JOURNAL OF CHINESE WRITING SYSTEMS

Article

Journal of Chinese Writing Systems
2018, Vol 2(3) 195–207

DOI: 10.1177/2513850218778825
journals.sagepub.com/home/cws
SAGE

An introduction to the Tangut *Homonyms*

Andrew West
Independent scholar, UK

Abstract

This paper discusses a printed edition of a hitherto unknown Tangut linguistic text called *Homonyms*, which was sold at auction in Beijing in 2014. The book is complete in 14 folios, and comprises two prefaces on three folios; an initial section on two folios showing the cursive forms of sample Tangut characters, ordered by components; and a main section on nine folios which lists 577 groups of homophone characters, listed under head characters that derive from the Tangut books *Synonyms* and *Mixed Characters*. The homophone groups do not closely correspond to homophones in other Tangut linguistic texts such as *Homophones* and *Sea of Writing*, and so this text may represent the phonetic system for a different dialect or historical stage of the Tangut language.

Keywords

Tangut cursive characters, Tangut homophones, Tangut phonetics, Tangut texts

Background

This paper introduces a woodblock-printed edition of a previously unknown Tangut linguistic text entitled [illegible] ^{1}tshi$_{4}$ ^{1}tser'$_{4}$ ^{2}u$_{3}$ ^{1}de'$_{1}$ ^{2}me'$_{4}$ ^{2}lew$_{1}$ ^{2}di$_{4}$ ^{1}dza$_{1}$ '*Essential Selection of Often-Transmitted Homonyms and Mixed Characters*', which I will refer to as *Homonyms* for short.[1] The provenance of this book is unknown, and it first appeared at an auction of ancient documents by Beijing Debao International Auction Company at Beijing Capital Library on 23 November 2014.[2] The book is now in private ownership.

I previously made a preliminary analysis of the contents of this book based on low-resolution images of the front cover and two and a half folios of the text (see West, 2015). The current paper is an extension of my preliminary analysis that is based on an examination of high-resolution photographs of the entire book that were made after conservation following the auction of the book.

Description

The book comprises a single fascicle of 14 folios in butterfly binding, with a yellow silk wrapper. A printed-paper title slip is present on the left side of the front cover, although unfortunately it is damaged and only the first four characters are complete. Part of the fifth character survives, but it is not easy to identify, and there is space on the title slip for two more characters which are now lost. The first four characters of the title slip are [illegible] ^{1}se$_{1}$ ^{2}dwewr$_{1}$ jy ^{2}o$_{1}$ which means 'entering the art of pure enlightenment'. The term 'pure enlightenment' (corresponding to Chinese *jìngjué* 净觉) is a Buddhist term, and it would seem to be the title for an unknown Buddhist text. As the term 'pure enlightenment' is not part of this book's end-title, and is not mentioned in the prefaces, I do not think this title slip refers to the linguistic text inside the cover. Probably this was originally the wrapper for a Buddhist work, and was later borrowed to cover the *Homonyms*.

The book is in relatively good condition, although all folios have an area of damage on the upper part of the central fold, which spreads out as a narrow tear halfway across the page on both sides. This has caused some localized damage to the text, but most of the characters on all folios are undamaged and legible.

The 14 folios are numbered '1' through '14' in Chinese characters on the lower part of the central fold (except for folios 1 and 10 where the numbers are missing due to damage). The upper part of the central fold, where the short title of the book would be expected, is damaged on all folios, but partial remains of one or more Tangut characters can be seen on folios 3 through 7 and folio 9. Unfortunately, the characters are too incomplete to identify.

Corresponding author:
Andrew West, 3 Abbot's Ride, Farnham, Surrey GU9 8HY, UK.
Email: babelstone@gmail.com

The last line on folio 14 gives the end-title of the book as [Tangut] ^{1}tshi$_{4}$ ^{1}tser'$_{4}$ ^{2}u$_{3}$ ^{1}de'$_{1}$ ^{2}me'$_{4}$ ^{2}lew$_{1}$ ^{2}di$_{4}$ ^{1}dza$_{1}$ ^{1}bu$_{1}$ a ^{1}deq$_{4}$ '*Essential Selection of Often-Transmitted Homonyms and Mixed Characters [with] Preface in One Category*'. The Tangut character [Tangut] 'end' would be expected after a space following the title, but is missing due to damage to this part of the page.

The book consists of the following contents:

- folios 1–2: first preface;
- folio 3: second preface;
- folios 4–5: initial section, comprising 30 letter signs followed by a list of Tangut characters in regular and cursive script, ordered by component; and
- folios 6–14: main section, comprising a mixture of large and small Tangut characters, where groups of small characters form homophone groups.

Prefaces

The first preface is on folios 1–2, and has nine columns of 20 characters per half-folio, except for the end of folio 2B where three columns of 23 characters are squeezed into the space of two columns. The total length of the preface is 728 characters. The preface does not have a title at the front or a date or author credit at the end, but it does provide some important information about the origins of the book, and several times mentions a book called *Sea of Characters* ([Tangut] ^{2}di$_{4}$ ^{2}ngon$_{2}$ = Chinese *Zìhǎi* 字海). The title of this book is very similar to that of the extant *Sea of Writing* ([Tangut] ^{2}wyr$_{4}$ ^{2}ngon$_{2}$ = Chinese *Wénhǎi* 文海), but probably refers to a different Tangut dictionary that has not survived. In one place the preface states

> [Tangut]。
>
> About two-and-half sevenths of [the characters in] *Sea of Characters* have been selected separately in *Often-transmitted Mixed Characters and Homonyms with Preface* in one volume.

This indicates that the *Sea of Characters* was the source text from which the oddly specific fraction of two-and-half sevenths of its characters were selected for inclusion in the *Homonyms*, presumably as the homophone characters in the main section (that would amount to 2142 characters if *Sea of Characters* had 6000 characters). The title of the book given here as '*Often-transmitted Mixed Characters and Homonyms with Preface*' is a close variation on the end-title given on folio 14B.

The second preface is on folio 3, and has nine columns of 19 characters per half-folio, making it 342 characters in length. This preface also lacks a title and final authorship credit. It mentions a text called [Tangut] *Synonyms with Preface* a couple of times (it is also mentioned twice in the first preface). The preface starts with the following words, which, although somewhat opaque in meaning, indicate the important role of *Synonyms with Preface*:

> [Tangut]，[Tangut]。[Tangut]。
>
> As to what are called letters, they are what whole characters are born from and can be constructed from. In order to easily write and study, letters called 'mothers of written characters' are collected from the stem meanings in *Synonyms with Preface* for use as letters.

Later in the preface there are some interesting statistics concerning how many characters in *Homonyms* ('[this] little volume') derive from *Synonyms with Preface*:

> [Tangut]□[Tangut]。[Tangut]。[Tangut]。
>
> There are about five thousand [?] hundred characters in *Synonyms with Preface*. In [this] little volume there are about two thousand eight hundred [characters]. There are about nine hundred large characters.

The main section of *Homonyms* comprises 913 large characters and about 1856 small characters – 2769 characters in total. This is an excellent match for the figures of about 900 large characters and about 2800 characters in total given in the preface, and is strong evidence that *Synonyms with Preface* was the source for Tangut characters in the main section of the *Homonyms* book. However, this contradicts the first preface, which indicates that the characters in the *Homonyms* were taken from a book called *Sea of Characters*.

