



# The Old Khotanese Metanalysis

## Citation

Hitch, Douglas A. 2016. The Old Khotanese Metanalysis. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493273>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

The Old Khotanese Metanalysis

A dissertation presented

by

Douglas A. Hitch

to

The Committee on Inner Asian and Altaic Studies

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Inner Asian and Altaic Studies

Harvard University

Cambridge, Massachusetts

April, 2016



This work is licensed under the  
Creative Commons Attribution 4.0 International License.

To view a copy of this license, visit

<http://creativecommons.org/licenses/by/4.0/>

or send a letter to

Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

## The Old Khotanese Metanalysis

## Abstract

A range of present stems in Khotanese which historically ended in a consonant behave synchronically as vowel final stems. Forms like *bvāre* ‘they know’ or *bvāne* ‘may I know’ imply a synchronic stem *bu-*, but most scholars have preferred a shape *bud-* because the Proto-Iranian antecedent is *\*baud-*. A metanalysis resolves the apparent contradiction between diachrony and synchrony.

Sometime in pre-Khotanese there was a morphophonological reanalysis wherein some consonant final verb stems were reinterpreted as being vowel final, and a series of suffixes historically beginning with single *\*-t-* were reinterpreted as beginning with double */tt/*. A word like 3Sp.m *butte* ‘he knows’ was originally analyzed as *\*bud-* + *-te* and later reanalyzed as *bu-* + *-tte* */βuttē/*. The vowel stems resulting from the metanalysis behave no differently from other vowel stems. All forms arising from the addition of vowel initial suffixes show contractions which are either attested elsewhere in the grammar or are phonologically natural. They show no trace of a final consonant on the synchronic level. Properly identifying the synchronic forms of the affected stems and suffixes enables a more regular and systematic description of the attested forms. It permits a reduction from Emmerick’s four present stem types, A, B, C, D, to just two, A and B.

## Table of Contents

Table of Contents .....	iv
List of Figures .....	viii
Acknowledgments .....	x
Chapter 1: Introduction and Background.....	1
1.0 Overview.....	1
1.2 Method.....	5
1.3 The Brāhmī Script.....	6
1.3.1 Sander’s Paleography .....	7
1.3.2 The New Sign <i>ä</i> .....	9
1.3.3 Other Features .....	9
1.4 Stages within Old Khotanese .....	10
1.5 Symbology.....	12
1.5.1 Orthographic Transliteration and Transcription .....	13
1.5.2 Phonemic Transcription.....	14
1.5.3 Morphological Rebracketing .....	16
1.5.4 Metrical Apparatus .....	16
1.5.5 Unicode fonts and characters.....	19
Chapter 2:Old Khotanese Consonants.....	20
2.1 Inventory .....	20
2.1.1 Emmerick on the consonants.....	21
2.1.2 <i>haththā</i> - ‘truth’ .....	22
2.2.0 Synchronic umlaut: evidence for aspiration, retroflexion and palatality ...	25
2.2.1 Chinese in Khotanese Brāhmī .....	25
2.2.2 Nine key segments in umlaut.....	28
2.2.2.1 AOKh and COKh <i>ky</i> , <i>c</i> , <i>gy</i> , and <i>j</i> .....	30
2.2.3 Synchronic umlaut rule: .....	32
2.2.4 Consonant behavior in umlaut .....	33
2.2.4.1 Palatality.....	34
2.2.4.2 Umlautability .....	34
2.2.4.3 Transparency.....	34
2.3 Retroflexion in Old Khotanese .....	35
2.3.1 Synchronic evidence for retroflexion .....	35
2.3.2 Typology and /ɹ~ɻ/ .....	37
2.3.3 Loanword evidence .....	39
2.3.3.1 History of the Khotanese rhotics .....	39
2.3.4 Diachronic evidence for retroflexion.....	41
2.3.4.1 <i>t</i> .....	41
2.3.4.2 <i>l</i> and <i>d</i> .....	41
2.3.4.3 <i>th(th)</i> .....	42
2.4.0 Retroflex affricate <i>kʂ</i> = /tʂ/ .....	45
2.5.0 Aspirated labiovelar glide <i>hv</i> = /h <sup>w</sup> / .....	49
2.5.0.1 <i>hv</i> ’ ( <i>hv</i> + apostrophe) .....	51
2.5.1 LKh <i>h</i> : in Old Turkic.....	52

2.6.0	Voiced <i>p</i> , <i>b</i> = /b/ .....	55
2.6.1	Distribution of voiced and voiceless stops.....	55
2.6.2	Old Turkic evidence for /b/.....	60
2.6.3	LMC and Tibetan evidence for OKh /b/ .....	64
2.6.4	Typology of inventory gaps .....	67
2.7.0	The Nasals .....	69
2.7.1	Anusvāra <i>ṃ</i> .....	69
2.7.2	Retroflex <i>ṇ</i> /ṇ/.....	70
2.7.2.1	The apostrophe and <i>ṇ</i> .....	71
2.8.0	OKh <i>ḍ</i> , stop or liquid? .....	75
2.8.1	OKh synchronic /ḍ > l/ .....	75
2.8.1.1	Suffix initial /t/ in the 3S of B stems .....	76
2.8.2	Inventory gap .....	79
2.8.3	Late Khotanese liquid <i>ḍ</i> .....	80
2.8.3.1	OT /l/ in LKh Transcription .....	81
2.8.3.2	Palatal <i>ḍ</i> ?.....	82
2.9.0	Meter.....	83
2.9.1	Double <i>tt</i> as /tt/ .....	83
2.9.2	<i>ṭṭṭ</i> and <i>ḍḍ</i> .....	83
2.9.3	<i>rr</i> and <i>r</i> .....	86
2.10.0	Old Khotanese allography.....	86
2.10.1	Dental.....	86
2.10.2	Phonemic contrasts and allography .....	88
2.10.3	Labial.....	89
2.10.4	Velar .....	90
2.10.4.1	Glottal Stop .....	92
2.10.5	The apostrophe and <i>g'~t'</i> .....	95
2.10.6	<i>g</i> allography complications .....	97
2.10.6.1	Archaizing spellings.....	97
2.10.6.2	Later forms in earlier texts .....	98
2.10.6.3	Sanskrit spellings .....	100
Chapter 3:	Old Khotanese Vowels.....	101
3.1	Vowel Symbols .....	101
3.1.1	The Monophthongs /i~ī, u~ū, a~ā/ .....	101
3.1.2	The Diphthongs <i>e</i> , <i>o</i> , <i>ai</i> , <i>au</i> .....	101
3.1.3	The New Signs <i>ä</i> and <i>ei</i> .....	102
3.2	Vowel Linguistics .....	103
3.2.1	EWKh Monophthongs.....	103
3.2.2	EWKh Diphthongs <i>ai</i> , <i>au</i> , <i>ei</i> .....	104
3.2.3	EWKh On-Glide Diphthongs .....	105
3.2.3.1	/ịu/.....	106
3.2.3.2	/ịo/.....	106
3.2.3.3	/ịa/.....	107
3.2.3.4	/ịe/.....	109
3.2.3.5	/ụo/ .....	110
3.2.3.6	/ụa/.....	110

3.2.3.7	/ue/.....	110
3.2.3.8	/ui/.....	111
3.2.3.9	On glide Diphthongs with present stems .....	111
3.2.4	Triphthongs .....	112
3.2.5	EWKh and AOKh overlong vowels.....	113
3.2.5.1	/ē/ .....	114
3.2.4.2	/ō/ .....	115
3.2.6	Vowels in Canonical Old Khotanese .....	115
3.2.7	Non final /ō/ in <i>tcohora</i> .....	115
3.2.7	Emmerick, and Emmerick and Maggi Vowels.....	118
3.3.0	The Six Alternations of <i>ä</i> .....	121
3.3.1	Alternations of <i>ä</i> in AOKh (Śgs).....	122
3.3.2	Alternations of <i>ä</i> in COKh (Z) .....	123
3.3.2.1	- <i>ä</i> ~ <i>e</i> alternation in the 3S.....	123
3.3.2.2	Other Endings Showing <i>-e</i> > <i>-ä</i> .....	127
3.3.2.3	Non final <i>-ä</i> ~ <i>-e</i> .....	127
3.3.2.4	Non final <i>-ä</i> ~ <i>-i</i> .....	128
3.3.2.5	<i>ä</i> > <i>i</i> in Final Position .....	129
3.3.2.6	<i>a</i> > <i>ä</i> .....	130
3.3.2.7	<i>kāḍana</i> > <i>kādāna</i> .....	130
3.3.2.8	Elision of <i>ä</i> next to <i>r</i> .....	131
3.3.2.9	Elision of <i>ä</i> in IASm <i>-āna</i> .....	131
3.3.2.10	Summary of <i>ä</i> alternations.....	131
3.4	The Alternations of <i>o</i> and <i>au</i> .....	132
3.4.1	EWKh non final <i>o</i> .....	132
3.4.2	EWKh <i>au</i> .....	134
3.4.3	EWKh <i>-o</i> and <i>-au</i> in Inflections .....	134
3.4.3.1	ASf of the <i>ā</i> -declension, EWKh <i>-o</i> .....	135
3.4.3.2.	ASm of the <i>aa</i> -declension, EWKh <i>-au</i> .....	135
3.4.3.3	IAPm and IAPf, EWKh <i>-yau</i> .....	135
3.4.3.4	ASf of the <i>aā</i> -declension, EWKh <i>-o</i> .....	136
3.4.3.5	LPM and LPf, EWKh <i>-uṣo</i> .....	136
3.4.3.6	LSm and LSf, EWKh <i>-o</i> .....	137
3.4.3.7	3Pop, EWKh <i>-īro</i> .....	137
3.4.3.8	3Psj.m., EWKh <i>-āro</i> .....	137
3.4.3.9	3Siv.m type A, EWKh <i>-āto</i> .....	137
3.4.4	Exceptions .....	138
3.4.5	Canonical Old Khotanese <i>o</i> ~ <i>au</i> .....	138
3.4.5.1	<i>o</i> and length in monosyllables.....	138
3.4.5.2	<i>ko</i> .....	140
3.5.0	Umlaut Evidence.....	141
3.5.1	Umlautable Vowels.....	141
3.5.2	Not Umlautable Vowels .....	142
3.5.3	Umlaut phonetics .....	143
3.6.0	Enclisis.....	144
3.6.1	3S enclitic pronoun <i>-i</i> .....	144

3.6.2	3S enclitic pronoun <i>-yā</i> .....	147
3.6.3	Another parallel, <i>e+u&gt;yau</i> and <i>o+i&gt;vai</i> .....	148
Chapter 4:	Stem and Suffix Metanalysis .....	150
4.1	Morphological Reanalysis .....	150
4.2	The two sides of the morpheme boundary .....	152
4.3	Skjærvø 2004 .....	153
4.4	Conjugated present stems showing <i>VttV</i> .....	156
4.5	The vowel contractions exhibited by the <i>VttV</i> stems .....	159
4.5.1	Contractions with stems ending in <i>-a-</i> .....	159
4.5.1.1	<i>a+a&gt;ā</i> in <i>VttV</i> .....	159
4.5.1.2	<i>a+ī&gt;ai</i> in <i>VttV</i> .....	160
4.5.1.3	<i>a+ä&gt;ei&gt;ai</i> in <i>VttV</i> .....	161
4.5.1.4	<i>a+ā&gt;ā</i> in <i>VttV</i> .....	163
4.5.1.5	<i>a+ū&gt;∅+ū</i> in <i>VttV</i> .....	165
4.5.1.6	<i>a+∅&gt;a</i> .....	165
4.5.2	Contractions with stems ending in <i>-i-</i> .....	166
4.5.2.1	<i>i+ī&gt;ī</i> .....	166
4.5.2.2	<i>i+ä&gt;ī</i> .....	167
4.5.3	Contractions with <i>VttV</i> stems ending in <i>-u-</i> .....	168
4.5.3.1	<i>u+a&gt;va&gt;uva</i> .....	169
4.5.3.2	<i>u+ā&gt;vā&gt;uvā</i> .....	170
4.5.3.3	<i>u+ī&gt;vī&gt;uvī</i> .....	172
4.5.3.4	<i>u+e&gt;ve&gt;uve</i> .....	174
4.5.4	Contractions with <i>VttV</i> stems ending in <i>-ū-</i> .....	174
4.5.4.1	<i>ū+ā&gt;vā&gt;uvā</i> or <i>u+ā&gt;vā&gt;uvā</i> .....	175
4.5.4.2	<i>ū+ā&gt;ūyā</i> .....	175
4.5.5	Contractions with <i>VttV</i> stems ending in <i>-ä-</i> .....	176
4.5.5.1	<i>ä+a&gt;ya</i> .....	176
4.5.5.2	<i>ä+ī&gt;ī</i> .....	176
4.5.5.3	<i>ä+ā&gt;yā&gt;iyā~äyā</i> .....	177
4.5.6	Contractions with <i>VttV</i> stems ending in <i>-e-</i> .....	179
4.5.6.1	<i>e+ī&gt;ai</i> .....	180
4.5.6.2	<i>e+ā&gt;yā</i> .....	181
4.5.7	Summary of the contractions exhibited by the <i>VttV</i> stems .....	183
4.6	<i>tt</i> -initial suffixes contain double / <i>tt</i> / .....	183
4.6.1	Past stems in <i>-Vda-</i> .....	184
4.6.2	<i>tara</i> -comparatives and <i>tāti</i> -nouns .....	187
4.6.3	Present stems in <i>-r-</i> conjugated with <i>tt</i> -initial suffixes .....	189
4.6.4	The odd case of <i>buḍḍamu</i> .....	192
4.6.5	<i>-ltt-</i> and <i>-ṃtt-</i> .....	194
4.6.6	The 3Sp.a suffix <i>-<sup>i</sup>ttā</i> and vowel final stems. ....	195
4.7	Other suffixes with the <i>VttV</i> exhibiting stems .....	201
4.7.1	The deverbal noun suffix <i>-na-</i> and vowel stems .....	201
4.7.2	The causative suffixes <i>-ñ-</i> and <i>-āñ-</i> with vowel stems .....	205
4.8	The historical evolution of the metanalysis .....	207
4.8.1	Glide resolution .....	208



4.8.2	Secondary declensions.....	209
4.8.2.1	*-ikV-.....	210
4.8.2.2	*-ukV-.....	211
4.8.2.3	*-akV-.....	212
4.8.3	Historically vowel final present stems.....	213
4.8.4	Hist. cons. final, but synchronically vowel final present stems.....	216
4.8.5	Parallel Historical Evolution.....	219
4.8.5.1	*uC+V > uw+V > u+V.....	219
4.8.5.2	*iC+V > iy+V > i+V and *äC+V > äy+V > ä+V.....	221
4.8.5.3	*eC+V > *ey+V > e+V.....	222
4.8.5.4	*ãC+V > *aĥ+V > a+V.....	223
4.8.6	Summary of the parallel evolution.....	224
4.9	Skjærvø's Proposal.....	225
4.10	The metanalysis.....	227
4.10.1	Cluster reduction (R)CCC > (R)CC.....	228
4.10.2	Possible role of <i>se-</i> 'appear' in the metanalysis.....	233
4.10.3	Fossils.....	234
4.11	Summary.....	236
	Abbreviations.....	238
	Bibliography.....	241

## List of Figures

<b>Table 1.1:</b>	Transmission of the Brāhmī from India to the Tarim Basin.....	7
<b>Table 2.1:</b>	Old Khotanese consonant inventory.....	20
<b>Table 2.2:</b>	Consonants in Emmerick 1981.....	21
<b>Table 2.3:</b>	Consonants in Emmerick 1989.....	22
<b>Table 2.4:</b>	Differences between Emmerick and Hitch.....	24
<b>Table 2.5:</b>	Late Middle Chinese aspirates and fricatives in LKh writing.....	26
<b>Table 2.6:</b>	Late Middle Chinese retroflex initials in LKh writing.....	27
<b>Table 2.7:</b>	Nine segments related by umlaut.....	29
<b>Table 2.8:</b>	Six natural classes in umlaut.....	29
<b>Table 2.9:</b>	Velar and palatal phonemes in umlaut.....	30
<b>Table 2.10:</b>	Consonant behavior in umlaut.....	33
<b>Table 2.11:</b>	Stem final -r- + -tt- and -rr- + -tt-.....	36
<b>Table 2.12:</b>	History of initial and intervocalic rhotics.....	40
<b>Table 2.13:</b>	Typical inventory with retroflex affricate.....	48
<b>Table 2.14:</b>	LKh Transcriptions of OT Velars.....	53
<b>Table 2.15:</b>	Initial <i>p, b, t, d</i> distribution.....	57
<b>Table 2.16:</b>	Single OKh Oral Stops and Affricates at the End of Present Stems.....	58
<b>Table 2.18:</b>	LOT /b/ spellings in LKh orthography.....	62
<b>Table 2.19:</b>	LOT initial /b/=[β] spellings.....	63
<b>Table 2.20:</b>	LOT medial /b/=[b, p] spellings.....	64
<b>Table 2.21:</b>	Theoretical labial stop inventory gaps.....	67

<b>Table 2.22:</b> Suffix initial /t/ after voiceless segments.....	77
<b>Table 2.23:</b> Suffix initial /t/ after most voiced segments .....	78
<b>Table 2.24:</b> Suffix initial /t/ after voiced /d, ɖ, l/ .....	78
<b>Table 2.25:</b> Gap in the retroflex inventory .....	80
<b>Table 2.26:</b> LOT /l/ spelled <i>ɖ</i> .....	81
<b>Table 2.27:</b> LOT /l/ spelled <i>l</i> .....	82
<b>Table 2.28:</b> LOT /l/ spelled <i>r</i> .....	82
<b>Table 2.29:</b> Labial and velar obstruent inventory gaps.....	88
<b>Table 2.30:</b> Labial and dental allography.....	90
<b>Table 2.31:</b> Archaic <i>g</i> ~ Canonical <i>t</i> for /ʔ/.....	91
<b>Table 2.32:</b> Archaic <i>g</i> ~ Canonical <i>t</i> as peg radical .....	92
<b>Table 2.3.3:</b> Evolution of glottal stop.....	94
<b>Table 3.1:</b> EWKh Monophthongs .....	103
<b>Table 3.2:</b> Old Khotanese onglide diphthongs .....	105
<b>Table 3.3:</b> Spanish onglide diphthongs .....	106
<b>Table 3.4:</b> Monophthongs in Emmerick 1979.....	118
<b>Table 3.5:</b> Monophthongs in Emmerick 2009.....	120
<b>Table 3.6:</b> The alternations of ä.....	131
<b>Table 4.1:</b> Persian metanalysis .....	151
<b>Table 4.2:</b> All known <i>VttV</i> categories with vowel stems.....	158
<b>Table 4.3:</b> <i>se-</i> , <i>*sed-</i> and the metanalysis .....	234

## Acknowledgments

My interest in Khotanese grammar owes its foundation to Ronald Emmerick. He assisted my stay in Germany in 1983-4 welcoming me into his home in Quickborn on a weekly basis. I had the enormous privilege of being the first student to work through his teaching grammar of Khotanese. We would begin with Ron marking my homework, and then read through the chapter he had prepared in the meantime on his typewriter with carbon paper. My background at that point was in linguistics, and this synchronic approach to learning the language was effective for me. We discussed many things and with his immense knowledge he was able to show immediately if an idea had any merit. I regret that due to his far too early passing I am unable to show him that his efforts bore fruit.

Richard Frye accepted me into the PhD program with Inner Asian and Altaic Studies at Harvard in 1985. Our interests intersected with Middle Iranian languages and Inner Asian history. In 1990 I went to the far north of Canada to work with aboriginal languages. Professor Frye retired in 1990 and P. Oktor Skjærvø arrived at Harvard, in 1991. By mail he kindly checked early drafts of the sections included here on consonants and vowels, making many helpful suggestions. In 1993 began the long interruption in the PhD process as northern activities became my main preoccupation. I never thought the interruption would last that long.

In the early 2000's Dieter Maue got me to comment on his work on Tumshuqese, a language on which I had done some research and publication. I promised him to spend more time on Tumshuqese after finishing my ideas on Khotanese phonology, morphology and writing. In 2008 I prepared a paper on a

theory I had developed at Harvard to explain why medieval Turfan was so cosmopolitan with documents in more than twenty languages attested. I was able to present this in Hamilton, Toronto, Copenhagen, Berlin and Leiden and then Victor Mair published it in 2009 as “The Special Status of Turfan” in the Sino-Platonic Papers. In 2010 Victor requested a paper for a catalog on an exhibition of Xinjiang mummies, which as appeared that year in SPP as “Aramaic Script Derivatives in Central Eurasia”. Dieter, Victor and John Colarusso from McMaster University (Hamilton), my first professor of linguistics from the late 1970’s all encouraged me to finish the PhD.

It can be challenging gathering materials in an obscure field without access to a major library. In the summer of 2010 I was in England for my daughter’s wedding. I was able to collect some materials from the SOAS library, then Nicholas Sims-Williams met me at the Ancient India and Iran Trust in Cambridge and helped me obtain a range of items now essential for Khotanese work. I smuggled a forty pound backpack of books and papers onto the airplane. In 2012, Keren Rice, Athapaskanist from the University of Toronto, kindly arranged for me to access the Robarts Library. I made several visits there, scanning materials that would later prove essential, and was able also to use the online resources.

In the summer of 2012 I took two months off my job in Yukon to resuscitate the dissertation and in November circulated a paper that I thought would go a long way to completing it, “Old Khotanese type B present stems ending in *-a-*”. This included a description of the metanalysis. I received useful comments back from Almuth Degener, Ilya Yakubovich, Dieter Maue, Hiroshi Kumamoto, Mauro Maggi

and Nicholas Sims-Williams. It was clear that for my synchronic approach to be understood and accepted I had to do a better job preparing the groundwork. That one paper has become six papers. The first five are published or in press (Hitch 2014, 2015a, 2015b, in-press-a, in-press-b). The sixth is here as the last chapter.

With his comments, Hiroshi Kumamoto also supplied an electronic version of most Khotanese texts which I think came originally from Ron Emmerick and had been in circulation for several years. Hiroshi also sent a word list which he had electronically compiled from the files of the texts. These materials allowed the efficient finding of forms and patterns. I have used them at every step and believe that without them many of the insights of recent work would not have been possible.

In 2014 I thought I could see an end to the research and writing. Two friends I have met in Whitehorse over the years, Michael Heine from the University of Western Ontario, and Wayne Horowitz from the Hebrew University of Jerusalem, both with years of experience getting graduate students to finish, rigorously encouraged me to finish. All of these supportive voices helped keep me persistent on the topic while being somewhat isolated from the related academic world. In late 2014 Michael Witzel agreed to act as advisor and the long road of rebirthing the PhD process began. Leonard van der Kuijp and Mauro Maggi also agreed to sit on the committee. Mauro had been the main commenter on my work since 2012, suggesting countless improvements. He has been in essence technical advisor on this highly specialized work which would be much poorer without his willing, able and painstakingly thorough assistance.

Finally, Marlene Chapelle has stood behind my dream of finishing since our first meeting. She really launched the project with an invitation to work at her lakeside home in the summer of 2012, and has been patient and supportive as the project stretched from months to years. She also took great photos at the defense and prepared an outstanding email report for family and friends.

# Chapter 1: Introduction and Background

## 1.0 Overview

A metanalysis is an historical linguistic process in which there is a shift in a morpheme boundary. For example, the Middle English *a napron* became Modern English *an apron* through the shift to the right in the boundary between article and noun. The Old Khotanese metanalysis is much more complex. It involves a shift to the left of the morpheme boundary in a class of present stems and a class of suffixes, which created new stem shapes, and new suffix shapes. For instance, *butte* 'he knows' was analyzed something like \*[*bud*][*te*] in Pre-Khotanese with a stem \**bud-* and suffix \**-te*. After the metanalysis the form is analyzed [*bu*][*tte*] with a stem *bu-* and a suffix *-tte*. Diachronically there is no doubt that *bu-* derives from Proto-Iranian \**baud-* 'know'. Synchronically, it is clear that *bu-* is vowel final. From its morphophonological interactions there is no synchronic trace of a stem final consonant.

This thesis began as a study of Old Khotanese phonology. The two chapters after this one were largely written in the early 1990's and never published. It is useful to present them here as my phonological analyses depend on my understanding of the Old Khotanese consonants (chapter 2) and vowels (chapter 3). My study of the vowels could not be completed until I understood the complex morphophonology of the present stems. I suspected a metanalysis played a role long ago but it was not until I tried to convince others that I realized much ground work was needed before the notion could receive general acceptance. In the end, this has involved the

publication of five articles, the content of which precedes the last chapter here specifically on the metanalysis (chapter 4).

**①: 'Meter in the Old Khotanese Book of Zambasta' (2014)**

This developed an apparatus for analyzing hemistichs of the great Buddhist poem. It identified several cadences which prove crucial to later arguments. It outlined several principles, the recognition of which allows more regular analysis of the meter which turns out to be violated much more rarely than appears at first glance. The meter reveals features of word structure which are not apparent from the orthography.

**②: 'tt in Old Khotanese' (2015a)**

This details the evidence showing that between vowels the digraph *tt* in Old Khotanese is ambiguous, standing for either single /t/ or double /tt/. Although several earlier scholars recognized this ambiguity, some did not. Using the metrical apparatus developed in ①, it is shown beyond doubt that *VttV* is ambiguous. There is double /tt/ in two categories: in words of Indian origin, and in conjugated forms (like *butte*) from a sizable group of present stems.

**③: 'Contracted Diphthongs in Old Khotanese' (2015b)**

This studies the contractions which take place when a nominal stem ending in /a/, /i/ or /u/, receives a suffix beginning with a vowel. The investigation is limited to the singular number and NAP. The oblique plural cases, GDP, IAP, VP, LP, behave differently and are dealt with in 4. Relying heavily on the evidence provided by the position of forms in the meter of Z, it is shown that all contracted syllables, no



matter how they are written, are two moras long. The vocalic entities resulting from contraction and written *ya, ye, yo, yu* or *va, ve, vo, vī* are all two-mora diphthongs /j̄a, j̄e, j̄o, j̄u, ʷa, ʷe, ʷo, ʷi/.<sup>1</sup>

#### ④: ‘Contracted Semivowels in Old Khotanese’ (in-press-a)

This begins with an examination of contracted diphthong resolution. An orthographic sequence like *Cya* or *Cva* containing a diphthong, may be resolved to *Cäya~Ciya* or *Cuva* without changing the mora count. The remaining nominal contractions, those affecting the oblique plurals, are then detailed. In contraction, a final /a/ is deleted, and final /i/ and /u/ semivocalize to /y/ and /w/ rather than diphthongize. When an orthographic sequence involving one of these semivocalized segments like *Cyā* or *Cvā* resolves to *Ciyā~Cäyā* or *Cuvā* there is a one mora increase in metrical count. In contrast, the diphthong resolution analyzed in 3 does not increase mora count. Finally it is shown how several present stems in /i/ and /u/ exhibit the same contractions as nominals in /i/ and /u/. The synchronic suffixes involved are the same as those found with type A consonant stems.

#### ⑤: ‘Old Khotanese type A stems in -a- and -e-’ (in-press-b)

This shows how Emmerick’s type D verbs may be synchronically analyzed as stems ending in -a- and -e-. Emmerick’s *pai-* ‘protect’ and *ttätsu-* ‘cross (a river)’ are *pa-* and *ttätsa-*, while his *dai-* ‘appear’ and *ysai-* ‘be born’ are *de-* and *yse-*. The final

---

<sup>1</sup> In Hitch 2015b what I now transcribe as /j̄e, j̄o, ʷe, ʷo/ were given with macrons as /j̄ē, j̄ō, ʷē, ʷō/ because those diphthongs were formed from phonemic /ē, ō/, and when the diphthongs resolve to orthographic -*ye, -yo, -ve, -vo*, the original phonemic /ē, ō/ resurface in /-iyē, -iyō, -uwē, -uwō/. But as diphthongs these behave no differently than those transcribed /j̄a, j̄u, ʷa, ʷi/ (orthographic -*ya, -yu, -va, -vī~uī*).

vowels of these stems exhibit contractions which are either attested among the contractions detailed in the earlier articles, or are phonologically natural extensions of those contractions. This analysis uses the same synchronic suffixes as are attested for the consonant final type A stems. There is no need to maintain a separate type D category. These verbs are now vowel final type A stems.

#### **Chapter 4 (= article ⑥ in the sequence) summary**

Chapter 4 demonstrates two things. It shows that all stems showing *VttV* in particular conjugations are now synchronically vowel final. And it argues that all these stems use the regular type B suffixes, both vowel and consonant initial, found with type B consonant final stems. Emmerick classified the stems showing *VttV* as either type B or type C. Here, all these stems are type B.

The key part of this description is the contractions exhibited by the *VttV* exhibiting stems when vowel initial suffixes are attached. The majority of these contractions are already attested in ③–⑤ or are easily understood as phonologically natural extensions of those contractions. By allowing for a metanalysis [Vt][tV]>[V][ttV] in certain inflections, the behavior of the stems under all other conditions receives a transparent and straightforward explanation. Emmerick's type C verbs are synchronically vowel final type B. Historical motivation for the metanalysis is suggested.

