly from it. This means, that there are none or only very few variants left that must be taken into consideration as possibly original readings. The trunk we have arrived at, is not the original text itself, but only the archetype of all extant textual witnesses. It is, however, the best possible text that can be reconstructed from the available material.

Here we have to ask: How do we come to know the original order of the scattered branches, how do we discern which branch stemmed from which other?—Even if we know the age of a textual witness we cannot draw from it any reliable conclusion as to its position in the pedigree: Manuscripts can reach an age of several centuries, and together with their age grows their esteem. The older they are, the more eagerly and conscientiously they will be copied. Even the youngest manuscript can therefore be a faithful copy of a very ancient text version and can be closer to the original than any older witness. Contrary to common assumptions the age of a textual witness alone is of no significance for the evaluation of its readings! The only valid criterion for the assessment of a witness and its variants is its position in relation to the stem of the stemma.

In order to determine this position, stemmatology takes advantage of the fact, that changes that once have occurred in a certain text version, will be usually copied together with the text itself, and are thus inherited by subsequent

manuscript generations. Such lines of inheritance correspond to the branches of the pedigree and enable us to reconstruct the tree. To trace these lines of inheritance stemmatology focuses on a few particularly unambiguous errors. They are called *significant errors*. We can further subdivide them into *conjunctive errors* that establish a conjunction between two or more textual witnesses, i.e. place them on the same branch of the tree, and *separative errors* that enable us to separate these branches of transmission from one another, i.e. to establish bifurcations between the branches. This distinction is, however, not too strict, because in many cases significant errors will function as both conjunctive and separative errors as well.

Common errors hint at common ancestors. Textual witnesses sharing a certain error therefore have to be placed on the same branch, whereas witnesses not bearing this error are either the direct ancestors of that branch or have to be placed on a separate branch. Such a separate branch has to be assumed, if the hitherto error free branch bears a separative error of its own that was not inherited by the error affected branch discerned first. By scrutinizing each single witness for significant errors, step by step the branches and bifurcations in the genealogical tree can be found and the positions of the textual witnesses relative to each other can be determined. We become thus able to reconstruct the genealogical tree.

But what do such *significant errors* look like?—Some of the changes that can occur to a text are not necessarily inherited by the next generation of copies and therefore cannot be used to identify the branches and bifurcations. Orthographic or punctuative features are likely to be copied into the next generation, but they can also be easily adjusted to the conventions of a scribe's scriptorium and may be readjusted by the next scribe according to the conventions of his own scribal tradition. Each scribe will thus follow a more or less consistent system of orthography or punctuation that can be introduced to the text at will, without relying on an earlier version of the text. Some of these features may even appear randomly without following any rule. Orthographic and punctuative features are therefore not directional. This means, that there is no way to infer from them which state of the text is more original and which one is derivative. Such features, therefore, do not give any hint as to the relative position of a witness in the pedigree and must be ignored by the stemmatologist. This is different with errors: By definition an error is a deterioration of an original state. Deterioration is a special kind of derivation and can be therefore used to trace the lines of descent. Not every error, however, is irreversible. Obvious scribal errors can be detected by an even mediocre copyist and will be corrected as a matter of course; in many cases—at least in western languages—

a scribe will do so even unconsciously by overlooking an inconspicuous orthographic or a grammatical mistake and thus restoring the correct reading. If an error once committed is corrected again, it can thoroughly mislead the textual critic in his attempt to reconstruct the genealogical tree. Judging from the absence of such a corrected error in a relatively young witness, the critic might erroneously place the corrected witness closer to the original than its faulty ancestors. To build the pedigree stemmatology has, therefore, to rely on irreversible errors, that cannot be expected to be corrected by the average scribe. Probably the best case of such an error is given, when the scribe has omitted a longer passage (that is preferably not a self-contained unit which could be a later inflation), be it because he inadvertently skipped a line or because the manuscript he was copying was damaged in that place. In this case no other scribe will be able to emend the corrupted passage without the help of another version of the text. A single significative error of this kind is already enough to establish a whole branch of textual transmission. Not all textual witnesses, of course, will show such clearly irreversible errors. In these cases we will have to resort to errors the irreversibility of which is less reliable, but we can try to compensate for their reduced reliability by relying on a larger number of errors.

