

CORPUS TEXTORUM TANGUTORUM

Volume 1

Gerard Clauson's
Skeleton Tangut (Hsi Hsia) Dictionary

A facsimile edition

With an Introduction by

Imre Galambos

With Editorial notes and an Index by

Andrew West

Prepared for publication by

Michael Everson



2016

Published by Evertime, 73 Woodgrove, Portlaoise, R32 ENP6, Ireland. www.evertime.com.

This edition © 2016 Michael Everson.
Introduction © 2016 Imre Galambos.
Editorial notes © 2016 Andrew West.
Dictionary © 1974 Estate of Gerard Clauson.

First edition 2016.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission in writing of the Publisher, or as expressly permitted by law, or under terms agreed with the appropriate reprographics rights organization.

A catalogue record for this book is available from the British Library.

ISBN-10 1-78201-167-6
ISBN-13 978-1-78201-167-5

Typeset in Baskerville by Michael Everson.

Cover design by Michael Everson. The cover image is folio 53b of the *Homophones* (*Tóngyīn* 同音) edition B collected from Khara-khoto by Aurel Stein in 1914. © The British Library Board. Or.12380/3116.

The frontispiece photograph of Gerard Clauson was taken in June 1945 by Walter Stoneman (1876–1958). © National Portrait Gallery, London.

Printed and bound by LightningSource.

EDITORIAL NOTES

Organization

Clauson's *Skeleton Tangut Dictionary* consists of two large notebooks. Part I (pp. 1–446 in this edition) originally comprised 222 folios: an originally unnumbered folio headed “Table of number of characters in the ‘Homophones’” followed by 221 folios numbered 1 through 221. A blank folio and a folio for Clauson's introductory note were inserted at the beginning of the notebook at a later date, presumably in 1963 when he wrote the introductory note, and the folio with the introductory note and the following folio with the table of number of characters were numbered as i and ii.⁵⁰ Part II (pp. 447–892 in this edition) comprises 222 folios numbered 222 through 441, with folios 284A and 284B, two folios numbered 315 (numbered 315A and 315B in this edition), two folios numbered 321 (321A and 321B here), and no folio 323, as well as an extra sheet inserted after folio 247 (the two being numbered 247A and 247B here). Folios originally numbered 352–370 in ink (preceded by 350, and followed by 370) were renumbered as 351–369 in faint pencil; in the headers the pencil numbers have been used. The dictionary entries are written on the recto side of each folio only; on the verso of each folio, supplementary data relevant to the following recto is given.

Clauson wrote out the draft of his dictionary into these two notebooks between October 1938 (note in top margin of *f.* 1r) and 27 March 1939 (note on *f.* 441r). Six months is surely far too short a period of time for him to have written the dictionary from scratch, and the fact that the entries appear to have been written out sequentially in component order with no gaps between them strongly suggests that there must have been earlier rough drafts or a set of index cards which allowed him to order the characters to be included in his dictionary before he wrote the entries for them into his notebooks. As discussed below, the major sources for Tangut characters used by Clauson were published in 1935 (and none of his sources are later than this date), so it is likely that he did preliminary work on the compilation and ordering of characters for two or three years before he started to write up the actual dictionary entries in October 1938.

The system for ordering Tangut characters that Clauson devised for his dictionary uses a set of components that occur on the top or left side of characters, similar to the radicals used in ordering Chinese characters. Characters sharing the same top or left side component are grouped together as described in his introductory note dated 29 July 1963 (see pp. xvi–xvii above and p. 1 below). Clauson lists 70 main component types: 12 covers, 16 verticals, 12 rectangles, 12 curves, and 18 miscellaneous. These components are listed in the table below.

50 The unnumbered blank and folio i both have 24 rules, while the rest of the pages in the volume have 25. Both sheets are somewhat cleaner than the other sheets, as well.