## 1.2 Method

Evidence of the phonological structure of Old Khotanese consonants is drawn from many sources:

- 1 The origin and structure of the Brāhmī writing system. Most values will remain constant when a writing system is borrowed. For instance, because *s* has the value [s] in Indian and other Brāhmī systems, it will likely have the same or similar value in Old Khotanese.
- 2 Synchronic morphophonemic variation. A segment in a stem or suffix may change in some fashion with suffixation. These changes are phonologically governed.
- 3 Khotanese metrics. The primary source is the rich metrical system exhibited by *The Book of Zambasta* (Z). This reveals aspects of word structure not visible in the orthography.
- 4 Synchronic spelling alternations. Two symbols in use for one segment will have related phonological value. Two spellings for a word may reveal morphophonological information.
- 5 Developments between stages of Khotanese. Systematic historical relationships between stages, from Earliest Written Khotanese, through Archaic Old Khotanese and Canonical Old Khotanese to Late Khotanese provide insight into sound and word structure.
- 6 Transcriptions of Late Khotanese in Tibetan script. This source, as well as the next two, can provide information about both languages.
- 7 Transcriptions of Late Middle Chinese in Late Khotanese writing.

- 8 Transcriptions of Old Turkic in Late Khotanese writing.
- 9 The form of Sanskrit and Prakrit loan words and foreign words in Khotanese. How these are adapted may tell us something about Khotanese and about the source language.
- 10 Patterns in other South-Eastern Iranian languages. Related languages may show similar features.
- 11 Language typology. Patterns found in a broad range of the world's languages can sometimes tell us which patterns are plausible for Khotanese.
- 12 The Proto-Iranian origin of segments and their development into OKh.

The primary source is of course the writing system itself (1). Next in importance has traditionally been the historical evidence of Proto-Iranian (12) but here much other information is given more weight. The danger of over emphasizing the role of Proto-Iranian is illustrated below (2.1.2) where a supposedly transparent etymology of *haṭhṭhā*- 'truth' misled the field about fundamental aspects of the phonology for more than a half century. The value of the other sources will become clear in the ensuing pages. In general, my approach to unraveling the puzzles of the synchronic grammar is to give most weight to the evidence presented by the language itself.

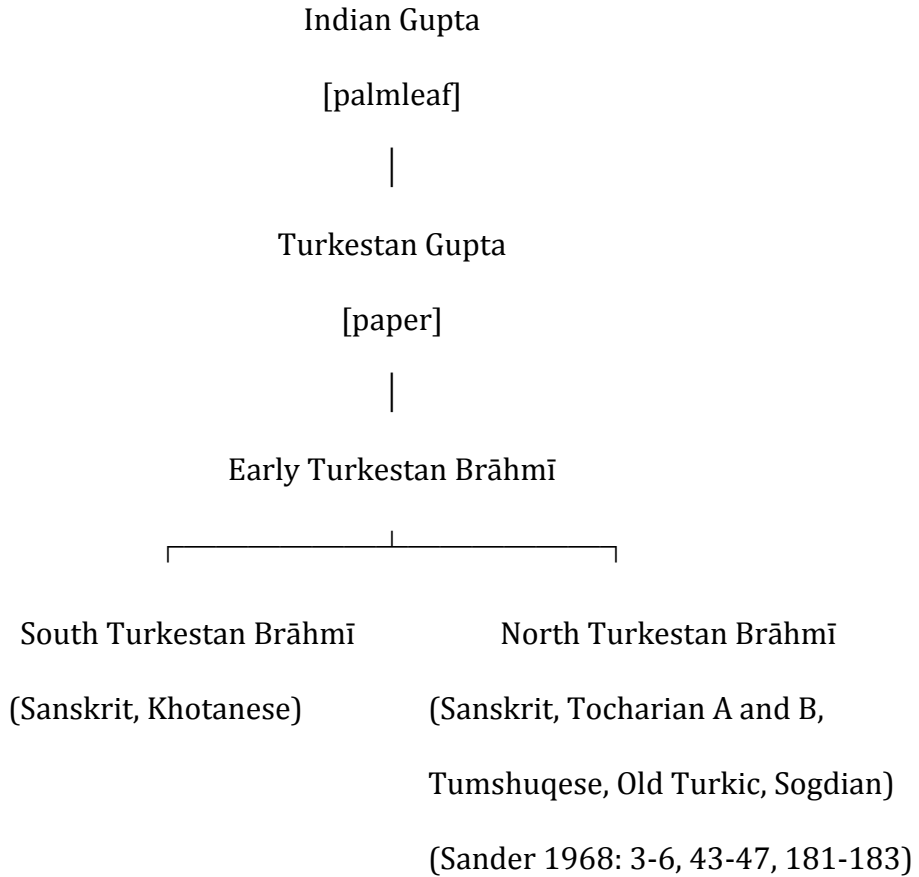
### **1.3 The Brāhmī Script**

It is not known whether the form of Brāhmī used to write Khotanese came directly from an Indian language. There are features of Khotanese writing which are not found in Indian practice such as the vowel signs *ā* and *ei*, the consonant *rr*, and the apostrophe.





### 1.3.1 Sander's Paleography

The shapes of the letters used in the script also point to an indirect transmission from India. Lore Sander, in her study of the paleography of the Sanskrit manuscripts in the Berlin Turfan collection (1968), proposes that two types of writing can be distinguished which intervene between Khotanese and Indian practice. The transmission is summarized in the chart below:

**Table 1.1:** Transmission of the Brāhmī from India to the Tarim Basin



Sander distinguishes two script types between the Gupta of India and the South Turkestan Brāhmī of Khotan. The Turkestan Gupta manuscripts are written mostly on paper, never on palmleaf, while Indian Gupta are mostly on palmleaf, never on paper. The character shapes are very similar but it is clear that the Turkestan Gupta

manuscripts were written in Inner Asia because there was no paper in India at this time, and there are no palms in Inner Asia. The Early Turkestan Brāhmī is distinguished from the older Turkestan Gupta most clearly by the shape of the diacritics -e and -ai. The older type bends down to the left, e.g., -e , -ai , while the younger type ends in a bend to the right, e.g., -e , -ai . Sander further distinguishes two types within Early Turkestan Brāhmī, an alphabet r which is the progenitor of the North Turkestan Brāhmī and an alphabet s which is the ancestor of the South Turkestan Brāhmī (ibid. 181-182).

It seems likely that the letter shapes underwent development outside India before being introduced to Khotanese. Sander identifies within the Indian Gupta type, a group A and a group B. The alphabets of group A are contemporaries of each other and represent regional styles. Those of B are chronological stages of one style (Sander 1968: 85). The relative chronology within group B is found in manuscripts from Bamiyan and also in those from Turfan. Sander concludes that Gupta manuscripts of this type reached the north of the Tarim Basin from Bamiyan (ibid. 135). Similar parallels in other groups lead her to conclude “that the majority of the Brāhmī manuscripts of the Berlin collection imported from India to Eastern Turkestan probably came from Bamiyan”<sup>2</sup> (ibid. 136). Bamiyan is a center of ancient Bactria. The people native to this area probably spoke the middle Iranian language Bactrian but used Sanskrit for religious purposes. It is possible that some of the non

---

<sup>2</sup> “... daß die meisten aus Indien nach Ostturkistan eingeführten Brāhmī-Manuskripte der Berliner Sammlung wahrscheinlich aus Bamiyan kamen ...”

Sanskritic features of Khotanese writing have their source in a local tradition, somewhere between Bamiyan and Khotan.

### 1.3.2 The New Sign *ä*

The Brāhmī scripts used for languages on the northern rim of the Tarim Basin also feature the vowel sign *ä*. The Brāhmī systems of Kuchean, Agnean, Tumshuqese and Old Turkic are clearly related. Their relationship to Khotanese writing is not clear. We cannot conclusively prove the origin and development of *ä*. It may have come from India, or Bactria or from the Tarim Basin itself. It may have been borrowed separately from a third party by both the northern group and Khotanese, or it may have been borrowed by either the northern group or Khotanese and then transferred to the other. The possibility is not unrealistic that influence on the origin of these scripts came from outside India.

### 1.3.3 Other Features

Even if it could be proven that the letter shapes and *ä* came to Khotan from outside India, this does not necessarily mean that some or all of the principles used to write the language could not have been borrowed directly from India. Oskar von Hinüber 1980 proposed to see in the scripts of Indian Prakrits the models for some of the special, non Sanskritic orthographic developments exhibited by Khotanese. Among these features are the values / $\gamma^3$ ,  $\delta$ ,  $d$ / of *g*, *d*, *t*, the use of *ky*, *gy* for palatals, the existence of double *tt*, and the ligature *ys* for voiced /*z*/ (124-125).

---

<sup>3</sup> As argued in the section on consonants, there was no [ $\gamma$ ] in Khotanese.

## 1.4 Stages within Old Khotanese

There is no unanimity among scholars about the linguistic stages of Khotanese. R. E. Emmerick 1987b has usefully summarized the views of Ernst Leumann and H.W. Bailey, and has presented his own understanding. They all would, however, probably agree that the difference between Old and Late Khotanese is that Old is more highly inflected and has freer word order, while Late has lost much of the inflectional morphology and has a more restricted word order. P.O. Skjærvø has proposed use of the term “Middle Khotanese” and presented criteria for defining his three stages (SuvI: lxxii–lxxiv).

For the present study, it is useful to distinguish several stages within the older language. The evolution of these stages parallels the evolution of the orthography. I use one set of terminology for both linguistic and orthographic stages, which may later require separation.

*Pre-Khotanese* refers to any period of the language before it was first written down, and after Old Iranian. It is unattested. It may be possible one day to further refine the early periods into Proto-Saka and Proto-Eastern Iranian.

*Earliest Written Khotanese* (EWKh) refers to the stage of the language when it was first recorded in the Brāhmī script. There are no documents from this stage. It is an abstraction which can be reconstructed from information in the writing system and from the language of the earliest texts. The concept is useful in defining the early development of the language and orthography. It has features which are not found in later stages.



*Archaic Old Khotanese* (AOKh) is attested in a range of manuscripts which have not been compiled into a comprehensive list. The longest and best preserved is the Śgs.<sup>4</sup> This is known from an edition by Emmerick 1970, and from other folios transcribed and translated in Skjærvø 2002 (listed p.608). Another important example is Manuscript 1 from the Sgh. Manuscripts 2–9 also have mostly archaic features. Canevascini organized the manuscripts of the Sgh in his edition (1993) from oldest to youngest. Lower numbers have more archaic features. A third important document consists of one folio of Z found at Šorčuk near Qarashahr on the northern rim of the Tarim Basin (Maggi 2004). The features of Archaic orthography include single ś, ṣ, g (instead of double śś, ṣṣ, gg) for /ś, ṣ, g/. The Archaic morphology consistently distinguishes Type A 3Sp.a -ätä from 3Sp.m -äte, the NSm -ä of *a*-stems from the GDSm -i, and the NSm -ei of *aa*-stems from the GDSm -ai, among other features.

*Canonical Old Khotanese* (COKh) is admittedly loosely defined. Linguistically it comes after Archaic and before Late Khotanese. It encompasses the majority of Old Khotanese texts. Orthographically a definition is clearer. These texts will have double śś, ṣṣ, gg for /ś, ṣ, g/ and then keep single ś, ṣ for /ž, z/. All surviving Khotanese texts of any stage are copies. Some of the texts in Canonical orthography are the result of many generations of copying over several centuries which can

---

<sup>4</sup> Emmerick at one time thought it would be unjustified to conclude that Śgs is older than the bulk of our OKh material. He saw other factors possibly coming into play, 'such as the style of an individual scribe involving certain spelling predilections, possibly also the practice of different schools of writing' (Śgs:xx; also cf. 1992:32). And the distinctions 'may reflect a tradition handed down from a time when the difference in origin [i.e., in the distinctions,] was reflected in a difference in spelling and pronunciation' (ibid.).

make the original stage of language difficult or perhaps sometimes even impossible to discern.

*Late Khotanese* is not under study here, but it is useful to recognize a Late text, or a text with heavy Late influence when studying morphology. There are both orthographic and linguistic diagnostic features which are generally effective in identifying a Late text. The documents, mostly from Dunhuang, appear to be from a different scribal tradition than the Canonical, perhaps independently descending from the Archaic. We find single *ś, ṣ, g* (still?) in use for /*ṣ, ś, g*/. For voiced /*ž, z*/ there are the single sibilant graphs used in conjunction with the subscript hook or apostrophe, i.e. *ś', ṣ'*. Another probably orthographic feature is the seemingly random use of anusvāra, the superscript dot, which earlier marked a nasal homorganic with the following consonant in a cluster. For linguistic diagnostics there are two distinctive nominal suffixes. For OKh GDP *-ānu* there is LKh *-āṃ*, and for OKh LP *-uvo'* there is LKh *-vā*. A text with all of these orthographic and morphological indicators will be Late. A text with some but not others may reflect transitional language or heavy influence of LKh copyists.

## 1.5 Symbology

Several different sets of symbols are necessary to effectively analyze Khotanese and other languages in these pages. Besides the metrical apparatus (described below) there are three sets of letters for orthographic, phonemic and phonetic representations. These sets largely overlap and there can easily be confusion. Orthographic representations are in italics, phonemic in slanted brackets //, and phonetic in square brackets [ ]. Choices in symbology have been made with several

principles in mind such as avoiding ambiguity, retaining tradition, and being internally consistent.

### 1.5.1 Orthographic Transliteration and Transcription

In a *transliteration* the intent is to reproduce every salient feature of a manuscript. In theory, a written page could be accurately recreated from a transliteration. In contrast, a *normalized transcription* reflects a norm established by the scholarly community. It may add features not found in the written text, or it may omit some written features which are linguistically unnecessary. An extra feature in Khotanese transcription is the separation of words by spaces. An omitted feature may be an indication that an akṣara has two vowel diacritics. For instance, the transcription *kuī* ‘where - to him’ would in Sanskrit practice imply two akṣaras, \**ku ī*. But in Khotanese sometimes two vowel diacritics may be attached to one radical. In this case, subscript *-u* and superscript *-ī*, are attached to the same consonant radical *ka*. Some scholars have transcribed the word as *kuī* or *kuī* to show this.

For the most part, the symbols used in transliterating or transcribing Khotanese are identical to those used for Sanskrit. The exceptions are: 1) *rr* which is here regarded as denoting the alveolar rhotic /ɹ/ (in contrast with retroflex *r* /ɻ/); 2) the subscript hook or *apostrophe* which is rendered with a latin apostrophe and regarded here as denoting a sequence of two vowels; 3) *ei*, a superscript diacritic resembling the initial *-e* diacritic stroke with a second stroke through it at close to a right angle, and here standing for the low to mid diphthong /ae/; 4) *ā*, the two dot superscript diacritic also found in Tocharian and here marking short /ě/; 5) a superscript *i* used in normalized transcription of suffixes which carry the umlaut

potential, e.g. GDS -<sup>i</sup>*e* of the *ā*-declension. These symbols all follow traditional and current practice, with the exception that the work of Ernst and Manu Leumann used the schwa symbol ə instead of ā for the two dot diacritic. All transliterated or transcribed words within the text are given in italics

### 1.5.2 Phonemic Transcription

All phonemic symbols are listed in slanted brackets, //. The symbols in large part match those used in the International Phonetic Alphabet (IPA) but there are also notable divergences. For retroflex segments, a dot under the letter is used as in the Indological tradition, instead of the IPA hooked letters. That is, instead of hooked /t̪ d̪ ɳ z̪ ʎ/ I use dotted /ṭ ḍ ṇ ṣ ṛ ʎ/. As it is preferable to use one letter for each segmental phoneme, for the alveopalatal affricates I use č instead of tʃ and ĵ instead of dʒ. Then for consistency for the corresponding sibilants I use š and ž instead of ʃ and ʒ. The symbols with haček are well established in both Americanist and European philological traditions. I avoid the symbol c as it is ambiguous. In Indology it is the transliteration for [č], some Americanists have used it for the alveolar affricate [tʃ], and in IPA it is a palatal plosive. Here it is used as in Sanskrit transliteration and not in phonemic or phonetic transcription. I use the phonemic symbols /kʸ/ and /gʸ/ for the segments written *ky* and *gy*. It is not clear if the segments are phonetically post-palatalized velars (IPA *k<sup>j</sup>* and *g<sup>j</sup>*) or palatal plosives (IPA *c* and *ɟ*). For the palatal semivowel or glide I use /y/ instead of the IPA palatal approximant *j*. Here, *j* is used in the Indological fashion to transcribe a Brāhmī akṣara denoting [j].

The phonemic symbols used for vowels are derived from the transliteration symbols where possible. Orthographic *i ī u ū a ā* are phonemically /i ī u ū a ā/. There is no advantage to using the IPA length symbol : instead of the macron. Orthographic *e* and *o* are long vowels in both Sanskrit (diphthongs in Vedic) and Khotanese so are phonemically /ē/ and /ō/. I consider *ä* to represent the short correspondent to *e* /ē/ so render it /ě/ with a breve to avoid ambiguity. It is admittedly inconsistent to use breve on short /ě/ while not using it on short /i u a/ however there is no ambiguity, it is simpler to write /i u a/ and the relationship to orthographic *i u a* is immediately obvious. The phonetic vowel symbols are the same as the phonemic except for the mid vowels. For the vowel written *e*, phonemic /ē/, I use [æ̃] while for *ä*=/ě/ I use [ɛ̃] because in this case there are crucial phonetic differences between the long and short phonemes which are not obvious from the phonemic representations. Also, final *e*=/ē/=[æ̃] is phonetically shortened at the end of most words, so in those instances we need phonetic [æ̃̄]. Similarly, *o*=/ō/=[ō̃] is phonetically shortened at the end of most words and so we need [ō̃̄].

The symbols used for diphthongs require special note. Familiar from Sanskrit are *ai* and *au*, and from traditional Khotanese scholarship *ei*. Phonemically these are /ai̯, au̯, ae̯/. I use the inverted breve symbol ˘ to show that the second element is an offglide. It is preferable to avoid simple /ai, au, ae/ as Khotanese also features sequences of two vowels (normally marked with the apostrophe) and such transcriptions would be ambiguous. It is also best to avoid formulas like /a<sup>i</sup>/ since the superscript <sup>i</sup> is already established to show the potential for umlaut. Ideally, the offglides should be superscripted, but there is no way to achieve that appearance



lead to a better understanding of the meter. After several years of working with both, each continues to provide new information.

Study of meter and morphophonology led me to develop a partly new apparatus in order to capture the salient details of the meter. For cadences I use the symbol H for a heavy and L for a light syllable. A heavy syllable is either closed or contains a long vowel including diphthongs. A light syllable is open and has a short vowel. The most common cadence, used for segment of seven moras is cadence 1 HLLHL. Previous scholarship has used symbology like  $\text{—}\sim\text{—}\sim\text{—}$ . The chief reason I adopted the new system was because there appears to be a third kind of metrical and linguistic syllable, X, which has a well defined distribution and peculiar linguistic properties. In the X position may only appear syllables derived from synchronic contraction of two vowels. Outside of a cadence a contracted syllable counts as if two moras or H. But the reverse is not true: in the metrical X positions a simple H never appears. So cadence 2 is defined as HXHL. A traditional formula like  $\text{—}\text{—}\text{—}\sim$  fails to capture this significant feature. And when defining the metrical behavior of the IASm morpheme *-yau*, the concept of a the X position is crucial: *-yau* is two moras in an X position and one mora elsewhere unless followed by *jsa* in which case it is two moras. In other words, *-yau* can be X, L or H depending on metrical and linguistic environments.

A second innovation is the use of the ballot box  $\square$  or similar box-shaped Unicode symbol to mark the graphic column breaks in the four columned main manuscript of Z. This is in essence a transliteration symbol reflecting graphic content. It is useful to mark because the composer of Z matched his meter to the columns in deliberate and

interesting ways. A line in Z consists of two metrical half lines or hemistichs. There is almost always a column break between them. Where there is not, some kind of copyist intervention may be suspected. In meter A a hemistich consists of twenty-four moras arranged in segments of 5+7+5+7 moras. There is almost always a column break between the third and fourth segments, that is, 5+7 □ 5+7. With meter A1, 5+9+3+7 the arrangement is usually 5+9 □ 3+7. Meter C, 7+5+5, is most often 7 □ 5+5 and C1, 9+3+5, is usually 9 □ 3+5. Any deviation from those patterns in column breaks can be evidence of copyist changes. With meter B this symbol has the most interesting implications. The segments are 5+6+7 moras and the column break almost always comes within the middle 6 mora segment. In addition, it almost always comes between words. The most common cadence for a type B 6 mora segment is HL+word boundary+3 moras (HL, LH or LLL). Marking the column break, this cadence, which I call B<sub>1</sub>, may be represented HL □ +3. There are also B<sub>2</sub> HLL □ +2 and B<sub>3</sub> HX □ +2. Together these make up 98% of a survey of 153 hemistichs reported on in Hitch 2014:7. When the column break does not correspond to a word boundary in a type B hemistich, it may be suspected that the composer's intentions have not been accurately followed by a copyist, and so the symbol □ provides useful information.

The other elements of the apparatus have origins in the symbology of others. The letters a, b, c, d refer to each of the four columns of the manuscript. These are given with the line numbers. Most often, the hemistich fully occupies two columns and is either ab or cd. Sometimes it may be ac or bd when the break between the second and third columns does not match the metrical break. The vertical bar symbol is



used between segments. The basic transcription and the translation are from Emmerick's 1968 edition unless otherwise noted.

In Emmerick's edition, Z 3.71 looks like:

71           kho ttāro auysu   puṣṣo paśśātaimä bi—  
                  śyo tta puṣṣo paśśimä   ttū varata auysu ysurrä

Here it would look like:

3.71ac       kho ttāro auysu □ | puṣṣo paśśā|taimä bi□śyo (C:7+5+5)

Just as I have completely given up that anger with all,

3.71cd       tta puṣṣo paśśimä □ | ttū varata | auysu ysurrä (C:7+5+5)

so I completely give up anger, wrath towards this one.

### 1.5.5 Unicode fonts and characters

This document was prepared entirely in Unicode. As far as possible, no character formatting was used to construct special characters. There is automatic superscripting associated with MS Word. Otherwise all super- and subscript characters but one are unique Unicode characters, enabling better searching. The one exception is superscript formatted  $\text{ṣ}$  in the phonemic representation / $\text{ṣ}^{\text{ṣ}}$ /. There is no unique Unicode  $\text{ṣ}$ . There is Unicode IPA  $\text{ṣ}$  (U+1DB3), which could serve but was eschewed in the interests of symbolic unity. Underlined characters like  $\text{ṭ}$  (U+1E6F) are also unique Unicode symbols. The MS Word character formats bold, italic and underline are used to offset text, not to build special characters.

## Chapter 2: Old Khotanese Consonants<sup>5</sup>

### 2.1 Inventory

Old Khotanese probably had the following inventory of consonant phonemes (standard Brāhmī transliterations in italics):

**Table 2.1:** Old Khotanese consonant inventory

	<i>/t/</i>	<i>tt, t</i>	<i>/t̥/</i>	<i>t̥</i>	<i>(/kʸ/ ky)<sup>①</sup></i>	<i>/k/</i>	<i>k</i>		
<i>/p<sup>h</sup>/</i>	<i>ph</i>	<i>/t<sup>h</sup>/</i>	<i>th</i>	<i>/t̥<sup>h</sup>/</i>	<i>t̥<sup>h</sup></i>		<i>/k<sup>h</sup>/</i>	<i>kh</i>	
<i>/b/</i>	<i>p</i>	<i>/d/</i>	<i>t, d</i>	<i>/d̥/</i>	<i>d̥</i>	<i>(/gʸ/ gy)<sup>①</sup></i>	<i>/g/</i>	<i>gg, g</i>	
		<i>/t<sup>s</sup>/</i>	<i>tc</i>	<i>/t̥<sup>s</sup>/</i>	<i>kʂ</i>	<i>/č/</i>	<i>c, ky<sup>①</sup></i>		
		<i>/t<sup>sh</sup>/</i>	<i>ts</i>			<i>/č<sup>h</sup>/</i>	<i>ch</i>		
		<i>/d<sup>z</sup>/</i>	<i>js</i>	<i>(/d̥<sup>z</sup>/ jʂ)<sup>②</sup></i>	<i>/j/</i>	<i>j, gy<sup>①</sup></i>			
<i>/β/</i>	<i>b</i>	<i>/δ/</i>	<i>d</i>						
		<i>/s/</i>	<i>s</i>	<i>/ʂ/</i>	<i>ʂʂ, ʂ</i>	<i>/š/</i>	<i>śś, ś</i>		
		<i>/z/</i>	<i>ys</i>	<i>/z̥/</i>	<i>ʂ</i>	<i>/ž/</i>	<i>ś</i>		
<i>/m/</i>	<i>m</i>	<i>/n/</i>	<i>n, ṃ, ñ</i>	<i>/ṇ/</i>	<i>ṇ</i>	<i>/ñ/</i>	<i>ñ</i>		
<i>/w/</i>	<i>v</i>	<i>/ɹ/</i>	<i>rr, r</i>	<i>/r̥/</i>	<i>r</i>	<i>/y/</i>	<i>y</i>	<i>/h/</i>	<i>h</i>
<i>/h<sup>w</sup>/</i>	<i>hv</i>	<i>/l/</i>	<i>l</i>					<i>[ʔ]<sup>③</sup></i>	<i>t, g, v</i>

#### Notes to table 2.1

- <sup>①</sup>It is possible that in Archaic Old Khotanese *ky* was distinct from *c*, and *gy* from *j*. They then merged sometime in Canonical Old Khotanese. See 2.2.2.1 below.

<sup>5</sup> Some of the ideas discussed here were earlier presented in ‘Old Khotanese Consonantal Palatality, Aspiration and Retroflexion’ at the November 16, 1987, *Middle Eastern Studies Association* meeting in Baltimore. P.O. Skjærvø made numerous useful comments on a 1993 version of a study on consonants.

②The possibility that a retroflex affricate written jṣ may have existed in Late Khotanese has been suggested by N. Sims-Williams (St.III:123–124, under *mijṣe* ‘woman’).

③As explained below in 2.10.4.1, the glottal stop [ʔ] is an allophone of /g/ in Archaic Khotanese, then of /t/ in Canonical Khotanese. It is never phonemic, but plays an important role in the evolution of orthographic devices and merits separate discussion.

### 2.1.1 Emmerick on the consonants

The most detailed reference on Old Khotanese consonantism may be *The Consonant Phonemes of Khotanese*, by R.E. Emmerick (1981). This lists much information about each segment in a systematic fashion and, although significant revisions have been made, it remains an essential work. Emmerick’s revisions to the 1981 analysis can be determined from Emmerick 1989:209. This lists the new values of the segments but does not defend them in a detailed way. Specifically, the 1981 view involved the following categories and values (my symbology):

**Table 2.2:** Consonants in Emmerick 1981

post-palatalized:	/θʸ/	/tʸ/	/dʸ/	/tˢʸ/	/čʸ/	/sʸ/	/zʸ/	/nʸ/	/lʸ/
	ṭh	ṭ	ḍ	ts	ch	śś	ś	ñ	l
voiceless fricative:	/φ/	/θ/	/θʸ/	/χ/	/χš/				
	ph	th	ṭh	kh	kṣ				
alveopalatal:	/š/	/ž/							
	ṣṣ	ṣ							

In the 1989 revision these segments were given the following values (my symbology):

**Table 2.3:** Consonants in Emmerick 1989

retroflex:	/ṭ <sup>h</sup> /	/ṭ/	/ḍ/	/ṣ/	/z/	—
	<i>ṭh</i>	<i>ṭ</i>	<i>ḍ</i>	<i>ṣṣ</i>	<i>ṣ</i>	<i>kṣ</i>
voiceless aspirate:	/p <sup>h</sup> /	/t <sup>h</sup> /	/t <sup>sh</sup> /	/ṭ <sup>h</sup> /	/č <sup>h</sup> /	/k <sup>h</sup> /
	<i>ph</i>	<i>th</i>	<i>ts</i>	<i>ṭh</i>	<i>ch</i>	<i>kh</i>
alveopalatal:	/č <sup>h</sup> /	/š/	/ž/	/ñ/		
	<i>ch</i>	<i>śś</i>	<i>ś</i>	<i>ñ</i>		
post-palatalized:	/l <sup>y</sup> /					
	<i>l</i>					

While Emmerick 1981 posited a set of post-palatalized segments and a series of voiceless fricatives, Emmerick 1989 regarded retroflexion and aspiration as distinctive features. Emmerick credited this significant change from 1981 to research by E.G. Pulleyblank into ‘the use of the Khotanese script to render Chinese’ (1989:209).

### 2.1.2 *haṭhṭhā*- ‘truth’

I began discussing Khotanese phonology with Professor Emmerick when I first met him at the Middle Iranian Studies conference in Leuven in 1982. I unsuccessfully proposed that the language had retroflexion and aspiration. Ultimately, the theoretical kingpin of the 1981 values was what seemed a very clear etymology: OKh *haṭh(ṭh)ā*- ‘truth’ < OIr *\*haθya*- ‘true’. This provided the seemingly unshakeable value /θ<sup>y</sup>/ for *ṭh* (first suggested in Konow 1932:8-9). The debate continued during my period of study with him in Hamburg 1983–84, always ending with ‘truth’.