As will be demonstrated below, the *Synonyms with Preface* mentioned in the preface is the same work as the extant manuscript text entitled *Synonyms* ([Tangut] ^{2}vo$_{1}$ ^{2}lew$_{1}$ = Chinese *Tóngyì* 同义), with a colophon dated 1188.[3] This gives a *terminus post quem* for the composition of *Homonyms*.

Initial section

The initial section on folios 4 and 5 comprises two parts, a list of 30 letter signs and a list of sample Tangut characters ordered by component.

Thirty letters

The first three columns of folio 4 are a list of 30 signs preceded by a heading cut in intaglio that reads [Tangut] ^{1}ta$_{4}$ ^{1}shwy$_{3}$ ^{1}soq$_{1}$ ^{2}ghaq$_{1}$ ^{2}di$_{4}$ ^{1}ma$_{4}$ (see Figure 1). The last four characters of the heading are unproblematic, and

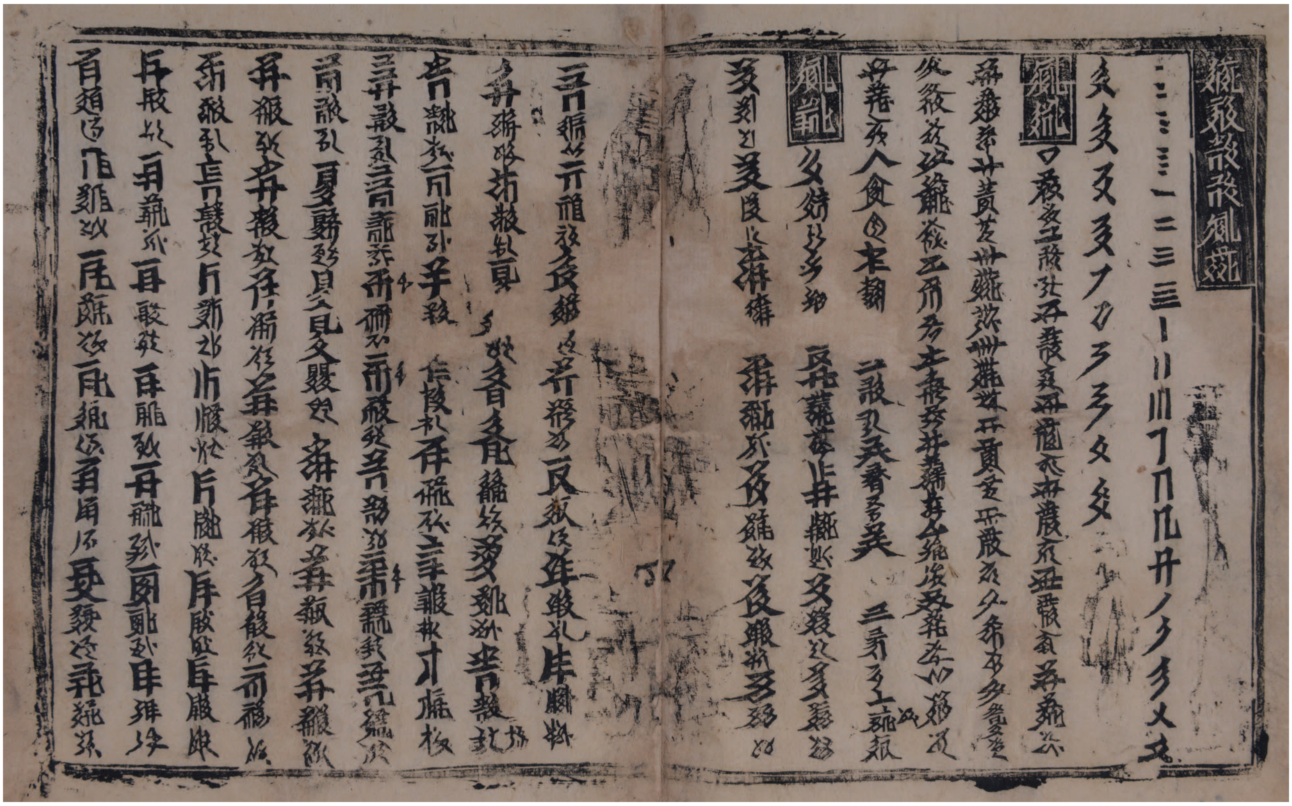


Figure 1. Folio 4 showing the first folio of the initial section.

mean ‘thirty letters’. However, the first two characters are difficult to interpret. The first character [illegible] ^{1}ta$_{4}$ is a topic-marking particle (similar to Classical Chinese *zhě* 者), and grammatically cannot occur in an initial position. The following character [illegible] ^{1}shwy$_{3}$ means ‘time’ (a borrowing from Chinese *shí* 时), but in extant Tangut texts the character almost always occurs reduplicated as [illegible] meaning ‘often’. Together, these two characters make no sense preceding the phrase ‘thirty letters’.

In my initial analysis of low-resolution images of this book, I read the second character as [illegible] ^{2}vi$_{1}$, ‘to cut’, which has an extended meaning of ‘rime’, and so I took these two characters to mean the ‘*ta* rime’; that is, a Tangut rime group named after the character *ta* [illegible] (West, 2015). However, from the high-resolution image of folio 4, it is clear that the character in question is [illegible] rather than [illegible]. Moreover, the two characters [illegible] occur together three times in the second preface, once in the exact same phrase [illegible] that occurs at the start of folio 4. It seems from this correspondence that the preface on folio 3 refers to the initial section on folios 4 and 5. Unfortunately, the presence of these two characters in the preface does not greatly help in elucidating their meaning.

Following the heading are a list of 30 signs, presumably the ‘thirty letters’ referred to here and in the preface. However, these signs are not letters of any known alphabet, but are composed of Tangut components or groups of horizontal, vertical or diagonal strokes resembling tally marks. The 30 signs can be divided into nine series of between two and four signs, with each series showing an incremental increase in stroke complexity (see Table 1). These nine series correspond quite neatly to the nine initial classes in *Homophones* ([illegible] ^{2}ghiq$_{2}$ ^{2}lew$_{1}$ = Chinese *Tóngyīn* 同音). In *Dissected Rimes of the Five Sounds* ([illegible] ^{1}ngwy$_{1}$ ^{2}ghiq$_{2}$ ^{2}vi$_{1}$ ^{1}bu$_{4}$ = Chinese Wuyinqieyun 五音切韵), these nine initial classes are subdivided into a total of 36 initials, and the 30 letter signs do broadly map to these 36 initials.[4]

These signs do not occur anywhere else in this book, so it is not clear how they were intended to be used. However, they may be derived from a version of *Synonyms* where these signs were used to indicate the phonetic classification of Tangut characters.

Table 1. The 30 letter signs.