Components used for ordering characters in Clauson's dictionary

Nº	Component type	Entries	Actual components
1	一	1-522	一, 二, 三, 四 [十], 五, 六, 七, 八, 九, 十, 十一, 十二, 十三, 十四, 十五, 十六, 十七, 十八, 十九, 二十
2	廿	523-536	廿, 卅, 卌, 卍
3	卅	537-558	卅, 卌, 卍
4	卌	559-689	卌, 卍, 华, 协
5	卍	690-695	卍
6	华	696-1222	华, 协, 卐, 卑, 卒, 卓, 協, 单
7	卑	1223-1264	卑, 卒, 卓
8	卒	1265-1282	卒
9	卓	1283-1293	卓
10	協	1294	協
11	单	1295-1302	单
12	卖	1303-1316	卖
13	南	1317-1519	南 (単, 卙, 博, 卛, 卜, 卝, 卞, 卟, 占, 卡, 卢, 卣, 卤, 卥, 卦, 卧, 卨, 卩, 卫, 卬, 卭, 卮, 卯, 危, 卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿, 卺, 卻, 卼, 卽, 卿)
14	単	1520-1656	単 [卙], 博, 卛, 卜, 卝, 卞, 卟, 占, 卡, 卢, 卣, 卤, 卥, 卦, 卧, 卨, 卩, 卫, 卬, 卭, 卮, 卯, 危, 卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿
15	卙	1657-1736	卙, 博, 卛, 卜
16	博	1737-1749	博, 卛, 卜
17	卛 + 卜	1750-1823	卛, 卜, 卝, 卞, 卟, 占, 卡, 卢, 卣, 卤, 卥, 卦, 卧, 卨, 卩, 卫, 卬, 卭, 卮, 卯, 危, 卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿
18	卜 + 卝	1824-1868	卜, 卝, 卞, 卟, 占, 卡, 卢, 卣, 卤, 卥, 卦, 卧, 卨, 卩, 卫, 卬, 卭, 卮, 卯, 危, 卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿
19	卝	1869-1892	卝, 卞
20	卞	1893	卞
21	卟	1894-1899	卟, 占
22	占	1900-2241	占 [卡] (卢, 卣, 卤, 卥, 卦, 卧, 卨, 卩, 卫, 卬, 卭, 卮, 卯, 危, 卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿)
23	卡	2242-2250	卡
24	卢	2251-2333	卢, 卣, 卤, 卥, 卦, 卧, 卨, 卩, 卫, 卬, 卭, 卮, 卯, 危, 卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿
25	卣	2334-2347	卣, 卤, 卥
26	卤	2348-2382	卤, 卥, 卦, 卧
27	卥	2383-2410	卥, 卦, 卧, 卨, 卩
28	卦	2411-2417	卦
29	卧 + 卨	2418-2727	卧, 卨, 卩, 卫 [卬], 卭 [卮], 卯, 危, 卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿
30	卩	2728	卩
31	卫	2729-2733	卫, 卬
32	卬	2734-2738	卬
33	卭	2739-2923	卭 (卮), 卯, 危 [卲], 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿
34	卮	2924-2927	卮
35	卯 + 危	2928-3143	卯 (危, 卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿), 危 (卲, 即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿), 卲 (即, 却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿), 即 (却, 卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿), 却 (卵, 卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿), 卵 (卶, 卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿), 卶 (卷, 卸, 卹, 卺, 卻, 卼, 卽, 卿), 卷 (卸, 卹, 卺, 卻, 卼, 卽, 卿), 卸 (卹, 卺, 卻, 卼, 卽, 卿), 卹 (卺, 卻, 卼, 卽, 卿), 卺 (卻, 卼, 卽, 卿), 卻 (卼, 卽, 卿), 卼 (卽, 卿), 卽 (卿)

EDITORIAL NOTES

36	𠄎	3144–3169	𠄎 (𠄎), 𠄎
37	𠄎	3170–3174	𠄎
38	𠄎	3175–3186	𠄎, 𠄎, 𠄎, 𠄎, 𠄎, 𠄎
39	巾	3187–3211	巾, 巾, 巾, 巾, 巾, 巾, 巾
40	巾	3212–3244	巾, 巾, 巾
41	爻	3245–3254	爻, 爻, 爻
42	爻	3255–3256	爻
43	爻	3257–3648	爻 (𠄎, 𠄎, 𠄎, 𠄎), 𠄎, 爻, 爻 (𠄎), 爻, 爻, 爻 [爻] (𠄎, 爻, 𠄎), 爻 [爻], 爻, 爻, 爻, 爻 (𠄎), 𠄎, 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻
44	爻	3649–3661	爻
45	爻	3662	爻
46	爻	3663–3664	爻
47	爻	3665–4620	爻 (𠄎, 𠄎, 𠄎), 𠄎, 爻, 爻, 爻, 爻 [爻], 爻, 爻, 爻, 爻, 爻, 爻, 爻, 爻
48	爻 + 爻	4621–4704	爻
49	爻	4705–4711	爻
50	爻	4712–4841	爻 (𠄎, 𠄎), 𠄎, 爻, 爻, 爻, 爻, 爻, 爻
51	爻 + 爻	4842–4882	爻, 爻, 爻, 爻, 爻, 爻
52	爻	4883–4898	爻, 爻, 爻
53	爻	4899–5083	爻 (𠄎, 𠄎, 𠄎, 𠄎, 𠄎, 𠄎, 𠄎, 𠄎, 𠄎), 𠄎
54	爻	5084–5097	爻
55	爻	5098–5266	爻, 爻
56	爻	5267–5422	爻 (𠄎, 爻), 爻, 爻, 爻 [爻], 爻 [爻], 爻, 爻, 爻, 爻, 爻, 爻
57	巾	5423	巾
58	𠄎	5424–5534	𠄎 (𠄎, 𠄎, 𠄎, 𠄎)
59	𠄎	5535–5561	𠄎, 𠄎
60	𠄎	5562–5569	𠄎
61	𠄎	5570–5607	𠄎, 𠄎, 𠄎
62	𠄎	5608–5619	𠄎, 𠄎
63	𠄎	5620–5646	𠄎, 𠄎, 𠄎, 𠄎
64	𠄎	5647–5660	𠄎, 𠄎, 𠄎, 𠄎
65	𠄎	5661–5663	𠄎
66	𠄎	5664–5716	𠄎, 𠄎, 𠄎, 𠄎, 𠄎, 𠄎, 𠄎, 𠄎
67	𠄎	5717–5718	𠄎
68	𠄎	5719–5721	𠄎
69	𠄎	5722–5723	𠄎
70	𠄎	5724	𠄎

There are a few anomalies in the component types listed in Clauson's notes:

Component type 14 (|=). In his introductory note, Clauson inserts |= after the component |, but this does not seem to be a concrete component type. It is perhaps intended to

represent components 𐰇 through 𐰇 where the vertical stroke is the leftmost part of a more complex component.