In 1985 I presented a paper on Old Khotanese umlaut for Professor Richard Frye's Middle Iranian Survey course at Harvard. This assumed retroflexion and aspiration in order to systematically describe and explain the rich umlaut process. I presented the same paper at the 1986 ICANAS in Hamburg, in a session organized and chaired by Professor Emmerick, and later it was published as Hitch 1990.

It seemed unavoidable to me that the segment written *ṭh* should historically derive from *\*rθ*, not *\*θy*. In early 1987 I had what I thought was an idea which might convince Emmerick, that Khotanese *haṭh(ṭh)ā-* was from an intermediate *< \*harθa-*, which was itself a blend of the two OIr words for 'truth', *\*haθya-* and *\*arta-*. In response to my letter of April 17, I received an envelope on May 12 containing one sheet of paper with an x in blue ink beside the footnote 48 at the bottom which read:

In a recent communication (15/12/1986) Emmerick, who now accepts the idea that Khotanese had retroflex phonemes, suggests that Old Khotanese *haṭhṭhā-* 'truth' may not directly reflect Avestan *haiθiia-* 'true', Old Persian *hašiya-*, but may represent a contamination of the more usual word for 'truth', Old Iranian *\*arta-*, Avestan *arəta-* or *aša-*, with *\*haθya-*.

Unfortunately I was never able to discuss the matter further with Professor Emmerick, but it seems clear that retroflexion and aspiration as features of Old Khotanese phonology are now beyond debate. But several other features are likely not.

Emmerick's final treatment of the consonants appeared in 2009 after his death. The consonant chart there has some errors. There is no symbol corresponding to /š/ or /ž/. The curious symbol *ṭṣ<sup>h</sup>* appears in place of *ṭṣ<sup>h</sup>*. There is a fricative *v* in addition to an approximant *w* without explanation. There is a dental trill *r*, a dental approximant *rr* and a retroflex approximant *r̥* with the explanation, "The exact

phonetic nature of the two *r*-sounds is unknown. ... Both *r*-phonemes may have had retroflex allophones” (2009:382–383). As well as can be determined, the inventory proposed here differs from those given in Emmerick 1981, 1989 and 2009 in at least the following respects:

**Table 2.4:** Differences between Emmerick and Hitch

Graph	Emmerick	Value here
<i>p</i>	/p/	/b/
<i>b</i> in groups	/b/ peripheral	/b/
<i>b</i> single	/b/	/β/
<i>g</i> in groups	/ɣ/	/g/
<i>g</i> single	/ɣ/ peripheral	[ʔ] EWKh /g/, OKh /d/
<i>l</i>	/lʲ/	/l/
<i>ky</i> and <i>c</i>	/č/	/kʲ/ and /č/ > /č/ <sup>6</sup>
<i>gy</i> and <i>j</i>	/ǰ/	/gʲ/ and /ǰ/ > /ǰ/
<i>r</i>	dental trill	/r̥/ retroflex
<i>rr</i>	dental approximant	/r̩/ alveolar
<i>ṅ</i>	/ŋ/ peripheral	/n/ [ŋ] allophone of /n/
<i>ḳ</i>	two segments	/ṭʰ/
<i>hv</i>	—	/hʷ/
<i>’</i>	breathed syllable (2009:381)	— vowel sequence

The discussion which follows will defend these values.

---

<sup>6</sup> *ky* appears to have been distinct from *c*, and *gy* from *j* in AOKh. See 2.2.2.1 below.

## 2.2.0 Synchronic umlaut: evidence for aspiration, retroflexion and palatality

One of the most widespread morphophonological processes in Khotanese is umlauting. In this process, a stem vowel or consonant may be modified when particular suffixes are attached. Compare, for example, the locative singular feminine forms *kīntha*<sup>7</sup> < *kanthā*- ‘city’, *ūca* < *ūtcā*- ‘water’, and the denominal verbs *parś*- ‘to serve’ < *parysa*- ‘servant’, *ātīm*- ‘to desire’ < *ātama*- ‘desire’. The umlaut data are complex. Some stems will show umlaut, others will not. Some stems will show vowel umlaut, others consonant umlaut.

The pervasiveness and intricacy of the process can be seen in *Saka Grammatical Studies* (SGS; Emmerick 1968a), which is still an essential reference on Khotanese morphology. The descriptions of umlaut there are partly diachronic and partly synchronic. The effects of umlaut are listed in different places, for different morphological categories. A unified rule is not recognized.

It is clear that there is a single synchronic umlaut process in Khotanese. Its understanding and description depend on the understanding and description of the phonological system. The analysis in Emmerick 1981 of the consonant inventory did not permit a simple, unified umlaut rule to be sketched.

### 2.2.1 Chinese in Khotanese Brāhmī

One of the manuscripts in Late Khotanese Brāhmī purchased by Stein in Dunhuang has turned out to be a transcription, in Brāhmī characters, of Buddhist materials in

---

<sup>7</sup> Khotanese forms given without reference may be found in the appropriate place in SGS.

Middle Chinese language. Beginning in 1937 with F.W. Thomas, the document has been studied by a range of scholars, each making contributions (summarised by Emmerick and Pulleyblank 1993:3–5). In 1993 R. E. Emmerick and E. G. Pulleyblank, a scholar of historical Chinese language, published what is now the definitive study (ibid.).<sup>8</sup> The chart below sketches the correspondences determined by Pulleyblank and Emmerick between Late Middle Chinese aspirate initials<sup>9</sup> and LKh graphs as well as between two sets of LMC fricative initials and LKh graphs (ibid.:31).

**Table 2.5:** Late Middle Chinese aspirates and fricatives in LKh writing

LKh		LMC	LKh		LMC
graph		aspirate	graph		fricative
<i>kh</i>	=	/k <sup>h</sup> , k <sup>h̄</sup> /	<i>h:</i>	=	/x, x <sup>h̄</sup> /
<i>ch</i>	=	/t <sup>sh</sup> , t <sup>r̄h̄</sup> /			
<i>th</i>	=	/t <sup>h</sup> , t <sup>h̄</sup> /			
<i>ph</i>	=	/p <sup>h</sup> , p <sup>h̄</sup> /	<i>hv:</i>	=	/f, f <sup>h̄</sup> /
<i>ts</i>	=	/t <sup>sh</sup> , t <sup>s̄h̄</sup> /			

The transcriptions of the LMC aspirate initials, both voiceless, e.g., /k<sup>h</sup>, t<sup>sh</sup>, t<sup>h</sup>, p<sup>h</sup>, t<sup>sh̄</sup>/, and voiced, e.g. /k<sup>h̄</sup>, t<sup>r̄h̄</sup>, t<sup>h̄</sup>, p<sup>h̄</sup>, t<sup>s̄h̄</sup>/, with the LKh aspirate graphs *kh*, *ch*, *th*, *ph*

---

<sup>8</sup> In 1984 I obtained a draft of Pulleyblank 1986, a study on this text and its implications for Late Khotanese phonology. That study was not published, being succeeded by the 1993 work with Emmerick. Pulleyblank's reanalysis of Late Khotanese values essentially did not change between 1984 and 1993, although some symbology did.

<sup>9</sup> To be consistent with the phonological transcriptions used throughout this study, the transcriptions of LMC differ in some minor ways from Pulleyblank's system: the retroflex sibilant is transcribed *ʃ* rather than the *s*-hook *ʂ* of the International Phonetic Alphabet, the high front vowel is rendered *ü* rather than the IPA *y*, and the double (or triple) symbols are given with the second (and third) symbols superscript, e.g., /t<sup>h</sup>, t<sup>r̄h̄</sup>/ instead of /th, trh/.



and the digraph *ts* suggest that these signs stood for aspirate sounds. The transcription of LMC /x, x<sup>h</sup>/ by *h*: and not by *kh*, as well as of LMC /f, f<sup>h</sup>/ by *hv*: and not by *ph*, proves that *kh* and *ph* did not mark fricatives. It seems unavoidable to conclude that the Late Khotanese aspirate graphs denote aspirates.

The following chart shows the spellings of LMC retroflex initials (*ibid.*).

**Table 2.6:** Late Middle Chinese retroflex initials in LKh writing

LKh		LMC		LKh
retroflex		retroflex		palatal
graph		segment		graph
<i>ṭ</i>	=	/t <sup>r</sup> /	=	<i>c</i>
<i>kṣ</i>	=	/t <sup>r</sup> ḥ/	=	<i>ch, c</i>
<i>ṣ</i>	=	/ṣ/	=	<i>ś</i>
—	=	/ṣ <sup>h</sup> /	=	<i>ś</i>
—	=	/t <sup>s</sup> /	=	<i>c</i>
—	=	/t <sup>s</sup> ḥ/	=	<i>ch</i>
—	=	/r/	=	<i>ś'</i> [ž]
<i>(d</i>	=	/l/ before [i, ü] elsewhere		LKh <i>l</i> for LMC /l/)

Three of the four retroflex LKh graphs in use denote only LMC retroflex consonants: LKh *ṭ* = LMC /t<sup>r</sup>/, LKh *kṣ* = LMC /t<sup>r</sup>ḥ/, and LKh *ṣ* = LMC /ṣ/. These spellings strongly suggest that Late Khotanese had phonetic retroflexion. The LMC retroflex sounds could also be transcribed with palatal symbols *c, ch, j, ś, ś'*. This is likely related to the fact that Early Middle Chinese palatals and retroflexes merge in LMC (Pulleyblank 1984:65-66, 162, 169, 171, 232). The LMC retroflex phonemes

would have had palatal allophones. Also, the LMC retroflex stops were probably somewhat phonetically affricated (ibid:67) which would contribute to the impression that they were similar to palatal affricates.

The evidence of the fourth retroflex graph, LKh *ḍ*, does not follow the pattern of the other three. It represents LMC /l/ before [i, ü], while LKh *l* = LMC /l/ elsewhere. This evidence suggests that LKh *ḍ* was a palatalized [lʲ] or a palatal [ʎ] while LKh *l* was a plain [l].

### 2.2.2 Nine key segments in umlaut

In 1985 I experimented with phonological values for Old Khotanese with an aim to finding a single rule for umlaut. By combining the lists in SGS of forms which umlaut and those which do not umlaut it was possible to see how each segment behaved with regard to the process. But this information by itself was not quite enough to argue for phonological values for segments. A turning point came with a rereading of Emmerick 1981. A form is listed there which did not appear in SGS: *ggarcha* LS < *ggarkha*- 'heavy' (195). A striking pattern emerged. It was clear that nine segments were related in an elegant fashion by the umlaut process: *k* and *tc* umlaut to *c*, *gg* and *js* umlaut to *j*, and *kh* and *ts* umlaut to *ch*:

**Table 2.7:** Nine segments related by umlaut

Examples				plain	umlauted
<i>byūka-</i>	>	<i>byūca</i>	LSm	'room'	<i>k</i>
<i>khārgga-</i>	>	<i>khārja</i>	LSm	'mud'	<i>gg</i>
<i>ggarkha-</i>	>	<i>ggarcha</i>	LSm	'heavy'	<i>kh</i>
<i>ūtā-</i>	>	<i>ūce</i>	GDSf	'water'	<i>tc</i>
<i>haṃgūjsa-</i>	>	<i>haṃgūj-</i>	denom.vb.	'meet'	<i>js</i>
<i>ggaṃtsa-</i>	>	<i>ggaṃcha</i>	LSm	'hole'	<i>ts</i>

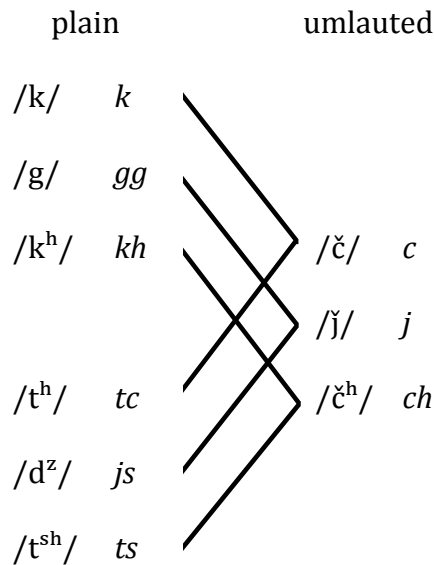
This pattern reveals that the segments fall into six natural classes:

**Table 2.8:** Six natural classes in umlaut

place:	velar	palatal	dental affricate
	<i>k, gg, kh</i>	<i>c, j, ch</i>	<i>tc, js, ts</i>
manner:	voiceless	voiced	?
	<i>k, c, tc</i>	<i>gg, j, js</i>	<i>kh, ch, ts</i>

It has long been recognized that Old Khotanese had velars, palatals, dental affricates, and voiced and voiceless segments. The class of segments represented by *kh, ch* and *ts* also must have a common distinguishing feature. The obvious explanation is that they are aspirates. No other value makes sense. They cannot be fricatives (*kh* is not \*[χ]) or stops with a palatal offglide (*ts* is not \*[t<sup>sy</sup>], and *ch* is not \*[ç<sup>y</sup>]):

**Table 2.9:** Velar and palatal phonemes in umlaut



### 2.2.2.1 AOKh and COKh *ky*, *c*, *gy*, and *j*

In Earliest Written Khotanese, the hypothetical linguistic stage when the writing system was first adapted to Khotanese, *ky* and *gy* were perhaps clearly distinguished from *c* and *j*. If so, the symbols *ky* and *gy* likely denoted either velars with a palatal offglide, i.e., /*k<sup>y</sup>*, *g<sup>y</sup>*/ or true palatals (as opposed to alveopalatals). It is hard to imagine a motivation for creating digraphs *ky* and *gy* if the sounds they represented were not distinct from *c* and *j*. Within OKh, *c* and *ky* merge in /č/, while *j* and *gy* merge in /ǰ/. The longest text with the most conservative or archaic orthography, the *Śūraṅgamasamādhisūtra* (Śgs), has *kye* ‘who’ with *ky*- all 18 times (*kye* 16×, *kyi* 2×; Śgs:121a; +11× in the new folios), while the later *Book of Zambasta* (Z) sometimes has *ce* (about 75×) for *kye* (occurs about 150×). Similarly, Śgs has *kyerä* with *ky*- all 3 times, while Z has *c*- 6 times and *ky*- 16 times. The merger of /ǰ/ with /*g<sup>y</sup>*/ is suggested by the alternants NSm *gyagarrä* Z 20.35~*jatarrä* Z 20.55

‘liver’; NAPf *gyahe* Z 3.42, 17.22~ GDSf *jahe* Z 20.5 ‘fountain, spring’. The word *padamgyā-* has consistent *gy* in Śgs 6× (124; + IOL Khot 187/2v2), while in Z it has *gy* 21× and *j* 5×. ‘The particle *ju* has *j-* consistently 5 times in Śgs (121; +4× in the new folios) while in Z there is one case of inverse spelling *gyu* 24.79 among the very frequent instances of *ju* (Glossar:427). The religious word *gyasta-* ‘divine; god’ and its derivatives have *gy-* in Z around 200 times and *jasta* just twice 4.31, 14.90, perhaps because the spelling *gyasta-* itself had iconic value, since it is frequent in the phrase *gyasta- balysa-* ‘deva Buddha’.

If /kʸ, gʸ/ were distinct from /č, ĵ/ in Earliest Written Khotanese, then in the umlaut process at that linguistic stage, *k* and *g* may have regularly become *ky* and *gy*, not *c* and *j*, cf. LSm *bilsamgya* Z 13.84, 15.41, 22.308, 24.463 2× < *bilsamgga-* ‘bhikṣu-saṃgha’. I find no Old Khotanese examples of *k*>*ky* umlaut but R.E. Emmerick pointed out LKh *pamākye jsa* IASf < *pamākā-* ‘measure’ (personal communication; cf. DKS:213a).

Thus there may originally have been two distinct processes, i.e., velar palatalization [k > kʸ] and dental palatalization, e.g., [s > š]. The velars may have become velars with a palatal offglide, or perhaps true palatals, while the dentals became alveopalatals. This would allow observation of two natural classes, undergoing two parallel types of umlaut: the velars becoming palatalized velars or true palatals, and all dentals which are not oral stops becoming alveopalatals.

O. von Hinüber noted the spellings *Kaligya* and *paralokikye* in the Aśokan inscription from Kalsi, and suggested a Middle Indian model for *ky*, *gy* as variants of

*c, j* (1981:124). However, he was not aware of the distribution in AOKh and was under the impression that *ky* and *gy* were always spelling variants of *c* and *j*.

### 2.2.3 Synchronic umlaut rule:<sup>10</sup>

It has been possible to organize most of the segments of Old Khotanese according to their behavior in umlaut and to define the umlaut rule itself. These advances help further determine natural classes. Consonants fall into three natural classes according to their behavior in umlaut.

When an umlaut-triggering suffix attaches to a morpheme, the potential to umlaut is drawn into that morpheme towards the stressed vowel. Morpheme final consonants behave in one of three ways:

*Umlautable* consonants are palatalized, e.g., *khāysa-* + *-i* > *khāśa* LSm ‘food’,

*khārgga-* + *-i* > *khārja* LSm ‘mud’, *byūs-* + *-tā* > *byūštā* 3S.pr.a. B ‘to dawn’;

*Transparent* consonants allow the umlaut potential to pass, e.g., *ūra-* + *-i* > *uīra*

LSm ‘womb’, *hvata-* + *-i* > *hvīte* infinitive ‘to speak’, *bāḍa-* + *-i* > *beḍa* LSm ‘time’;

*Absorbing* consonants neutralize the umlaut potential so that there is no trace of it

on the surface, e.g., *drūjā-* + *-i* > *drūje jsa* IASf ‘lie’, *paśś-* + *-i* > *paśśā* 2S.pr.a.

‘to let go, release’, *cā’yi-* + *-i* > *cā’ye* GDSf ‘magic’.

---

<sup>10</sup> cf. Hitch 1990:182.

## 2.2.4 Consonant behavior in umlaut<sup>11</sup>

**Table 2.10:** Consonant behavior in umlaut

Umlautable	Absorbing or Umlauted	Neutral
/k/ <i>k</i>		/t/ <i>tt, t</i>
/g/ <i>gg, g</i>		/d/ <i>t, d</i>
/k <sup>h</sup> / <i>kh</i>		/t <sup>h</sup> / <i>th</i>
/t <sup>h</sup> / <i>tc</i>	/č/ <i>c, ky</i>	/t/ <i>t̥</i>
/dz/ <i>js</i>	/j/ <i>j, gy</i>	/d̥/ <i>d̥</i>
/t <sup>sh</sup> / <i>ts</i>	/č <sup>h</sup> / <i>ch</i> <sup>①</sup>	/t <sup>h</sup> / <i>t̥h, t̥th</i>
	(umlauted only)	
/s/ <i>s</i>	/š/ <i>śś, ś</i>	/s/ <i>śś, ś</i>
/z/ <i>ys</i>	/ž/ <i>ś</i>	/z/ <i>ś</i>
/n/ <i>n, ŋ</i>	/ñ/ <i>ñ</i>	/nd/ <i>nd</i> <sup>②</sup>
		/nt <sup>h</sup> / <i>nth, ŋth</i> <sup>②</sup>
	/y/ <i>y</i> <sup>①</sup>	
	(absorbing only)	/r/ <i>rr, r</i>
		/r̥/ <i>r</i>
		/l/ <i>l</i>
		/w/ <i>v</i>
		/m/ <i>m</i>
		/h/ <i>h</i>

[Category unknown: /δ/ *d*; /β/ *b*; /b/ *p, b*; /p<sup>h</sup>/ *ph*; /h<sup>w</sup>/ *hv*; /t̥<sup>s</sup>/ *k̥s*]<sup>③</sup>

### Notes to table 2.10

- ① /č<sup>h</sup>/ *ch* is only attested as an umlauted segment. /y/ *y* is only attested as an absorbing segment.
- ② /n/ *n, ŋ* when followed by a dental stop, /nd/ *nd* or /nt<sup>h</sup>/ *nth, ŋth*, remains assimilated to the stop and so, likewise, is transparent.
- ③ The synchronic behavior of these segments has not been observed. Diachronically, *ph*, or its historical antecedent, is transparent, cf., *paltcīmph-* ‘to check’ < \**pari-sčambaya-* (SGS:76).

<sup>11</sup> cf. Hitch 1990:183.

#### 2.2.4.1 Palatality

As can be seen on the chart, the consonants that result from umlauting, *c, j, ch, śś, ś, ñ*, are the same as those which neutralize it, e.g., *ñ* is the result of umlaut in *hīñe jsa* < *hīnā- + 'e jsa* GDSf 'army', while it absorbs umlaut in *armūtañe* < *armūtañā- + -'e* GDSf 'congratulation'. It is highly likely that *ñ* and the other segments in this class are palatal or alveopalatal. By the same token, *şş* and *ş* can not be regarded as palatal because they do not belong to this class.

#### 2.2.4.2 Umlautability

Four types of consonants are umlautable, the velar stops, *k, gg, kh*, the dental affricates, *tc, js, ts*, the dental sibilants, *s, ys*, and the dental nasal *n~ṇ*. These segments do not appear to form a single, coherent, readily definable class in COKh. In contrast, in EWKh and perhaps AOKh there may have been two distinct processes, velar palatalization, e.g. [*k > kʲ*], and dental palatalization, e.g., [*s > š*]. The velars may have become velars with a palatal offglide, or perhaps true palatals, while the dentals became alveopalatals (see 2.2.2.1). This would allow observation of two natural classes, undergoing two parallel types of umlaut: the velars becoming palatalized velars or true palatals, and all dentals which are not oral stops becoming alveopalatals.

#### 2.2.4.3 Transparency

Fourteen consonants are observed as being transparent to umlaut: *tt, t, th, ʈ, ɖ, ʈh, şş, ş, rr, r, l, v, m, h*. The basic phonological character of some of these has been clear since the beginning of Khotanese studies: *tt, t, th* are non-nasal dental obstruents while *v = /w/, m = /m/* and *h = /h/*. The value of the retroflex graphs, *ʈ, ɖ, ʈh, şş, ş*,



and of the liquid graphs *rr*, *r*, *l*, has not been clear. Nevertheless, the umlaut pattern allows certain things to be established. As mentioned, the retroflex graphs *ṣṣ*, *ṣ* cannot mark palatals [*\*š*, *\*ž*] as in Emmerick 1981 because these sibilants are transparent to umlaut. If palatal, they would absorb the umlaut potential, and they would be produced by umlauting. It is also unlikely that the segments denoted by *ṭ*, *ḍ*, *ṭh* would indicate palatalized [*\*tʲ*, *\*dʲ*, *\*θʲ*] as in Emmerick 1981 because they too are transparent to umlaut.

Emmerick 1981:206-207 and 1989:209 regard *l* as /*lʲ*/. As *l* is transparent to umlaut, it cannot have a palatal offglide. It is also typologically unlikely for a language to have /*lʲ*/ without a corresponding plain /*l*/. In Old Khotanese, *l* must be simply /*l*/.

### **2.3 Retroflexion in Old Khotanese**

The Khotanese Brāhmī ultimately comes from an Indian language which had phonemic retroflexion. Because of this origin, the mere use of the retroflex graphs *ṭ*, *ḍ*, *ṭh*, *ṣṣ*, *ṣ* suggests by itself that Old Khotanese also had phonemic retroflexion. The use of *ṭ*, *ṣ*, *kṣ* to write Middle Chinese cerebrals in Late Khotanese transcription (see 2.2.1 and table 2.6) suggests that retroflexion may have been a feature of Late Khotanese and perhaps Old Khotanese. But there is other evidence at hand.

#### **2.3.1 Synchronic evidence for retroflexion**

In the third person singular of the present indicative of type B verbs, there is a synchronic morphophonological process which suggests that retroflexion was a

feature of Old Khotanese. Stem final *-r-* blends with suffix initial *-tt-* to give *-ḍ-* (/ḍḍ/) while stem final *-rr-* + *-tt-* results in *-rrd-* or *-rd-* (/ɹd/).

**Table 2.11:** Stem final *-r-* + *-tt-* and *-rr-* + *-tt-*

	Middle	<i>-tte</i>	(no umlaut)	Active	<i>-<sup>i</sup>ttä</i>	(umlaut)
<i>-r-</i>	<i>bar-</i>	> <i>baḍe</i>	ride	<i>bar-</i>	> <i>bīḍä</i>	carry
	<i>handār-</i>	> <i>handāḍe</i>	care for	<i>bār-</i>	> <i>beḍä</i>	rain
<i>-rr-</i>	<i>nāhvarr-</i>	> <i>nāhvarrde</i>	long for	<i>purr-</i>	> <i>purḍä</i>	overcome
(LKh	<i>naṣkirr-</i>	> <i>naṣkirrdä</i>	cut off)	<i>hamurr-</i>	> <i>hamurrdä</i>	crash

Diachronically, the development of the past passive participle (ppp) shows similar consonantism. Stems in *-r-* have a ppp in *-ḍ-*, e.g., *bār-:bāḍa-* ‘rain’, *bar-:buḍa-* ‘carry’, *haphār-:haphaḍa-* ‘be distracted’. Stems in *-rr-* have a ppp in *-rrd-*, e.g., *baṣṭarr-:baṣṭarrda-* ‘bestrew’, *nāhvarr-:nīhvarrda-* ‘long for’, *purr-:purrdä-* ‘overcome’.

The contrast between the assimilation *-r- + -t- > -ḍ-* and the preservation *-rr- + -t- > -rrd-* suggests that *rr* is not phonemically equivalent to *r + r* as suggested in Emmerick 1981:209 and apparently followed in Emmerick 1989:209, but not by Emmerick 2009 where he stated, “The exact phonetic nature of the two r-sounds is unknown” (383). Rather, the assimilation implies that *r* and *ḍ* share a feature which is not shared by *t* or *rr*. The simplest explanation may be that *r* and *ḍ* are retroflex, i.e. with present stems, /r + tt > ḍḍ/, while *rr* and *t~d* are dental (or alveolar), i.e. /ɹ + tt > ɹd/. In the assimilated *ḍ*, the retroflexion of the first segment, *r*, is preserved.

In this way, information about the original structure of the stem is retained in the inflected form.

### 2.3.2 Typology and /ɹ~r̥/

Some other South-Eastern Iranian languages containing retroflex obstruents also have retroflex liquids. Pashto has two *r*-sounds or rhotics, dental and retroflex (Grjunberg-Edel'man 1987:20), as does the lower dialect of Munja (Grjunberg 1987a:166). Ishkashmi has one *r*-sound but both *l* and *l̥* (Pakhalina 1987b:494). Other languages with retroflex obstruents and two *r*-sounds, dental/alveolar and retroflex, include Kariera-Ngarluma, Arabana-Wanganura and Maung from Australia, Kaliai from New Britain, the Dravidian Kota and the Indic Bengali and Punjabi (Maddieson 1984: s.v.). It is typologically plausible that Old Khotanese had dental or alveolar /ɹ/ and retroflex /r̥/.

Besides casting light on place of articulation, typology may also help illuminate the manner of articulation of the two Khotanese rhotics. Peter Ladefoged and Ian Maddieson note that, "The most prototypical members of the class of rhotics are trills made with the tip or blade of the tongue (IPA r)" (1996:214). If all other evidence were lacking, Khotanese dental~alveolar *rr* would most likely be a trill. Some languages may be typological models for the Khotanese rhotics. Hausa contrasts an alveolar trill with a post-alveolar flap or approximant (ibid:237). Australitan Arrernte contrasts an alveolar trill with a post-alveolar approximant (ibid:238-9). Djapu seems to have an alveolar tap and a retroflex approximant. If Khotanese *rr* is a trill and *r* is an approximant, this may explain the *rr*-device as a phonetic approximation of a trill.

It may be that the contrast between the Khotanese dental~alveolar trill and the retroflex approximant is neutralized in some consonant clusters as suggested by spelling alternations, e.g., 3Ppf.tr.f *tcabriye* Z 2.44 ~ 3Ppf.tr.m *tcabbrīya* Z 20.46, *drrai* Z 16× ~ *drai* 5× ‘three’, *purrdu yan-* Z 2.206 ~ *purdu yan-* Z 22.236 ‘able to overcome’, NSf *baṣṭarrgya* Z 3.41 ~ ASf *baṣṭargyo* Z 2.49 ‘carpet’, NAPm *cakrra* Z 4.60 ~ *cakra* 5.74 ‘discus’, 3Pp.a *prravaindi* Z 22.205 ~ *pravaindä* Z 22.207 ‘undertake *pravrajyā*’, NAPm *baṣṭarrda* Sgh[10,24] 66.3,4,6,6,7 ~ *baṣṭarda* Sgh[10] 66.5, Z 22.136 ‘spread’. The contrast is possibly maintained before *n*. Although there are alternations like IASm *baṣṭarna* Sgh[9] 125.3 ~ *baṣṭarrna* Sgh[18] 128.3 ‘blanket (?)’ and NSf *urrna* Z 23.40 ~ IASf *urñe* jsa 2.56 ‘ūrñā’, some words seem to have more or less consistently one or the other. The word *pūrna-* ‘arrow’ occurs eight times in Old Khotanese, always with *rn* (6× in Z, Sgh[17] 243.27, IOL Khot 17/9v2). The word *ysarrnaa-* ‘golden’ appears in Z nine times with *rr* and once with *r* which might be copyist influence (NSm *ysarnai* Z 2.72).<sup>12</sup> At least one cluster with a rhotic has restricted distribution. In Z, a cluster *(t)tr* occurs 613 times while *\*(t)trr* does not occur at all. The latter appears to be restricted to Late Khotanese.