### 3. Application

After these theoretical considerations we can try to test the applicability of the stemmatological method to a Chinese textual tradition. The example chosen consists of three manuscripts of the well-known *Analects of Confucius* (*Lunyu* 論語) with the commentary of He Yan 何晏 (190–249). All of them were found in Dunhuang and are today part of the Pelliot collection belonging to the *Bibliothèque Nationale* in Paris:

- A: Pelliot 2664. A fragment from the middle of a scroll, both ends are torn away. The paper was not ruled, the lines are shaky, written with a stylus in a clumsy hand. Corrections can be found, some of them being obviously done by the scribe himself. Most probably it is a writing exercise by a school boy. The fragment comprises *Lunyu* XII.7–12.
- B: Pelliot 2620. An almost intact scroll, which is, however, frayed at the beginning. The handwriting is dense, not very elaborate but orderly. Comprises *Lunyu* XI.3–XII.24

- C: Pelliot 4732, 3402. A scroll the beginning of which has been torn away. It consists of two fragments that can be neatly fitted together. The paper has been ruled by plying. The handwriting is firm, leaving ample space between the characters and lines. At the end there is a colophon that unfortunately gives a date without a year. Verso we find a medical text and a line in Tibetan. Both fragments comprise together *Lunyu* XI.12-XII.24.

None of the three manuscripts can be dated exactly. Most of the datable Dunhuang manuscripts, however, come from Tang dynasty (618–907). Judging from the handwriting all three manuscripts seem to be rather from mid or late Tang. On manuscripts *B* and *C*, moreover, the character *min* ‘people’ which was part of the name of Li Shimin 李世民 (599–649), the founder of the Tang dynasty, is tabooed by replacing it with *ren* ‘men’ or by leaving out the character’s last stroke. Only manuscript *A* does not taboo this character, but this is not a conclusive proof of a date later than the Tang, because the character appears only in a few instances on the manuscript in a clearly legible form and none of the other manuscripts observes the taboo consistently. The use of a stylus as writing instrument on *A*, however, hints towards the end of Tang dynasty or even to the 10<sup>th</sup> century. Manuscripts *B* and *C* are of clearly earlier date and the line in Tibetan on the reverse of *C* may go back to the period of Tibetan rule 781 – ca. 840.

Because stemmatology consists mainly in a comparison of variant readings, we have to confine our investigation to the text passage that all three manuscripts have in common, i.e. the length of the shortest manuscript *A* covering chapters 7 through 12 of book XII of the *Analects*. To simplify our example, we ignore furthermore the commentary and examine only the main text. This short passage covers but 240 characters. Nonetheless, we find 21 instances of variant readings that are more significant than simple orthographic variants. In chapter XII.8 we can easily spot a conspicuous omission:<sup>3</sup>

3 Following conventions have been followed in the transcriptions and translations:

The symbol ‘●’ is used to indicate not lacunae but the absence of characters that are present in other versions.

The Chinese text has been transcribed into standard full characters. There are many orthographic variants in the manuscripts that might be of interest, but are beyond the scope of the present investigation. Most conspicuous in the example given above would be that *A* and *C* use the loan character 熟 where *B* has 孰. Note, however, that their distribution is not indicative of the bifurcation we are going to establish.

The translations are mostly adopted from Raymond Dawson: *Confucius—The Analects*, Oxford & New York 1993.

Table 1

A	百姓足，君孰●不足？百姓不足，●●●●。
	If the hundred surnames have plenty, then who among the lords will have not plenty. The hundred surnames do not have plenty.
B	百姓足，君孰與不足？百姓不足，●●●●。
	If the hundred surnames have plenty, then with whom will your lordship share insufficiency. The hundred surnames do not have plenty.
C	百姓足，君孰與不足？百姓不足，君孰與足。
	If the hundred surnames have plenty, then with whom will your lordship share insufficiency. But if the hundred surnames do not have plenty, <i>then with whom will your lordship share plenty?</i>

Obviously only manuscript *C* gives a truly satisfactory reading and therefore the missing of the final clause must be taken as an erroneous omission. Manuscript *C* which has no omission represents a more original version of the text than *A* or *B*, and has to be placed on a branch stemming directly from the archetype  $\alpha$ . The omission separates both *A* and *B* from the branch of *C*, and is called therefore a *separative error* against *C*. Simultaneously the omission functions also as a so-called *conjunctive error* between *A* and *B*, because it establishes a conjunction between the two manuscripts. Both have to be placed on the same branch of the tree. We do not know, however, in which order they have to be placed on that branch. *A* may be an ancestor of *B*, or *B* an ancestor of *A*. Both may be also derived from a third source  $\beta$ . We can resolve this problem by checking *A* and *B* for separative errors against each other:

Table 2

A	受之，欲其生；惡之，欲其死。
	If you <i>receive</i> someone, you want him to live. If you hate someone, you want him to die.
B	愛而，欲其生；惡之，欲其死。
	Out of <i>love</i> for someone, you want him to live. If you hate someone, you want him to die.
C	愛之，欲其生；惡之，欲其死。
	If you <i>love</i> someone, you want him to live. If you hate someone, you want him to die.