Component type 19 (𐰇). The component 𐰇 is not found as a component on the left side of any Tangut character, and I think it must be a mistake for the component 𐰇 or 𐰇, as there are 24 characters with either of these components at the end of the entries for characters with the preceding component type (𐰇 or 𐰇).

Component type 30 (𐰇). The component 𐰇 only corresponds to a single character in the dictionary (𐰇), which does not include the component as written by Clauson.

Component type 62 (𐰇). The component 𐰇 corresponds to characters with 𐰇 or 𐰇 on their left, and it would seem to be a mistake.

Each component type corresponds to one or more actual components, as listed in the above table. In some cases, the actual components include additional interpolated components. For example, there are 53 characters corresponding to the component 𐰇 (5267–5316), but these include three characters with 𐰇 on the left (5286–5288), and one character with 𐰇 on the left (5306). In the table above, such interpolated components are given in parentheses.

The component ordering method used by Clauson is well-designed, and it prefigures the various radical systems used in all later works of Tangutological scholarship. However, there are a few problems caused by Clauson's incomplete knowledge of the correct glyph forms of Tangut characters. Clauson carefully distinguishes some minor glyph differences, such as those between 𐰇 and 𐰇 and between 𐰇 and 𐰇, which are not always distinguished in the works of later scholars. Unfortunately, some significant glyph differences that should be distinguished are not. The most serious issue is that Clauson treats the elements 𐰇 and 𐰇 as being the same, both when they occur as top components and when they occur as elements of left side components (such as 𐰇 and 𐰇) and elements of right side components (such as 𐰇 and 𐰇). Clauson uniformly writes 𐰇 as 𐰇, and only distinguishes them in a few cases where two characters are otherwise identical (e.g. 2538 𐰇 and 2539 𐰇). Another pair of components that Clauson conflates together are 𐰇 and 𐰇, which are both written as 𐰇. In the table above, a component that has been conflated with another component by Clauson is given in square brackets after the component form used by Clauson for both components.

Numbering

Main entries in the dictionary were numbered sequentially from 1 through 5724 in the first draft completed on 27 March 1939. In addition to these main entries, the first draft included about 200 subsidiary entries for characters that Clauson considered to be uncertain or mistaken. These subsidiary entries are mixed in with the main entries, but are not given a unique sequential number. Instead, a subsidiary entry has the same number as the preceding main entry, followed by the letter A. Where there are two or three subsidiary entries in succession, then they share the same number, followed by the letter A, B, or C. For example, on *f.* 64r the main entry 849 𐰇 “cork” is followed by the subsidiary entries 849A (noted as an error for 1240 𐰇), 849B (noted as an error for 1242 𐰇), and 849C (noted as doubtful, possibly an error for 857 𐰇, although actually it is a miswritten form of 𐰇 which is unclear in the original *Homophones* text).

Clauson continued to work on the dictionary sporadically for many years after the completion of the first draft in 1939. He notes on the final page that he finished cross-referencing entries on

14 September 1946, and his final introductory note was added on 29 July 1963. During this time Clauson made many corrections, additions, and deletions to the original entries, and renumbered many entries to account for additions and deletions. The draft in its final form consists of main entries numbered 1 through 5724, with two entries for 1438 (1438 𠄎 is the last entry on *f.* 108r, and 1438 𠄎—given here as 1438* in the header—is the first entry on *f.* 109r), and no entries for seven numbers:

- 224 is crossed out as a duplicate of 179 𠄎;
- 2388 is missing;
- 3363 𠄎 is crossed out as a mistake for 4688 (3363 is actually the correct form);
- 3833 is crossed out as a mistake for 3764bis 𠄎;
- 4689 (first entry on *f.* 362r) is crossed out as a duplicate of 4688 𠄎 (last entry on *f.* 361r);
- 5018 is crossed out as a mistake for 3423 𠄎;
- 5688 (mistake for 𠄎) is renumbered as 5674A.

Additional main entries, added in the upper or lower margin, take the number of the preceding entry suffixed by *bis*. There are seventeen *bis* entries, of which sixteen are additions in the margins (267bis 𠄎, 330bis 𠄎, 781bis 𠄎, 832bis 𠄎, 962bis 𠄎, 1473bis 𠄎, 2083bis 𠄎, 2203bis 𠄎, 3271bis 𠄎, 3764bis 𠄎, 3782bis [mistakenly written as 3872bis] 𠄎, 4269bis 𠄎, 4581bis 𠄎, 4939bis 𠄎, 5103bis 𠄎, and 5202bis 𠄎), while one entry (695bis 𠄎) is an original main entry to which Clauson added a *bis* suffix at a later date because renumbering had inadvertently resulted in two entries numbered 695. There is also one deleted *bis* entry on *f.* 14r (158bis, which is a duplicate of 422 𠄎). Most of the *bis* entries are for variant forms of characters listed elsewhere in the dictionary, but seven of the entries are for main characters in the *Homophones* that Clauson must have missed the first time round. There are also a few inserted main entries with a unique number due to renumbering of adjacent entries (e.g. 771 𠄎 which is inserted in the bottom margin of *f.* 59r).