It is common for rhotics to have allophones. Persian /r/ is a voiced trill in initial position, a flap between vowels, and a voiceless trill at the end of a word. In Fula the rhotic is an approximant before a consonant and a trill elsewhere. In Palauan one rhotic is an approximant in initial position and a tap in intervocalic and postvocalic

---

<sup>12</sup> An interesting case is IASm *turrāna* Sgh[8,10] 215.4, [2] 254.2 ‘mouth’ which has the ä-elided form *turrna* in Z 2.234, 13.109, 24.515 and once *turrrna* Z 20.57 with diacritic *r* above radical *rr*. This might also show a copyist trying to write *r* in this cluster.

environments while the other, written *rr* “is most commonly an approximant with some frication, but its range of variation encompasses trills” (Ladefoged and Maddieson 1996:237). Palauan approximant *r* and trill *rr* might provide something of an orthographic and linguistic parallel to the Khotanese. The commonness of rhotic allophony in the world’s languages makes it possible that the Old Khotanese rhotics also had allophones which may partly explain the spellings in clusters. Perhaps the distinction between the two was neutralized in certain clusters and the phonetic nature of the sound was not clearly closer to the initial or intervocalic allophones of either.

### **2.3.3 Loanword evidence**

As noted by Emmerick, “In initial position *r* occurs only in loanwords e.g. *ratana-* ‘jewel’, cf. Pali *ratana-*, and in the enclitic *ro* reduced from *rro* ‘also (1981:207). The single *r ro* appears to be LKh influence. In Z, *rro* occurs 127 times and *ro* 16 times. Late Khotanese texts have just *r-*. The LKh *Jātakastava* has just single *r* forms *ra*, *rā* (Dresden 1955:483), and the *Bhadracaryādeśanā* has just *ra*, *ri* (Asmussen 1961:77). The fact that initial *r* in Old Khotanese occurs only in words of Indian origin may be informative. In Sanskrit, as is well known, *r* is retroflex, so an initial approximation of the *r* in Khotanese might be retroflex.

#### **2.3.3.1 History of the Khotanese rhotics**

Apparently, Old Iranian initial *\*r-*, whatever the phonetic implications, became an alveolar rhotic, likely a trill in Khotanese. It may have then been the massive influx of Indian words with an initial retroflex rhotic which helped give rise to the contrast

in rhotics. When the orthography was being developed, *r-* was kept in Indian loans, and the new device *rr-* created for inherited vocabulary.

In intervocalic position the origin of the two rhotics is more complex. There is *VrV* in both inherited and borrowed vocabulary, e.g. inherited *ysāra-* ‘1000’ cf. Av *hazaŋra, ggara-* ‘mountain’ cf. Av *gairi-*, borrowed *āśiria-* ‘teacher’, cf. Skt *ācārya, maraṇa-* ‘death’, cf. Gdh *marana*. In contrast intervocalic *rr* is found only in inherited vocabulary where it corresponds to *\*rn* or *\*rm*, e.g. *bastarr-* ‘bestrew’ cf. Av *stārānā, purr-* ‘overcome’ cf. Av *pārānā-* ‘fight’, *tarra-* ‘grass’ cf. OInd *tṛṇa-* ‘herb’, *ysurri-* ‘anger’ cf. Av *zar-* ‘to anger’ and Skt *hṛṇī-* ‘anger’, *pārra-* ‘feather’ cf. Av *parāna-, ārra-* ‘fault’ cf. Av *arāna-, ārra-* arm, palm cf. Av *arāma-* ‘arm’, *kārra-* ‘deaf’ cf. Av *karāna* ‘deaf’.

The development of the initial and intervocalic rhotics is summarized in the table below.

**Table 2.12:** History of initial and intervocalic rhotics

<b>Initial</b>		<b>Intervocalic</b>		
PIr <i>*r</i>	Ind <i>r-</i>	PIr <i>*r</i>	Ind <i>r</i>	PIr <i>*rn</i>
		∨		
<i>rr</i>	<i>r</i>	<i>r</i>		<i>rr</i>
<i>rraha-</i>	<i>rūva-</i>	<i>ggara-</i>	<i>āśiria-</i>	<i>pārra-</i>
‘chariot’	‘ <i>rūpa</i> ’	‘mountain’	‘teacher’	‘feather’
(Cf.:) Av <i>raθa-</i>		Av <i>gairi-</i>	Skt <i>ācārya</i>	Av <i>parāna-</i>
‘chariot’				

### 2.3.4 Diachronic evidence for retroflexion

#### 2.3.4.1 *ṭ*

An Old Iranian *\*r* in contact with OIr *\*t* or *\*θ* gives Old Khotanese *ḍ* or *ṭh*, but not, apparently, *ṭ*. In OKh, *ṭ* is extremely rare except in the combination *ṣṭ* which derives from *\*št*, e.g., *rraṣṭa* ‘right’, cf. Av *rašta-*; *haṣṭa* ‘eight’, cf. Av *ašta* (Emmerick 1981:190). Bailey in DKS lists three words with intervocalic *-ṭ-* which he derives from *\*-xt<sup>i</sup>-*, *bekhaṭe* ‘in digging in’ < *\*abi-kauxyā-* (302a), *rrauṭā-* ‘desire’ < *\*rauxtyā-* (369b), *ṣkūṭa* ‘in the throat’ < *\*skauxti-* (or *\*skauti*) (414-415). Emmerick 1981 repeats the first two etymologies as evidence ‘of the palatal character of *ṭ*’ (190).

An explanation of the intervocalic *ṭ* in these words may be related to that of the origin of *ṣṣ*. The ‘most common source of Kh *ṣṣ* is OIr *\*sr-* e.g. *ṣṣūni-* ‘thigh’, cf. Av *sraoni-*’ (Emmerick 1981:197). The cluster contracts, with the retroflexion of the original liquid being retained in the resulting sibilant. But *ṣṣ* also develops from *\*xš*, either before *a*, *ā* or intervocalically (Emmerick 1989:215, §20) e.g. *ṣṣavā-* ‘night’ < *\*xšapā-*, cf. OInd *kṣapā-*, Av *xšap-*; *huṣṣ-* ‘to grow’ < *\*uxša-*, cf. Av *vaxš-*, *uxšya-* (and possibly from *\*fš*; Emmerick 1981:196). It may be that the development of intervocalic *ṭ* in *rrauṭa-*, *bekhaṭe*, *ṣkuṭa* is parallel to this development of intervocalic *ṣṣ*. Just as *\*xš* > /ṣ/, so too *\*xt* > /ṭ/. In both cases the presence of the *\*x* may be related somehow to the retroflexion of the following segment.

#### 2.3.4.2 *l* and *ḍ*

The historical origins of OKh *l* and *ḍ* have always been well understood.

OKh *-l-* < OIr *\*-rd-*

Khotanese		cf. Avestan
<i>salīā-</i>	‘year’	<i>sarəd-</i>
<i>kamala-</i>	‘head’	<i>kamərəδa-</i>

OKh *-ḍ-* < OIr *\*-rt-*

Khotanese		cf. Avestan
<i>sāḍa-</i>	‘cold’	<i>sarəta-</i>
<i>ysāḍa-</i>	‘old’	<i>zarəta-</i> (Emmerick 1981:191)

This pattern is only partly helpful for determining values. While the retroflex value of *ḍ* was established above, it is unlikely that *l* is retroflex. Its development may be parallel to that for *rr* between vowels, i.e., as OIr *\*VrnV* > OKh *VrrV* (dental), so OIr *\*VrdV* > OKh *VlV* (dental).

### 2.3.4.3 ***ṭh(ṭh)***

The historical origin of OKh *ṭh* has not been so clear. Sten Konow proposed that OKh *haṭhā-* ‘truth’ was derived from OIr *\*haṭhya-* ‘true’ (1932:8-9). This etymology appeared rather obvious and inevitable. It understandably led Konow and later Emmerick to the assumption *\*ṭy* > *ṭh*, and thus that *ṭh* was [θʸ]. Verb stems with final *-ṭh-* were thought to require an etymology with *\*-ṭy-*, e.g., *bīṭh-* ‘to twist (tr.)’ < *\*varṭhya-* < *\*vart-* (cf. SGS:100; DKS:281b). In this theory, because the *\*y* caused the retroflexion in *\*ṭy* > *ṭh*, the historical *\*r* then would leave no trace. This would be unlikely given the developments *\*rd* > *l*, *\*rt* > *ḍ*, *\*rn* > *rr* and *\*sr* > *ṣṣ* listed above. A more likely evolution of *ṭh* would be *\*rṭ* > *ṭh*. Some reasonably straightforward examples of this are available:



OKh *-tḥ-* < OIr *\*-rθ-*

*baṭha-* ‘breastplate’ < *\*varθa-*, cf. Av *varəθa-*. Bailey gives < *\*varθra-* with *\*θr* instead of *\*θy* giving *tḥ* (DKS:266b).

*paṭha-*: *paṭhuta-* ‘to burn (tr.)’ < *\*parθu* < *\*pari-θu/θau-*, cf. Christian Sogdian *prθwty* ‘burned, singed’ (Sims-Williams 1985:220), Manichean Sogdian *prδ’w* ‘flame’ (DKS:203a).

*\*guvaṭhu-*: *guvaṭhuta-* as above with < *\*vi-*. *guvaṭhuta* NAPm Z 2.43, Emmerick: ‘half-burnt’; Bailey: ‘burnt in various parts’ (DKS:88b).

*suthṭha-* Bailey ‘bird of prey, rapax’ (DKS:427a). Connected with Sogdian *swrty* and Persian *ṣurad* (N. Sims-Williams apud Emmerick St.III:155).<sup>13</sup>

LKh *bīṭh-* ‘writhe, twist’, LKh *hambīṭh-* ‘retain’ < Pre-Khotanese *\*(ham)bart<sup>h-i</sup>* < OIr *\*uarθya-* < *\*uart-* ‘turn’ (EDIV:423–424).

*hasamīṭh-* ‘destroy’ has unclear etymology (cf. SGS:151, 252, DKS:473–474, but note EDIV:179 *\*Hmard-* ‘crush’)

This gives us three parallel developments from Old Iranian to Old Khotanese involving *\*r* and a following dental obstruent: *\*rd > l*, *\*rt > ḍ*, and *\*rθ > tḥ*. The last two parallels are even clearer in intransitive~transitive verb pairs which feature Pre-Khotanese *\*-art-* > *-aḍ-* in the intransitive and *\*-arθy-* > *-īṭh-* in the transitive. In the following examples, the *\*r* causes retroflexion in both forms while the *\*y* in the transitive causes *\*a > ī* historical umlaut, not retroflexion.

---

<sup>13</sup> Attested eight times, always with double *tḥtḥ*, and always in meter. As shown below in 2.9.2, the meter solidly demonstrates that the double *tḥtḥ* denotes a geminate.

<i>baḍ-</i>	writhe (intr.)	< * <i>vart-</i> , cf. Av <i>varət-</i> ‘turn’
<i>bīṭh-</i>	twist (tr)	< * <i>varθya</i> < * <i>vart-</i>
<i>ggaḍ-</i>	lie about	< * <i>gart-</i> , cf. Pahlavi <i>gartītan</i> (DKS:78b)
<i>pajsīṭh-</i>	store away	< * <i>pajarθya-</i> < * <i>pa-gart-</i> (SGS * <i>grθya-</i> )
<i>haṃbaḍ-</i>	be compacted	< * <i>ham-vart-</i> (DKS 461b)
<i>haṃbīṭh-</i>	obstruct	< * <i>ham-varθya-</i> < * <i>ham-vart-</i>
* <i>pasamaḍ</i> - <sup>14</sup>	*be rubbed, anointed	< * <i>pa-sam-mart-</i> , cf. Skt <i>mardana-</i> adj. ‘crushing; rubbing; anointing’ (SGS LW < Pkt <i>maḍḍ-</i> < Skt <i>mard-</i> ; DKS:224a compares Wakhi <i>mānd-</i> , <i>māṇḍ-</i> ).
<i>hasamīṭh-</i>	destroy	< * <i>fra-sam-marθya</i> < * <i>fra-sam-mart-</i> , cf. Av <i>marəd-</i> ‘destroy’ (DKS:474a < * <i>fratsa-mṛθya-</i> < * <i>mart-</i> ‘crush’; SGS:242 ‘unknown origin’) <sup>15</sup>

Another source of *ṭh* /*ṭʰ*/, is after *ṣ* in *ṣṭh* /*ṣṭʰ*/ < \**ṣtʰ* < \**šθ*. The retroflex *ṭh* in *naṣṭhrrīta-* ppp ‘pushed out’ (cf. SGS:51) or ‘drawn’ (DKS:177b), is due to assimilation to the preceding *ṣ* (< \**š*), just as *ṣṭ* < \**ṣt* < \**št* in *haṣṭa* ‘eight’ (cf. Av *ašta*). *naṣṭhrrīta-* shows *ṣṭhrr* < \**šθr*, i.e., < \**niš-θrak-* (SGS:51; EDIV \**θrač-*). The retroflexion is not due to the following \**r* as proven by *haṃthraj-* ‘to oppress’ < \**ham-θrak-* (SGS:140).

---

<sup>14</sup> Attested only in the LKh necessitative participle *pasamaṇḍāñä* ‘is to be rubbed’ which may show nasal assimilation *ḍ* > *ṇḍ* between nasal-initial syllables.

<sup>15</sup> Cf. Skt. *mṛd-* ‘destroy’, *saṃmṛd-* ‘kill’, *pramardana-* ‘destroying’.

#### 2.4.0 Retroflex affricate $k\text{ṣ}$ = /tṣʰ/

Emmerick did not include  $k\text{ṣ}$  in the phoneme list either in 1981 or 1989. In 1989 he appeared to regard it as a sequence of two segments /tṣ/. His view can be seen in the transcription of  $k\text{ṣustā}$ - ‘serum’ with /tṣ/ and in the historical rule OIr \*xš > /tṣ/ except before  $a, \bar{a}$  (p. 215, §19). The value /tṣ/ represents an innovation over the 1981 /xš/, but is not fully explained by Emmerick. In Emmerick 2009 the phoneme chart contains the symbology “ $t\text{ṣ}^h$ ” in the retroflex affricate slot, presumably representing an aspirated retroflex affricate /tṣʰ/, but there is no accompanying explanation. Here it is argued that  $k\text{ṣ}$  has the value /tṣ/, an unaspirated retroflex affricate, in Old Khotanese.

The meter in the *The Book of Zambasta* confirms the single consonant status of most digraphs. There are many instances of  $ys=/z/$ ,  $js=/d^z/$ ,  $tc=/t^s/$ ,  $ts=/t^{sh}/$ ,  $ky=/k^y/$ ,  $\check{c}/$  and  $gy=/g^y/$ ,  $\check{j}/$  in meter showing a single consonant. To prove that  $k\text{ṣ}$  in native vocabulary is a digraph, it must occur in meter in a (non-Sanskritic) word with a short vowel before  $k\text{ṣ}$ . I find only one candidate, but the position is ambiguous. In the hemistich below,  $nikṣūtā$  ‘he promotes’ occurs as LHL in a cadence 1, which is nominally HLLHL.

24.43 trāmu mā|ñamdu □ kye pharu | dātu **nikṣūtā** (B:5+6+7)

Similarly, in the case of one who greatly promotes the Law,

However, at the end of seven mora segments like cadence 1, we often find words shaped HHL, read as if LHL. I have labelled this the *uysnora*-effect (Hitch 2014:15–17). The meter in this case is thus not helpful.

Most words of Indian origin clearly show that the *kṣ* in them counts as two consonants. In the first hemistich below, *yakṣa* is HL in cadence 1 HLLHL, and in the second, *lakṣaṇa* is HLL in cadence 1 HLLHL:

24.487ab bāśśä gyasta | nāga □ aysura | dīvate **yakṣa** (B:5+6+7)

At the gods, Nāgas, Asuras, Devataās, Yakṣas

24.234ab hā yä dīsta | nāte □ biśśī | **lakṣaṇa** spāṣte (B:6+6+7)

He took him up in his hand. He saw all his *lakṣanas*.

But in at least one word ultimately of Indian origin, *akṣara*- ‘syllable’, *kṣ* counts as a single segment. In both hemistichs below *akṣara* counts LLL in cadence 4 LLLLHL:

13.113cd ttīyi vā | **akṣara** hāmī□ru | kho rro jaḍye | hva’ndi nā gāvu

(A:5+7+5+7)

And would the syllables arise then as even from an ignorant man? Not for a moment!

20.58ac **akṣara** hāmāre □ | salāva ye | pyūṣde . □ rraysā (C:7+5+6)

Syllables arise: one hears empty sounds.



The word *akṣara*- may have been borrowed from Gāndhārī as a single segment [t̪ʰ] (see below), or perhaps it is from an early loan layer which has been nativized while later loans from Sanskrit have a [k] + [ṣ] pronunciation. It appears that in Khotanese texts the spelling *kṣ* was ambiguous.

There are several indications that *kṣ* had a retroflex affricate [t̪ʰ] value in Old Khotanese. Historically *kṣ* can be from either \*xš or \*fš, e.g. *kṣustā*- ‘serum’, cf. Av *xšusta*- ‘fluid, liquid’ (DKS:69–70); *kṣāta* ‘six’, cf. Av *xšvaṣ*; *kṣārma*- ‘shame’, cf. Av *fšarāma*- (Emmerick 1981:197); *kṣundaa*- ‘husband’, cf. Av *fšuyant*- ‘husbandman’

(DKS:69b).<sup>16</sup> Since the OIr fricatives \*χ, \*f give OKh aspirates /k<sup>h</sup>, p<sup>h</sup>/, and the sibilant \*š gives OKh /š/ we may have the parallel developments \*χš > \*k<sup>h</sup>š > \*k<sup>s</sup> > t<sup>s</sup> and \*fš > \*p<sup>h</sup>š > \*p<sup>s</sup> > t<sup>s</sup>.

Tumshuqese writing, besides featuring the combination kš, has a new sign corresponding to Kh kš which Konow transliterated χš:

Da aber die Lautverbindung kš in unseren Texten anders geschrieben wird, muß es sich um einen etwas verschiedenen Laut handeln, und vom iranischen Standpunkt kann nur χš in Frage kommen. Kh. kš bezeichnet wahrscheinlich auch diesen Laut. (1935:777)

But from a strictly graphic point of view, the new sign is clearly derived from regular Brāhmī ś by means of the addition of a tail, cf. χša  < śe . This was first

noted by J. Harmatta more than forty-five years ago (1969:106), but the insight has generally been overlooked (cf. Hitch 1984:199 incorrectly suggesting Tq χš < Gāndhārī Kharoṣṭhī kš). Harmatta also proposed that the phonetic value of χš was [tʃ] and compared the value of Gāndhārī c̄h (Brough khkš; 1962:72–73) which denotes [tʃ] as shown by Bailey 1946:770-775.<sup>17</sup>

Typologically, languages with dental and palatal affricates, as well as dental, palatal and retroflex sibilants, may also have retroflex affricates. This is true of the

---

<sup>16</sup>Emmerick 1989 has apparently identified the phonetic conditioning active behind the long-standing puzzle of why OIr \*xš sometimes gives kš and sometimes śš, e.g., śšavā- < \*xšapā- ‘night’; kšāta ‘six’, cf. Av xšvaš. OIr \*xš becomes kš in initial position, but not before \*a or \*ā, while \*xš becomes śš elsewhere (that is, medially, and in initial position before \*a or \*ā; p. 215, §19-20).

<sup>17</sup> Bailey 1946:770-775 undertook an extensive study on Gāndhārī kš involving many languages and scripts. He concluded that the symbol represents ‘a retroflex unaspirated tʃ or the aspirate tʃʰ’ (774). Brough additionally pointed out that ‘the modern related languages maintain separately a retroflex affricate (tʃ, usually transcribed ç) where Sanskrit has kš’ (1962:72, fn. 3).

South-Eastern Iranian languages Munja (Grjunberg 1987a:164), Wakhi (Pakhalina 1987a:420) and Ishkashmi (Pakhalina 1987b:494). These are genetically close to Khotanese. Some languages which are geographically close to Khotanese also feature this pattern, e.g., Dardic Šiṅā (Bailey 1946:773) and Burushaski (Maddieson 1984:s.v.). And then there are languages which share neither geography nor genetic affiliation but have the inventory, like Pekingese, Basque, and Acoma in New Mexico (ibid:s.v.).

**Table 2.13:** Typical inventory with retroflex affricate

	dental	palatal	retroflex
sibilants	/s/	/š/	/ʂ/
affricates	/t <sup>s</sup> /	/č/	/t <sup>ʂ</sup> /

Khotanese is typologically ordinary with regard to possessing /t<sup>ʂ</sup>/.

In Late Khotanese, the OKh affricate *kʂ* /t<sup>ʂ</sup>/ becomes an aspirate [t<sup>h</sup>] and so merges with *th* /t<sup>h</sup>/. This produced a situation where *kʂ* and *th* were spelling equivalents. We therefore find OKh *haṭh(th)ā*- spelled *hakṣā*- in LKh although the pronunciation of the medial consonant in this word probably remained unchanged. This explains the use of LKh *kʂ* = /t<sup>h</sup>/ to write Late Middle Chinese /t<sup>rh</sup>/ (Emmerick and Pulleyblank 1993:20–21). The transcription is phonetically accurate. Symbols for a Khotanese retroflex aspirate are used to write the LMC retroflex voiced aspirate. The corresponding LMC voiceless segment, /t<sup>rh</sup>/ in Pulleyblank 1986 but /tʂ'/ in Emmerick and Pulleyblank 1993, occurs five times but is transcribed by *ch* not *kʂ* (Emmerick and Pulleyblank 1993:22). As mentioned above, this may have to do with a merger of retroflex and palatal segments in late LMC.

Other proof of LKh  $k\check{s}$  = /tʰ/ comes from the spellings of the LKh numeral ‘6’ in Tibetan script. Several Tibetan manuscripts which are paginated with LKh numerals in Tibetan writing show a correspondence LKh  $k\check{s}$  = Tib *thr*, e.g. LKh  $k\check{s}a$  ‘six’ = Tib *thra*. In his monograph on these numerals, Mauro Maggi concludes that ‘Tib. *thr* in ‘6’, ‘16’, ‘26’, ‘36’ regularly stands for /tʰ/ [tʰʰ/] and that the OKh sound written  $k\check{s}$  developed to LKh /tʰ/’ (Maggi 1991:s.v. ‘6’).

In Late Khotanese there was an effort to distinguish native  $k\check{s}$  (*th*) /tʰ/ from a learned Sanskrit pronunciation of  $k\check{s}$  [kʃ] through the use of *h*: or *h*, e.g., *ttīśi’rahi:kṣya* ‘Tiṣya-rakṣitā’, *ttahikṣa’śīlai* ‘Takṣaśīlā’, *ārahā:kṣam* ‘āraḥṣa’, *cahā:ṣa* ‘caḥṣuḥ’ (Bailey 1938:588).

### 2.5.0 Aspirated labiovelar glide $hv$ = /hʷ/

The use of digraphs in OKh for single consonants is well established, cf. *tt*, *tc*, *js*, *ts*, *ys*, *śś*, *ṣṣ*, *gg*, *kṣ*, *rr*, *ky*, *gy*, and probably  $k\check{s}$  in native words. Metrical evidence shows that  $hv$  also denotes a single consonant. In the hemistich below, *ahvato* ‘unspoken’ makes up three moras in a five mora segment:

6.55cd     **ahvato** bal|ysūstu bvāre □| kyai salāv|yau jsa hvatāndā (A:5+7+5+7)

(who) realize an unproclaimed *bodhi*, and who have proclaimed it by means of discourses.

A first guess as to its value can be based on the values of its components,  $h$ =/h/ and  $v$ =/w/. Since aspiration was a distinctive feature of Old Khotanese, it is plausible that the *h* in the digraph is used to modify the  $v$ =/w/ so that it shows aspiration. This could logically be transcribed /wʰ/ but the usual phonemic

transcription of this sound is /h<sup>w</sup>/, and this also symbolically echoes the orthographic *hv*.

Some dialects of English have a similar sound in words written with *wh*. For these speakers, the words *wails* ~ *whales*, *weather* ~ *whether* and *witch* ~ *which* are not homonymic. In these dialects the voiced and voiceless labial approximants have not yet merged in /w/ as they have in my dialect. Words written with *wh* are descended from words written in Old English with *hw*, e.g. *whale*, cf. OEng *hwil*; *whether*, cf. OEng *hwether*; *which*, cf. OEng *hwilc*. The development is probably both graphically and phonetically similar to that in Khotanese.

Historically, OKh *hv* derives mostly from OIr *\*h<sub>u</sub>*. With verbs: *hvañ*- ‘be called’ < *\*h<sub>u</sub>an<sub>i</sub>a-* (SuvII:373a); *hvar*- ‘consume’ < *\*h<sub>u</sub>ar-*, cf. Av *x<sup>v</sup>ar-*; *hva*- ‘strike’ < *\*h<sub>u</sub>ahaja-*, cf., Av *paiti x<sup>v</sup>an<sub>h</sub>-* ‘thresh’ (EDIV:141–148; SGS:156).<sup>18</sup> With nominals: *hvāšta*- ‘best, chief’ cf. Av *h<sub>u</sub>oīšta-* (SuvII:374a); NS *hvī* ‘sweat’ cf. OIr verb *\*h<sub>u</sub>aid-* ‘sweat’ (EDIV:143); *hvar*- ‘sister’ cf. Av *x<sup>v</sup>an<sub>h</sub>ar-* (SuvII:373a).

The development *\*h<sub>u</sub>* > /h<sup>w</sup>/ is similar to *\*rd* > /l/, *\*rt* > /d/, *\*rθ* > /t<sup>h</sup>/, *\*sr-* > /s-/ , *\*xš* > /s/, *\*xt* > /t/, and *\*fš, \*xš* > /t<sup>s</sup>/, where original sequences of two consonants are contracted to one in Khotanese.

---

<sup>18</sup> The stem *hva*- ‘strike’ was transcribed as *hvah-* by both Emmerick (SGS:156) and Bailey (DKS:508b) and as *\*hvay-* by Skjærvø who compared Av *x<sup>v</sup>an<sub>h</sub>aiia-* (SuvII:373a). Earlier, the Leumanns transcribed *hvad-* (Glossar:528b) and commented “Das Verbum urar. *suad* ‘stampfen’ (im Avesta bloß das Partizipium *χvasta*) lautet im Sanskrit *sūd* (*sūday*)” (E. and M. Leumann 1933–36:8 footnote 16). Konow listed the present stem first as *hvah-* (1932:145), then with *-d-*, as “*hvada-*, *hvasta-*, to beat” (1949:132) without etymological comment. Cipriano discussed the stem extensively (1998:210–235, especially 224–227) and proposed a Proto-Iranian *\*xwad-* “battere” (108).



In Late Khotanese, /h<sup>w</sup>/ probably remained. In transcriptions of Late Middle Chinese, LKh *hv:* is used for LMC /f, f<sup>h</sup>/ (Emmerick and Pulleyblank 1993:25). This probably does not indicate a pronunciation [f] for LKh *hv*. The colon ‘:’ was used above *h* to mark foreign sounds. The use of *h:* or *h* to distinguish Sanskrit *kṣ* from native *kṣ* has been mentioned above (4.0). LKh *h:* is also used for LMC /x, x<sup>h</sup>/, and twice for glottal stop (ibid:16, 40). With the trigraph *hv:*, the colon showed that the intended sound was not the same as Khotanese *hv*=/h<sup>w</sup>/ but similar. A modified /h<sup>w</sup>/ sign served to write /f, f<sup>h</sup>/.

### 2.5.0.1 *hv'* (*hv* + apostrophe)

When the subscript hook or ‘apostrophe’ is written with *hv*, i.e., *hv'*, the meaning is /hu/, not /h<sup>w</sup>/. The main use of the apostrophe in OKh is to mark a sequence of two vowel phonemes. In Skt writing, such sequences are systematically blurred by sandhi. It is technically possible in Khotanese writing to show the second vowel by means of a vocalic akṣara but this device appears to be reserved for vowel initial words. That is, vocalic akṣaras served to mark word boundaries. Because intervocalic \*š is lost in OKh, sequences of two vowels arise, e.g., *hvā'ñ-* /huāñ-/ ‘to dry’ < \**hušānya-* causative < \**hauš-*, cf. Av *haoš-* ‘to become dry’ (DKS:506a; SGS:157, 153 s.v. *huš-*); *hva'nd-* /huand-/ ‘man’ < \**aušavant* ‘mortal’, cf. Av *aošaṅ<sup>v</sup>hant-* (DKS:503b). The vowel hiatus is marked *v'* in *huve'* /huē/ NS ‘man’ (SGS:334). Because of this use of the apostrophe, the DKS entry *hve'* cannot be infinitive to *hvan-* as Bailey suggests (508; also rejected by Emmerick 1982:132) nor is *hvai'de* (DKS:508b) likely related to *hvāñ-* to speak.

### 2.5.1 LKh *h*: in Old Turkic

In the LKh transcriptions of Old Turkic cavalry terminology (Emmerick and Róna-Tas 1992), *h*: is used for OT /g/. It is useful here to show the spelling system used in this document for /g/ as it is revealing about both languages.