Also here it is not difficult to distinguish the correct reading: The scribe of *A* has mistaken the character *ai* (愛) 'to love' for the obviously very similar character *shou* (受) 'to receive'. If *A* had been the source for *B*, *B* would have most likely copied the error. We can therefore exclude that *B* is derived from *A* and are left with the two possibilities that either *A* is derived from *B*, or both *A*



and  $B$  are derived from a common ancestor  $\beta$ . Again, this question can be resolved by checking  $B$  for separative errors against  $A$ :

Table 3

A=C	片言可以折獄者 one who is able <i>to settle</i> a law suit on the basis of a partial submission
B	片言可●折獄者 one who is able <i>to be settled</i> a law suit on the basis of a partial submission

The missing of the particle *yi* (以) on  $B$  turns the sentence into a passive clause that, strictly speaking, makes no sense. The error is simultaneously inconspicuous enough to pass unnoticed and therefore uncorrected by the average scribe. The reading on  $A$  must therefore be more original than that on  $B$  and we can now also exclude that  $A$  is derived from  $B$ . Because of the conjunctive error between  $A$  and  $B$  found in the first example, we have to assume a common ancestor  $\beta$ . This common ancestor is—by definition—the very place where this conjunctive error took place. Because  $C$  does not share this error, it is neither identical with the common ancestor  $\beta$  of  $A$  and  $B$ , nor can it be derived from them.  $C$ 's only two possible positions in the stemma are either as the direct ancestor of  $\beta$  or on a separate branch stemming from the archetype  $\alpha$  that is not only the ancestor of  $C$  but also the ancestor of  $\beta$ . The solution to this problem gives a separative error of  $C$  against both  $A$  and  $B$ :

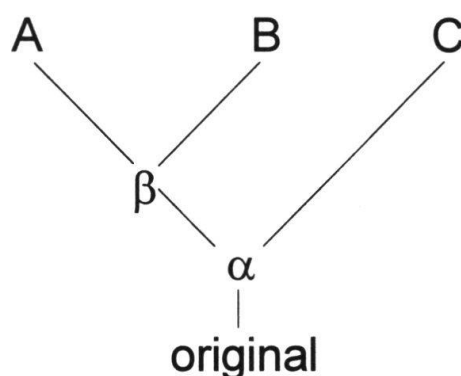
Table 4

A=B	君不君；臣不臣；父不父；子不子。 <i>The ruler is not a ruler, the subject not a subject, the father not a father, and the son not a son.</i>
C	●●君臣不臣。父不父；子不子。 <i>Ruler and subject are not subject. The father is not a father, the son not a son.</i>

The reading on manuscript  $C$  is clearly an erroneous omission and the common reading of  $A$  and  $B$  is therefore more original. This more original reading shows that  $\beta$ , the common ancestor of  $A$  and  $B$ , cannot be derived directly from  $C$ , because in that case the error on  $C$  would have been inherited by  $A$  and  $B$  through  $\beta$ . Therefore we must assume a witness  $\alpha$  that was free of the error and the common ancestor of  $\beta$  (along with its derivatives  $A$  and  $B$ ) on the one hand, and  $C$  on the other hand. Due to the foregoing examination of four separative errors we are now in the position to establish the following stemma:



Figure 1



This looks nice and neat but, after all, what is it good for?—Hitherto we have considered only how errors are inherited, but original readings are inherited along the same lines. Where *A* and *B* have the same reading they have it in common because it goes back to their common ancestor  $\beta$ . We can therefore safely reconstruct  $\beta$  in all instances where *A* and *B* do not differ from one another. For the same reason we can also reconstruct  $\alpha$  where  $\beta$  can be reconstructed and coincides with *C*. To put it in plain words: Where there are no variants, the common reading goes back to  $\alpha$ . This is but a trivial corollary. Much less trivial, however, is the conclusion, that where *A* coincides with *C* but not *B*, its reading must go back to  $\alpha$  and was inherited through  $\beta$ . From the common readings of *A* and *C* only, we are still able to reconstruct  $\alpha$ . The same applies *mutatis mutandis* to the analogous case in which *B* coincides with *C* and *A* differs.

Figure 2

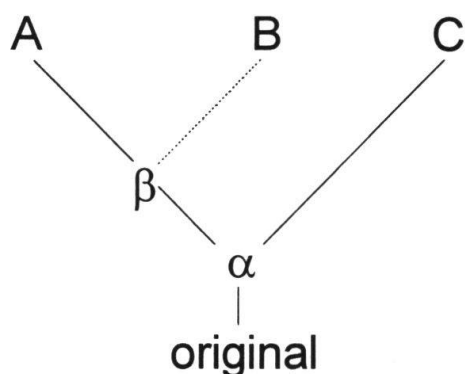
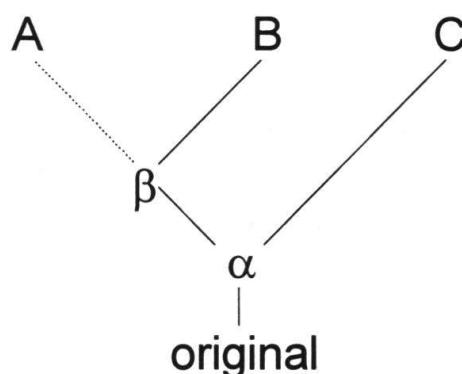