During Clauson's revisions he added about 56 additional subsidiary entries suffixed A or B in the margin or on the facing page. He also demoted a few main entries by adding an A suffix and renumbering adjacent entries (e.g. 138 was renumbered as 136A, and the reader directed to the correct form of the character at 611 𠄎). In the final draft there are a total of 258 subsidiary entries:

- 237 numbers suffixed A;
- 2 numbers suffixed AA (5306AA before 5306A; and 4545AA after 4545A);
- 18 numbers suffixed B;
- 1 number suffixed C (849C).

In addition to the numbered entries, there are also seven unnumbered entries in the top margin, which are all errors for characters listed elsewhere:

- f.* 57r: mistake for 738 𠄎;
- f.* 82r: mistake for 1060 𠄎;
- f.* 89r: mistake for 685 𠄎 (Clauson suggests it is a mistake for 1222 𠄎, but the meaning *qiān* 慳 “avarice” indicates that 685 is correct);

f. 172r: mistake for 2123 𠄎;

f. 304r: mistake for 4649 𠄎;

f. 427r: mistake for 5547 𠄎 or 5578 𠄎 (neither character fits the meaning, *xīn* 辛 “pungent”, although Li Fànwén gives the character 𠄎 in the same source quotation);⁵¹

f. 436r: mistake for 5701 𠄎.

As he revised the dictionary, Clauson crossed out quite a few entries which are duplicates or mistakes for characters listed elsewhere, and with one exception these deleted entries are not included in the radical index. The one exception is 779A 𠄎, which was added in the upper margin on f. 59r. Clauson glosses this character as “chisel” based on *Handful of Pearls*, but he crosses out the entry, noting that it is an error for 798 𠄎, which he glosses as “a chisel, punch”. In fact, the definition of 798 is wrong (it means “to hold”), and despite only differing by one stroke, it is not the same character as 779A.

In total Clauson's dictionary comprises 5,993 numbered entries (5,718 main numbered entries, 17 *bis* entries, 237 A entries, 18 B entries, 1 C entry, and 2 AA entries). These represent a total of approximately 5,744 distinct Tangut characters (including the one crossed-out entry 779A for 𠄎).

The numbered dictionary entries are written on the recto side of each folio. The verso side of each folio is used to list pairs and chains of characters for the entries on the facing recto page, as well as supplementary information for which there was no room on the recto page.

Pairs of characters are two characters that mutually gloss each other in the *Homophones*. So for example, on the page facing folio 1r (f. ii v) is a note for 10 that states “0010 5342 pair”, which means that the *Homophones* entry for 10 𠄎 “foolish” (*H.* 54B13)⁵² has the gloss character 𠄎 (5342), and the *Homophones* entry for 5342 𠄎 “foolish” (*H.* 10B24) has the gloss character 𠄎 (10). In fact, unknown to Clauson at the time he wrote his dictionary, the gloss character is not a gloss *per se*, but should be read together with the main character to form a single word, starting with the gloss character if it is below right of the main character (as is the case for 𠄎), or starting with the main character if the gloss character is below left of it (as is the case for 𠄎). Thus, in this case the entries for 𠄎 and 𠄎 in the *Homophones* both give the word 𠄎𠄎 meaning “foolish”.

Chains of characters are a chained list of main characters and corresponding gloss character in the *Homophones*. As an example, on the facing page of folio 2 is the following chain of characters for 22 𠄎 “snake”:

0022	
2444	3661
1063	1168
	2056
3036	
3666	
4841	
5117	5121

51 Li Fànwén 李範文, *Xià-Hàn zìdiǎn* 夏漢字典 (Beijing: Zhōngguó shèhuì kēxué chūbǎnshè, 2008) entry #1079.

52 References with *H.* in this format are to the A edition of the *Homophones*.

This means that:

- 22 𪛗 is a gloss character for 2444 𪛗 “frog” (*H.* 9A45)
 2444 𪛗 is a gloss character for 3661 𪛗 “turtle” (*H.* 48A77)
- 22 𪛗 is a gloss character for 1063 𪛗 “dragon” (*H.* 9B37)
 1063 𪛗 is a gloss character for 1168 𪛗 “dragon” (*H.* 52B42) and for 2056 𪛗
 “dragon” (*H.* 27A41)
- 22 𪛗 is a gloss character for 3036 𪛗 “to crawl” (*H.* 14B17)
- 22 𪛗 is a gloss character for 3666 𪛗 “snake” (*H.* 46A64)
- 22 𪛗 is a gloss character for 4841 𪛗 “snake” (*H.* 26B47)
- 22 𪛗 is a gloss character for 5117 𪛗 “fish” (*H.* 52B25)
 5117 𪛗 is a gloss character for 5121 𪛗 “tortoise” (*H.* 31B56)

Sources

Clauson’s sources for his dictionary were of necessity very limited, and with no access to the Tangut material from Khara-khoto that was held in Leningrad (or indeed, it seems, to the Tangut material brought back by Aurel Stein, and held at the British Museum), he had to rely on a very few published editions and studies of Tangut texts that had been published during the 1920s and 1930s. The major sources used by Clauson are listed below.