In front vowel words /k, g/ have velar or postpalatal allophones [k, g] while in back vowel words they have uvular allophones [q, ɣ].<sup>19</sup> The distribution of the signs used for these sounds can vary with the position within a word. The distribution may be summarized as follows:

---

<sup>19</sup> The traditional transcription method for these segments is followed here. It might be preferable to write [k', g'] in front vowel words. With regard to the voiced uvular fricative in back vowel words, the symbol ʁ or ʁ̥ is preferred in American and IPA traditions, while ɣ is reserved for the velar.

Phonetic symbols for Turkic uvulars and velars:

	here	traditional	IPA
voiced uvular continuant	ɣ	ɣ	ʁ
voiceless uvular stop	q	q	q
voiceless (pre)velar stop	k	k	k/k'
voiced (pre)velar stop	g	g	g/g'
voiced (pre)velar continuant	ǰ	—	ɣ

**Table 2.14:** LKh Transcriptions of OT Velars<sup>20</sup>

	Initial	Medial	Final
[q] = k	<i>kapa'kā</i> [qabaq] 'eyelid'	<i>sakalādrūkā</i> [saqaldruq] 'sub-mental strap on the harness'	<i>yasīkā</i> [yasīq] 'bow case'
	<i>karnai</i> [qarni] 'its belly'	<i>ttākā</i> [toqo] 'belt buckle'	<i>kūrñālūkā</i> [qurŋoluq] 'bow case'
[k] = k	<i>kīrāpīkā</i> [kirpik] 'eyelash'	<i>aupāka</i> [öpka] 'lung'	<i>tterkākā</i> [terkök] 'saddle strap'
	<i>kyešä</i> [kēš] 'quiver'	<i>kākuysä</i> [köküz] 'breast'	<i>bīḍakä</i> [biläk] 'wrist, forearm'
[ɣ] = h:	---	<i>sapäh:akä</i> [sabyaq] 'waist'	<i>aysaih:ä</i> [aziɣ] 'canine tooth'
		<i>kapäha':kä</i> [qabyaq] 'cover of the quiver'	<i>kasai'hä:</i> 'inside of the cheeks'
[g] = g	---	<i>yūgünä</i> [yügün] 'bridle'	---
		<i>ttīysgīnä</i> [tizgin] 'bridle'	
[g] = h:	---	<i>ehī':nä</i> [egin] 'shoulder'	<i>yaihä':</i> [yig] 'bit [of the bridle]'
		<i>bühä:sakä</i> [bügsäk] 'upper chest'	<i>īḍaih:ä</i> [ilig] 'attachment (to the harness)'

Some observations can be made about the patterns. The phoneme /k/ is consistently written with *k*, whether for the back allophone [q] or for the front [k], and no matter whether the segment is initial, medial or final. The phoneme /g/ is

<sup>20</sup> I follow the phonological interpretation of the OT forms as given by Emmerick and Róna-Tas with these exceptions: 1) the labial stop is [b] not [p] in most positions; 2) the voiced velar stop in front vowel words is always given as [g] never as [ɣ] even when written with *h:*; 3) In back vowel words [i] not [ɨ] is written.

slightly more complex. It does not occur in initial position. Finally it is always (10×) written with *h*: Medially in back vowel words the allophone [ɣ] is always written with *h*: The only inconsistent spelling is medially for the allophone [g] in front vowel words where it is most often (6× in identified words) also written with *h*: There are three instances of medial [g] written with *g*, the two in the list, and the unidentified *yūrāgakä* /yürgäk/ ‘?’.

About these segments Emmerick and Róna-Tas remark,

According to Hovdhaugen intervocalic [ɣ] and [g] were most probably fricatives in Old Turkish, but it would appear to be more accurate to say that in the Turkish language of Khotan the phonetic difference between the voiced gutturals in front and back vocalic words was in the process of disappearing. (1992:236)

Their conclusion seems plausible. I would note that within later Old Turkic there is medial [k > g] voicing (see Erdal 1979; Hitch 1989) and at the same time older [g] may be becoming a continuant to increase its phonetic distance from the newer [g]. That is, as [k] voices towards [g], older [g] may tend to fricativize to [ǰ]<sup>21</sup>. The fricative [ǰ] would have been a foreign sound for the Khotanese, hence the *h*: spellings.

For our purposes, it is most significant that the allophone [ɣ] is consistently marked by *h*: This shows the lack of a voiced back continuant \*[ɣ] in Late Khotanese. It was a foreign sound.

---

<sup>21</sup> Because the symbol  $\gamma$  is used in the Turkic philological tradition for a uvular voiced continuant, IPA and American *R*, I use the symbol  $\check{g}$  for a voiced velar continuant, IPA and American  $\gamma$ . Compare the phonetic symbol chart in footnote 19.

## 2.6.0 Voiced *p*, *b* = /b/

There is a gap in the OKh labial obstruent inventory. While there are four signs *th*, *tt*, *t*, *d*, for dental obstruents, there are only three labial *ph*, *p*, *b*. It seems clear that *ph* is aspirated /p<sup>h</sup>/ and *b* is the voiced continuant /β/ (Emmerick 1981:202). The remaining *p* ought to be either a plain voiceless /p/, corresponding to *tt*=/t/, or a voiced /b/, corresponding to *t*=/d/.

A preliminary indication that *p* could be considered /b/ and not \*/p/ comes from the symbols used for the dentals. Just as the voiced Brāhmī character *d* marks a fricative /δ/ and the voiceless character *t* marks a voiced stop /d/, the voiced sign *b* shows /β/ and the voiceless sign *p* could show a voiced stop /b/, i.e., *d:t::b:p* /δ:d::β:b/.

### 2.6.1 Distribution of voiced and voiceless stops

Bilabial plosives do not occur intervocalically in inherited vocabulary because between vowels both OIr \**p* and \**b* result in OKh *v* /w/ (Emmerick 1981:185). But the symbols *p* and *b* do occur intervocalically in secondary situations. The graph *p* between vowels is reasonably common in Indian loans like NSm *upaguttā* Z 4.8 ‘Upagupta’, or *karmapaha* Z 13.62 ‘course of action, *karmapatha*’. Otherwise I notice it in Old Khotanese only in *hoparedärsä* ‘37’ Z 7.6, literally ‘seven-beyond-thirty’.<sup>22</sup>

---

<sup>22</sup> Tumshuqese similarly shows *p* in the ‘seven-beyond-’ words between vowels as in *hoparsana* ‘17’ but *w* in other ‘beyond’ numerals between vowels as in *śowarsana* ‘11’ (cf. DKS:217b). Numerals tend to be irregular. Perhaps here is a fossil reflecting an earlier cluster involving the *d* /δ/ in *hauḍa* ‘7’, i.e. /haḍbarē < haḍbarē < hauḍ- + -barē-/. Alternatively, the forms may contain double /bb/ < /δb/, which may be supported by the Late Khotanese spelling in Tibetan script *hob-be-re-ḥbes* for LKh *hauparibistā* ‘27’ (Maggi 1995:439).

This is unusual because otherwise there is *-pare-* after consonants and *-vare-* after vowels, e.g., NAP *pusparenautä* ‘95’ Sgh[17] 31.3, NAP *haṣparedärsä* ‘38’ Sgh[16] 80.8, NAP *dvāvaredärsi* ‘32’ Z 4.12, NAP *nauvarenautä* ‘99’ Sgh[16] 81.1.

In Old Khotanese compounds a morpheme initial *p* becomes *v* between vowels. The word *paa-* ‘foot’ (SGS *pāa-*) is found in GDPm *śivānu* ‘biped’ Z 22.257 < *śivaa-* (Emmerick 2009:409). A Late Khotanese example involvess *pāḍa-* ‘document’: *gārya-vāḍa-* ‘document of purchase’, *perma-vāḍa-* ‘document of adoption’, *draṃmā-vāḍa-* ‘document regarding money (?)’ (Skjærvø 2002:lxxvii-lxxviii). A *p* also becomes *v* where the noun *prahaṇa-* ‘clothes’ is the second member of a compound as in *rrusta-vrahaṇa* ‘red garbed’ Z 22.309, *āṣeiṇ’a-vrrahoṇe* ‘blue-garbed’ Z 14.34, *śīya-vrrahaṇä* ‘white-garbed’ IOL Khot 160/4v5.

Intervocalic *b* is of course seen in loanwords like ASm *pratābim̐bau* Z 23.30 ‘image’ cf. Skt *pratibimba* ‘reflection’. Besides the child language form *pāba* ‘Papa’ Z 5.109 it is common with prefixes, e.g., *hubārūñandei* ‘very brilliant’ Śgs 2.5v1, *pabastago* ‘continuously’ Śgs 3.2.5 (DKS:212b), NSm *pabanā* ‘rebirth (continuation)’ Śgs 3.7v3, NAPm *abasta* ‘not bound’ Z 2.152, NAPm *ābei’sa* ‘whirlpools’ Z 17.13 (cf. NAPm *bei’sa* ‘whirlpools’ Z 24.238).

The historical development of OIr *\*p* and *\*b* intervocalically contrasts with the development of OIr *\*t* and *\*d* intervocalically. These remain distinct as *t* /*d*/ and *d* /*ḍ*/ (Emmerick 1981:187).

The phonemes marked by the symbols *p* and *b* appear to contrast only in initial position. With the dentals, there seems to be a similar but not parallel contrast in initial position, between *tt* /*t*/ and *d* /*ḍ*/, involving just two of the dental obstruents.

Initial /d/ *t* may not occur at all. Emmerick noted that ‘Initial *t* is rare’ (1981:188) but a stronger statement is possible. As mentioned in Hitch 2014:683, phonemically there is no initial \*/dV-/ written \**tV*- in Old Khotanese. The absence of initial \*/dV-/ is disguised in the normalized transcription where it appears that some words begin with *tV*- /dV-/, but these are actually all enclitics and could be transcribed with a leading hyphen. Their enclitic nature is revealed by spelling variation. The -*t*- may be written -*g*- or -*y*-. The Leumanns’ Glossar to Z, for example, lists *tan*- enclitic form of *yan*- ‘make, do’ (also *gan*- in some documents, see fn. 50), -*taraṇa*- compound form of *karaṇa*- ‘deed’, *tātu~tāvu~gāvu* ‘time, while’, *tāmu~gāmu* ‘-wise, -ways’, *tā~yā* GDSm enclitic pronoun ‘him’ (Glossar:430a). The GDS enclitic *tā* to the pronoun *thu* ‘you’ appears in Z to have consistent clitic initial -*t*- (Glossar:439b) but there is no question that it is an enclitic.

**Table 2.15:** Initial *p, b, t, d* distribution

Initial position	labial obstruents	—	/b/ <i>p</i>	/β/ <i>b</i>
	dental obstruents	/t/ <i>tt</i>	—	/δ/ <i>d</i>

The apparent lack of Kh initial \*/d/ does not match precisely the general lack of Kh \*/p/. But generally speaking it is parallel.

Voiceless, unaspirated plosives of all kinds had restricted distribution. One place is at the end of present stems. In OKh present stems inherited from OIr and which end in a single consonant, that consonant is never a voiceless unaspirated stop. It can be a voiced stop, a voiceless aspirate, or several other consonant types.<sup>23</sup> Table 2.16

---

<sup>23</sup> Verb stems in Old Khotanese of Iranian origin can end synchronically in the consonants or consonant clusters *mj, lj, ñ, ṣt, d, ṇ, t, lt, śt, st, th, mth, d, ṃd, śd, n, rn, (ph?), ṃph, rb, m, y, tcy,*

below compares the distribution of stem final plosives and affricates. It contains three examples given in parentheses. There are no examples of *p* /*b*/ because OIr *\*b*, *\*p* between vowels becomes Kh *v*. There are no examples of single dental and labial aspirates probably through accident rather than pattern. Examples of nasal plus consonant are given. The distribution of this configuration appears to be similar to that of single consonants (see table 2.17).

**Table 2.16:** Single OKh Oral Stops and Affricates at the End of Present Stems

	voiced		voiceless		voiceless aspirate	
velar	<i>gg, g</i>	/g/	<i>k</i>	/k/	<i>kh</i>	/k <sup>h</sup> /
	—	—	—	—	—	—
palatal	<i>j</i>	/j/	<i>c</i>	/č/	<i>ch</i>	/č <sup>h</sup> /
	<i>daj-</i>	burn	—	—	<i>pach-</i>	be cooked
retroflex	<i>ḍ</i>	/ḍ/	<i>ṭ</i>	/ṭ/	<i>ṭh</i>	/ṭ <sup>h</sup> /
	<i>baḍ-</i>	writhe	—	—	<i>bīṭh-</i>	twist (tr.)
dental	<i>t</i>	/d/	<i>tt, t</i>	/t/	<i>th</i>	/t <sup>h</sup> /
	<i>pat-</i>	fall	—	—	( <i>maṃth-</i>	agitate)
dent. affr.	<i>js</i>	/dʒ/	<i>tc</i>	/tʃ/	<i>ts</i>	/tʃ <sup>h</sup> /
	<i>pajs-</i>	cook	—	—	<i>patäts-</i>	abandon
labial	<i>p</i>	/b/	—	—	<i>ph</i>	/p <sup>h</sup> /
	( <i>ttav-</i>	be hot)	—	—	( <i>paltcīṃph-</i>	check)

---

*ly, vy, hy, r, rr, l, v, rrv* (Late Kh *m̄jv, yv, şv, sv, lysv, hv*), *ś, rś, lś, śś, ş/ş'/t'/v', rş, şş, s, js, m̄js, ljs, ts, ys, lys, rs, ls, h* (SGS 165-171).



Stems which are loanwords can end in single voiceless stops: *āyāc-* ‘request’ < BHS *āyācati* ‘supplicates’, *pakuṭ-* cf. BHS *ākoṭayati* ‘strike’. There are no examples of stems ending with single voiceless dentals (-*Vtt-*) or dental affricates (-*Vtc-*). There are loans in -*rtt-*, e.g. *anuvartt-* ‘conform to’ < BHS *anuvart-*, and there is /t/ or /ṭ/ after a sibilant in -*st-*, -*śt-*, -*ṣṭ-*, e.g., *phast-* ‘flutter’ (DKS ‘move’), *viśt-* ‘place, establish’, *paṣṭ-* ‘arise, set out’. An Old Iranian \**p* at the end of a stem becomes *v* in Old Khotanese, e.g., \**abi-Hap-aya-* > *byev-* ‘obtain’, \**tap-* > LKh *ttav-* be hot. Only borrowed stems can end in *p* /*b*/ in OKh, e.g., *nyāp-* ‘be known’ cf. Skt *jñāpyate*.

Voiceless unaspirated plosives do not occur after nasals in inherited vocabulary except in rare exceptional circumstances. This is illustrated in the chart below. The clusters \*/*ṅt*/, \*/*nt<sup>s</sup>*/ and \*/*mp*/ do not exist at all. As far as I can tell the clusters /*ñc*/ and /*nt*/ occur in native words only in the one morpheme listed for each. The cluster /*nk*/ (phonetically [ṅk]) as well as retroflex /*ṅḍ*/<sup>24</sup> and /*ṅṭ<sup>h</sup>*/ are only found in words of Indian origin (loans in parentheses).

---

<sup>24</sup> Bailey included *khāṇḍa-* ‘way, manner, etc.’ in DKS (74b) because he assumed an Iranian etymology. That this word is borrowed is shown by -*ṅḍ-* which would have to derive from an OIr \**nrt-*, and by the fact that the cluster alternates with -*nd-* (ibid.). A not exact match is BHS *khaṇḍa* ‘quantity, mass, large number’ used after a noun (Edgerton 1953b:203a). Degener also includes *khāṇḍa-* as an *anda-p.ptc.a.m*. It is the only such participle with -*ṅḍ-* instead of -*nd-* in her ten page list. Even those from stems ending in -*r-* like *gvīranda-* ‘talking’ or *hvaranda-* ‘eating’ do not show retroflexion (KhSuf:33–42)

**Table 2.17:** Post-nasal distribution of plosives and affricates

	voiced	voiceless	voiceless aspirate
velar	/ŋg/ <i>kaṃggaṅ-</i> 'dig'	(/ŋk/) ( <i>saṃkalpa</i> Z ) ( <i>saṃkalpa</i> )	/ŋk <sup>h</sup> / <i>saṃkhal-</i> 'be tainted'
palatal	/ñj/ <i>usthaṃj-</i> 'pull out'	/ñč/ <i>-aṃcā-</i> p.ptc.a.f	/ñč <sup>h</sup> / <i>ggaṃcha</i> Z 2.29 LSm 'hole'
retroflex	(/ṇḍ/) ( <i>ggaṃḍyo</i> Z 2.102) (ASf 'gong')	*/ṇṭ/ <i>*ṃṭ, *ṇṭ</i> ---	(/ṇṭ <sup>h</sup> /) ( <i>saṃṭhānāna</i> Z 23.45) (IASm 'appearance')
dental	/nd/ <i>kṣunda-</i> 'husband'	/nt/ <i>bihaṃtta-</i> ppp < <i>bihan-</i> 'smile'	/nt <sup>h</sup> / <i>nuvaṃth-</i> 'be removed'
dent. affr.	/ndz/ <i>paṃjsa</i> 'five'	*/nt <sup>s</sup> / <i>*ṃṭc, *ntc</i> ---	/nt <sup>sh</sup> / <i>haṃtsa</i> 'together'
labial	/mb/ <i>haṃber-</i> 'fill'	*/mp/ <i>*ṃp, *mp</i> ---	/mp <sup>h</sup> / <i>paltcīṃph-</i> 'check'

## 2.6.2 Old Turkic evidence for /b/

The Turkish-Khotanese wordlist from Dunhuang has been studied by H.W. Bailey (1944), E. Hovdhaugen (1971), G. Clauson (1973), and R.E. Emmerick and A. Róna-Tas (1992). None of these studies adequately addresses the labial spellings. The latest study makes no mention of them (cf. Emmerick and Róna-Tas 1992:235–236).

In ‘Old Turkic “b” and “p”’ (Hitch 1989) I examined the use of the various symbols used to write labial consonants in the many Old Turkic scripts. The changes in spellings over time is best explained through a phonemic reanalysis.

Old Turkic (OT) underwent a process of consonant weakening which saw voiceless segments voice, and voiced stops probably become fricatives. Marcel Erdal 1979 used this process as it is reflected in spelling to chronologically classify OT texts. Group I has consistent spelling, group II has rare *d~t* and *z~s* interchange, group III has frequent *d~t*, *z~s*, and, in Brāhmī and Manichean texts, frequent *g~k* interchange, and group IV has occasional passage of /d/ to /y/.

In Hitch 1989 I mentioned that /d/ to /y/ ‘presumably entailed an intermediate fricative stage [δ]’ (137). This paralleled the development of medial Earlier Old Turkic (EOT) /b/ > Later Old Turkic (LOT) /v/, which led to the phonemic reanalysis eliminating phonemic \*/p/ from the later inventory (ibid.:132-133).<sup>25</sup> I did not bring the LKh transcription evidence into the discussion because the general view at the time was that Khotanese *p* stood for /p/.

OT initial /b/ did not undergo phonemic reanalysis in the consonant shift. However, there is internal evidence which suggests that it also was subject to phonetic fricativization, but not to as great a degree as its medial counterpart, and possibly only in certain dialects. Parallel to the /b/ > /v/ shift in medial position, there is a /b/ > /m/ shift in words with a following nasal, e.g., Orkhon *bän~män* ‘I’, *tarban-qa~tarman-qa* ‘geographical name’ (ibid.:138-139). EOT /b/ was weakening

---

<sup>25</sup> In some OT dialects, in certain clusters, older /b/ remained a stop but in other dialects it became /v/ (Hitch 1989:140-141).

in some dialects in all positions, probably to [β], which then became susceptible to regressive nasalization. Medially, the voicing in EOT /p/ > [b], caused the further shift of EOT /b/ [β] > [v] in order to increase the phonetic distance in this position. In LOT there was a phonemic reanalysis, with newly phonemic /v/ arising from EOT medial /b/, and with the LOT phoneme /b/ now reflecting both EOT initial /b/ and EOT medial and final /p/. Phonetically LOT /b/ in initial position may have still been pronounced [β] in some dialects but if so, this pronunciation was levelled to [b] to match the stop pronunciations, [b] or [p] elsewhere.<sup>26</sup> These developments are sketched in the following chart.

**Table 2.18:** LOT /b/ spellings in LKh orthography

	initial	medial	medial	final
EOT phonemic	/b/	/b/	/p/	/p/
EOT phonetic	[b]	[b]	[p]	[p]
general weakening	[b], [β]	[β]	[b]	—
distance increase	—	[v]	—	—
LOT phonetic	[b], [β]	[v]	[b]	[p]
LOT phonemic reanalysis	/b/	/v/	/b/	/b/

---

<sup>26</sup> This discussion does not take into account the development of the EOT labial plosives in consonant clusters. This is affected by the voicing of the other consonant, and the effects vary with dialect. For instance, in the Old Turkic texts from Turfan, the phrase ‘the passion of anger’ appears as *öpkä nızbanı* in one text and as *övkä nızvanı* in another (Clauson 1972:9a). The word *öpkä* ‘lung, anger’ appears below in the list of cavalry terms in LKh script. *nızbanı* is a loanword from Sogdian *nizβ’ny* (Nadelyaev et al. 1969:359b) but may still show cluster evolution within Old Turkic.

The LKh transcriptions of LOT labials confirm these values. LKh *b* /β/ [β] is used for LOT [β], initial /b/ (table 2.19), while LKh *p* /b/ [b, p] is used for LOT [b, p], medial /b/ (table 2.20). The transcription was phonetically precise. In the tables below of the certainly identified terms, I give the Khotanese spelling, a phonetic transcription, the phonemic or normalized transcription given by Emmerick and Róna-Tas 1992, their gloss, and the paragraph in which they discuss the term.

**Table 2.19:** LOT initial /b/=[β] spellings

LKh script	phonetic	Emmerick and Róna-Tas 1992
<i>baha':rai</i>	[βayari]	bayari 'a wide leather strap for suspending the quiver from the belt' 8
<i>bahai':rä</i>	[βayir]	bayir 'liver' 76
<i>bahä':räsäähä:</i>	[βayirsoq]	bayirsoq 'entrails' 86
<i>bakañākä</i>	[βaqañoq]	baqañoq 'the frog of the horse's hoof' 72
<i>bīḍakä</i>	[βiläk]	biläk 'wrist, forearm' 61
<i>bai'kiñä</i>	[βiqin]	biqin 'hip' 96
<i>bauhū':nä</i>	[βoyun]	boyun 'knuckle' 70
<i>bauhq:nä</i>	[βögän]	böyän 'caecum' 82
<i>bau'hū:rä</i>	[βögür]	böyür 'kidney' 92
<i>bühä:sakä</i>	[βügsäk]	büysek 'upper chest' 64

**Table 2.20:** LOT medial /b/=[b, p] spellings

<i>aupäkä</i>	[öpkä]	öpke ‘lung’ 74
<i>kapa’kä</i>	[qabaq]	qapaq ‘eyelid’ 37
<i>kapäha’:kä</i>	[qabyaq]	qapyaq ‘cover of the quiver’ 5
<i>kiräpikä</i>	[kirbik]	kirpik ‘eyelash’ 38
<i>sapäha:kä</i>	[sabyaq]	sapyaq ‘waist’
<i>ttupī</i>	[tübi]	tüpi ‘its bottom’ 2
<i>yapī</i>	[yabi]	yapi ‘saddle blanket’ 19

### 2.6.3 LMC and Tibetan evidence for OKh /b/

The LKh transcriptions of Late Middle Chinese, and the Tibetan transcriptions of LKh numerals are consistent with the view that  $p = /b/$ . We have the equations LMC /m/ = LKh *b* = Tib *hb*, and LMC /p/ = LKh *p* = Tib *p*.

The fricative value of *b* [β] is clear in both transcriptions. LMC /m/ is transcribed primarily with LKh *b* but also with LKh *m*. This shows the tendency in LMC to denasalize initial nasals to fricatives. The tendency is illustrated by the phonemic change of EMC /ñ/ to LMC /r/ (Pulleyblank 1984:171) and of EMC /m/ probably before non-nasal finals to LMC /β/<sup>27</sup> (cf. *ibid*:232). LMC /ŋ, n/ were allophonically denasalized to [ɣ, δ] before non-nasal finals. This is why LMC /ŋ/ = LKh *g* or *h:*, why LMC /n/ = LKh *d* or *n* and why LMC /m/ = *b* or *m*. In the modern Amoy dialect, [b, l, g] before oral vowels are in complementary distribution with [m, n, ŋ] before

---

<sup>27</sup> Where I use β for Middle Chinese, Pulleyblank uses the IPA labio-dental approximant symbol v.

nasalized vowels and are considered members of the same phonemes (Pulleyblank 1986:22). This allophony is likely derived in a way similar to that shown in LMC.

Traditionally, the LMC allophones [ɣ, δ, β] have been described as pre-nasalized stops, i.e., \*[ŋg, nd, mb]. This is because of transcriptions of /ŋ, n, m/ which alternate between signs for nasals and voiced stops or fricatives, such as those in LKh script, and because of Tibetan transcriptions with 'a-chuñ, i.e., ḥg, ḥd, ḥb, in which the 'a-chuñ (ḥ) is thought to show a nasal as in Tibetan. But the use of alternating stop and nasal signs probably has to do with the oral~nasal allophony in LMC mentioned above, while the 'a-chuñ shows that a foreign segment was voiced and perhaps fricative or affricate. The Tib transcriptions of LKh numerals show this practice. LKh /δ, β, d<sup>z</sup>/ are transcribed ḥd, ḥb, ḥj as shown in the examples below. Where the Late Khotanese vocalism is unclear to me, a hyphen stands in for a vowel in phonemic transcription. All data is from Maggi 1991.

**Tib ḥd=LKh d /δ/** Tib ḥdva = LKh dva /δwa/ '2'; Tib ḥdra-ya = LKh drraya /δraya/ '3'; Tib ḥda-so = LKh daso /δaso/ '10'; Tib śu-ḥdas = LKh śūdāsä /śuδas/ '11'; Tib ḥdvos = LKh dvāsä /δw-s/ '12'; Tib ḥdres = LKh \*drraisä (cf. LKh ordinal NS drraisam) = /δr-s/; Tib ḥdes, ḥders = LKh dirsä, dairsä = /δ-rs/ '30' (and '31' to '36'). Exception: Tib ho-dus = LKh haudūsa = /hoδus/ '17'

**Tib ḥb=LKh b /β/** Tib ḥbes = LKh bistä /β-st/ '20'; Tib ḥdvor-hbes = LKh dvārabistä /δw-rβ-st/ '22'; Tib śur-ḥbes = LKh śśūvaribistä /š-r-β-s/ '21'; (and all the 20's);

**Tib *ḥj* = LKh *js* /dʒ/** Tib *pan-ḥjah* = LKh *paṃjsa* /bandʒa/ '5'; Tib *pan-ḥjus* = LKh *paṃjsūsä* /bandʒus/ '15'.

The use of LKh *b* [β] for LMC [β] and of Tib *ḥb* for LKh [β] seems clear enough. The stop spellings are a little more complicated. There is one intervocalic labial plosive:

**Tib *Vb-bV* = LKh *VpV* /VbV/** Tib *hob-be-re-ḥbes*, LKh *hauparibistä* '27'

This one example fairly clearly shows an effort to write [b] or perhaps [bb].<sup>28</sup>

In initial position LKh *p* /b/ is all three times written with Tib *p*:

**Tib *p-* = LKh *p-* /b-/** Tib *pan-ḥjah*, *pan-ḥja* = LKh *paṃjsa* '5'; Tib *pan-ḥjus* = LKh *paṃjsūsä* '15'

At some point in the history of the Tibetan written language single initial voiced stops and affricates devoiced, causing the following vowel to take on a low tone.

This resulted in the situation still valid today in Tibetan orthography in which both original voiced and voiceless graphs are used in writing initial voiceless stops and affricates. The change in Tibetan may have taken place before the Khotanese numerals were recorded. Initial devoicing may also be visible in transcriptions of LKh *tc* /tʰ/. This is four times as expected with *ć*: Tib *ćo-ra* = LKh *tcaura* '4', Tib *ćo-las* = LKh *tcaulasä* '14', Tib *ćor-ḥbes* = LKh *tcaurabistä* '24', and Tib *ćor-ḥdes* = LKh *\*tcauradirsä* '34' (Maggi 1991, s.v.). But it is once with *j* in Tib *jo-ra* = LKh *tcaura* '4' (ibid.) which shows that initial *j* could write a voiceless sound. If this sound change had already happened, then Tibetan had only one initial labial stop, written either *p* or *b*. The use of Tib *p* for LKh initial *p* does not indicate a [p] pronunciation of Khotanese initial *p* /b/.

---

<sup>28</sup> As mentioned in footnote 22, this word may contain double /bb/ in OKh.



Labial stop spellings in clusters may also not be terribly revealing. In LKh, sibilant clusters may have had a tendency to voice between vowels. This is noticeable in the hesitation between *št* and *śd*, e.g., LKh *gvašte/gvaśde* ‘it is digested’ (SGS:32); LKh *dašte/daśde* ‘it ripens’ (SGS:43); OKh *padaśdä/LKh padaštä* ‘he burns (tr.)’ (SGS:68 and pattern summary SGS:192). Consequently, there are two different Tib transcriptions for LKh *p /b/* after a sibilant: (1) intervocalic [z̥b], *har-be-re-hbes*, LK *haşparibistä* ‘28’; and (2), initial [sp], *sper-hdes* ‘35’, cf, LKh ordinal, gen.-dat. sing. masc. *sparadirsamyē* ‘35th’; Tib *sper-hbes*, LKh *sparibista* ‘24’ (Maggi 1991:s.v.).