Figure 3



In both cases the particular variants that we find in *B* alone (respectively *A* alone) have arisen on the way from  $\beta$  to *B* (or *A* respectively) either when *B* was copied directly from  $\beta$  or on an intermediate copy between  $\beta$  and *B*. These

variants are doubtlessly deviations from the original and can be safely discarded as errors without any further examination. It is, nonetheless, worthwhile to take a closer look at these particular errors of *B* and *A* respectively in our example:

Table 5

No.	$\alpha$ reconstructed	Particular readings	Translation
	from A=C	of B	
1.	亦祗以異	亦知以異	<i>merely</i> for a change / <i>I know</i> it's for a change
	from B=C	of A	
2.	必不得已而	必不德已而	having no alternative / [?]
3.	去於斯三者	去於斯二者	give up one of these <i>three</i> / give up one of these <i>two</i>
4.	●●曰	●●●	[Zi Gong] <i>said</i> / ●
5.	惜乎夫子之 說君子也	惜乎夫子之 說君子●	It is a pity you said that, sir, about the gentleman.
6.	盍徹乎	盍●乎	Why not <i>take one tenth as tax</i> ? / [?]
7.	君孰與不足	君孰●不足	<i>With whom</i> will your lordship share plenty? / <i>Who among</i> the lords will have not plenty?
8.	徙義	從義	<i>Move towards</i> righteousness. / <i>Follow</i> righteousness.
9.	愛之欲其生	愛而欲其生	<i>If you love</i> someone, you want him to live. / <i>Out of love</i> for someone, you want him to live.

In addition to the four errors we used to build the stemma, we can eliminate the nine errors listed above (and we could eliminate even more, if we would examine the commentary too). We have discarded all these readings 'blindly', i.e. only on grounds of their position in the stemma, without examining whether they make sense or not. This may seem rather bold, but we can still verify the result by analysing the discarded readings. First, we find, that there is no reading reconstructed for  $\alpha$  on the left hand side that would be clearly inferior to one of the discarded readings on the right. Moreover, most of the discarded readings can be identified as obvious scribal errors:

- 1 A mistake motivated by the obvious phonetic similarity of the two characters. Typically a rare character *zhi* (祗) ‘merely’ has been mistaken for a frequent homophone *zhi* (知) ‘to know’.<sup>4</sup>
- 2 Another phonetic mistake and one of the most common ones. The two characters are more or less equally frequent, so that we cannot infer the correct reading from the frequency of the characters involved. The context however is a set phrase *bu de yi er* (不得已而) ‘having no alternative,’ which allows only one reading.
- 3 A mistake motivated by the obvious graphical similarity of the characters *er* (二) ‘two’ and *san* (三) ‘three’. We can clearly determine that the latter is the correct reading from the wider context not cited above.
- 6 The sentence makes no sense without the omitted word *che* (徹) ‘take one tenth as tax’.
- 7,9 Both discarded sentences seem to make sense at first glance, but the parallelisms in the wider context suggest that they offer inferior readings.

Considering that all these text-critical decisions were made ‘blindly,’ the result is amazing. If the method had failed, we would have had to expect a much more ambiguous result. Most of the errors we have identified by the application of stemmatology could have also been identified without it, as shown in the discussion above. The true efficacy of the method becomes only visible when we turn to the cases we did not yet discuss. We find the omission of a final particle (5.) or of the explicit introduction of direct speech (4.). Such cases are otherwise extremely difficult or even impossible to decide upon. Similarly in the last case (8.) we find two more or less equally satisfactory readings: (A:) *to follow righteousness* and (α:) *to move [literally: to migrate] towards righteousness*. The latter may sound awkward and therefore inferior to many readers, but stemma-

- 4 I deliberately do not use the well known term *lectio difficilior* that one might expect here. Due to the ambiguities of Chinese characters and Classical Chinese grammar, it is almost always possible to read a *difficult* meaning into a corrupted sentence. The *difficulty* of a reading must be therefore carefully weighed against its plausibility. The preference for such *lectiones difficiliores* may lead to a liking for the unlikely and is not a reliable basis for sound text-critical decisions. (See Table 11.)