Xīxià Guóshū Zìdiǎn Yīntóng (西夏國書字典音同). [Lüshun], 1935. This is a lithographic edition of the *Homophones* published in China by Luó Fúchéng (羅福成, 1885–1960), eldest son of Luó Zhènyù (羅振玉, 1866–1940). It reproduces a hand-copied facsimile made by Luó Fúchéng from photographs of a woodblock edition held at the Asiatic Museum in Leningrad. *Homophones* (Tangut title: 𪛗𪛗) is the largest surviving source of Tangut characters, listing nearly 5,800 Tangut characters by rime, with all characters sharing the same pronunciation grouped together. Each entry consists of a single Tangut character with one or two small-sized gloss characters beneath it. There are two known editions of *Homophones*: Edition A, which only survives as a single, incomplete copy (missing most of folios 37B and 38A, and parts of folios 54A through 56B); and Edition B, which survives as several incomplete copies of at least two printings from different printing blocks. The 1935 edition used by Clauson represents the A edition of *Homophones*.

Fān-Hàn Héshí Zhǎngzhōngzhū 番漢合時掌中珠. [Tianjin]: Yí’āntáng jīngjīpū (貽安堂經籍鋪), 1935. This is a lithographic edition of the Chinese-Tangut glossary, *Handful of Pearls*, published in 1935 by Luó Fúchéng. It is not the same as the lithographic edition of *Handful of Pearls* that was published by Luó Fúchéng in 1924. The earlier edition reproduces a hand-copied facsimile made by his father, Luó Zhènyù, from photographs of a woodblock edition held at the Asiatic Museum in Leningrad, and provided to Luó in 1922 by Aleksei Ivanovich Ivanov (Алексе́й Ива́нович Ива́нов, 1878–1937). However, the 1935 edition states that the text was hand-copied by Nikolai

Nevsky, and it includes an appended table of corrections not found in the 1924 edition. There are some significant differences between the text of the 1924 and 1935 editions, and in some cases (as discussed below) the text of the 1935 edition is inferior to that of the 1924 edition. Clauson's copy is now held at SOAS library (CL495.49 / 84332). *Handful of Pearls* (Tangut title: 級級形繼繼形辭絞) is the only surviving bilingual Chinese-Tangut glossary, and was the main key to the decipherment of the Tangut language. There are several incomplete copies of *Handful of Pearls*, representing two slightly different editions, and the 1924 and 1935 editions are both composites of several different copies. In Clauson's dictionary *Handful of Pearls* is referenced as "C.C.C." (standing for the Wade-Giles transliteration *Chang Chung Chu* 掌中珠).

Bulletin of the National Library of Peiping (國立北平圖書館館報) vol. 4 May–June 1930 (issued January 1932) No. 3: "A volume on Tangut (Hsi Hsia) Studies" (西夏文專號). This is a collection of articles by Wáng Jìngrú (王靜如, 1903–1990) and others on Tangut texts discovered at Khara-khoto and elsewhere in China. In Clauson's dictionary this work is referenced as "B".

Luó Fúcháng (羅福菴, 1896–1921), *Xīxià Guóshū Lüèshuō* (西夏國書略說). Kyoto: Higashiyama gakusha (東山學社), 1914. Introductory handbook on Tangut script written by Luó Fúcháng, second son of Luó Zhènyù. Clauson's copy is now held at SOAS library (CL495.49 / 84331). In Clauson's dictionary this work is referenced as "Lo Handbook".

Nicolas Nevsky, "A Brief Manual of the Si-Hia Characters with Tibetan Transcriptions" (西藏文字對照西夏文字抄覽); *Research Review of the Osaka Asiatic Society* No. 4 March 15, 1926. This is a list compiled by Nikolai Aleksandrovich Nevsky (Никола́й Алекса́ндрович Не́вский, 1892–1937) of 334 Tangut characters with corresponding Tibetan phonetic glosses. In Clauson's dictionary this work is referenced as "T".

Wáng Jìngrú (王靜如, 1903–1990), *Hsi Hsia Studies* (*Xīxià yánjiū* 西夏研究). Beijing: 1932. This is collection of articles on Tangut studies by Wáng Jìngrú, in three volumes. Clauson's copy is now held at SOAS (CL495.49 / 28564). In Clauson's dictionary this work is referenced as "W. I", "W. II" and "W. III" for vols. I through III respectively.