#### 2.6.4 Typology of inventory gaps

The phonemic segment inventory 1 below with a gap in the voiceless labial stop category, like that argued here for Old Khotanese, is typologically quite ordinary, while inventory 2 with a gap in the voiced labial stop category, like that suggested in Emmerick 1989 (213), is not attested.

**Table 2.21:** Theoretical labial stop inventory gaps

<u>1</u> <u>common</u>	<u>2</u> <u>unattested</u>
-      /t/    /k/	/p/    /t/    /k/
/b/    /d/    /g/	-      /d/    /g/

The facts of these inventories, and the typological constraints that define them, were first illustrated by I. Melikishvili (1970, 1972) in work reported on by Thomas V. Gamkrelidze (1978:14–17). The patterns are confirmed by examining the UCLA

Phonological Segment Inventory Database (known as the UPSID).<sup>29</sup> The first edition of the database contained inventories of 317 languages (Maddieson 1984:6, 263–422). The current online edition (Maddieson and Precoda ongoing, accessed through Reetz undated) has 451 languages<sup>30</sup>. The selection is intended to be a representative sample, geographically and genetically.<sup>31</sup>

The UPSID lists 27 languages with inventory 1: Aghem, Amele, Arabic, Beja, Chuave, Dizi, Ewondo, Feʔfeʔ, Hausa, Koiari, Kullo, Kunama, Maba, Nera, Noni, Nubian, Nyimang, Shilha, Socotri, Somali, Songhai, Tama, Tamasheq, Tigre, Tiruray, Yareba, Yoruba.<sup>32</sup> . Later Old Turkic was also of this type (Hitch 1989). Another nine languages follow probably related /p/-less patterns: Irarutu, Kwaio, Kwoma, Mor, Yessan-Mayo have prenasalized /<sup>m</sup>b, <sup>n</sup>d, <sup>ŋ</sup>g/ with /t/ and /k/ or /x/, Efik and Ket have /b, d/ and /t, k/, and Seneca and Warao have /b/ with /t, k/ (Maddieson and Precoda ongoing).

---

<sup>29</sup> The UPSID distinguishes between dental and alveolar stops. In my data collection no distinction is made between those positions. A /t/ or /d/ in a particular language may be either dental or alveolar. We also likely cannot tell whether OKh /d, t/ were dental or alveolar.

<sup>30</sup> <http://www.linguistics.ucla.edu/faciliti/sales/software.htm#upsid>; March 20, 2016.

<sup>31</sup> The languages included in UPSID have been chosen to approximate a properly constructed quota sample on a genetic basis of the world's extant languages. The quota rule is that only one language may be included from each small family grouping, for example, among the Germanic languages, one is included from West Germanic and one from North Germanic ... (Maddieson 1984:5, Maddieson and Precoda ongoing).

<sup>32</sup> Maddieson 1984:appendix b, consulted by hand, then supplemented by the online UPSID, Maddieson and Precoda ongoing, accessed through the online interface by Henning Reetz, undated.

There appear to be no languages in the database with inventory 2 above. The UPSID lists 129 languages (28.60%) which do not have a 'voiced bilabial plosive' sound. I checked half of these. From that sample a generalization may be made. If a language has no /b/, then it will also have no /d/ or /g/, so it will not reflect inventory 2 above. I notice one possible exception. Gadsup has /p, t, k/ and /d/ and no /b/ or /g/. But it has only eight consonants in total. The other four are /β, m, n, ʔ/ (ibid). With only eight consonants it is a marginal exception at best.

In Old Khotanese, the single bilabial stop phoneme was certainly voiced.

## 2.7.0 The Nasals

It is clear that OKh *m*, *n*, and *ñ* represented the phonemes /m, n, ñ/ (cf. Emmerick 1981:200-202). Khotanese Brāhmī inherited three other symbols for nasals, *ṃ*, *ṅ*, *ṇ*.

### 2.7.1 Anusvāra *ṃ*

The first of these, *ṃ* or *anusvāra*, a dot above an akṣara, does not show vowel nasalization, but is used to write a nasal stop of the same articulation point as the following plosive.<sup>33</sup> As such, it can alternate with a nasal radical. This is readily apparent with the dentals, e.g., ppp *biysānda~biysāṃda* < *biysān* 'to wake up' (SGS:97), and with the velars, e.g. *pāstuṅgā~pāstumgga~pāstumgā* 'inverted, headlong' (DKS:235b). I notice just one example of retroflex *ṅḍ~ṃḍ* alternation within Old Khotanese: NSm *daṃḍā* Z 12.128 ~ ASm *daṅḍu* Z 12.128 'punishment'. But both spellings for /ṅḍ/ are found in various words such ASf *ggaṃḍyo* Z 2.101

---

<sup>33</sup> The *-ṃh~-nh-* alternation in ASm *baṃhyu* Z 2.142~*banhyu* Z 5.106, NAPm *baṃhya* Z 2.112~*banhya* Z 2.45 'tree' shows that /n/ is the least marked nasal.

‘gong’ and IASm *śsuṇḍina* Z 24.194 ‘(elephant’s) trunk’. All words in Old Khotanese with /ṇḍ/ appear to be loans, cf. Skt *daṇḍa*, *gaṇḍikā*, *śuṇḍā*. The spelling *mb* is exceedingly rare with just one example in Old Khotanese in a word which also has *ṃb* spellings: NSm *pratābimbai* Z 6.6 ~ *pratābiṃbai* Z 2.224 ‘image’. I find no examples of \**ñj* or \**ṅgy* spellings in Khotanese at all. The cluster /*ñj*/ is always written *ṃj*.

### 2.7.2 Retroflex ṇ /ṇ/

For many years I thought that the retroflex nasal sign *ṇ* was merely an allograph for /*n*/. Emmerick 1981 does not discuss *ṇ* nor include a phoneme /*ṇ*/. Emmerick 1989 and 2009 did list a phoneme /*ṇ*/ but without discussion. Most frequently *ṇ* is used in a fashion similar to Sanskrit in words where an *r* or *ṣ* precedes, such as NAPm *graṇe* Z 3.61 ‘garlands’ or NSm *ṣṣāṇaumā* ‘favor’ Z 2.240. The effect is clearest with suffixes containing /*n*/. About the IASm suffix *-āna* Emmerick wrote, “In the vicinity of *r* or *ṣ*, it is sometimes spelled *-āṇa* due to Sanskritization” (SGS:259). Examples are *sūtrāṇa* ‘sūtra’ Z 12.91, *veṣāṇa* ‘clothing’ Z 23.167, *sumīrāṇa* ‘Sumeru’ Sgh[10] 65.4. In this use it is sometimes not consistent as in *vīrāna* Z 12.111 ~ *vīrāṇa* Z 3.76 ‘hatred’. The GDP suffix *-ānu* is usually written *-āṇu* after *r*, e.g., *uysnorāṇu* ‘being’ Z 23.90, *puraṇu* ‘son’ Z 24.483, *handarāṇu* ‘other’ Z 24.274. But occasionally it is *-ānu* after *r*: *uysnaurānu* Sgh[22] 253[75], *haṃdarānu* IOL Khot 149/1r4 (=Neb 128[423]). The 3P and 1P oblique enclitic pronoun *nā* ‘(of/from/to/for) them/us’ can be *ṇā* after *r* in the preceding word, e.g., *pharu ṇā pajsamu yanīndā* ‘they will do them great honour’ Z 22.297, *duru ṇā* ‘far from them’ Z 14.37. The present participle middle *-āna-* (KhSuf:78–83) maybe written *-āṇa-*

when following *-r-*, e.g., NSm *barāṇā* ‘riding’ Suv[Or] 6.2.35. From the preceding instances there would be no reason to propose the existence of /ṇ/. They are all predictable.

Emmerick listed a present stem *haraṇ-* ‘throw’ (SGS:149), based on the 3Pp.a *haraṇīndā* Z 22.265. However, the 3Sop.a *haraña* Z 13.72 shows that the word has underlying /n/, i.e. /haraña < haran- + <sup>-i</sup>a/. If the nasal were retroflex, it would be transparent to umlaut as are retroflex /t, th, d, ṣ, z, r/ (see 2.4.3 above). Bailey listed the shape *haran-* (DKS:467a). This *ṇ* is also predictable.

An indication that *ṇ* may be orthographic rather than phonological, is its use after *rr*, the dental-alveolar /ɹ/, e.g., IASm *darroṇa* Z 13.21 ~ *darraunai* (+ *-i*) Z 2.56 < *darrau* ‘daring, valour’ (SGS:329, DKS:153a); *ysurrä ṇa jyāte* ‘may anger disappear from them’ (Z 3.63); IAS *ārrāṇa* ‘fault’ Suv[Or] 6.2.77 ~ *ārrnai* (+ *-i*) Suv[Or] 11.9. For *n* after *rr* also compare IASm *nyūrrāna* ‘harness’ Z 24.405, *erra nā nāte* ‘he took them in his arm’ Z 24.511. **RRĪNA- LKH RRĪNĀ-** It seems unavoidable to conclude that in these cases of *ṇ* it is the graph *r* (in *rr*), not the sound [ɹ] which triggers the use of the symbol *ṇ*.

In spite of all of the evidence pointing to allophonic or orthographic *ṇ*, I now think that there was a phonemic /ṇ/.

### 2.7.2.1 The apostrophe and *ṇ*

In the synchronic derivation of Old Khotanese, an underlying intervocalic *ṣ* /z/ is lost. The resulting sequence of two vowels is marked by the use of the apostrophe. The numeral ‘six’ is underlyingly /ṭṣḗza/ (cf. Av *xšuuas*). On the surface it is /ṭṣḗa/ which may be spelled *kṣāṣa*’ Suv 5.7[Or], *kṣāta*’ Z 4.12, *kṣā’t’a* Z 13.12 (Leumann’s

reading), *kṣāta* Z 6.20, *kṣei* Śgs 3.11r2, or *kṣei* Sgh[10] 161.9. Most forms of ‘six’ have a peg akṣara *t* of empty consonantal content and usually with apostrophe showing /ěa/. Sometimes the peg akṣara is *ṣ* which has no consonantal value here even though it reflects historical derivation.<sup>34</sup> Sometimes the scribe chose to use the sign *ei* (which was originally introduced to write the diphthong /aē/), usually in conjunction with the apostrophe, to write the sequence /ěa/.

Some words showing apostrophe do have related forms showing *ṣ*. For instance, the present stem *pyūṣ-* ‘listen’ is visible in the common 3Sp.m *pyūṣde* Z 3.147 /byūzḍē < byūz- + -ttē/. When vowel initial suffixes are attached, the *ṣ* /z/ is lost as in 3Pp.m *pyūvā're* Z 1.52 /byūāṛē < byūzārē < byūz- + -āre/. The peg akṣara *v* plus apostrophe shows that there is a vocalic sequence. The *v* is the usual choice when the first vowel is /u/ or /ū/. When the suffix contains /n/, that consonant appears in writing as *ṇ*. The form p.ptc.m NAPm *pyūvā'ṇa* Sgh[24] 91.1 shows the p.ptc.m suffix *-āna* and is derived /byūāṇa < byūzāṇa < byūzāna < byūz- + -ān- + -a/. The sequence of vowels is again indicated with the peg *v* + apostrophe. In this case there is no surface phonetic or orthographic trigger for *ṇ*. It might be argued that in this word there is knowledge of *ṣ* /z/ from elsewhere in the paradigm. But there are words for which no synchronic *ṣ* /z/ exists anywhere on the surface in any related words.

---

<sup>34</sup> Emmerick once thought that the “archaic pronunciation was still known” in forms like this:

Another peculiar feature of the script is the use of a subscript hook to indicate the recent loss of an internal sound.... Thus, Khotanese *gguva* ‘ears’ from \**gūža-* from \**gauša-* (cf. Avestan *gaoša-*). The archaic pronunciation was still known. (1979:9). While that hypothesis is tempting, it then needs to be explained why the peg akṣara is then not always *ṣ*.

Old Khotanese *gguṣa*-<sup>35</sup> ‘ear’ (cf. Av *gaoša*-) is underlyingly /guz-/ and the /z/ is lost in declined forms between vowels as for instance, NSm *gguvä*’ Z 8.35 /guë < gužë < guž- + -ë/, NAPm *gguva*’ Z 5.1 /gua < guža < guž- + -a/. In those spellings the peg akṣara *v* + apostrophe indicates that a sequence of vowels is present. With the IASm suffix *-āna*, the spelling *ggū’ṇa* Z 2.22 reflects the derivation /guṇa < gužṇa < gužṇa < guž- + -ṇa/. The long *ū* + apostrophe is used to write /uë/. The /ṇ/ is created in the derivation by echoing the retroflexion of /z/. The /z/ is subsequently lost leaving /ṇ/ on the surface with no visible phonetic or orthographic trigger.

A similar case is shown by *bāṣa*- ‘poison’ (cf. Av *viša*-). Compare the derivations NSm *bei*’ Z 3.76 /βëë < βžë < βž- + -ë/ with /ëë/ written *ei* + apostrophe, and ASm *bātu*’ Z 7.47 /βëu < βžu < βž- + -u/ with peg akṣara *t* + apostrophe helping to write /ëu/. With IASm *-āna* again a surface /ṇ/ is created in *bei’ṇa* /βëṇa < βžṇa < βžṇa < βž- + -ṇa/.

The frequently occurring word *āṇa*-, a p.ptc.m meaning ‘sitting, dwelling’ always has *ṇ* in Old Khotanese, and never apostrophe. It is a curious formation about which much may be said elsewhere. Synchronically there is a stem *āṣ*- ‘sit, remain, dwell’ which is suppletive with *ās*-. Traditionally the stem is listed as *āh*-. The stem *āṣ*- is synchronically underlying in:

3Pp.m *ā’re* Śgs 3.1r3 • /āārē < āžārē < āž- + -ārē/

3Ssj.m *ā’tē*’ Z 2.222 • /āādē < āžādē < āž- + -ādē/

---

<sup>35</sup> Previous scholarship list the stem as either *gguva*’- (the Leumanns Glossar:419a; Bailey DKS:88b) or *ggua*’- (Emmerick SGS:332; Skjærvø SuvII:260b; Canevascini Sgh:260b). The stem is an abstract, underlying shape and in this case should be *gguṣa*-.

*āñā*-p.nec ‘to be dwelled in’ NSm *ā’ñä* SI P48r4 • /āāñě < āzāñě < āz- + -āñ- + -ě/  
*āmatā*-noun ‘dwelling’ ASf *ā’mato* Suv[Or] 8.59 • /āāmado < āzāmado < āz- + -  
 āmad- + -o/

The stem *āṣ-* is surely contained in *āñā-*, as indicated by the retroflex *ṇ*: /āāñ- < āzāñ- < āzāñ- < āz- + āñ-/. The absence of the apostrophe may have one of two explanations. Either the word is so frequent that it has a conventionalized spelling without apostrophe, or it has irregularly contracted the two vowels, i.e. /āñ- < āāñ-/. The present participle *āñā-* is the only one which in Old Khotanese is attested as being reinforced by the *-andaa-* (m) or *-amkyā-* (f) present participle suffix (KhSuf:29b). The forms ASm *āñamḍau* Sgh[15] 70.2 and NSf *āñamkyā* Suv[Or] 1.13 suggest that the origin of *āñā-* as a participle from *āṣ-* has been forgotten and that *ṇ* here had phonemic status.<sup>36</sup>

The sound [ṇ] was likely not phonemic in pre-Khotanese. After intervocalic *ṣ* /z/ was lost in those words with a following /ṇ < n/, those cases of /ṇ/ developed phonemic status. It may be preferable to describe the status as marginally phonemic, since most cases of *ṇ* are predictable. Certainly, by the time of Late Khotanese, the /ṇ/ is fully phonemic. The graph *ṇ* is found where Sanskrit imitation is not to be considered, e.g. *khūñā-* ‘hole’ (DKS:75a); *gāñāṇ* ‘plant name’ (DKS:82).

This scenario may explain the spellings of ‘queen’ in OKh *rrīñā-* with *-ñ-* and in LKh *rrīnā-* with *-n-*. The consistent (13× in Z) but likely orthographic *ṇ* in OKh is

---

<sup>36</sup> The unique OKh word *oñä* Z 13.35, E. and M. Leumann ‘gewaltig’ (Glossar:406a), Emmerick ‘strength’, Bailey ‘powerful’ (DKS:46a), should in theory reflect historical \*ṣ. There is *oñā-* in the Khotanese index to EDIV (522d) but no discussion on the indicated page (429).



replaced by phonemic *n* in LKh. That the *ŋ* is orthographic is perhaps proven by GDSf *rrīñe* twice in Suv[Or] (6.2.78, 6.3.4) beside NAPf *rrīñe* (6.3.21).

## 2.8.0 OKh *ḍ*, stop or liquid?

There is some question as to whether OKh *ḍ* was phonologically a stop or a liquid. The evidence seems to suggest that it was a single phoneme with stop and liquid allophones. In Sanskrit, *ḍ* is a stop. But in the language from which Khotanese borrowed the Brāhmī, *ḍ* may have marked a retroflex [ʎ]. In Pali, intervocalic *ḍ* has become *ḷ*. And in BHS, *ḍ* is very commonly found for Skt intervocalic *ḷ* (Edgerton 1953a:§2.46). Bailey 1938:586 compared Khotanese *ḍ* < \**rt* to Ishkashmi *ḷ* < \**rt*. For example, Ishkashmi *keḷ* ‘knife’, cf. Av *karəti-*; *mūḷūk* ‘muž(čina)’ < \**mṛta-* + *-aka-*, cf. Old Persian *martiya*; *kūḷ-* past stem ‘do’ < \**kṛta-* (Pakhalina 1987b:494).

### 2.8.1 OKh synchronic /ḍ > l/

In Old Khotanese there is a synchronic process of /ḍ > l/ which suggests but does not prove a liquid value for *ḍ*.

There are five present stems ending in *-ḍ-*, which show /ḍ > l/ before /t/ in the 3S present. These show type B (cf. SGS:177) and, as far as can be determined, middle morphology. The *-ḍ-* is historically from Pre-Khotanese \**-rt-*:

<i>baltte</i>	< <i>baḍ-</i>	move; writhe	(cf. 3P <i>baḍāri</i> )	* <i>vart-</i>
<i>nyūltte</i>	< <i>nyuḍ-</i>	rush down	(cf. 3P <i>nyūḍāre</i> )	* <i>-vart-</i>
<i>ggaltte</i>	< <i>ggaḍ-</i>	lie about	(cf. 3P <i>ggaḍāre</i> )	* <i>gart-</i>
<i>haṃggaltte</i>	< * <i>haṃggaḍ-</i>	result; develop	—	* <i>-gart-</i>

*patälttä* < \**patäd*-<sup>37</sup> cut off — \**-kart-*

Traditionally, the type B, 3Sp.a. suffix has been given in the form *-tā* and the middle as *-te*. These transcriptions might be interpreted as showing that the dental stop in these suffixes is underlyingly voiced, since single *t* is /d/.<sup>38</sup> An inflection like *baḍ-* > *baltte* would show /ḍ+d > lt/ with suspicious suffix initial *t*=/d/ devoicing to *tt*=/t/ after /ḍ/. But a review of the 3S B inflectional evidence suggests that the dental stop in the suffixes is voiceless and thus that transcriptions *-ttā* and *-tte* may be preferred. Later, in discussing the metanalysis, it is further argued that the suffixes underlyingly, synchronically contain double /tt/.

### 2.8.1.1 Suffix initial /t/ in the 3S of B stems

There are three synchronic behaviors of 3S present B suffix initial /t/ in Old Khotanese:

- 1) It remains voiceless after stem-final voiceless segments:

---

<sup>37</sup> This stem is listed as *patält-* in SGS and DKS but note that \**-rt-* gives *-ḍ-* while \**-rd-* gives *-l-*; see above, 2.3.4.2 and cf. Av *kart-* ‘cut’. There is insufficient evidence to determine if the verb is middle or active.

<sup>38</sup> The transcriptions date from before a rigorous distinction between diachronic and synchronic analyses was recognized. Probably no specific phonological content was intended.

**Table 2.22:** Suffix initial /t/ after voiceless segments

<i>s</i>	/s/	<i>grūs-</i>	a.	<i>grūstā</i>	call	/s+ <sup>i</sup> tt > št/
<i>ṣṣ</i>	/ṣ/	<i>huṣṣ-</i>	a.	<i>huṣṭā</i>	grow (intr.)	/ṣ+tt > ṣṭ/
<i>śś</i>	/ś/	<i>panaśś-</i>	m.	<i>panašte</i>	perish	/ś+tt > št/
<i>ts</i>	/t <sup>sh</sup> /	<i>patäts-</i>	m.	<i>patäste</i>	give up, abandon	/t <sup>sh</sup> +tt > st/
<i>ch</i>	/č <sup>h</sup> /	<i>gvach-</i>	m.	<i>gvašte</i>	be digested	/č <sup>h</sup> +tt > št/
<i>th</i>	/t <sup>h</sup> /	<i>nuvaṃth-</i>	–	<i>nuvaṃtte</i>	be removed	/t <sup>h</sup> +tt > tt/
<i>h</i>	/h/	<i>ggīh-</i>	m.	<i>ggitte</i>	help	/h+tt > tt/ <sup>39</sup>

I find no examples after *tc* /t<sup>s</sup>/, *c* /č/, *kṣ* /t<sup>s</sup>/, *ṭh* /ṭ<sup>h</sup>/, *ph* /p<sup>h</sup>/.

- 2) It voices to /d/ (or /ḍ/) after voiced sibilants and affricates, and after sonorants<sup>40</sup> except /l/

---

<sup>39</sup> The same synchronic development is shown by *jah-* m. > *jatte* LKh ‘be cleaned, cured’, but not by *\*hamäh-* m. > *hamättä* ‘change (intr.)’ as this stem was metanalyzed and is synchronically *hamä-* (cf. 3Pp.m. *hamyāre* Z 6.17). The two other type B stems listed with final *-h-* in SGS, *ah-* a. ‘be’ and *āh-* m. ‘sit; remain’, show a stem in *-s-* in the 3S present: *as-* > *aštā*, *ās-* > *āste*.

<sup>40</sup> *byaude* < *byau-* ‘to be found; to be’ shows irregular *-d-* instead of *-t-*. Verbs meaning ‘to be’ frequently show irregular morphology.

**Table 2.23:** Suffix initial /t/ after most voiced segments

<i>rr</i>	/r/	<i>purr-</i>	a.	<i>purdä</i>	overcome	/ɹ+tt > ɹd/
<i>r</i>	/r̥/	<i>handār-</i>	m.	<i>handāde</i>	care for	/r̥+tt > d̥d/
<i>n</i>	/n/	<i>ysān-</i>	m.	<i>ysānde</i>	shine	/n+tt > nd/
<i>ñ</i>	/ñ/	<i>hvañ</i>	m.	<i>hvīnde</i>	be called	/ñ+tt > <sup>i</sup> nd/
<i>ys</i>	/z/	<i>gvays-</i>	m.	<i>gvaysde</i>	be separated	/z+tt > zd/
<i>ṣ</i>	/z̥/	<i>nijaṣ-</i>	m.	<i>nijaṣde</i>	show	/z̥+tt > z̥d/
<i>ś</i>	/ʒ/	<i>baś-</i>	a.	<i>*baśdä<sup>41</sup></i>	be suitable	*/ʒ+tt > ʒd/
<i>j</i>	/j̥/	<i>uskuj-</i>	m.	<i>uskuśde</i>	rise up against	/j̥+tt > ʒd/
<i>js</i>	/dʒ/	<i>bihījs-</i>	m.	<i>bihīysde</i>	increase	/dʒ+tt > zd/

I find no example after *m*, *v*, *y* or a vowel. SGS lists type B present stems ending in *-uv-*, *-ūv-*, *-iy-*, *-īy-* and vowels. These are all type A vowel stems as shown in Hitch in press a and b (04 and 05).

3) It remains voiceless after voiced /d, d̥, l/<sup>42</sup>

**Table 2.24:** Suffix initial /t/ after voiced /d, d̥, l/

<i>t</i>	/d/	<i>pat-</i>	a.	<i>pītā</i>	fall	/d+ <sup>i</sup> tt > <sup>i</sup> tt/
<i>d̥</i>	/d̥/	<i>baḍ-</i>	m.	<i>baltte</i>	move, writhe	/d̥+tt > lt(t)/
<i>l</i>	/l/ <sup>43</sup>	<i>saṃkhāl-<sup>44</sup></i>	m.	<i>saṃkhiltte</i>	be tainted	/l+tt > lt(t)/

<sup>41</sup> There appear to be no type B stems ending in *-ś-* in Old Khotanese. Emmerick identified two instances of LKh *baś-* ‘be suitable’ and noted, “*baśtā* is L.Kh. for *\*baśdā*” (SGS:94).

<sup>42</sup> Type B stems like *bu-* ‘know’ or *nā-* ‘sit down’ which have been traditionally transcribed with final *-d-*, e.g., *bud-*, *nād-*, were involved in the metanalysis and are synchronically vowel final stems. At an earlier period of the language they ended in a voiced obstruent, /δ/ or perhaps /d/, and then had a 3S derivation involving /δ+t > tt/. This is the historical explanation behind 3Sp.m *butte* and 3Sp.a *nättā*.

Stems ending in /d/ in Old Khotanese show a cluster /tt/ in the -*tt*- of the 3S present (Hitch 2015a:670–677). Stems ending in /l, ɖ/ show a cluster written *ltt* which probably reflects /lɖt/. The historical explanation of the voicelessness in the suffix initial is straightforward. In *pīttā* the -*tt*- is from *\*t+\*t*, in *baltte* the *tt* is from *\*rt+\*t*, and in *saṃkhiltte* it is from *\*rd+\*t*. But it is not obvious what the precise value of *ɖ* is in the synchronic grammar. On the one hand it could be argued that *ɖ* is a liquid since it patterns with *l*. If so, then *baɖ-* > *baltte* shows /lɖ+tt > ltt/ and simply assimilation from retroflex to dental position, a process involving one feature. On the other hand, if *ɖ* is viewed as a stop, then the process, /ɖ+tt > ltt/, involves not only assimilation to position, but also dissimilation from stop to liquid. Further, if this were true, then it should be asked why the stop /d/ does not also synchronically dissimilate to a liquid, i.e., *pat-* ‘fall’ > *\*pīttā* with /lt/ instead of *pīttā* with /tt/. This evidence by itself suggests a liquid value for *ɖ*.

### 2.8.2 Inventory gap

There appears to be a gap in the voiced retroflex segment inventory of either a stop or a liquid:

---

<sup>43</sup> There are just two examples of B stems ending in -*l*- in OKh. Both show evidence that they may actually end in -*ly-*: *saṃkhalyāre* 3P LKh; *pātāl-* > *pātālyāre* 3Pm. (DKS lists *pātāly-*).

<sup>44</sup> DKS:417 and SGS:130 give the stem as *saṃkhal-*, but note the *āmata*-noun NSf *saṃkhālāmata* ‘defiling’ Vim 3.39.3.9.

**Table 2.25:** Gap in the retroflex inventory

	sibilant	stop	lateral	r-sound
dental	<i>ys</i> /z/	<i>t</i> /d/	<i>l</i> /l/	<i>rr</i> /ɹ/
retroflex	<i>ṣ</i> /z̥/	— <- <i>ḍ</i> /ʈ/ -> —	—	<i>r</i> /ɹ̥/

One explanation might be that the language had two phonemes /ḍ/ and /l̥/ but that the script had no way of distinguishing them. That can be rejected as all segments written *ḍ* have clearly unified historical and synchronic origins. Instead, the explanation is probably that there is a single phoneme, which for convenience is best rendered /ḍ/. This has two phonetic allophones, a stop [ḍ] in the clusters /ṇḍ, ṣḍ, ḍḍ/ or a liquid [l̥] between vowels. It may be that the consonant written *ḍ*, which clearly is voiced and retroflex, is phonemically unmarked for either laterality or continuance and may be phonetically lateral or non-lateral and continuant or non-continuant depending on the environment.

### 2.8.3 Late Khotanese liquid *ḍ*

In Late Khotanese, *ḍ* is certainly a liquid. As Bailey 1938:585-6 pointed out, it can represent a foreign *l*, e.g., *rrgyaḍäsūṃmā* = Tib *rgyal-sum*; *aḍpā* = OT *alp*; *kaḍāṇa* = Skt *kalyāṇa*. Similarly, a Prakrit *ḍ* (= Skt *ṭ*) can be written with *l*, e.g., *nālai*, Skt *nāṭaka*, *kūla*, Skt *koṭi*. Emmerick 1981 and 1989 treated Kh *ḍ* as [lʲ], an [l] with a palatal off-glide. This value is unlikely for OKh but something similar may be correct for LKh. This is suggested by the evidence of Late Middle Chinese and Old Turkic words in LKh transcription.