The frequency of a word, on the contrary, can be easily understood as constituting the likeliness with which the scribe anticipates a certain word to appear, and the scribe’s anticipation is the very reason for the mistake. Unlike the *difficulty* of a reading, the frequency of a word can be objectively attested, at least in the text itself. We have, however, also to take into account the word frequencies in the scribe’s contemporary language, which has also a strong impact on the scribes errors, but can only be roughly estimated.

tology helps us to abstract from our stylistic biases, which are mostly influenced by our own mother tongue.<sup>5</sup>

We have seen that stemmatology has enabled us to emend a considerable number of errors with great certainty. However there are still two possible cases left that we cannot decide upon:

- 1 If the readings of all three witnesses differ from each other, we have, of course, no coincidences that would help us to reconstruct their ancestor by means of stemmatology.

This situation occurs very rarely and does not appear in our example. The second unsolvable case is:

- 2 When *A* and *B* coincide, but *C* has a different reading, we reconstruct  $\beta$  from *A* and *B*, but are not able to decide between  $\beta$  and *C*. From the view point of stemmatology both readings are concurrent, because both may stem directly from the common ancestor  $\alpha$ .

In fact, we find a number of such cases in our example:

Table 6

	$\beta$ from A=B	C	Translation
1.	自古皆有死	自故皆有死	Since <i>olden times</i> everyone has met his death. / Since <i>before</i> everyone has met his death.
2.	民無信不立 ●	民無信不立也	If the people have no trust, one will fall.
3.	駟不及舌 ●	駟不及舌也	Even a team of four horses cannot catch up with one's tongue.
4.	文猶質也	文猶也質	Refinement is as important as raw material. / [?]
5.	虎豹之鞶	虎豹 ● 鞶	the shaven skin of a tiger or a leopard / the shaven skin of tiger and leopard
6.	崇德也	崇德 ●	<i>this is</i> exalting virtue / exalting virtue
7.	無宿諾 ●	無宿 ● 也	<i>agreed</i> without hesitation / without hesitation
8.	無宿諾 ●	無宿 ● 也	<i>agreed</i> without hesitation / without hesitation

- 5 Note that the *Analects* apply metaphors of settlement in connection with virtues also in another instance: Chapter IV.1 speaks of *li ren* (里仁) 'taking one's abode in benevolence' and *chu ren* (處仁) 'settling in benevolence'. Moreover we can observe that the character *xi* (徙) 'to migrate' seems to be less frequent than the character *cong* (從) 'to follow' and therefore *cong* is more likely to be mistaken for *xi* than the other way round.

If we have no further textual witnesses, these variants cannot be eliminated by means of the stemmatological method. We have to decide between them one by one, judging their plausibility. Usually only the variants rejected in this last stage are listed in the *apparatus criticus*, whereas those that could be eliminated stemmatologically can be neglected.

If we have further textual witness, the whole procedure begins again: We determine its place in the stemma by looking first for conjunctive and then separative errors. If we find that the fourth witness is not a direct offspring of one of the witnesses already examined, it will enable us to eliminate further variants. The process is repeated again and again, until all witnesses have been examined and the stemma has been completed.

To make the discussion complete, it has to be mentioned, that there are rather rare cases where a direct common ancestor has to be assumed not only for two but for three (or more) witnesses. Such a stemma has to be drawn when all of these three (or more) witnesses have common separative errors against the other witnesses, and each has its own separative errors against the other two (or more), but no two of them have a common separative error against the remaining one (or more). It goes without saying that our reasoning will then be analogous to what we have already seen in the example above. This means we will eliminate variant readings relying either on coinciding readings of at least two of the group of three (or more), or relying on coincidences of one of the three (or more) with a concurrent branch of the stemma.

#### 4. Methodological Criticism

The main advantage of the stemmatological method is that, based on only very few especially clear significative errors, it enables us to eliminate a majority of variants without judging their correctness or faultiness in detail. This means that our judgment will be as little as possible influenced by our prejudices as to content, style, or even grammar of the text. Mathematically conclusive and infallible as the method may seem, it has however its specific weak points. The very presupposition of the method, namely the assumption that errors are inherited, can turn out to be a fallacy.

First of all we have to ask whether the coincidence of the conjunctive errors we use to identify two witnesses, as belonging to the same branch of the stemma, is really always the result of inheritance. Is it not possible that the same errors arise independently?—Of course it is, but the question is not whether it

can happen at all, but with what probability this can occur. Based on the data of our example we can try to estimate this probability:

Apart from omissions or additions,<sup>6</sup> the most frequent errors are the phonetic mistakes. We find four of them in our sample:

*Table 7*

witnesses	B=C	A = B	A = C	B=C
original character	得	古	祗	之
pronunciation	[teǎk]	[kuǎʻ]	[tʂi]	[tʂi]
witness	A	C	B	A
erroneous character	德	故	知	而
pronunciation	[teǎk]	[kuǎʻ]	[tri]	[ri]

The pronunciations given above follow the reconstruction of standard Late Middle Chinese by E. G. Pulleyblank. For the two cases on the right that are not true homophones, we have to assume that the initials of the syllables have converged in the dialect spoken in Dunhuang. This can be seen from similar cases frequently found on other Dunhuang manuscripts.