Series	30 letters	Class	Initial name	36 initials
1	[illegible]	I	[illegible] 'heavy lip sound' (bilabials)	[illegible]
2	[illegible]	II	[illegible] 'light lip sound' (labio-dentals)	[illegible]
3	[illegible]	III	[illegible] 'tongue tip sound' (dentals)	[illegible]
4	[illegible]	IV	[illegible] 'tongue top sound' (alveolars)	[illegible]
5	[illegible]	V	[illegible] 'tooth sound' (velars)	[illegible]
6	[illegible]	VI	[illegible] 'tooth tip sound' (dental affricates and fricatives)	[illegible]
7	[illegible]	VII	[illegible] 'true tooth sound' (palatal affricates and fricatives)	[illegible]
8	[illegible]	VIII	[illegible] 'larynx sound' (laryngeals)	[illegible]
9	[illegible]	IX	[illegible] 'wind only sound' (resonants)	[illegible]

Sample characters

The remainder of folios 4 and 5 list Tangut characters in regular and cursive script, arranged by component. There are a total of 230 components, divided into five sections with a heading cut in intaglio:

- [illegible] ^{2}di$_{4}$ ^{1}ghu$_{2}$: head components (31 components);
- [illegible] ^{2}di$_{4}$ ^{1}pha$_{1}$: left-side components (127 components);
- [illegible] ^{2}di$_{4}$ ^{1}ne'$_{4}$: central components (11 components);
- [illegible] ^{2}di$_{4}$ ^{2}ber'$_{4}$: right-side components (19 components); and
- [illegible] ^{2}di$_{4}$ ^{2}ngorn$_{1}$ ^{1}pha$_{1}$ ^{2}vi$_{1}$: whole characters that are used as lateral components (38 components).

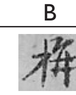
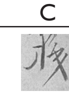
Appended to the last section, without a heading or any explanation, are four left-side components that extend underneath to the right.

For the first four sections, each component is followed by a Tangut character in regular script that incorporates that component, and then by the cursive script form of that character. The whole characters used as lateral components are only followed by the cursive script form of the character. The four left-and-under components do not have a sample character or cursive script form, but there is a short description for each component. Many of the left-side components can occur on either side, and some have a sample character where the component is actually on the right side. In contrast, all of the right-side components only ever occur on the right side.

The sample characters for the first four sections do not share the same rime group or initial class, but represent a mixture of phonetic syllables, so have not been chosen on phonetic grounds. Some of the sample characters show a semantic relationship (e.g. 'land', 'water', 'fire' and 'wind' are together, and 'finger' and 'hand' are next to each other), but on the whole there seems to have been no particular logic behind the choice of sample characters. The 230 components and sample characters are listed in Table 2.

Considering that the cursive characters are written in a very small space, and this book may have been printed from recarved blocks, most of the cursive forms are recognizably correct. Table 3 compares some of the cursive script characters from folio 4 with the corresponding cursive form characters that occur in the Tangut manuscript translation of the *Classic of Filial Piety* ([illegible] ^{2}iq$_{4}$ ^{2}lhe$_{1}$ ^{1}vy'$_{1}$ ^{2}ge$_{4}$ ^{1}de'$_{1}$ = Chinese Chongyi Xiaojing jingzhuan重译孝经传), as identified by Eric Grinstead (1972: 300–376).

There is generally a fairly good match with the same cursive character in the *Classic of Filial Piety*, and in a few cases the match is extremely close; for example, the complex left-hand component of [illegible] is written with exactly the same simple cursive form as used in the *Classic of Filial Piety*. In cases where there is not a good match, this may be because there is more than one way of writing the cursive form of the same character. A good example is [illegible], where the cursive form of the character (A) does not match the form given in the *Classic of Filial Piety* (B), but is very similar to the form of the character found in the Tangut translation of the *Mahāprajñāpāramitā-sūtra* (C).[5]

A	B	C
		

Main section

The main section of the book occupies folios 6 through 14, and has seven columns per half-folio, although folio 6 only has nine columns, with most of the second half of the folio left blank. There is no title or heading at the start of folio 6, but there is an end-title on the third column of the second half of folio 14 (see earlier).

Table 2. Tangut components and sample characters.

Position	Components (sample characters)
Top	[illegible]
Left	[illegible]
Middle	[illegible]
Right	[illegible]
Whole	[illegible]
Left and under	[illegible]

Table 3. Tangut cursive characters.

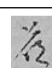
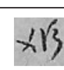
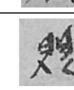
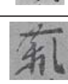
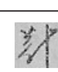
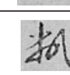
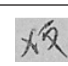
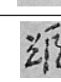
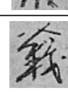
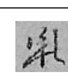
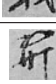
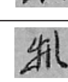
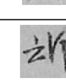
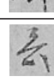
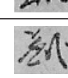
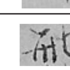
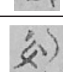
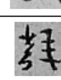
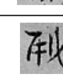
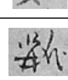
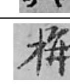
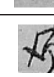
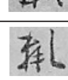
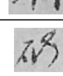
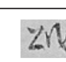
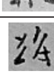
Character	Cursive	*Xiaojing*	Character	Cursive	*Xiaojing*	Character	Cursive	*Xiaojing*
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
[illegible]	[illegible]							

Figure 2. Folio 8 showing part of the main section.

The text comprises large-sized characters intermixed with small-sized characters written in two small columns (see Figure 2). There are a total of 913 large characters and about 1856 small characters. The small characters are divided into 577 groups of between 1 and 24 characters. Each group of small characters are homophones or near homophones of the preceding large character. Thus, 577 large characters are head characters for homophone groups, and 337 large characters are not associated with homophone groups. The reason why not all large characters head homophone groups is that the large characters are derived from texts that are unrelated to homophones, so they are not ordered phonetically and they are not all phonetically unique. Thus, there may be several large characters with the same pronunciation, and it seems that in most such cases only the first large character with the same pronunciation is followed by a group of small character homophones, and subsequent large characters with the same pronunciation are not followed by any small characters. For example, [illegible] ^{2}kha$_{4}$ on folio 8A is followed by a list of six small-sized homophone characters, whereas [illegible] ^{2}kha$_{4}$ on folio 14B is not followed by any small-sized characters. None of the 337 large characters with no associated small characters also occur as small characters under other large characters.

There is only a single small homophone character which occurs under two different large characters: [illegible] ^{1}o$_{1}$ occurs in a group of four small characters under [illegible] a on folio 7A, and then by itself under [illegible] a$_{1}$ on folio 7B. In addition, [illegible] occurs twice as a small character, but with two different readings and with somewhat different glyph forms, suggesting that the author considered them to be two separate characters: [illegible] ny$_{1}$ 'fox' under [illegible] ^{2}nwy$_{4}$ on folio 9A; and ^{1}ta$_{1}$ [illegible] 'tail' under [illegible] ^{1}ta$_{1}$ on folio 11B.

Large characters

If we ignore the small homophone characters, and just read the large characters sequentially, we discover that the large characters are not random, but are semantically related to each other. That is to say, in very many cases pairs of adjacent large characters, or sometimes longer sequences of large characters, form words or phrases. The large characters on folio 6 and on folios 7 through 14 have different origins, and so are discussed separately in the following sections.

Folio 6. Folio 6 only has nine columns, with space for the remaining five columns on folio 6B left blank, which is a strong indication that it is separate from the following eight folios. There are 52 large characters on folio 6, ending with the character [illegible] a$_{1}$ which is used to transcribe the Buddhist syllable *oṃ* at the start of mantras. As the first six large characters of folio 7 are [illegible] ^{1}tha$_{1}$ ^{1}tsir'$_{4}$ ^{1}sin$_{1}$ ^{1}e$_{4}$

^{1}dzwyq$_{4}$ ^{1}tshwew$_{1}$ ‘show respect to the Buddha, the Dharma, and the Sangha’, perhaps the *oṃ* at the end of folio 6 should actually precede ‘Buddha, Dharma and Sangha’ at the start of folio 7. The penultimate two large characters on folio 6 are [Tangut] ^{1}cha$_{3}$ ^{1}jwa$_{3}$ ‘end of ceremony’, which is perhaps a form of wording that indicates the end of the text composed from the preceding 49 characters.