### 2.8.3.1 OT /l/ in LKh Transcription

In Old Turkic words in Late Khotanese writing, OT /l/ is written in disparate ways in various texts as observed by Hovdhaugen 1971:177. In some documents it is written with *ḍ*, in some with *l*, and in others with *ḍ* or *l*, apparently without any system. In the list of cavalry or equestrian terms, the so-called *Turkish-Khotanese Vocabulary*, /l/ is written *ḍ* when preceded or followed by /i, ī/ (8 times) and is written *l* elsewhere (6 times) except for one instance of *r*. In the lists below, I give the Khotanese spelling, a phonetic transcription, the phonemic or normalized transcription given by Emmerick and Róna-Tas 1992, their gloss, and the paragraph in which the term is discussed.

**Table 2.26:** LOT /l/ spelled *ḍ*

LKh script	phonetic	Emmerick and Róna-Tas 1992
<i>aḍḍinā</i>	[alɪn]	/alɪn/ 'forehead' 34
<i>bīḍakä</i>	[biläk]	/biläk/ 'wrist, forearm' 61
<i>īḍaihä:</i>	[ilig]	/iliy/ 'attachment [to the harness]' 17
<i>īḍäpacākä</i>	[ilpačoq]	/ilpačoq, ilpäčök/ '?' 90
<i>saḍī</i>	[salī]	/sali/ 'small leather or felt pieces stuck into the quiver to fix the arrows' 9
<i>ttīḍī</i>	[tili]	/tili/ 'its tongue [of the buckle]' 24
<i>ttāḍai</i>	[tolī]	/toli, töli/ '?' 32
<i>yīḍädi</i>	[yildi]	/yildi/ 'descended (?)' 51

**Table 2.27:** LOT /l/ spelled *l*

<i>kāmülädrühä:</i>	[kömüldrüg]	/kömüldrüy/	'breast strap' 21
<i>kālūnä</i>	[qolun]	/qolun/	'saddle [girth]' 22
<i>kūrñälükä</i>	[qurñoluq]	/qurñoluq/	'bow case' 3
<i>sakalädrükä</i>	[saqaldruq]	/saqaldruq/	'sub-mental strap on the harness' 16
<i>tulūnä</i>	[tulun]	/tulun/	'temple' 35
<i>ülūnä</i>	[ulun]	/ulun/	'arrow shaft' 30

**Table 2.28:** LOT /l/ spelled *r*

<i>ttirakä</i>	[tīlaq]	/tīlaq/	'clitoris' 77
----------------	---------	---------	---------------

On the basis of this evidence, the rule seems to be that the /l/ was written *ḍ* before or after high, unrounded vowels. This might suggest a difference between LKh *ḍ* and *l* like that between Polish 'clear' *l* and 'dark' *ł*, the 'dark' variety being associated with labiovelarity. In modern Uyghur, [i, i̯] have merged in /i/ which may be pronounced as back [i̯] in the vicinity of back consonants or syllables containing back vowels (Nadzhip 1971:49). It is conceivable that the merger was already underway in the late stage of Old Turkic recorded in LKh script. In this case, the rule in the Turkish-Khotanese vocabulary may be that OT /l/ tends to be written *ḍ* beside /i/.

### 2.8.3.2 Palatal *ḍ*?

In LMC words, as mentioned above (2.1), *ḍ* denotes /l/ before [i, ü], and *l* is used elsewhere. This accords with the second interpretation of the OT evidence above (8.3.1). It appears plausible that LKh *ḍ* is a palatalized or palatal segment



corresponding to dental  $l = [l]$ . Because of structural considerations it is possible that  $\check{d}$  is a palatal lateral  $/\lambda/$ , with the same phonemic point of articulation as  $y, c, ch, \tilde{n} /y, \check{c}, \check{c}^h, \tilde{n}/$ . It is unlikely that the segment would stand for  $[l^y]$ , a dental with a palatal off-glide since there are no other such dentals in the language.

If LKh  $\check{d}$  was a palatal lateral  $/\lambda/$ , it remains unexplained how an originally retroflex segment could develop into a palatal one. It would be useful to find a parallel development in another language showing  $[\check{d} > \check{l} > \lambda]$ .

## 2.9.0 Meter

### 2.9.1 Double *tt* as */tt/*

In Old Khotanese the orthographic sequence *tt* when intervocalic, *VttV*, is ambiguous, standing for either single */t/* or double */tt/*. Essentially, orthographic *VttV* occurs in four kinds of words, (1) Indian loans, (2) conjugated present stems, (3) past participles, and (4) words reflecting an adjective ending in a consonant and a suffix historically beginning with *\*t-*. In “*tt* in Old Khotanese” (Hitch 2015a:668–680) I examined the metrical distribution of these categories in detail and determined that (1) and (2) have double */tt/* while (3) and (4) have single */t/*.

### 2.9.2 *ṭhṭh* and *ḍḍ*

The spellings *ṭhṭh* and *ḍḍ* also show double consonants. There are eighteen instances of *ṭhṭh* which may be examined with the meter of Z, and one from Sgh. There are seven instances of *sṭhṭha-* ‘a bird of prey’ in Z. Of these, there are five, NAPm *sṭhṭha* Z 2.46, 20.30, 24.412, GDSm *sṭhṭhi* Z 21.20, 22, in which the mora count is three, confirming that *ṭhṭh* denotes two consonants and that the first

syllable is heavy (H) . For example in the first hemistich below *suṭṭha* is three moras in a six mora segment, and in the second, *suṭṭhi* is three moras in a five mora segment.

24.412ab uskāna nu | bendā **suṭṭha** | bei'sśa västāta (B:5+6+7)

Vultures came high above them in a whirl.

21.20ab cūḍai vaysña | **suṭṭhi** □ paśśā | thīye puṣṣo (C:7+5+5)

Why do you now leave it to the raven to pull out competely?

The two instances of GDPm *suṭṭhānu* Z 21.26, 21.30 do not clearly show a heavy initial syllable. They both occur in a segment which is normally nine moras, *suṭṭhānu śvānānu*, but here it will count as ten if the first syllable is heavy. E.g.,

21.30cd **suṭṭhānu** śvānānu □ | rrāśa | vaysña biśśā (C1:10+3+5)

They are all now in the control of vultures, dogs.

The strongest metrical evidence for the structure of *suṭṭha*- actually comes from a metrical passage in the *Saṅghāṭasūtra*. Here NAP *suṭṭha* is HL in cadence 2 **HXHL**:

Sgh[17] 243[33a] bārga rrū|vāsa ṣundā | śvānū **suṭṭha** (B:5+6+7)

wolves, jackals, ravens, dogs and vultures

The two instances of GDSf *hīṭṭhe* Z 6.59 (2×) from *haṭṭhā*- 'truth' provide no evidence about the structure of *ṭṭh* since the *ī* makes the syllable long. There are seven<sup>45</sup> cases where the first syllable is spelled *haṭṭh*-. All seven are best analyzed with a heavy first syllable. The two most decisive examples are below. The first has NSf *haṭṭha* as HL in a cadence 1 **HLLHL**, and the second has NAPf *haṭṭhe* as HL in **HLLHL**:

---

<sup>45</sup> NSf *haṭṭha* 7.28, 22.316, NAPf *haṭṭhe* 6.29, 10.10, 28, 29, ASf *haṭṭho* 22.241.

22.316ab ggāṭhā nā | ttīyi ni bvā□re | kāmānu | **haṭṭha** u dātā (A:5+7+5+7)

Then the householders will not know who have the truth and the Law,

10.29cd pātco vari | ṣṭāni paysendā □ | tcohorī | handare **haṭṭhe** (A:5+7+4+7)

then at once he recognizes four other truths.

The final two *ṭṭh* spellings in Z are in the word NAPf *śśaṭṭhe* n. ‘deception’, adj. ‘deceitful’. The first hemistich below has *śśaṭṭhe* as HL in a cadence 1 HLLHL, and the second has *śśaṭṭhe* as HL in a cadence 2 HXHL:

23.126cd strīye atā | īrate **śśaṭṭhe** □ | kye nā hotte | uairā sañānu

Women are very cunning, deceitful. Who is capable of plans to match them?

23.172 tvī prajña | pamāta stārña □ | avamāte | īrye **śśaṭṭhe** . (A:5+7+5+7)

(you have the limited wisdom) of a woman, the unlimited wiles (and) deceptions,

The word is spelled once with single *ṭh* but the meter shows an underlying two consonants. Below *śśaṭṭhe* counts as three moras in the five mora segment:

24.268cd cvī **śśaṭṭhe** | byūgga □ drūje | īrye yole (B:5+6+7)

trickeries, disparagements, lies, wiles, evils

As shown below (Metanalysis), the spelling *ḍḍ* stands for a double consonant.

From *bar-* ‘ride’, the 3Sp.m *baḍe* Z 2.95, 13.26, 24.120, 249 is shown by the meter to be underlyingly /βaḍḍē/ and is in fact once spelled *baḍḍe* Z 13.149 3Sp.m. We can assume that all type B *-r-* stems which show single orthographic *ḍ* in conjugated forms with /tt/ initial suffixes have underlying /ḍḍ/.

Recognizing underlying double /ḍḍ/ in the orthographic *-VḍV-* in conjugated type B stems permits an interesting generalization about the oldest stage of the language: Whenever a /tt/ initial suffix attaches to a B stem ending in a consonant, the result is always a consonant cluster.

### 2.9.3 *rr* and *r*

Emmerick 1981:207 noted that in meter,

-*rr*- did not make position. Thus, *ttare* ‘grass’, *purra* ‘moon’, and *ysurrä*- ‘anger are found in iambic endings along with words like *pharu* ‘many’ and *ysärä* ‘heart’.

This evidence alone should be enough to show that *rr* is a single segment, and not phonemically equivalent to *r + r* as Emmerick at one time suggested (1981:209, 1989:209). He later thought the language had “two *r*-sounds” (2009:383). The metrical evidence that *rr* denotes a single segment is clear. In the first hemistich below NSm *uspurri* ‘complete’ is HLL in cadence 1 **HLLHL** while in the second, NSf *aysurra* ‘wrathless’ is LLL in cadence 4 **LLLLHL**:

2.212ab    avāyā ne | ysam̐thu ne nā□ste | indryau jsa | **uspurri** ysaite  
(A:5+7+5+7)

He does not take birth in Apāya. He is born complete in senses.

3.77ab    **aysurra** avīra □ | maitra bhā|vāñña mamä (C:7+5+5)

I must meditate on love without wrath, without hatred.

### 2.10.0      **Old Khotanese allography**

It has long been recognized that OKh writing contains a systematic and complex allography where phonemes have different spellings in different contexts. At least a part of this was likely borrowed, with the script, rather than invented, and, as such, the allography partly reflects phonological developments within Middle Indian.

#### 2.10.1      **Dental**

From the earliest texts, single *t* represents voiced /d/ while double *tt* shows voiceless /t/. When written in clusters with other consonant graphs, the group is

voiceless, e.g. *st* /st/, *śt* /śt/, *ṣṭ* /ṣṭ/. It seems that *t* in clusters is always voiceless. It is likely that this set of relationships was borrowed from use in Prakrit. In Hitch 2015a I outline the orthographic evolution of *t* (and *d*) from Middle Indic to Old Khotanese (682–685).

In contrast, Middle Indic does not appear to have had a phonemic voicing contrast among the palatal and retroflex sibilants. At the least a contrast is not readily apparent from the records.<sup>46</sup> The Old Khotanese device of writing double *śś* and *ṣṣ* to show voicelessness is introduced within Old Khotanese. Unlike *tt*, *śś* and *ṣṣ* are not a regular part of the archaic orthography reflected, for instance, in the *Śūraṅgamasamādhisūtra* (Śgs:xix). And unlike single *t* in clusters which is always voiceless, single *ś* and *ṣ* in clusters may be either voiceless or voiced. The voicing of the sibilant matches the voicing of the other consonant(s) in the cluster, e.g., *śt* /śt/ ~ *śd* /žd/, *ṣṭ* /ṣṭ/ ~ *ṣḍ* /zḍ/.

Parallel to the voiceless~voiced allophony shown by *t* is a stop~continuant allophony demonstrated by *d*. When single it marks a continuant /δ/ and in clusters it may show /d/, e.g., *nd* /nd/, *rrd* /rd/, *ysd* /zd/. This results in a complicated but effective graphic system in which three phonemes /t, d, δ/ are written with two graphs *t* and *d*. The system reflects the fact that within the Middle Indian language

---

<sup>46</sup> Many of the Indian loans in Khotanese show *ś* [ž] corresponding to Sanskrit *y*, e.g., OKh GDSm *ttāryaśūni* Z 4.87 cf. Skt *tiryagyoni* 'animal', GDSm *śamā* Z 24.45 'Yama', *śāś-* 'ask for' (p.nec NSm *śāśāñi* Z 13.24) cf. Skt *yācati* 'asks for' and Gdh *yayadi* 'begs'. Possibly there was *y*=[ž] in some cases in one or more Prakits. If so, then this offers an explanation for the development of the *ys*=/z/ device we see in Khotanese but known earlier in India (cf. von Hinüber 1980:125–126 and footnotes 21–23). Just as dental *s* modifies palatal affricate *j* [j] to dental affricate *js* [dʒ], dental *s* modifies palatal sibilant *y* [ž] to dental sibilant *ys* [z].

which was the source of the Brāhmī, there had been intervocalic weakening of /t, d/ to /d, δ/ and a partial consonant shift (Hitch 2015a:682–685).

### 2.10.2 Phonemic contrasts and allography

The dental allography just discussed is well understood. What has not been made clear is that similar kinds of allography affected the spellings of labials and velars as well. These points of articulation may also have been affected by the Middle Indian shift. That is, between vowels we might expect the Middle Indian to have had [g>ɣ, k>g] and [b>β, p>b]. These developments ought also to be reflected in Khotanese writing. But here a complication arises. While there is in OKh a three way phonemic contrast /t~d~δ/ among the dentals, there seems to be just a two way phonemic contrast among the velars and labials. The velar contrast is voicing /k~g/ while the labial is continuance /b~β/. This creates an inventory with two gaps.

**Table 2.29:** Labial and velar obstruent inventory gaps

—	/t/	/k/
/b/	/d/	/g/
/β/	/δ/	—

While such an inventory may look suspicious due to the asymmetry, it is typologically ordinary. As established by Melishkvili, among plosives the combination of labiality and voicing is unmarked while the combination of velarity and voicing is marked. Conversely the combination of labiality and voicelessness is marked while that of velarity and voicelessness is unmarked (Gamkrelidze

1978:15). Gamkrelidze notes, “Systems with gaps in the class of plosives, opposed on the feature of voicing/voicelessness, have basically the following forms:

(a) b —	(b) b p	(c) b —
d t	d t	d t
g k	— k	— k” (ibid.:15–16)

This markedness principle may be extended to the voiced fricatives. Many languages show a gap in the voiced velar fricative slot. Some languages that have either /p, t, k/ or /b, d, g/ or both, and, like Old Khotanese, also have /β, δ/ but no \*/ɣ/ are Iai (East Malayo-Polynesian), Dahalo (Cushitic), Tacuna (Macro-Panoan), Cubeo (Central Tucanoan), Mixtec (Oto-Manguean), and Fijian. Languages with /p, t, k/ and /β/ but lack \*/δ, ɣ/ are Amuzgo (Oto-Manguean), Ejagham (Bantoid), and Pashto. I notice no language with precisely the same inventory as Old Khotanese, but Dizi (Omotiic) comes close. It is identical to the Old Khotanese except that it lacks a /δ/ as well as /ɣ/, like Amuzgo, Ejagham and Pashto mentioned above (Brasington 2012 s.v.).

### 2.10.3 Labial

Among the labials, the symbol *b* has the same dual function as the graph *d*. That is, single *b* is the continuant /β/ while in clusters *b* may show a stop, e.g. *mb* /mb/ or *rb* /rb/ as in IASm *tcārbina* ‘fatty’ Z 22.147, or 3Sp.a *sarbätä* ‘rise’ Z 2.43.<sup>47</sup> What has

---

<sup>47</sup> While the instances of *rb* in Old Khotanese are rare (most commonly perhaps in *harbiśśa* ‘all’), there do not seem to be any of *rp* or *rrp* at all in that stage of the language. There are both *pr*~*prr* and *br*~*brr* in native Old Khotanese words, e.g., NSm *prrahaṇä* Z 2.63 ~ ASm *prahaṇu* Sgh[23]216.1 ‘clothing, IASf *brītye jsa* Z 13.59 ~ *brrītye jsa* Z 19.19 ‘love’. Because *pr(r)* must be /bɿ/, then *br(r)* must contain the labial fricative /βɿ/. This raises the

not been recognized until now is that single *p*, like single *t*, marks a voiced phoneme. It is shown in 6.0-6.4 that *p*=/b/ and that OKh lacks a voiceless \*/p/ phoneme. However, it seems likely that /b/ had a voiceless allophone [p] in voiceless clusters, e.g., *sp* [sp], *ʃp* [ʃp] (see 6.3). On a phonetic level, the dental and labial allographies are largely parallel.

**Table 2.30:** Labial and dental allography

	voiced graph		voiceless graph	
	single [continuant]	cluster [stop]	single [voiced]	cluster [voiceless]
labial	<i>b</i> =[β]	<i>mb</i> =[mb]	<i>p</i> =[b]	<i>sp</i> =[sp]
dental	<i>d</i> =[δ]	<i>ysd</i> =[zd]	<i>t</i> =[d]	<i>st</i> =[st]

#### 2.10.4 Velar

At first glance the state of the velar spelling in Khotanese seems to suggest that the Middle Indian velars did not undergo the same kind of weakening as the labials and dentals. OKh *k* seems always to mark /k/ and *g* or *gg* seem always (except where *g* marks /ʔ/) to show /g/ (cf. Emmerick 1981:188-89). But probably the opposite is true. That is, the weakening was more advanced here with intervocalic [ɣ] disappearing altogether leaving just two phones [k, g] and two characters, *k*, *g*.

The archaic orthography of the *Śūraṅgamasamādhisūtra* features just *k* and *g*. It does not have double *gg* at all (*Śgs*:xix). This was a later invention. Compare the following cases of single *g* in *Śgs* with double *gg* in the later orthography of Z: LSm

---

possibility that *rb* is /ɹβ/, which would make sense in *harbiśśa*- 'all' as it is related to *biśśa*- 'all' which begins with /β/.



*guvo*’ ~ *gguvo*’ ‘ear’ (Śgs 3.2v2 ~ Z 8.35); ASm *agūnau* ~ ASm *aggūnaina* ‘without characteristics’ (Śgs 2.7v2 ~ Z 6.8); 3Spf.tr.m *gurṣṭe* ~ *ggurṣṭe* ‘he called’ (Śgs 3.4r5 ~ Z 20.50).

There are two other differences between the Śgs orthography and that of later Old Khotanese with respect to the use of *g*. One is, we tend to find intervocalic *g* in Śgs where later there is a tendency to write *t*. In the list below the relative frequency of intervocalic *g* and *t* in words with historical \*[g] is compared for Śgs and Z. This shows that *g* is preferred in AOKh and *t* in COKh. The spellings of *nāga-* in Z go against the trend, probably because the Sanskrit origin was recognized as shown by the one spelling with double *gg*.

**Table 2.31:** Archaic *g* ~ Canonical *t* for /ŋ/

	Śgs	Z
<i>agāṣṭa-</i> ~ <i>atāṣṭa-</i> ‘inconceivable’	<i>g</i> 6×, <i>t</i> 1× <sup>48</sup>	<i>g</i> 0×, <i>t</i> 3× <sup>49</sup>
<i>bāga-</i> ~ <i>bāta</i> ‘root’	<i>g</i> 2× <sup>50</sup> , <i>t</i> 0×	<i>g</i> 0×, <i>t</i> 3×
<i>āgāśa-</i> ~ <i>ātāśa-</i> ‘sky’	<i>g</i> 2× <sup>51</sup> , <i>t</i> 0×	<i>g</i> 1×, <i>t</i> 39×
<i>kāḍāgāna-</i> ~ <i>kāḍātāna-</i> ‘evil deed’	<i>g</i> 1×, <i>t</i> 0×	<i>g</i> 0×, <i>t</i> 1×, <i>y</i> 10×
<i>nāga-</i> ~ <i>nāta-</i> ‘serpent, Nāga’	<i>g</i> 2× <sup>52</sup> , <i>t</i> 0×	<i>g</i> 14×, <i>gg</i> 1×, <i>t</i> 4×

<sup>48</sup> The Bailey and Emmerick reading 2.2v2 [a]hutāṣṭā was corrected by Skjærvø to *atāṣṭā*, see St.I:129 s.v. *hutāṣṭa-* “well-thought.”

<sup>49</sup> NSm *atāṣṭā* Z 2.151, 14.87, NAPm *hutāṣṭe* ‘well thought’ Z 22.277.

<sup>50</sup> NAPm *bāga* Śgs 2.4v1, NSm *nasta-bagā* ‘root (of merit)’ Śgs 3.9v3.

<sup>51</sup> NSm *āgāśā* 3.5v4, *āgāśalstu* ‘skyward’ 3.9v1.

The *g* and *t* in these forms is sometimes described as a hiatus element or hiatus filler, but is here regarded as phonemic glottal stop /ʔ/ (see 2.10.4.1).

The other difference involves forms based on the stem *pāṣa-* ‘power’. In these items in Śgs we find *g* used as a peg akṣara with the apostrophe nine times and *t* once<sup>53</sup>, while in Z we find peg akṣara *t* with apostrophe more than thirty times, and no instances of *g* with apostrophe (Glossar:462b)<sup>54</sup>. Some examples:

**Table 2.32:** Archaic *g* ~ Canonical *t* as peg radical

<i>Śūraṅgamasamādhisūtra</i>	<i>Book of Zambasta</i>
<i>pāga</i> ’ (3x), <i>pā’ga</i> (2x), <i>pā’ga</i> ’ (Śgs:125a)	<i>pāta</i> ’ NAPm 10.28 ‘power’
<i>pā’gañyau jsa</i> IAP 3.12v2	<i>pāta’ñyau jsa</i> 13.145 IAP ‘power’
<i>pā’gajsa</i> NAPm 3.8v3	<i>pāta’jsa</i> NAPm 10.2 ‘powerful’
<i>pā’gajsānu</i> GDPm 3.12v1	<i>pāta’jsu</i> ASm 22.126 ‘powerful’

These three differences with regard to *g* between the archaic orthography of Śgs and the later OKh system may reflect interdependent phonological and orthographic developments within OKh.

#### 2.10.4.1 Glottal Stop

In Earliest Written Khotanese, *g* everywhere stood for the phoneme /g/. At some point, intervocalic /g/ was weakened to glottal stop [ʔ]. At this point, the phoneme

---

<sup>52</sup> NAPm *nāga* Śgs 3.4v5, IASm *nāgīñāna* ‘of a *nāga*’ IOL Khot 189/4r5.

<sup>53</sup> Śgs has *g* with apostrophe 9× (Śgs:125a) and has *t* with apostrophe 1× in IAP *pā’tañyau jsa* IOL Khot 188/3v2.

<sup>54</sup> There is also one very conservative spelling with *ṣ* and apostrophe in NAPf *pāṣa’ñi* Z 13.35 ‘power’.

/g/ had [g] and [ʔ] allophones. The graph *g* continued to be written for the phoneme /g/ with no ambiguity. This is the stage perhaps attested in *Śgs*.

Later, intervocalic /d/ written *t* also develops to [ʔ] and in this position merges with the [ʔ] from /g/. There is a phonemic reanalysis so that [ʔ] is everywhere regarded as the intervocalic allophone of /d/. This reanalysis has direct consequences on the script. The device *gg* is introduced for /g/ since single intervocalic *g* is now an allograph for *t* (both mark /d/ [ʔ]). The doubling of *gg* eliminates ambiguity and since *t* marks the [ʔ] allophone of /d/, it can now be used in places where the older system had *g*.<sup>55</sup>

These developments may be summarized in a chart. Phonetic, phonemic and orthographic representations are given for each of the three stages.

---

<sup>55</sup> The *-g-* in 3Ppf.tr. *hagāṣṭāndä* ‘send’ *Śgs* 2.4r2 for later *-t-* in 3Spf.tr.m *hatāṣṭe* Z 2.237 is interesting. It suggests that ProtoKhotanese had /g/ in this word, which was pronounced [ʔ] in the Archaic period, and which later merged with the [ʔ] of intervocalic /d/. The stem in EWKh may have been *\*hagīs-*: *hagāṣṭa-*. Cheung proposes the existence of an Old Iranian stem *\*kaih-* ‘move, set in motion’ and lists stems from seven New West Iranian languages beginning *kī-* and meaning ‘send’ (EDIV:229). Thus, possibly < *\*fra-kais-aya*. The understanding of the *t~g* alternation also implies that *gāvu* and *gāmu* are enclitics rather than independent particles: *-gāvu* emphatic enclitic, *ne-gavu* ‘not at all’ *Śgs* 3.8v3, *ne-ni-tāvu* ‘they...not at all’ Z 23.134; for *-gāmu~-tāmu* as an indefinite enclitic cf. Emmerick 1981:190. Also note that the *-gan~-tan-* forms of *yan-* ‘make, do’ may reflect an enclitic stem *\*-gan-* < *\*-k-* (cf. *ibid*; SGS 111). The initial [ʔ] is spelled four ways, *y-*, *t-*, *g-*,  $\emptyset$ , in 3Pp.a in Suv[Or] 12: *yīndä* 12.6, *tīndä* 12.17, *gīndä* 12.52, *īndä* 12.49.

**Table 2.3.3:** Evolution of glottal stop

	Pre-Khotanese */g/		Pre-Khotanese */d/	
	Intervocalic	Elsewhere	Intervocalic	Elsewhere
Earliest Written Khotanese				
phonetic	[g]	[g]	[d]	[d]
phonemic	/g/	/g/	/d/	/d/
script	<i>g</i>	<i>g</i>	<i>t</i>	<i>t, d</i>
Archaic Khotanese ( <i>Śūraṅgamasamādhisūtra</i> )				
phonetic	[ʔ]	[g]	[d]	[d]
phonemic	/g/	/g/	/d/	/d/
script	<i>g</i>	<i>g</i>	<i>t</i>	<i>t, d</i>
Canonical Khotanese (Suv[Or], Z)				
phonetic	[ʔ]	[g]	[ʔ]	[d]
phonemic	/d/	/g/	/d/	/d/
script	<i>t, g</i>	<i>gg</i>	<i>t</i>	<i>t, d</i>

This phonological and script evolution explains why within Old Khotanese an earlier *g* will sometimes correspond with a later *gg* and sometimes with a later *t*.

The spelling alternation between *g'* and *t'* is related to this evolution. The apostrophe is one of the most interesting features of Khotanese writing and merits a separate study.<sup>56</sup> For the moment, it is necessary to discuss some aspects of it.

---

<sup>56</sup> Emmerick 1992b devoted twelve pages to a discussion of the apostrophe or subscript hook. His final opinion appears to be that it denoted a “breathed syllable” (2009:381).

### 2.10.5 The apostrophe and *g'~t'*

Intervocalic *ṣ* /*z*/ does not occur in writing by the time of the earliest known texts. It stopped appearing phonetically on the surface probably within Earliest Written Khotanese.

The loss of intervocalic *ṣ* /*z*/ gave rise to sequences of vowels within words.<sup>57</sup> There was no convenient way to mark this in the Brāhmī so a new sign, the apostrophe, was introduced. In Old Khotanese this almost always indicates a sequence of two vowels. Also known as the subscript hook, the apostrophe appears under an akṣara. Its usage seems complex because the sign interacts with other elements in the orthography (see below). And sometimes it appears under the placeholder or “peg” radical as in *pāga'*, sometimes under the preceding akṣara as in *pā'ga* and sometimes under both as in *pā'ga'* NAPm ‘power’ (Śgs:125a). The oldest use may be that in conjunction with *ṣ*. That is, a word which was originally, in Earliest Written Khotanese, written with intervocalic *ṣ*, kept the *ṣ* spelling after the /*z*/ was lost<sup>58</sup> and the hook was written to show that the *ṣ* had no consonantal value. Some examples of *ṣ'* for older /*z*/ are:

---

<sup>57</sup> It is possible that the phonetic change was [V*z*V > V*h*V > VV] and that the apostrophe was introduced to write [i]. This might tantalizingly link the apostrophe with Tibetan script ‘*achuñ*’ in form and function.

<sup>58</sup> Initially in OKh the *ṣ* /*z*/ disappears during the synchronic derivation. That is, at that time the segment still existed in the underlying phonemic forms. At a later stage in the language it is possible that the segment was lost even from the underlying forms. Then, in those cases where the *ṣ* /*z*/ existed elsewhere in the paradigm for a stem, a synchronic relationship could be maintained. Where there was no paradigmatic support, the loss of the segment was complete. At that point, the existence of spellings with apostrophe would be due to knowledge of those spellings in older texts.