Next come the graphical mistakes. We find three of them in our sample:

*Table 8*

witnesses	A=C	B=C	B=C
original character	三	愛	徙
witness	B	A	A
erroneous character	二	受	從

Finally we have one transposition:

6 I have deliberately excluded the omissions and additions from the following considerations, because their distribution is not equal. E.g. final particles are omitted much more often than whole sentences; whole sentences are omitted mostly when they appear in parallelisms etc. The distribution of errors like transpositions and phonetic mistakes, on the other hand, depends much less on the context or the meaning or function of a character. This makes such errors more suitable for an assessment of frequency.



Table 9

witnesses	A=B
original wording	質也
witness	C
erroneous wording	也質

Each error we have found appears only once within the three manuscripts we have examined. Each manuscript covers 240 characters, which makes together 720 characters. When we find 4 phonetic mistakes within 720 characters, this means that the frequency of phonetic mistakes is  $4 \div 720 = 0.0056$ . In the same way the frequency of graphical mistakes can be calculated as being  $3 \div 720 = 0.0042$  and that of transpositions as  $1 \div 720 = 0.0014$ . The frequency with which a certain type of error can be found on the manuscripts corresponds to the probability with which the same type of error may occur on an additional manuscript in exactly the same place accidentally.<sup>7</sup> In our example this means, that the probability that a certain type of conjunctive error on a certain manuscript has come about by mere chance, and is not the result of inheritance, is below 1%. In other words, the reliability of the method is above 99%. If this is not enough, we can still use two or more conjunctive errors to make our decision safer. The overall probability of accidental failure of the method will decrease with each additional error we use by the factor of the specific probability of this additional error. To simplify our estimate, we assume an equal frequency of 0.005 for all error types. The probability of failure, when using two conjunctive errors, is then calculated as  $0.005 \times 0.005 = 0.000025$ , which is one chance in forty thousand cases. With three errors we will get  $0.005 \times 0.005 \times 0.005 = 0.000000125$ , which is one in eight million. It is very obvious that the reliability of the method increases exponentially with the number of conjunctive errors used to build a certain branch of the stemma.

So far, this estimate is only valid for the case of two textual witnesses, which is a very limited case that does not allow the method to be fully applied.

7 Strictly speaking, characters can be misinterpreted in different ways. A phonetic misreading of one and the same character might result, therefore, in any one of its many homophones. We would have then to take into account not the types of errors, but also the number of possible errors of the same type. This would further reduce the chance of an accidental occurrence of exactly the same error. In other words, we are calculating the worst case. This worst case is, however, realistic, because a given character mostly results in the same erroneous character, due to the above mentioned tendency to mistake rare characters for frequent ones.

Mostly there will be more witnesses involved, and the more there are, the greater the chance that the conjunctive errors we are looking for might occur in one or more of them accidentally. With each additional witness that has to be compared, the reliability of a stemmatological decision decreases by the factor of the reliability of that additional witness. The effect of this reduction will be however not dramatic, because such reliabilities, as has been shown above, are clearly above 99%. Without going into further details of such probability calculations, we can conclude that the probability that conjunctive errors arise independently, and are not the result of inheritance, may seriously challenge the reliability of our stemma only when one conjunctive error is used in combination with a large number of textual witnesses. By using two or more errors we can reduce this risk to a negligible size and attain a degree of reliability that is rarely to be found even in natural sciences.

The principle of inheritance can be questioned not only because of the independent occurrence of the same mistake. Much more serious is the problem of emendation. This means, that an error once committed may be corrected again. As we have mentioned in our theoretical introduction, such a corrected error may make a late witness seem more original than its ancestors. Moreover, almost every scribe will take it as his duty to correct the errors he finds. Some of the errors we used to build our stemma could be corrected by analogy to parallel sentences we find in the wider context. Let us for example take a closer look at chapter XII.11 as we find it on manuscript C:

Table 10

齊景公問政於孔子。	Duke Jing of Qi asked Master Kong about government.
孔子對曰：「君君，臣臣，父父，子子。」	Master Kong replied: <i>Let a ruler be a ruler, a subject a subject, a father a father and a son a son.</i>
公曰：「善哉！信如君臣不臣，父不父，子不子，雖有粟，吾焉得而食諸？」	The duke said: Excellent! Indeed, <i>ruler and subject not being subject</i> , fathers not being fathers and sons not being sons, how could I manage to eat anything, even if there were food around.