In fact, the large characters on folio 6 up to the words ‘end of ceremony’ derive directly from the *Synonyms* text discussed in the second preface. The first 10 large characters form a complete sentence: [Tangut] ^{2}wyr$_{4}$ ^{2}di$_{4}$ ^{2}rar$_{4}$ ^{2}si$_{4}$ ^{1}soq$_{1}$ ^{2}ghaq$_{1}$ ^{1}gi$_{4}$ ^{1}nen$_{2}$ ^{2}sew’$_{1}$ ^{1}niq$_{1}$ ‘Examine and speak the thirty clear and turbid mothers of written characters’. The term [Tangut] ‘mothers of written characters’ is a longer synonym for [Tangut] ‘letters’, and the terms ‘clear’ and ‘turbid’ refer to ‘voiceless’ and ‘voiced’ letters, respectively. The phrase [Tangut] ‘thirty clear and turbid mothers of written characters’ mirrors the term [Tangut] ‘thirty letters’ given in the second preface and in the heading to the 30 letter signs on folio 4, and evidently refers to the division of *Synonyms* into 30 sections based on clear and turbid phonetic distinctions. I suggest that this sentence prefaced the main text of *Synonyms*, and may have been in the first column of the first folio, which is lost due to manuscript damage. The 39 following large characters correspond to the 30 section-header characters in *Synonyms* as well as the clear and turbid designations for the four parts of the *Synonyms*, as shown in Table 4.

Folios 7–14. The large characters on folios 7 through 14 are not copied from *Synonyms*, but form words and phrases that appear to be items in a vocabulary list. Most of the characters form two-character words, but there are some four-character phrases, and the longest phrase is the six-syllable mantra *oṃ maṇi padme hūṃ* on folio 7B: [Tangut] a$_{1}$ ^{1}ma$_{4}$ ^{2}ni$_{4}$ ^{1}pa$_{4}$ ^{1}me$_{4}$ ho$_{1}$.

The end-title for this book is ‘*Essential Selection of Often-Transmitted Homonyms and Mixed Characters*’, and ‘mixed characters’ is the same as the title of the Tangut printed book *Mixed Characters* ([Tangut] ^{2}di$_{4}$ ^{1}dza$_{1}$ = Chinese *Zázì* 雜字), which lists Tangut vocabulary in semantic categories without any glosses. It seems very likely that ‘mixed characters’ in the title of the current book refers to the large characters in most of the main section being derived from a Tangut vocabulary book entitled *Mixed Characters*. However, this *Mixed Characters* cannot be the same as the extant book with the same title, as only a few words constructed from the large characters in the main section also occur in the extant *Mixed Characters*, and those words are not in the same order in the two texts (e.g. the only word on folio 8 that also occurs in the printed edition of *Mixed Characters* is [Tangut] ‘young and strong’). Moreover, this text has a relatively high proportion of words and phrases related to Buddhism, whereas the extant *Mixed Characters* text has almost no Buddhist vocabulary. Thus, we can conjecture that the large characters on folios 7–14 were derived from an otherwise lost Tangut vocabulary list entitled *Mixed Characters*. This may have been a constituent part of the book entitled *Sea of Characters* that is mentioned several times in the first preface.

All the large characters on folios 7–14 only occur a single time, and only a single large character also occurs on folio 6 ([Tangut] occurs both on folio 6B and on folio 7B). Moreover, none of the large characters also occur as small-character homophones under other large characters. This is quite unexpected, as there are many individual characters that occur in multiple words composed from two or more characters, and we would expect any vocabulary list to include some words that are composed from at least one character that occurs in another word in the list. This can be seen in the extant *Mixed Characters* where 318 of the 1734 identifiable Tangut characters occur in more than one word (e.g. the character [Tangut] ‘tree’ occurs in 16 words, [Tangut] ‘tree’ occurs in 12 words and [Tangut] ‘man’ occurs in 10 words). That is to say, more than one-quarter of the 1186 words in the extant *Mixed Characters* are not composed of unique characters. On this basis, the absence of any duplicate characters on folios 7 through 14 cannot be a coincidence, and we have to conclude that duplicate large characters were deliberately removed, as well as any large characters that had already been listed as small-character homophones under a previous large character.

The 107 large characters on folio 8 (plus one large character at the start of folio 9) can be analysed as follows, where sequences of two or more characters are joined to form a word or phrase. Those words marked with an asterisk are not attested in any of the sources I have checked, but they are plausible words based on the meanings of the individual characters.

[Tangut] ^{2}seq$_{4}$ ‘to write’
[Tangut] ^{1}ku$_{1}$ ‘so, then’
[Tangut] ^{2}bi$_{2}$ ^{2}lhew$_{1}$ ‘to release’
[Tangut] ^{2}lhuq$_{4}$ ^{1}ror$_{4}$ ‘to capture’ + ‘to get’ = * ‘to capture’
[Tangut] ^{1}me$_{1}$ ^{1}nu$_{4}$ ^{2}ni’$_{4}$ ^{1}lhwa$_{4}$ ‘eyes, ear, nose, tongue’
[Tangut] ^{2}luq$_{3}$ ‘body’
[Tangut] ^{1}phi$_{4}$ ‘scheme, idea’
[Tangut] ^{2}ne$_{1}$ ^{2}zher$_{2}$ ‘to live in peace’
[Tangut] ^{1}lo$_{3}$ ^{1}lhu$_{1}$ ‘good fortune’ + ‘to increase’ = * ‘to increase good fortune’
[Tangut] ^{1}du’$_{1}$ ^{2}dzyq$_{4}$ ‘to save up’
[Tangut] ^{2}hwin$_{1}$ ^{2}bi’$_{1}$ ‘demon’
[Tangut] ^{2}hoq$_{1}$ ^{1}khwa’$_{1}$ ‘calamity’ + ‘to curse’ = * ‘to curse with calamity’
[Tangut] ^{1}peq$_{4}$ ^{2}cha’$_{2}$ ‘to pray to avert misfortune’ + ‘suffering’ = * ‘to pray to avert suffering’

Table 4. Correspondence between *Synonyms* and folio 6 large characters.