In his latest revisions of the dictionary in the early 1960s, Clauson also adds a few references in blue ballpoint to Nevsky's draft manuscript Tangut dictionary that had been published posthumously in 1960 under the title *Тангутская Филология: Исследования и Словарь* (*Tangut Philology: Research and Dictionary*). However, this work was not used as a source for any additional entries to Clauson's dictionary. Actually, Nevsky's dictionary, with 5,237 entries, only includes about sixty distinct characters that are not present in Clauson's dictionary, of which about half are from the B edition of *Homophones* that are missing in the extant A edition.

The two major primary sources for Clauson's dictionary were the facsimile editions of the *Homophones* and the *Handful of Pearls*. However, as these were both based on hand-written copies of the original xylograph editions rather than being photographic reproductions, they were not entirely reliable. Characters that were unclear or damaged in the original text were commonly written incorrectly in the hand-written facsimile, and even clearly-written characters were sometimes miswritten in the facsimile edition. This means that, through no fault of his own, the character forms given in Clauson's dictionary are not always exactly correct, although in most

cases the correct form is not difficult to determine. Another issue with his sources is that the character forms used in *Homophones* and *Handful of Pearls* sometimes differ from what modern scholars consider to be the standard forms, so even where a character is clearly written in the original xylographic edition, correctly copied in the hand-written facsimile, and carefully transcribed by Clauson, it may not exactly match the form given in modern Tangut dictionaries. As an example, Clauson's entry 275 𗞑 faithfully reflects the character form given in *Homophones A* folio 16A53. The form of this character in modern dictionaries is 𗞑 (e.g. *Lǐ Fànwén* 2008 #4849), which reflects the form given in *Homophones B* folio 17A33.

The *Homophones* is the main source for the repertoire of Tangut characters listed as main entries in Clauson's dictionary, and most main entries start with a reference to the position of the character in the *Homophones*, and list the gloss character(s) given under the character in the *Homophones*. The A edition available to Clauson would originally have listed between 5,778 and 5,784 main characters, but between 133 and 139 characters are partially or wholly missing in the extant copy (see Clauson's summary at the start of the dictionary). Clauson includes all of the 5,645 surviving complete main characters, as well as two partial characters (4363 𗞑 which is the left side of 𗞑, partially legible on *H.* folio 54A; and 0821 (f. 62r) which is a mistaken reconstruction of 𗞑, partially preserved on *H.* folio 38A). These characters form the bulk of the 5,718 main entries in Clauson's dictionary, although seven were added later as *bis* entries (781bis 𗞑, 962bis 𗞑, 3271bis 𗞑, 3764bis 𗞑, 4269bis 𗞑, 4939bis 𗞑, and 2203bis 𗞑), and three are subsidiary entries: 3195A 𗞑 (which Clauson mistakenly considered to be a misspelling of 3192 𗞑), 3991A 𗞑, and 5674A 𗞑 (renumbered from 5688).

There are about 80 main entries in Clauson's dictionary which do not have a reference to a main character in the *Homophones*, most of which are found in *Handful of Pearls*, or occur as gloss characters in the *Homophones*, or are sourced to *Bulletin of the National Library of Peiping* or *Hsi Hsia Studies*. Some of these characters are actually variants of main characters in the *Homophones* that are listed under a different entry (e.g. 1189 𗞑 = 675 𗞑, and 1748 𗞑 = 3250 𗞑). Eight main entries give no source and no details of their usage or meaning, and it is not clear where Clauson found them:

- 595 𗞑: occurs as a gloss character for 679 𗞑 in the *Homophones*.
- 1205 𗞑: occurs as a phonetic gloss character in *Handful of Pearls*.
- 1247 𗞑: a rare transliteration character that does not occur in *Handful of Pearls* or as a main character in the *Homophones*. The original entry for 1247, which was later crossed out, gives the same glyph, but defines it as a main character in *Homophones* glossed with 4561 𗞑, which would actually be 1230 𗞑.
- 4344 𗞑: occurs as a phonetic gloss character in *Handful of Pearls*.
- 4706 𗞑: occurs as a phonetic gloss character in *Handful of Pearls*.
- 5272 𗞑: This is an error for 5292 𗞑, which occurs as a phonetic gloss in *Handful of Pearls* folio 21b. The character is written correctly in the original text and in the 1924 lithographic edition, but 𗞑 has been miswritten as 𗞑 in the 1935 lithographic edition used by Clauson.
- 5276 𗞑: Clauson does not give a source or meaning for this character, but Nishida Tatsuo gives this character in his Tangut dictionary, where it is glossed as “lamp wick” *dēngxīn* 燈心, and sourced to *Handful of Pearls* folio 23 6th column (i.e. last

column of folio 23b).⁵³ This character does not occur in *Handful of Pearls*, but the penultimate entry on folio 23b gives the Tangut word 蕪翹 “grass wick” against the Chinese word *dēngcǎo* 燈草 “lamp wick”, and the unknown character 蕪 is composed from the left side of 翹 “grass” and the right side of 蕪 “wick”. The 1924 lithographic edition of *Handful of Pearls* correctly shows 蕪翹 for “lamp wick”, but in the 1935 lithographic edition this word is miswritten as 蕪蕪. Therefore, 5276 蕪 is a mistake for 5271 翹 in both Clauson’s and Nishida’s dictionaries, due to them both relying on the 1935 edition of the *Handful of Pearls*.