The corrupted sentence is *Ruler and subject not being subject*. The first two words have been omitted and it should read 君不君，臣不臣 *The ruler not being a ruler and the subject not being a subject*. Basically it is possible to infer the correct reading by analogy from Master Kong's parallel sentences above (*Let a ruler be a ruler, the subject be a subject ...*). The solution seems simple and clear. Experience shows, however, that the simplest solutions are often the most difficult to find. Moreover, we have to consider that the *Analects* enjoyed

a half-canonical status from Han dynasty (203 B.C. – A.D. 220) on. It would therefore be somewhat bold for an average scribe to introduce new readings to such an old and venerable text. One thousand years had elapsed since the text had been composed when the scribe copied it. Its language had become obsolete and difficult to understand. Simultaneously readers had become accustomed to encounter many obscure passages in their classics. For their understanding they relied not on their own invention but mainly on commentaries. If we scrutinize such commentaries themselves, in order to know how they apply textual criticism, we will discover that mostly they do not do it at all. They rather rely on semantic approaches to elucidate obscurities. Text-critical approaches are to be found only sporadically and the *repertoire* of supposed errors that are proposed by the commentators differs considerably from what we find in real manuscripts. Though omissions are the most frequent error, as can be seen even from our short example, the commentators almost never propose them to solve textual problems. Their favorite device is just the opposite—they suggest so-called ‘superfluous characters’ *yanzi* (衍字). I cannot resist the temptation to imagine such a commentator dealing with the above example. Most probably he would cancel the word *ruler* from the corrupted sentence. The emended text would then read:

The duke said: Excellent! Indeed, the subjects not being subjects, fathers not being fathers and sons not being sons, how could I manage to eat anything, even if there were food around.

Our hypothetical commentator might even claim, that his version is a true *lectio difficilior*, because the sense it makes is more subtle than ours. It is namely not very suitable for the duke, being a ruler himself, to suggest, that a ruler might not be a ruler, and therefore he does not mention this possibility, though Confucius suggested it.—Such may be the fallacies of textual criticism, as soon as interpretative judgment is involved. Our example shows, moreover, that Classical Chinese language possesses an amazing interpretive flexibility. Due to the extreme ambiguity of Chinese characters and syntactical structures a meaningful sentence can almost always be fabricated from any given sequence of characters. This reduces the necessity of emendation not only for commentators but also makes it much more difficult for the average scribe to identify errors.

We can conclude from the above considerations, that emendations to the text by means of mere conjecture are possible, but in more complex cases not very probable. The fallacies that such emendations may bring about can also be avoided by wisely choosing the errors we use to build the stemma. The case we

have just discussed was actually not a perfect separative error, because there is a certain chance that the original reading may be restored again. We had, however, little choice due to the extreme shortness of our example.

Errors may be corrected not only by means of conjecture but also by comparing one copy of the same text with another. The result is a contamination of the text. This means that readings from one line of transmission are introduced to another. If this happens not only the principle of inheritance is violated but also the implicit assumption that the branches of the tree will always grow apart and never together. If this happens, stemmatology will fail completely.

The more popular a text is, the more copies will be produced and circulated. The greater, therefore, the chance becomes that two copies will be compared. The *Analects* were one of the most widespread texts of Chinese literature, and already the preface to the oldest transmitted commentary to the *Analects* by He Yan<sup>8</sup> states that two of the three versions of the text that had appeared in early Han time were collated by the imperial tutor Zhang Yu 張禹 (†5 B.C.). Another collation of all three versions he mentions, was undertaken by Zheng Xuan 鄭玄 (127–200). When we read He Yan's preface carefully, we will notice, that he does not make a distinction between text versions and the commentarial traditions that they were accompanied by. Most probably He Yan's *Collected Explanations* (*Jijie* 集解), which he compiled from a number of earlier commentaries, implied a merging of the text versions these commentaries were transmitted with. The practice of collation was of course continued also after He Yan. Some dynasties like the Han or the Tang had the Confucian classics, and together with them also the *Analects*, engraved in stone steles, which were erected in front of the imperial academy. These authoritative text versions were, of course, the result of careful collation by the most renowned scholars of their day. With the promulgation of xylographic printing in the 10<sup>th</sup> century identical text versions spread quickly throughout the country and became readily available. Hardly anyone who embarked on the sumptuous enterprise of a new edition would do so without making sure that the text version he cut into wood was in accordance with earlier prints. This led very quickly to a levelling out of different lines of transmission and to the emergence of almost uniform text versions. This process can be exemplified by the following three early editions of the *Analects*, two from Japan and one from China:

- D: Shōhei 正平 edition. Contains the text with the commentary of He Yan. A copy of this edition came through Korea to China and became known as

8 *Shisan Jing Zhushu* (fn. 1), 2454–2456.