Synonyms headings	Folio 6	Reading	Meaning
1. [illegible]			Part 1: completely clear, seven sections
1.1 (0112.01) : □	[illegible]	^{1}chen$_{3}$	correct
1.2 (0127.01) : □	[illegible]	^{2}chha$_{3}$	virtue
1.3 (0311.01) : □	[illegible]	^{2}seq$_{4}$	wisdom
1.4 (0317.01) : [illegible]	[illegible]	^{1}zhyr$_{3}$	wisdom
1.5 (0516.01) : [illegible]	[illegible]	^{2}sew'$_{1}$	to survey
1.6 (0614.01) : [illegible]	[illegible]	^{2}se'$_{4}$	thought
1.7 (0817.01) : [illegible]	[illegible]	^{2}tse$_{4}$	to know
	[illegible]	^{2}zi$_{4}$ ^{1}gi$_{4}$	completely clear
2. [illegible]			Part 2: partially clear, seven sections
2.1 (0825.01) : [illegible]	[illegible]	^{1}gi$_{4}$	clear
2.2 (1016.01) : [illegible]	[illegible]	^{1}rer$_{4}$	skilful
2.3 (1022.01) : [illegible]	[illegible]	^{1}ge$_{4}$	special
2.4 (1111.01) : □	[illegible]	^{2}u$_{3}$	to seek
2.5 (1227.01) : [illegible]	[illegible]	^{1}swew$_{1}$	bright
2.6 (1323.01) : [illegible]	[illegible]	^{1}du$_{4}$	to have
2.7 (1527.01) : □	[illegible]	^{2}le$_{4}$	to see
	[illegible]	^{2}bu$_{3}$ ^{1}gi$_{4}$	partially clear
3. [illegible]			Part 3: half clear and turbid, eight sections
3.1 (1712.01) : [illegible]	[illegible]	^{2}kur$_{4}$	will
3.2 (1811.01) : [illegible]	[illegible]	^{2}kwe$_{2}$	true
3.3 (1817.01) : [illegible]	[illegible]	^{1}pho$_{4}$	dexterous
3.4 (1911.01) : [illegible]	[illegible]	^{2}rar$_{1}$	to flow
3.5 (2023.01) : [illegible]	[illegible]	^{2}dzo$_{4}$	poetry
3.6 (2217.01) : [illegible]	[illegible]	^{2}ngaq$_{1}$	good
3.7 (2312.01) : [illegible]	[illegible]	^{1}la$_{4}$	to fall
3.8 (2327.01) : [illegible]	[illegible]	^{1}nwy$_{1}$	to know
	[illegible]	^{1}khwy$_{1}$ ^{1}gi$_{4}$ ^{1}nen$_{2}$	half clear and turbid
4. [illegible]			Part 4: completely turbid, eight sections
4.1 (2424.01) : [illegible]	[illegible]	^{1}khon$_{4}$	strong
4.2 (2612.01) : [illegible]	[illegible]	^{1}biq$_{1}$	to pass through
4.3 (2624.01) : [illegible]	[illegible]	^{1}ver$_{1}$	luxuriant
4.4 (2712.01) : [illegible]	[illegible]	^{1}ur$_{4}$	to breed
4.5 (2824.01) : [illegible]	[illegible]	^{1}gwyr$_{4}$	to rise
4.6 (2912.01) : [illegible]	[illegible]	j?	oblique
4.7 (3111.01) : [illegible]	[illegible]	^{1}lwo$_{1}$	dim
4.8 (?) :	[illegible]	^{1}mur$_{1}$	dark
	[illegible]	^{2}zi$_{4}$ ^{1}nen$_{2}$	completely turbid
	[illegible]	^{1}cha$_{3}$ ^{1}jwa$_{3}$	end of ceremony

[Tangut] ^{1}shi’$_3$ ^{1}khwaq$_1$ ‘doubt’ + ‘far, distant’ = * ‘forethought’
[Tangut] ^{1}zhaq$_3$ ‘between’
[Tangut] ^{1}pha’$_4$ ‘to forbid’
[Tangut] ^{1}chy$_3$ ^{2}le’$_3$ ‘to destroy’
[Tangut] ^{1}jon$_3$ ^{2}ka$_1$ ‘to separate’
[Tangut] ^{2}vy$_3$ prefix
[Tangut] ^{2}phe’$_1$ ‘to divide’
[Tangut] ^{1}la$_3$ lho$_1$ ‘extraordinary’
[Tangut] ^{1}ge’$_4$ ^{1}dzeq$_4$ ‘to surpass’
[Tangut] ^{1}non’$_2$ ^{1}tsy$_4$ ‘and then’
[Tangut] ^{1}my$_4$ ^{1}puq$_4$ ‘without measure’
[Tangut] ^{2}ir$_4$ ^{1}tuq$_1$ ‘hundreds and thousands’
[Tangut] ^{1}ku$_4$ ^{2}chi$_3$ transliteration of Sanskrit *koṭi* ‘ten million’
[Tangut] ^{2}ner$_4$ ^{2}nenq$_4$ ^{2}a$_4$ ^{1}chha$_2$ ‘Yakṣa with a wrathful expression’ (Chinese *pínméi yàochā* 顰眉藥叉)
[Tangut] ^{1}hew$_1$ ‘kind of grass’
[Tangut] ^{2}kha$_4$ ‘to draw (water)’
[Tangut] ^{1}kiq$_2$ ^{2}ja$_3$ ‘vajra’
[Tangut] ^{1}naq$_4$ ‘god’
[Tangut] ^{2}kho$_2$ ^{2}nwi$_4$ ‘skilful’
[Tangut] ^{1}dzyr$_4$ ^{2}lwiq$_1$ ‘speed’ (literally ‘fast and slow’)
[Tangut] ^{1}kir$_4$ ^{1}ga’$_4$ ‘brave, violent, sudden’ + ‘to jump’ = * ‘to jump out suddenly’
[Tangut] ^{1}thwiq$_4$ ^{1}myr$_4$ ‘young and strong’
[Tangut] ^{2}lu’$_3$ ^{1}tho’$_4$ ‘beautiful’
[Tangut] ^{1}tsen$_1$ ^{2}khwe$_1$ ‘size’ (literally ‘small and large’)
[Tangut] ^{1}dza$_1$ ^{1}lwu$_1$ ‘mixed up’
[Tangut] ^{1}chon$_3$ Chinese transliteration (*zhong*)
[Tangut] ^{1}ghwan$_4$ Chinese transliteration (*yuan*)
[Tangut] ^{2}chhe$_2$ ^{1}lhor’$_1$ ‘stocked village’ + ‘market’ = * ‘village market’
[Tangut] ^{2}kuq$_1$ ‘inner’
[Tangut] ^{2}khew$_1$ ‘opening’
[Tangut] ^{2}pho$_2$ ‘window’
[Tangut] ^{2}kyr’$_4$ ‘house’
[Tangut] ^{2}giq$_4$ ‘wide’
[Tangut] ^{1}khu$_4$ ‘below’
[Tangut] ^{2}o$_1$ ‘to hold, to enter’
[Tangut] ^{2}dwu’$_1$ ^{2}lwu$_1$ ‘secret’ + ‘to hide’ = * ‘to hide in secret’
[Tangut] ^{2}khu’$_4$ ‘to watch’
[Tangut] ^{1}chhwen$_3$ ‘to prohibit’
[Tangut] ^{1}chuq$_3$ ^{2}wer$_1$ ‘to guard’
[Tangut] ^{2}va$_1$ ‘how, what’
[Tangut] ^{1}dzu$_1$ ^{1}ngwi$_1$ ‘to respectively love’
[Tangut] ^{2}gi’$_4$ ‘hope’
[Tangut] ^{2}sha’$_2$ ‘at will; target’
[Tangut] ^{1}pyq$_1$ ‘to hit’
[Tangut] ^{1}ma’$_4$ ^{2}tsha$_4$ ‘fruits of retribution’
[Tangut] ^{2}ngorn$_1$ ‘whole’
[Tangut] ^{1}duq$_4$ ‘to meet’
[Tangut] ^{1}sy$_1$ lhy$_1$ ‘complete’
[Tangut] ^{1}chy’$_3$ ^{1}my$_1$ ‘order’ + ‘sky, heaven’ = * ‘order of heaven’
[Tangut] ^{2}bi$_1$ ^{2}lhiq$_4$ ‘sun and moon’
[Tangut] ^{2}gyq$_4$ ^{1}geq$_4$ ‘constellation’
[Tangut] ^{2}vyq$_3$ ^{1}zyr$_4$ ^{2}li$_3$ ^{1}laq$_3$ ‘east, south, west, north’

There are 25 attested two-character words and nine conjectural two-character words, as well as three four-character phrases. That leaves 28 isolated characters. Some or most of these isolated characters may originally have formed words or phrases in the source text, but the characters they were paired with have been removed as duplicates.