5426 蕪: occurs as a phonetic gloss character in *Handful of Pearls*.

As to the 258 subsidiary entries in Clauson’s dictionary, most are variant, mistaken or corrupt forms of characters listed elsewhere in the dictionary, but a significant number are gloss characters attested once in the *Homophones*. Some of these are variant or mistaken forms of a character with a main entry in the dictionary, but 35 are distinct characters that are not listed elsewhere (13A 𠄎, 13B 𠄏, 92A 𠄐, 359A 𠄑, 629A 蕪, 710A 蕪, 792A 蕪, 977B 蕪, 1009A 蕪, 1029A 蕪, 1054A 蕪, 1546A 蕪, 1550A 蕪, 2026A 蕪, 2118A 蕪, 2123A 蕪, 2123B 蕪, 2133A 蕪, 2380A 蕪, 2745A 蕪, 2888A 蕪, 3316A 蕪, 3386A 蕪, 3519A 蕪, 3697A 蕪, 4304A 蕪, 4308A 蕪, 4502B 蕪, 4510A 蕪, 4610A 蕪, 4620A 蕪, 4712A 蕪, 4788A 蕪, 5174B 蕪, and 5345A 蕪), and probably correspond to a main character that was in the damaged part of the A edition of *Homophones*. Characters that were used as gloss characters in at least two place in the *Homophones*, or which were also attested in another source, are usually given as main entries by Clauson, but in most cases where a character was only attested in a single place as a gloss character he errs on the side of caution, and gives it as subsidiary entry. The only other subsidiary entries that are for known characters not listed elsewhere in the dictionary are: 677A 蕪 and 1164A 蕪, which are both transliteration characters that Clauson sourced to secondary sources; and 4235B 蕪, which Clauson mistakenly considered to be the correct form of 4176 蕪, when they are actually two distinct characters.

For most subsidiary entries for variant, corrupt, or mistaken character forms, Clauson provides a cross-reference to the main entry with the correct character form, although occasionally the character form in the subsidiary entry is actually the correct form. For example, 4779A 蕪 is noted as an error for 3315, when in fact 4779A is the correct form, and the form given under 3315 has an incorrect left side component (𠄎 instead of 𠄏). Some 52 subsidiary entries for characters with incorrect or variant glyphs either have no cross-reference to the correct character, or the cross-reference is doubtful. These are listed below.

437A: Error for 442 蕪. Gloss character for 3191 蕪 in the *Homophones* (H. 4A17).

471A: Error for 464 蕪. Gloss character for 4432 蕪 in the *Homophones* (H. 50B53).

596A: Error for 528 蕪. Gloss character for 947 蕪 in the *Homophones* (H. 44A13).

686A: Unidentified error in Nevsky’s *Manual* (#331 *dbi*’).

735A: Error for 872 蕪. Clauson’s misinterpretation of the character as given in *Bulletin of the National Library of Peiping*.

849C: Error for 730 蕪. Gloss character for 726 蕪 in the *Homophones* (H. 41B28).

870A: Probably an error for 926 蕪. Sourced to *Hsi Hsia Studies*.

889A: Unidentified error in Nevsky’s *Manual* (#317 *ko*).