*Gaoli Ben* 高麗本 or ‘Korean edition’. In fact it is the Japanese *editio princeps* or rather its reprint. It was initiated by the scholar-monk Michisuke 道祐 in the 19<sup>th</sup> year of the Shôhei era (1364).

- E: Kouton 篁墩 edition, also called Ashikaga 足利 edition. An early Japanese print with the He Yan commentary from the era Keichô 慶長 (1596–1614).
- F: An early Chinese print from Song dynasty that has been preserved in Japan. It contains not only the commentary by He Yan but also the sub-commentary by Xing Bing 邢昺 (932–1010) and the glosses by Lu Deming 陸德明 (550–630). It is probably the earliest extant predecessor of our modern edition of the Thirteen Classics.<sup>9</sup>

The variants given below cover the passage we have already discussed (*Lunyu* XII.7–12) with the Dunhuang manuscripts, except for the last example.

Table 11

D: Shôhei 1364	E: Kouton 1596–1614	F: Song ca. 1200	Translation
●●曰		子貢曰	<i>He said / Zi Gong said</i>
民不信不立		民不信無立	If the people have no trust, one <i>will fall</i> . / If the people have no trust, one <i>has</i> <i>nowhere to stand</i> .
棘子城		棘子成	Ji Zi Cheng [a name]
何以文爲矣	何以文爲●		Why take account of refinement?
虎豹之鞞猶犬羊之鞞也	虎豹之鞞猶犬羊之鞞●		The shaven skin of a tiger or a leopard is not different from the shaven skin of goat or a dog.
愛之欲其生也	愛之欲其生●		If you love someone, you want him to live.
惡之欲其死也	惡之欲其死●		If you hate someone, you want him to die.
既欲其生又欲其存是惑●	既欲其生又欲其存是惑也		If having wanted him to live, you also want him to die, this is confusion.
吾豈得而食諸		吾●得而食諸	<i>How could I manage to eat it? / Shall I manage to eat it?</i>
爲之●難也	爲之也難也	爲之也難●	Carrying it out is difficult. (XIV.20)

9 D and F are reproduced in facsimile, for E the edition of Yoshida Kangan 吉田漢官 was used. See *Wuqiubei zhai Lunyu jicheng* 無求備齋論語集成, ed. Yan Lingfeng 嚴靈峰, 308 vols., Taipei 1966, vol. 12–14, 32–35, 273.



It is quite evident that the readings of the youngest edition in the middle vacillate between the earlier editions to the right and the left. Obviously, the editor based his own edition on the earliest prints from both Japan (on the left), China (on the right), or editions that were closely related to them. We can even detect a certain strategy he followed when choosing between the variants of the two editions. For variants that seem to rather affect the meaning, he followed his own indigenous tradition (on the left), whereas for particles that do not affect the meaning at all he preferred the Chinese tradition (on the right).

The last example in the table is from chapter XIV.20 and is of special interest, because we can learn from it, why such contaminated traditions pose an extreme threat to stemmatology. The texts that our editor used differed by a transposition of the particle *ye* (也). In the Chinese edition it appears before the predicate. This is a rather rare, but nevertheless grammatically correct position, the function of which is marking the topic of a sentence. Mostly, however, the particle *ye* is found at the end of sentences, and this is the case in the Japanese edition. Both positions seemed, therefore, somehow plausible to our editor and he decided to retain both.

From the view point of stemmatology the Kouton edition is completely useless, because all its readings are derived from its earlier sources, which are still extant and should be examined themselves. Unfortunately, this state of affairs can be determined with certainty only when the age of the witnesses involved is known, or if the textual critic becomes aware of a clear strategy followed by the collator. If we did not know the age of the prints, we would have to conclude that the reading of the Kouton edition must be the most original one, because it explains easily how the readings of both the Shôhei and the Song editions came into being. Both readings can be understood as omissions of the particle *ye*, which is by far the most common scribal error in ancient texts, as could be already seen from the Dunhuang manuscripts we discussed above. Neither the reading of the Shôhei nor of the Song edition, on the other hand, is able to explain the readings of the remaining two witnesses. The case that the branches of the tree grow together again is not provided for by the stemmatological method. The contaminated text of the Kouton edition turns the genealogical tree upside down and leads to a wrong text-critical decision. Instead of eliminating the late Kouton edition, we would eliminate its earlier sources.

Another difficulty we meet when dealing with printed editions is also visible from the variants listed above: Most readings we find differ only in particles. These are called *xuci* 虛詞 ‘empty words’ in Chinese, because they do not contribute substantially to the meaning of the sentence. For most readers of



Classical Chinese they are more or less negligible and have a status comparable to punctuation in western languages. This makes it extremely difficult, to decide whether a given variant is an error or not. Consequently, with printed editions it generally becomes impossible to find significative errors that would enable us to apply stemmatology in cases where no contamination is to be suspected.