For example, the adjacent characters [Tangut] ‘at will; target’ and [Tangut] ‘to hit’ are not attested as a word, but the two characters do occur together in the phrase [Tangut] ^{2}ir$_4$ ^{2}sha’$_2$ ^{2}ir$_4$ ^{1}pyq$_1$ ‘a hundred shots and a hundred hits’. As the character [Tangut] meaning ‘hundred’ already occurs earlier on folio 8A in the expression ‘hundreds and thousands’, we can conjecture that the two characters meaning ‘hundred’ in the phrase ‘a hundred shots and a hundred hits’ were removed as duplicates, leaving just the two characters [Tangut] ‘target’ and [Tangut] ‘to hit’.

Likewise, the two adjacent characters [Tangut] ^{1}hew$_1$ ‘kind of grass’ and [Tangut] ^{2}kha$_4$ ‘to draw (water)’ do not at first sight appear to have any connection with each other, but they are actually also used as transliteration characters for the Sanskrit word *mahoraga* (Chinese 摩睺羅伽) meaning a great serpent protector: [Tangut] ^{2}mo$_1$^{1}hew$_1$ ^{1}lo$_1$ ^{2}kha$_4$. The character [Tangut] ^{2}mo$_1$ already occurs as a small-character homophone on folio 7A, and the character [Tangut] ^{1}lo$_1$ already occurs on folio 7B in the word [Tangut] ^{1}lo$_1$ ^{1}han$_1$ ‘Arhat’, so we can assume that both of these characters were removed from the word [Tangut] as duplicates, leaving the two characters [Tangut] by themselves.

Small characters

There are approximately 1856 small characters (localized damage means that in a few places it is not certain how many small characters are present), divided into groups of between 1 and 24 small characters under 577 large characters. One of the small characters ([Tangut]) is placed to the right of the large character [Tangut] on the first column of folio 12A, rather than with the other four small characters under this character, and so was probably added in as an afterthought.

According to our current understanding of Tangut phonology, based on Tangut philological texts such as *Sea of Writing* and *Homophones*, the small characters in each group are all homophones or near homophones of each other and of the large character under which they are listed.

As the large characters under which homophone groups are listed are not ordered phonetically, there is no phonetic sequence to the groups of homophone characters, with the result that the homophone groups cover a seemingly random range of syllables. To give an idea of the range of phonetic variation in homophone groups, the three largest homophone groups are shown below.

The large character [illegible] ^{1}rer$_4$ (IX 1.74) on folio 6A has 24 small characters with five readings:

^{1}rer$_4$ (IX 1.74): [illegible]
^{2}rer$_4$ (IX 2.68): [illegible]
^{2}rir$_4$ (IX 2.272): [illegible]
^{1}rur$_4$ (IX 1.76): [illegible]
^{2}len$_4$ (IX 2.37): [illegible]

The large character [illegible] ^{2}vi$_1$ (II 2.7) on folio 9A has 16 small characters with three readings and one character damaged beyond recognition:

^{2}vi$_1$ (II 2.7): [illegible]
^{1}vi$_1$ (II 1.8): [illegible]
^{2}ve$_1$ (II 2.30): [illegible]

The large character [illegible] ^{1}e$_4$ (VIII 1.36) on folio 7A has 16 small characters with eight readings:

^{1}e$_4$ (VIII 1.36): [illegible]
^{2}e$_4$ (VIII 2.33): [illegible]
^{1}i$_3$ (VIII 1.10): [illegible]
^{1}i$_4$ (VIII 1.11): [illegible]
^{2}iq$_4$ (VIII 2.60): [illegible]
^{1}y$_4$ (VIII 1.30): [illegible]
^{2}en$_4$ (VIII 2.37): [illegible]
^{2}er'$_4$ (VIII 2.86): [illegible]

In the following sections, I attempt to make some general observations about the phonetic features of homophone groups.

Circles. Most homophone groups with more than one small character are divided into two subgroups, with a small circle placed between them. For example, the homophone group [illegible] ^{1}rer$_4$ shown above is divided into two subgroups of 11 and 14 characters:

Subgroup 1	
0	[illegible] ^{1}rer$_4$
1	[illegible] ^{2}rir$_4$
2	[illegible] ^{2}rer$_4$
3	[illegible] ^{2}rer$_4$
4	[illegible] ^{1}rur$_4$
5	[illegible] ^{2}rir$_4$
6	[illegible] ^{2}rer$_4$
7	[illegible] ^{2}len$_4$
8	[illegible] ^{2}len$_4$
9	[illegible] ^{2}len$_4$
10	[illegible] ^{2}len$_4$

Subgroup 2	
11	[illegible] ^{1}rer$_4$
12	[illegible] ^{2}rer$_4$
13	[illegible] ^{2}len$_4$
14	[illegible] ^{2}rer$_4$
15	[illegible] ^{2}rer$_4$
16	[illegible] ^{2}rer$_4$
17	[illegible] ^{2}len$_4$
18	[illegible] ^{2}len$_4$
19	[illegible] ^{2}rer$_4$
20	[illegible] ^{2}rer$_4$
21	[illegible] ^{2}len$_4$
22	[illegible] ^{2}rer$_4$
23	[illegible] ^{2}len$_4$
24	[illegible] ^{1}rer$_4$

In some cases, a small circle is placed after the large character in the group and before the small characters in the group, and there is no other small circle in the group. In these cases, the large character belongs to one subgroup and all the small characters belong to a different subgroup. For example, on folio 8B the homophone group under [illegible] ^{1}ma'$_4$ has a small circle between the large character and the following small characters:

Subgroup 1	
0	[illegible] ^{1}ma'$_4$

Subgroup 2	
1	[illegible] ^{1}ma'$_4$
2	[illegible] ^{1}ma'$_4$
3	[illegible] ^{1}ma'$_4$
4	[illegible] ^{2}ma'$_4$
5	[illegible] ^{1}ma'$_4$
6	[illegible] ^{2}ma'$_4$

Where a homophone group is not subdivided by a small circle, or in the few cases where a small circle only occurs after the last small character in the group, we can assume that all characters belong to the same subgroup.

The obvious explanation for dividing homophone groups into two subgroups is that each subgroup contains characters with same tone (Tone 1, level; Tone 2, rising). However, in the majority of homophone groups, both subgroups have a mixture of Tone 1 and Tone 2 characters, so the distinction cannot be tonal unless this text represents a dialect of Tangut with a very different distribution of tones compared with *Sea of Writing*, which seems highly unlikely. There are no other obvious phonetic differences between characters in each subgroup, and both subgroups commonly have characters with the same readings. Therefore, at present I have no explanation for the division of homophone groups into two with a small circle.

A small circle is also frequently placed to the right of large characters, and less commonly to the right of small characters. The meaning of these small circles is also unclear at present.

Tones. Tones are not distinguished within homophone groups. That is to say, in most cases the same group of homophone characters includes a mixture of both Tone 1 (level tone) and Tone 2 (rising tone) characters.