53 Nishida Tatsuo 西田龍雄 (1928–2012), *Seikabun Shōjiten* 西夏文小字典 (Tokyo, 1966) entry 185-065.

INTRODUCTION

- 1234A: Unidentified error in Nevsky's *Manual* (#318 *g.yi*).
 1353A: Error for 1487 𣎵 in *Bulletin of the National Library of Peiping*.
 1383A: Error for 1401 𣎵. Gloss character for 3284 𣎵 in the *Homophones* (H. 5A15).
 1425A: Possibly an error for 1455 𣎵 (as a transliteration of Chinese *guāng* 光). Sourced to *Hsi Hsia Studies*.
 1598A: Error for 1598 𣎵. Gloss character for 3293 𣎵 in the *Homophones* (H. 6A73).
 1686A: Probably an error for 1691 𣎵. No source given.
 2153A: Error for 2151 𣎵. Gloss character for 2159 𣎵 in the *Homophones* (H. 36A71).
 2428A: Error for 2427 𣎵. Gloss character for 3303 𣎵 in the *Homophones* (H. 6B72).
 2529A: Error for 2439 𣎵 in *Hsi Hsia Studies*.
 3018A: Error for 3107 𣎵 in *Hsi Hsia Studies*.
 3105B: Error for 2339 𣎵. Gloss character for 2159 𣎵 in the *Homophones* (H. 19B17).
 3137A: Error for 3134 𣎵. No source given.
 3147A: Unidentified error in Nevsky's *Manual* (#329 *bla'*).
 3274A: Error for 3281 𣎵. Gloss character for 2571 𣎵 in the *Homophones* (H. 22B57).
 3324A: Error for 3331 𣎵. Gloss character for 617 𣎵 in the *Homophones* (H. 18A18).
 3343A: Probably an error for 3870 𣎵. No source given.
 3574A: Unidentified error in Nevsky's *Manual* (#232 *'dzwa*).
 3575A: Unidentified character in *Bulletin of the National Library of Peiping*. Probably not 3580 𣎵 as suggested by Clauson.
 3614A: Error for 3454 𣎵. Gloss character for 3864 𣎵 in the *Homophones* (H. 40A71).
 3674A: Error for 4216 𣎵. Gloss character for 4673 𣎵 in the *Homophones* (H. 39B58).
 3700A: Error for 3709 𣎵. Gloss character for 329 𣎵 in the *Homophones* (H. 52A78).
 3758A: Unidentified error in Nevsky's *Manual* (#308 *nī*).
 3769A: Clauson glosses this character as *zàng* 藏 “Tibet”, and sources it to Luó Fúcháng's 1914 *Handbook* (*Xīxià Guóshū Lùeshuō* 西夏國書略說). On folio 20a of the *Handbook* there is indeed a list of five languages: Sanskrit (𣎵𣎵 = *fàn yǔ* 梵語), Chinese (𣎵𣎵 = *hàn yǔ* 漢語), Tibetan (𣎵𣎵𣎵 = *xīzàng yǔ* 西藏語), Khitan (𣎵𣎵𣎵 = *qīdān yǔ* 契丹語), and Tibetan (𣎵𣎵 = *fān yǔ* 番語). 3769A 𣎵 occurs as the middle character in the first term for Tibetan, but this term is problematic as Tibet was not known as Xīzàng during the Western Xia period. In fact, Luó Fúcháng's list of languages derives from a fragment of a Tangut manuscript collected by Count Ōtani Kōzui 大谷光瑞 (1876–1948), which lists the following seven languages on one line: Sanskrit (𣎵𣎵), Uighur (𣎵𣎵𣎵), Tangut (𣎵𣎵), Tibetan (𣎵𣎵), Chinese (𣎵𣎵), Khitan (𣎵𣎵𣎵), and Jurchen (𣎵𣎵). Luó's “Tibetan” (𣎵𣎵𣎵) is evidently a misreading of “Uighur” (𣎵𣎵𣎵), so 3769A is an error for 4048 𣎵.⁵⁴
 3884A: Error for 3839 𣎵. Gloss character for 1147 𣎵 in the *Homophones* (H. 10B27).
 3887A: Error for 3846 𣎵 in *Hsi Hsia Studies*.
 3935A: Error for 3284 𣎵. Gloss character for 4977 𣎵 in the *Homophones* (H. 10B18).

54 The Ōtani fragment is held at the library of Ryūkoku University (龍谷大学図書館) in Kyoto, and is reproduced in *Ribēn Cáng Xīxiàwén Wénxiàn* 日本藏西夏文獻 [*Xixia documents collected in Japan*] (Beijing: Zhōnghuá shūjū, 2011), vol. 1: 不明論典 03-03 (the list of languages is on column 3 of folio 8a). I am grateful to Viacheslav Zaytsev for helping to explain this error.

- 3948A: Error for 3702 𪛗. Gloss character for 578 𪛗 in the *Homophones* (H. 51A13).
- 4188A: Error for 4677 𪛗 “hump” [of a camel] in *Bulletin of the National Library of Peiping*.
- 4201A: Error for 4197 𪛗. Gloss character for 2925 𪛗 in the *Homophones* (H. 26B18).
- 4227A: Error for 4132 𪛗. Gloss character for 3922 𪛗 in the *Homophones* (H. 22A44).
- 4235A: Error for 4694 𪛗. Gloss character for 5167 𪛗 in the *Homophones* (H. 40A72).
- 4249A: Error for 4692 𪛗. Gloss character for 125 𪛗 in the *Homophones* (H. 49A48).
- 4311A: Error for 4302 𪛗. Gloss character for 3698 𪛗 in the *Homophones* (H. 23A12).
- 4412A: Error for 4528 𪛗. Gloss character for 4529 𪛗 in the *Homophones* (H. 45A14).
- 4430A: Error for 4494 𪛗. Incomplete gloss character for 3569 𪛗 in the *Homophones* (H. 56A21).
- 4441A: Variant of 4468 𪛗. Gloss character for 4410 𪛗 in the *Homophones* (H. 54A24).
- 4545AA: Probably an error for 4546 𪛗. No source given.
- 4545B: Error for 4555 𪛗. Gloss character for 5496 𪛗 in the *Homophones* (H. 40B23).
- 4571A: Error for 4578 𪛗. Gloss character for 4583 𪛗 in the *Homophones* (H. 46A36).
- 4697A: Error for 4626 𪛗. Gloss character for 4675 𪛗 in the *Homophones* (H. 19A51).
- 4985A: Probably an error for 4899 𪛗 in *Bulletin of the National Library of Peiping*. Second *fǎnqiè* character for 3793 𪛗.
- 5306AA: Error for 5309 𪛗 or 5323 𪛗 in *Bulletin of the National Library of Peiping*.
- 5491A: Error for 5458 𪛗. Gloss character for 5118 𪛗 in the *Homophones* (H. 44A74).
- 5697A: Error for 5694 𪛗 in *Bulletin of the National Library of Peiping*.

Andrew West