Where a homophone group consists of characters that all have the same tone, there is not normally a corresponding homophone group for characters with the same reading but a different tone. For example, the homophone group of seven characters headed by [Tangut] ^{2}mer$_4$ on folio 14A all have the same reading and tone, but there is no homophone group for ^{1}mer$_4$ (the only character with this reading and tone is [Tangut], which does not occur). Likewise, the homophone group of seven characters headed by [Tangut] ^{1}ma$_4$ on folio 7B all have the same reading and tone, but there is no homophone group for ^{2}ma$_4$ (the only character with this reading and tone is [Tangut], which does not occur).

Initials. Homophone groups generally share the same reconstructed initial, and in most cases where initials do vary within a group they still belong to the same initial class:

- homophone group [Tangut] on folio 8A has characters with d- and t- initials: [Tangut] ^{1}du'$_1$, [Tangut] ^{2}du'$_1$ and [Tangut] ^{1}tu'$_1$ (all Class III);
- homophone group [Tangut] on folio 8B has characters with g- and k- initials: [Tangut] ^{2}gi'$_4$, [Tangut] ^{1}gi'$_4$, [Tangut] ^{1}gi$_4$ and [Tangut] ^{1}ki'$_4$ (all Class V);
- homophone group [Tangut] on folio 8B has characters with ts- and dz- initials: [Tangut] ^{1}tsen$_1$, [Tangut] ^{2}tse$_1$ and [Tangut] ^{1}dzen$_1$ (all Class VI); and
- homophone group [Tangut] on folio 8A has characters with h-, gh-, w- and ø- initials: [Tangut] ^{2}hoq$_1$, [Tangut] ^{1}ghon$_4$, [Tangut] ^{1}woq$_1$ and [Tangut] ^{2}on$_4$ (all Class VIII).

However, there are a few homophone groups that have characters with two or more different initial classes. Some examples are given in the following.

Most Class VIII characters form homophone groups together, but in a few cases characters with Class VIII initials share a homophone group with Class II, Class V or Class IX characters:

- homophone group [Tangut] on folio 9B has characters with ø- (Class VIII) and v- (Class II) initials: [Tangut] ^{2}o$_1$, [Tangut] ^{1}o$_1$, [Tangut] ^{2}u$_1$, [Tangut] ^{1}von$_1$ and [Tangut] ^{1}voq$_2$;
- homophone group [Tangut] on folio 14A has characters with w- (Class VIII) and v- (Class II) initials: [Tangut] ^{1}vi$_3$, [Tangut] ^{2}vi$_3$, [Tangut] ^{1}viq$_3$, [Tangut] ^{2}viq$_3$, [Tangut] ^{1}wyr$_4$, [Tangut] ^{1}wy$_4$ and [Tangut] ^{1}win$_4$;
- homophone group [Tangut] on folio 6A has characters with gh- (Class VIII) and k- (Class V) initials: [Tangut] ^{2}ghaq$_1$, [Tangut] ^{1}kaq$_1$ and [Tangut] ^{2}kaq$_1$; and
- homophone group [Tangut] on folio 7A has characters with ø- (Class VIII) and r- (Class IX) initials: [Tangut] ^{1}rur$_4$, [Tangut] ^{2}rur$_4$ and [Tangut] ^{2}ur$_4$.

Class IV initials, which modern linguists have struggled to understand, form homophone groups either by themselves or with Class III or Class VII characters:

- homophone group [Tangut] on folio 6A has characters with n- (Class IV) and j- (Class VII) initials: [Tangut] ^{1}nen$_2$, [Tangut] ^{2}nen$_2$ and [Tangut] ^{2}je$_2$;
- homophone group [Tangut] on folio 8A has characters with n- (Class IV) and j- (Class VII) initials: [Tangut] ^{1}jon$_3$, [Tangut] ^{1}jwo$_3$, [Tangut] ^{2}jwo$_3$ and [Tangut] ^{1}non$_2$; and
- homophone group [Tangut] on folio 8A has characters with n- (Class IV) and n- (Class III) initials: [Tangut] ^{1}non'$_2$, [Tangut] ^{1}nwu$_1$, [Tangut] ^{1}nu$_1$ (Class III) and [Tangut] ^{2}non$_2$ (Class IV).

There is one homophone group that mixes Class III and Class VI initials:

- homophone group [Tangut] on folio 12B has characters with t- (Class III) and ts- (Class VI) initials: [Tangut] ^{1}tu$_4$, [Tangut] ^{2}tu$_4$ and [Tangut] ^{1}tsu$_4$.

In addition to the above examples, there are a few examples where individual characters have a totally unexpected reading for their homophone group. Some of these may be due to errors made by the editor of the text. For example, [Tangut] ^{2}ni'$_4$ under homophone group [Tangut] ^{2}li$_3$ on folio 13A is probably a mistake for [Tangut] ^{2}liq$_3$; and [Tangut] ^{1}na$_1$ under homophone group [Tangut] ^{1}my$_1$ on folio 8B is probably a mistake for [Tangut] ^{1}myq$_1$.

Rimes. Homophone groups do not always correspond to a single rime group in *Sea of Writing* or belong to the same homophone group in *Homophones*, but frequently include characters from several different but related rimes in *Sea of Writing*.

In most homophone groups all characters have the same grade (1–4), and in those cases where characters in the same homophone group have different grades this is almost always not the only phonetic distinction:

- [Tangut] ^{2}u$_3$ = [Tangut] ^{1}ghu$_4$ (folio 6A);
- [Tangut] ^{2}uq$_1$ = [Tangut] ^{1}on'$_2$ (folio 7B);
- [Tangut] ^{1}gi$_2$ = [Tangut] ^{1}ky'$_4$ (folio 8A);
- [Tangut] ^{1}gu$_1$ = [Tangut] ^{1}gwy$_4$ (folio 14A); and
- [Tangut] ^{1}vi$_3$ = [Tangut] ^{1}wyr$_4$ (folio 14A).

Many homophone groups include one or two characters with a different reconstructed vowel from the other characters in the group, suggesting vowel mergers. Most of the examples show common phonetic variation, but the a ~ i variation in several cases is unusual.

e ~ i
- [illegible] ^{1}me$_{1}$ = [illegible] ^{2}mi$_{1}$ (folio 8A)
- [illegible] ^{2}be$_{1}$ = [illegible] ^{2}bi$_{1}$ (folio 8B)
- [illegible] ^{1}ver$_{1}$ = [illegible] ^{1}vir$_{1}$ (folio 6B)

e ~ i ~ y [ə]
- [illegible] ^{1}e$_{4}$ = [illegible] ^{1}i$_{4}$ = [illegible] ^{1}y$_{4}$ (folio 7A)
- [illegible] ^{1}she$_{3}$ = [illegible] ^{1}shi$_{3}$ = [illegible] ^{2}shy$_{3}$ (folio 7A)

i ~ y [ə]
- [illegible] ^{2}dziq$_{4}$ = [illegible] ^{2}dzyq$_{4}$ (folio 8A)
- [illegible] ^{2}tsi$_{4}$ = [illegible] ^{1}tsy$_{4}$ (folio 8A)
- [illegible] ^{2}ir$_{4}$ = [illegible] ^{2}yr$_{4}$ (folio 8A)
- [illegible] ^{1}syq$_{4}$ = [illegible] ^{1}siq$_{4}$ (folio 13B)

e ~ y [ə]
- [illegible] ^{2}gew$_{4}$ = [illegible] ^{2}gyq$_{4}$ (folio 8B)
- [illegible] ^{1}be$_{1}$ = [illegible] ^{2}by$_{1}$ (folio 13B)

u ~ i
- [illegible] ^{2}lwu$_{4}$ = [illegible] ^{1}lwi$_{4}$ (folio 7B)
- [illegible] ^{2}gu'$_{4}$ = [illegible] ^{1}gi'$_{4}$ (folio 7B)

o ~ u
- [illegible] ^{2}o$_{1}$ = [illegible] ^{2}u$_{1}$ (folio 8B)
- [illegible] ^{2}ror$_{1}$ = [illegible] ^{2}rur$_{1}$ (folio 12A)
- [illegible] ^{1}po$_{1}$ = [illegible] ^{1}pu$_{1}$ (folio 14A)
- [illegible] ^{1}hon$_{1}$ = [illegible] ^{1}hun$_{1}$ (folio 14B)

o ~ y [ə]
- [illegible] ^{2}nwo$_{4}$ = [illegible] ^{1}nwy$_{4}$ (folio 9A)

u ~ y [ə]
- [illegible] ^{2}ngwu$_{1}$ = [illegible] ^{1}ngwy$_{1}$ (folio 7B)
- [illegible] ^{1}phu$_{1}$ = [illegible] ^{1}phy$_{1}$ (folio 11B)

a ~ o
- [illegible] a$_{1}$ = [illegible] ^{1}o$_{1}$ (folio 7B)

a ~ i
- [illegible] ^{2}a$_{4}$ = [illegible] ^{1}an$_{4}$ = [illegible] ^{1}in$_{4}$ (folio 8A)
- [illegible] ^{1}man$_{1}$ = [illegible] ^{1}min$_{1}$ (folio 7B)
- [illegible] ^{1}ga'$_{4}$ = [illegible] ^{1}gin$_{4}$ (folio 8B)
- [illegible] ^{1}phar$_{4}$ = [illegible] ^{1}phi$_{4}$ (folio 8A)

Some homophone groups suggest loss in secondary phonetic distinctions, such as tenseness (-q), nasalization (-n), retroflexion (-r) and final glides (-w), as shown in the following examples:

- [illegible] ^{1}gi$_{4}$ = [illegible] ^{2}gi'$_{4}$ (folio 8B)
- [illegible] lho$_{1}$ = [illegible] ^{2}lhoq$_{1}$ (folio 8A)
- [illegible] ^{2}a$_{4}$ = [illegible] ^{1}an$_{4}$ (folio 8A)
- [illegible] ^{2}kha$_{4}$ = [illegible] ^{1}khan$_{4}$ (folio 8A)
- [illegible] ^{2}kho$_{2}$ = [illegible] ^{1}khon$_{2}$ (folio 8B)
- [illegible] ^{1}phi$_{4}$ = [illegible] ^{1}phin$_{4}$ (folio 8A)
- [illegible] ^{1}jwo$_{3}$ = [illegible] ^{1}jon$_{3}$ (folio 8A)
- [illegible] ^{2}hwo$_{1}$ = [illegible] ^{1}hon$_{1}$ (folio 14B)
- [illegible] ^{2}ka$_{1}$ = [illegible] ^{2}kar$_{1}$ (folio 8A)
- [illegible] ^{1}dza$_{1}$ = [illegible] ^{1}dzar$_{1}$ (folio 8B)
- [illegible] ^{1}dzeq$_{4}$ = [illegible] ^{2}dzen$_{4}$ (folio 8A)
- [illegible] ^{2}gy$_{4}$ = [illegible] ^{2}gyq$_{4}$ = [illegible] ^{2}gew$_{4}$ (folio 8B)
- [illegible] ^{2}dew'$_{1}$ = [illegible] ^{2}di'$_{1}$ (folio 14B)
- [illegible] ^{2}ne$_{1}$ = [illegible] ^{1}ner$_{1}$ = [illegible] ^{2}neq$_{2}$ (folio 8A)

In summary, it can be seen that homophone groups in this text do not closely match phonetic reconstructions based on *Sea of Writing*, *Homophones* and other existing Tangut phonetic texts, and the characters in each homophone group have a much wider and erratic range of phonetic values than we would expect given our current understanding of Tangut phonology. As there is little or no point in grouping together characters that have significantly divergent phonetic properties (such as different initials or different vowels), it seems probable that the characters in each homophone group in this text were considered to be homophones (excluding tonal differences) by the author of the text.

There are two possible explanations for why the homophone groups in this text do not accord with other Tangut phonological texts: either this text represents a different dialect of Tangut from that used by the authors of classic phonetic texts such as *Sea of Writing* and *Homophones*, or this text represents a later stage in the evolution of the Tangut language. I prefer the latter explanation, and suggest that the homophone groups in this text reflect changes that occurred in the Tangut language some time later than the compilation of the *Sea of Writing* and *Homophones*. As such, this text reflects a simplification of the earlier phonetic system, with a reduction in the number of initials, some vowel mergers and some loss of secondary phonetic distinctions such as tenseness, nasalization, retroflexion and final glides. A more comprehensive understanding of the phonetic system underlying the *Homonyms* awaits a detailed study of the complete text.

Declaration of conflicting interests

None declared.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Notes

1. Tangut readings given in this paper are the complex phonetic transcriptions in Marc Miyake's Tangut phonetic database version 1.3.1 (http://www.amritas.com/Tangut/tangutdb-1-3-1.xls). This is a system of phonetic transcription where 'letters and numbers symbolize phonetic distinctions but do not necessarily precisely represent them' (Miyake, 2015). Final -q, -r and -n are not consonants, but respectively indicate tension,

retroflexion and nasalization of the preceding vowel. The final apostrophe indicates an unknown phonetic quality. The 'y' represents a schwa [ə]. Initial numbers 1 and 2 represent the tone (level and rising respectively), and final numbers 1 through 4 represent the four grades.

2. See http://pmgs.kongfz.com/special/534/ and http://www.dbpm.cn/auction/detail.asp?cid=143.
3. IOM Tang. 24 (Inv. No. 2539). See Lǐ and Hán (2005) for a comprehensive study of the *Synonyms*.
4. See IOM Tang. 22/1 (old Inv. No. 620, new Inv. No. 86) folio 4a–4b. The table of 36 initials is transcribed in Lǐ (2006: 33–35).
5. Tang. 334/248. See Endangered Archives Programme, EAP140/1/60 (https://eap.bl.uk/archive-file/EAP140-1-60) image 201. The cursive character for 𗹦 'two' is part of the original Tangut pagination number '62' written on the right of the first column of text on the recto side of sheet 51.

References

Grinstead E (1972) *Analysis of the Tangut Script*. Lund: Studentenlitteratur. (Scandinavian Institute of Asian Studies Monograph Series, No. 10.)

Lǐ F 李范文 (2006) Wuyinqieyun yu Wenhaibaoyun bijiao yanjiu. *Xixia yanjiu* 西夏研究, No. 1. Beijing: Zhongguo shehuikexue chubanshe.

Lǐ F 李范文 and Hán X 韩小忙 (2005) 'Tongyi Yanjiu' 同义研究. *Xixia yanjiu* 西夏研究, No. 1. Beijing: Zhongguo shehuikexue chubanshe.

Miyake M (2015) Tangut Phonetic Database Version 1.0. Amaravati: Abode of Amritas. Available at: http://www.amritas.com/150207.htm#02012357 (accessed 23 April 2018).

West A (2015) Preliminary analysis of a newly-discovered Tangut wordbook. Available at: http://www.babelstone.co.uk/Blog/2015/08/preliminary-analysis.html (accessed 23 April 2018).