N<sub>Sm</sub> *nājsā'ṣā'kā* Śgs 2.7v3 'revealer' < *nājsaṣ-* 'show' (SGS:53) • /nēd<sup>z</sup>ēākē <

nēd<sup>z</sup>ēzākē < nēd<sup>z</sup>ēz- + -āk- + -ē/

I<sub>ASf</sub> *nāṣa'skye jsa* 3.5v1-2 'end' Śgs < *nāṣa's-* 'end' (SGS:57) • /nēask<sup>y</sup>ē <

nēzask<sup>y</sup>ē < nēzas- + -k<sup>y</sup>- + <sup>i</sup>ē/

3Ppf.tr. *nāṣā'stāndā* Suv[Or] 12.16 < \**nāṣā'y-* 'establish; appoint' (cf. SGS:57) •

/nēāstāndē < nēzāstāndē < nēzāst- + -ānd- + -ē/

LP *hāruṣo'* Rk 95.5 'thing' • /hēruō < hēruzō < hēr- + -uzō/

A full description of the apostrophe is reserved for elsewhere. The following forms show the range of orthographic devices used with the apostrophe to show a sequence of two vowels: *hve'* /huē/ N<sub>Sm</sub> 'man' Śgs 2.6r2, etc.; *bva'ñi* /buañi/ NAP 'perfumes' Śgs 2.4r3; *pyū'* /pyūu/ 2S imperative m. < *pyūṣ-* 'hear' Śgs 2.2r5; *pyūve'mate jsa* /pyūēmadē d<sup>z</sup>a/ I<sub>ASf</sub> 'hearing' Śgs 3.3v3; for /pyūāre/ 3Pp.m. < *pyūṣ-* 'hear' cf. the Z spellings *pyūvā're*, *pyu'vāre*, *pyu'vā're*, *pyuā're*, *pvyā're* (for \**pyvā're?*; SGS:87); *tcei'maṇi* /t<sup>s</sup>aēmaṇi/ (with the *ei* /aē/ symbol for /aē/) NAPm 'eye' Śgs 2.6r2; *guvo'* /guuō/ L<sub>Sm</sub> 'ear' Śgs 3.2v2; *gāṣo'* /gēō/ A<sub>Sf</sub>, *ggāṣā'nu* /gēānu/ GDPf 'group' DKS:84a; *kei'mā* /kēime/ 1Sp.a. < *kāṣ-* 'think' Z 2.215; *kā'mate* /kēāmatē/ NAPf 'thought' \* < *kāṣāmate* < *kāṣ-* 'think' Z 2.154; *uī'* /uē < uṣē < uṣ- + <sup>i</sup>ē/ NAP 'wits' Z 2.140; *kṣei'* /ṭ<sup>s</sup>ēa/ (with *ei* /aē/ symbol for /ēa/) 'six' Śgs 3.11r2.

In the unattested Earliest Written Khotanese, ṣ' was used to indicate a sequence of two vowels. In manuscript Or. of the *Suvarṇabhāsottamasūtra* this older practice is reflected in spellings of the stems *pāṣa-* 'power', *pāṣajsa-* 'powerful', *pāṣagaria-* 'powerful', and *pāṣajsattara-* 'stronger'. In seven cases ṣ' is used: N<sub>Sm</sub> *pāṣā'* 12.45, NAP *p[ā]ṣa'* 10.24, *pāṣa[']* 14.36, GDP *pāṣā'nu* 10.42, IAP *pāṣa'ñyau'* 10.31,

*p[ä]ṣy[au]'* 10.25, NSf *päṣa'jsättara* 10.38. In four cases *t'* is used: GDP *pätā'nu* 6.3.41, NSm *päta'jsättarä* 6.2.80, IAPm *päta'jsyau* 14.34 NAPm *päta'garya* 10.27. In Z which has later spellings than Suv[Or], we find *t'* in stems related to *päṣa-* more than thirty times (Glossar:462b), while *ṣ'* occurs once in NAPf *päṣa'ñi* Z 13.35. From the evidence of just those two texts, we might assume the spelling evolution *päṣa-* > *päṣa'-* > *päta'-*. However, Śgs has older morphology and orthography than either Suv[Or] or Z. In this manuscript the *ṣ'* device is known, as in NSm *näjsä'ṣā'kā* 2.7v3 'revealer', but it is not used in *päṣa-*. Instead, as mentioned above, there is *g* with apostrophe nine times, NAPm *päga'* 3×, *pä'ga* 2×, *pä'ga'* 1×, IAPm *pä'gañyau jsa* 1×, NAPm *pä'gajsa* 1×, GDPm *pä'gajsānu* 1× (Śgs:125a), and *t* with apostrophe once in *pä'tañyau jsa* IOL Khot 188/3v2. Thus, the evolution in spelling for this word seems to be *päṣa-* > *päṣa'-* > *päga'-* or *päta'-* > *päta'-*. The explanation for the use of *g* as a peg akṣara has to do with the development of /g/ > [ʔ] between vowels. After that happened, the devices for writing /VV/ and /VʔV/ became conflated so that any symbol for glottal stop could be used with the apostrophe to show a sequence of two vowels.

## 2.10.6 *g* allography complications

Several factors complicate the description or observation of the phonological and orthographic patterns described above.

### 2.10.6.1 Archaizing spellings

Older spellings may sometimes appear in later documents. That is, the author or copyist will use an older spelling which reflects an earlier stage of the language

rather than a spelling which reflects the phonological structure of the author's or copyist's language.

Skjærvø describes the archaizing phenomenon in texts he defines as “Middle Khotanese”, noting “hybrid forms, such as *himyetä* for *himye* + OKh *hämätä* or *kṣa'ta'* for *kṣa'* + OKh. *kṣäta*, *kṣäta'* (SuvI:lxxiii). He regards Z as written in “Archaizing MKhot.” (cf. IOL Khot 155/5). In Z we find ‘sky’ with *-t-*, i.e., *ätāśa-* 41 times (Leumann 1936:395; Emmerick 1968c:26) and with older *-g-* once in *ägāśo* 22.296 and perhaps a second time in NSm *ägāśä* IOL Khot 7/7v6 which seems to be from Z (Skjærvø 2002:174). Otherwise we find *-g-* almost exclusively in archaic documents: NSm *ägāśä* Śgs 3.5v4, *ägāśalstu* ‘skyward’ Śgs 3.9v1, AS *ägāśu* Suv[A] 2.57, GDSm [*ā*] *gāśi* Suv[C] 6.6.8, *ägāś[-]* Sgh[2] 253[84]. A case could be made that the author's language had *ätāśa-* but he inserted *ägāśa-* to archaize the composition.

#### 2.10.6.2 Later forms in earlier texts

Later copyists sometimes inserted later forms into older texts. Chapter 2 of Z in particular seems to have experienced this. Later Khotanese influence is shown by the use of an unnecessary extra anusvāra *ṃ* between *a* and a nasal, at least 11 times, e.g. *hamamṅgu* 2.5, *naramṅda* 2.104 (further examples in lines 103, 133, 134, 169, 183, 189, 197, 218, 236). As both the Archaic Old Khotanese orthography and the Late Khotanese literary tradition used single *g* not *gg* for /g/, we sometimes cannot tell if a single *g* is an archaic fossil or Late Khotanese copyist influence. There are several cases of *g=/g/* in Chapter 2 as in *gurṣte* 3Spf.tr.m <*grūs-* ‘call’ 2.81/130 for expected *ggurṣte* (cf. 20.50, etc.) and *guvašte* 3Sp. ‘it splits’ 2.179 for expected *gv-* or *gguv-*. There are three cases of *bilsaṅga-* or *bilsamṅga-* ‘community of monks’ with



*V(m)ṅgV* (2.123, 169, 183) as opposed to just one instance of a *ṅ* spelling outside of chapter 2, GDSm *bālsaṅgi* 22.291. The more usual *bīlsaṅga-* or *bālsaṅga-* with *VṅgV* occurs sixteen times in Z.

There are a couple of peculiar uses of *g* as a hiatus consonant in Chapter 2 which may show LKh influence. Instead of the usual Z nativized<sup>59</sup> form *vīvāta-* for /wīwā?/ ‘vipāka, fruition’, there is the unique *vātāga* NSm 2.144. Although *g* for [ʔ] is an old feature, it is also found in Late Khotanese. But the use of *-t-* for *-v-* where it is not [ʔ], and the NSm *-a* for *-ä* or *-i* are purely Late Khotanese features. Also, the meter is broken here, with the eight mora *ttatvata vātāga* overflowing a regularly seven mora segment. The same word occurs once in archaic form with *g* in Chapter 2: *vivāgä* NSm Z 24.650.

The second peculiar hiatus *-g-* in Chapter 2 is in *pahaiga* 3Ppf.intr.m <*pahīs-* ‘flee’. The normal OKh is *pahiya*, Z 20.33, 24.421, while LKh has *pahaiya* (SGS:80). The vowel spelling *-ai-* shows late influence so presumably the *-g-* is as well. In this instance the *g* would appear to stand for [y] and not for [ʔ].

There are a few cases of later forms in *Śgs*. It is normal for *g*, *ś*, and *ṣ* to be written single, but there are three cases of *śś* in NAPm *harbiśśä* ‘all’ 2.3r3, GDP *harbiśśānu*

---

<sup>59</sup> The oldest form of this word in Khotanese is probably *vīvāga-*, borrowed from Prakrit (cf. Gdh *vivaka*, *vivaveṇa*, *vivaḷe*, *vivao*; Baums and Glass s.v.). In Z, this spelling occurs 3× 4.20, 8.33, 24.433. *vivāga-* with short *i* occurs 1× 24.650 but meter suggests an original long *ī*. In Z, the favored spelling is *vīvāta-* 11x. The nine *vivāta-* and two *vāvāta-* spellings also could have originally had long *ī*. Most occur where the *uysnora*-effect makes the moraic value of the syllable ambiguous. In Z 4.23 *vivātā* fills a five mora segment suggesting original *vīvātā*. There are also spellings with *-pāk-* showing Sanskrit influence: *vīpāka-* 11×, *vīpāka-* 7×, *vāpāka-* 1×. Even though in Classical Sanskrit the word is *vīpāka-* with short *i*, a case could be made that the spellings in Z with *i* and *ä* disguise an original *ī*.

IOL Khot 188/3r4, and [kī]śśana 2.7v1 NAPm ‘abounding’ (cf. Śgs:xix). This may represent influence of a copyist.<sup>60</sup> The exceptions do not necessarily show that the composer was familiar with the convention of doubling *gg*, *śś*, *ṣṣ* as Emmerick once suggested (Śgs:xxi). The use of *tt* is a different matter (see above, 10.1).

Another possibly later feature in Śgs is the use of *t'* for *ś'* in *nāta'skya* NSf 3.2r5 beside older *nāṣa'skye jsa* IASf 3.5v1/2, and in *gā'to* LSf ‘group’ 3.6v1 for expected \**gāṣo'* (cf. ASf *gāṣo'* Suv[Or] 6.3.21).

### 2.10.6.3 Sanskrit spellings

Intervocalic single *g* can appear in words borrowed in written form from Sanskrit. In Śgs, *nāga* NAPm ‘serpent’ which occurs twice (1.1v4, 3.4v5) is the only example. We cannot be sure if the word was pronounced [nāʔa] as if native or [nāga] as if the foreign origin was recognized. The Khotanese who followed this period were also not sure how to treat this kind of word. In Z we have *nāga* NAPm 3.145 showing the Sanskrit spelling, and *nāta* NAPm 22.246 clearly showing nativized [nāʔa]. Possibly *nāggū* 23.159, if correctly identified as NAPm by Leumann and Emmerick, unambiguously shows the [g] of Sanskrit. Similarly, in proper names, there is *Nāgasenu* ASm 22.94, *Nātapuṣpī* GDSm (-i + -i ‘his’) 22.173, and *Nāggārjuni* NSm 11.32.

---

<sup>60</sup> I have not explored the suggestions made by earlier scholars (cf. E. and M. Leumann 1933–36:xxxii) and based on metrical evidence that the *śś* in *biśśa-*, *harbiśśa-* ‘all’ sometimes may show a consonant cluster in Z. If correct then the *śś* in these words may originally have been for phonological /šš/ and was not orthographic for single voiceless /š/.

## Chapter 3: Old Khotanese Vowels

### 3.1 Vowel Symbols

Without clear indication to the contrary, it makes sense to tentatively assign the Sanskrit or Prakrit values to the vowels written in OKh. In Sanskrit, *i*, *u*, *a* are short vowels and *ī*, *ū*, *ā* are long. The letters *e*, *o*, *ai*, *au* are technically diphthongs: *e* = *a+i*, *o* = *a+u*, *ai* = *ā+i*, *au* = *ā+u*. In Prakrit *e* and *o* have monophthongized, presumably to long /ē/ and /ō/.

#### 3.1.1 The Monophthongs /i~ī, u~ū, a~ā/

There is general agreement that there were three short cardinal vowels /i, u, a/ contrasting with three long /ī, ū, ā/. These values are consistent with those of Sanskrit. Emmerick offers minimal pairs to prove phonemicity of the length contrast: *bisa* 'house' ~ *bīsa* 'servants'; *mura* 'birds' ~ *mūra* 'coin'; *bata-* 'small' ~ *bāta-* 'wind' (1979:239).

#### 3.1.2 The Diphthongs *e*, *o*, *ai*, *au*

The Sanskrit short diphthongs *e*, *o* had probably become monophthongs early in Prakrit. No one has assumed a diphthongal value in Khotanese. Here it is argued that these consistently represented long mid vowel phonemes in Earliest Written Khotanese. Emmerick and Maggi propose that the symbols were ambiguous for long and short mid vowels as discussed below.

The Sanskrit long diphthongs *ai*, *au* probably remained diphthongs in Prakrit. Here they are assigned diphthongal value in EWKh, with /aṽ>ō/ monophthongization in Canonical Old Khotanese. Emmerick 1979 seems to have

proposed that the monophthong and diphthong signs are equivalent, i.e.,  $e=ai$ ,  $o=au$ . But in 2009 he wrote, “the diphthongs  $ai$ ,  $au$ , and  $aə$  of the older language were monophthongised [in the later language]” (382).

### 3.1.3 The New Signs $\ddot{a}$ and $ei$

The new diacritic  $\ddot{a}$  is assigned essentially the same phonetic value here and by Emmerick 1979. But phonemically there is an important difference. Here  $\ddot{a}$  is / $\text{ě}/$ , the short partner of  $e/\text{ē}/$ , while for Emmerick and Maggi 1991, if I have understood correctly,  $e$  represents two low or low-mid vowels, whose height is lower than that of  $\ddot{a}$ . Emmerick 2009 treats  $\ddot{a}$  as / $\text{ə}/$  and  $e$  as denoting long and short / $\text{ē}$ ,  $e/$  (382).

E. Leumann in 1912 assigned the transcription  $ei$  to the cross diacritic without clear reasoning. According to Emmerick it was because, “the cross sometimes appeared to alternate with  $ai$  and in the later language could be replaced by  $\bar{e}$  (1998:94). Emmerick 1979 did not assign a value to  $ei$  but noted that it was distinct from  $ai$  in the Śgs. In a separate article on this diphthong in 1998, “Khotanese  $El$ ”, Emmerick noted that in the Śgs  $ei$  represented a contraction of the two vowels written  $a$  and  $\ddot{a}$  in the NSm of the  $aa$ -declension. He concluded, “It would accordingly be simplest to transcribe it as  $a\ddot{a}$  were it not for the fact that the transcription  $ei$  by now is a hoary tradition so that it is perhaps better to continue to write  $ei$  but regard it merely as a graphic convention for  $a\ddot{a}$  or more precisely  $aə$ ” (94). Here it is argued that  $ei$  was a low to mid diphthong / $\text{aę}/$  in EWKh where it merited being marked with a new sign, and in AOKh where it is consistently written. It merged with  $ai$  in the Canonical period, hence the alternation then with  $ai$ .

## 3.2 Vowel Linguistics

### 3.2.1 EWKh Monophthongs

EWKh probably had five short vowels and four long vowels. It is useful to compare the spellings, the phonemic status and the phonetic nature of these segments.

**Table 3.1:** EWKh Monophthongs

orthographic	phonemic	phonetic
$\bar{i}$ $i$ $u$ $\bar{u}$	$/\bar{i}/$ $/i/$ $/u/$ $/\bar{u}/$	$[\bar{i}]$ $[i]$ $[u]$ $[\bar{u}]$
$e$ $\ddot{a}$ — $o$	$/\bar{e}/$ $/\check{e}/$ — $/\bar{o}/$	$[\varepsilon]$ — $[\bar{o}, \check{o}]$
$a$	$/a/$	$[\ddot{a}, \check{a}]$ $[a]$
$\bar{a}$	$/\bar{a}/$	$[\bar{a}]$

Each of the orthographic monophthongs denotes a phonemic monophthong. There are five long monophthongal symbols and phonemes, and four short monophthongal symbols and phonemes. There is a gap in the phonemic inventory for a short correspondent to long  $/\bar{o}/$ . While there are just four short vowels phonemes, phonetically there are six short vowel phones because  $/\bar{e}/$  and  $/\bar{o}/$  have short allophones in final position.

The symbols  $i$ ,  $\bar{i}$ ,  $u$ ,  $\bar{u}$ ,  $a$ ,  $\bar{a}$  denote long and short vowels, with values similar to Sanskrit and have been properly understood since the first studies in Khotanese. The phonemic and phonetic symbols are the same as the orthographic ones in these pages for simplicity. In contrast, the mid vowels, written  $e$ ,  $\ddot{a}$  and  $o$  exhibit complex phonology and have not been well understood. The symbology for these is also complex. The graph  $\ddot{a}$ , two dots above a vowel, is not known in Indian tradition but

is found in Tocharian and Tumshuqese from the northern rim of the Tarim Basin. It has often been given the value schwa [ə]. It is here regarded as phonemic short /ě/, corresponding to phonemic long /ē/. To show it is likely a phonetic mid lax vowel it is given as epsilon [ɛ]. The symbols *e* and *o* in non-final position have the expected long mid phonemic values from Prakrit. As phonemes, the vowels are always underlyingly long so are phonemically transcribed /ē, ō/ with macrons for clarity. Word finally, these symbols mostly, but not always, reflect phonetically short vowels, hence there are long and short allophones in the chart. The symbology for long and short back mid [ō, ǒ] is straightforward. That for long and short “mid” [ǣ, ǣ̃] is not. An *e* /ē/ in final position is phonetically short, but contrasts there with *ä* /ě/, so here *e* is regarded as phonetically lower than *ä*. In final position, *e* /ē/ [ǣ̃] contrasts with *ä* /ě/ [ɛ].

### 3.2.2 EWKh Diphthongs *ai*, *au*, *ei*

There were three low-high or low-mid diphthongs in EWKh, written *ai*, *au*, and *ei*. The first two symbols, *ai* and *au*, are familiar from Indian practice and have the familiar low-high values /aᵢ, aᵤ/. The third symbol, a superscript cross, possibly derived from the *e*-diacritic with a stroke through at right angles, or perhaps a geometric contrivance, has the unfortunate transliteration *ei*. As shown below, the value of *ei* is low-mid /aᵉ/. Presumably a new sign was introduced because /aᵉ/ was relatively common in EWKh and a new sign helped readability. In Canonical Khotanese it seems *ai* and *ei* have begun to merge.

Ladefoged and Maddieson report on research by Lindau, et al. 1985, that diphthongs occur in about a third of the world’s languages and of these, 75 percent

have “*ai*-type” and 65 percent have “*au*-type” diphthongs (1996:321). I find no statistics relating to an *ae*-type like that proposed for Khotanese, but at least one language, Estonian, has all three:

*ai*-type: *lai* ‘wide’      *au*-type: *laud* ‘table’      *ae*-type: *aed* ‘fence, garden’

### 3.2.3 EWKh On-Glide Diphthongs

The Khotanese *ai* /aṯ/, *au* /aṯ/ and *ei* /aṯ/ may be described as off-glide diphthongs as they end in a glide. EWKh also has several on-glide diphthongs which begin with a high front /i/ or high back /u/ on-glide. These mostly, if not exclusively, arise through synchronic vowel contraction. They could perhaps theoretically be regarded as underlying sequences of two monophthongs, but they are written more like diphthongs. In the list below, indirectly attested diphthongs are in curly brackets {}:

**Table 3.2:** Old Khotanese onglide diphthongs

{ <i>yu</i> /iṯu/}	<i>yo</i> /io/	<i>ya</i> /iṯa/	<i>ye</i> /iṯe/	—
—	{ <i>vo</i> /uo/}	<i>va</i> /uṯa/	<i>ve</i> /uṯe/	<i>vī</i> , <i>uī</i> /ui/

These were only identified recently (Hitch 2015b), perhaps partly because the spellings are all ambiguous. They can also reflect consonant+vowel sequences: *yu*=/yu/, *yo*=/yō/, *ya*=/ya/, *ye*=/yē/, *vo*=/wō/, *va*=/wa/, *ve*=/wē/, *vī*=/wī/. These diphthongs have been extensively discussed in Hitch 2015b and in in-press-a. In meter, these diphthongs, like all contracted vowels, are always two moras long. For instance, if the spellings *ya*, constitutes a single mora, it represents /ya/. But if it counts as two moras, it stands for contracted /iṯu, iṯa, uṯa/. Another disguising factor is that these diphthongs may be resolved into sequences of *iyV* or *uvV*. The Leumann’s

transcribed these diphthongs (*i*)*yV* and (*u*)*vV* assuming that the *yV* and *vV* spellings were abbreviations for *iyu* and *uva*.

These diphthongs are typologically rather ordinary. Spanish has an essentially identical set:

**Table 3.3:** Spanish onglide diphthongs

/ịu/	viuda	/ịo/	radio	/ịa/	hacia	/ịe/	tierra	—	
	‘widow’		‘radio’		‘towards’		‘earth’		
—		/ụo/	cuota	/ụa/	cuadro	/ụe/	fuego	/ụi/	fuimos
			‘quota’		‘picture’		‘fire’		‘we went’

### 3.2.3.1 /ịu/

*i+u>yu* /*i+u>ịu*/: *ia*-decl ASm (1).

(1) In theory, this contraction may occur with the nominals only in the ASm of the *ia*-declension. There are no attested spellings of the expected *Cyu* type. There are two examples where the diphthong is resolved to *Ciyu* which imply original *Cyu*: ASm *briyu* ‘dear’ Z 20.22, ASm *indriyu* ‘sense’ Suv[Or] 5.9. The more usual ASm ending with the *ia*-decl in COKh is *Cī*, presumably transferred from the NS and GDS, e.g., ASm *jadī* ‘folly’ Z 5.58, *padī* ‘way, manner’, *dātī* ‘lawful’ Z 6.4, *hivī* ‘own’ Z 5.7 (SGS:311, 314–315).

### 3.2.3.2 /ịo/

*i+o>yo* /*i+ō>ịō*/ including *i+<sup>i</sup>o>yo* /*i+<sup>i</sup>ō>ịō*/: *ia*-decl LSm (1), *iā*-decl ASf (2), LSf (3)

(1) LSm *naryo* < *nari-* ‘hell’ is the only example. It occurs seven times in Z but there is no non-ambiguous metrical occurrence. Below, the metrical structure /*na·ṛịo*/ is presumed and consequently that *naryo* is HL in cadence 1 HLLHL:



17.33      **naryo** bihiyu □ | atā stora | māsta dukha (C:7+5+5)

In hell are extremely fierce, great woes

(2) ASf *ggaṃḍyo* < *ggaṃḍiā*- ‘gong’, -*ḍyo* is X in cadence 2 HXHL:

2.101cd    ttai hvāña | badṛ ma pva’ttu □ | āvulātu | **ggaṃḍyo** bāḍā (A:5+7+5+7)

Speak thus to him: “Let Bhadra fear not. Let him strike the gong. (It is) time.”

An example not found in SGS is ASf *trisaḥasryo* SI M13.10v2 < *trisaḥasria*- ‘triple-thousand’.

(3) LSf *paḍāmjsyo* < *paḍāmjsiā*- (adj.) ‘former’ fills a five-mora segment suggesting that -*jsyo* is H in LHH:

6.7bd      paḍām□**jsyo** | tcalco aysātā | hālysdā □ vaṣ|ṭāmata niśtā .

(A:5+7+5+7)

There is no present arrival for one unborn at the previous end.

### 3.2.3.3      /ḷa/

*i+a>ya* /i+a>ḷa/ including *i+<sup>i</sup>a>ya*: *ia*-decl VSm (1), NAPm (2), LSm (3), *iā*-decl NSf

(4). With short *a* umlaut and diphthongization, *a+<sup>i</sup>a>ya* /a+<sup>i</sup>a>ḷa/ : *aā*-decl LSf (5),

*aa*-decl LSm (6). With Short *u+<sup>i</sup>a* diphthongization and onset simplification (only

*aysmya* attested), *u+<sup>i</sup>a>\*vya>ya* /u+<sup>i</sup>a>ḷa/ : *ua*-decl LSm (7).

(1) VSm *brya* < *bria*- ‘dear’ (adj.) is X in cadence 2 HXHL:

2.177ab    tterā kṣamo|vī **brya** balysa □ | samu kho śśan|dā ne ju oysa .

(A:5+7+5+7)

So forgiving are you, beloved Buddha: like the earth, you do not get angry.

(2) NAPm *indrya* < *indria*- ‘sense’, with -*drya* X in cadence 2 HXHL:

1.85ab kho ni ṣṣadda | **indr**ya hota □ | kho nā kuśśala-|mūla paysānde .  
(A:5+7+5+7)

As is their faith, senses, power, as he recognized their merit-roots ...

(3) LSm *narya* < *nari-* 'hell'. This is the only example. It occurs 11 times in Z but never in a cadence. It counts three moras. The structure /na·r̥ja/ is presumed:

24.452ab pharu **narya** | dāruṇa dukha | bīḍā jsañaulysā (B:5+6+7)

Many severe woes will the causer of death bear in hell.

(4) NSf *lovyā* < *loviā-* 'worldly' (adj.), -*vyā* is X in cadence 2 HXHL:

13.125cd tcohore-haṣṭātā ysāre ka□lpa | jsīna biśśa ṣa | **lovyā** jsīna . (A:12+5+7  
if *ysāre* an insertion.)

(one has) life for eighty-four thousand kalpas. All this is *lokika* life.

(5) LSf -*tyā* in *brītyā* 'love' is X in cadence 2 HXHL:

2.169ab balysā kāḍe | **brītyā** spāṣṭe □ | palimjvī | pāttro vistāte . (A:5+7+5+7)

The Buddha, greatly in love, looked (at him). He placed a bowl for him in *panyānka-* position,

(6) LSm *drau-mūjsya* < *drau-mūjsaa-* 'hair pore' is six moras in the hemistich below where -*jsya* is X in cadence 2 HXHL:

6.4ab paniña drau-|**mūjsya** balysa .□ | avamāta | āysāta balysa (A:5+7+5+7)

In every hair-pore of the Buddha, unlimited, adorned Buddhas,

(7) LSm *aysmyā* < *aysmua-* 'mind'. In the hemistich below -*myā* is X in cadence 2 HXHL:

2.9ab ttārthānu | mästā arātā □ | kāḍe nu dukhā | **aysmyā** saittā (A:5+7+5+7)

There was very great envy on the part of the heretics. It seemed in their mind very much a misfortune ...

### 3.2.3.4 /ie/

*i+e>ye* /i+ē>je/ including *i+<sup>i</sup>e>ye* /i+<sup>i</sup>ē>je/: *ia*-decl NAPm (1), *iā*-decl NAPf (2), GDSf (3), IASf (4), *ii*-decl GDSf (5). Also, with short *a* umlaut and diphthongization, *a+<sup>i</sup>e>ye* /a+<sup>i</sup>ē>iē>je/: *aā*-decl GDSf (6), IASf (7).

(1) NAPm *indrye* < *indria*- ‘senses’, IOL Khot 163/11v2, not in meter.

(2) NAPf *lovyē* < *loviā*- ‘worldly’, -*vyē* is X in cadence 2 HXHL:

23.111     vari ṣṭānā | lovyē ttīyā □ | biśśe paṃṣa | rruste abhijñe (A:5+7+5+7)

At once he lost all the five *abhijñās* of this world.

(3) GDSf *hävye* < *haviā*- ‘own’ (adj.) is three moras in the hemistich below

14.91cd    cu ne rro vā | balysä hävyē | irdi u hota (B:5+6+7)

How much greater are the *ṛddhis* and power belonging to the Buddha!

(4) IASf *ggaṃḍye jsa* < *ggaṃḍiā*- ‘gong’, -*ḍye* is H in cadence 1 HLLHL:

2.123ab    cītā ttāte | ggāha badṛ pyū □ ṣṭe | cu ggaṃḍye jsa naranda

(A1:5+9+3+7)<sup>61</sup>

When Bhadra heard these verses, which came out of the gong,

(5) GDSf *saḷye* is LH in cadence 1<sup>A</sup> HLLHHL<sup>62</sup>, indicating *sa·ḷye* /sa·līē/:

22.125ab   hatārā ke|rīndi saḷye ttīma □ | daso-gyūṃnau rrvītā . (A1:5+9+3+7)

---

<sup>61</sup> Meter A1 5+9+3+5 is a subtype of meter A 5+7+5+7. It was not recognized in Hitch 2014. It is not a new discovery. Already in 1912 Ernst Leumann listed a subtype of meter A with structure 5+9+3+7 (E. Leumann 1912:18). In 1919 he noted that alongside regular type C meter, 7+5+5, there was a C Nebenform of 9+3+5 (E. Leumann 1919:28). In 1967 Manu Leumann defined a subtype of meter A, Nebenform A1, 5+9+3+7 (1967:366). Partly following Manu Leumann I call the subtype A1, and the corresponding C subtype C1.

<sup>62</sup> This may be the most common cadence type for the nine mora segments of meter types A1 and C1 (see preceding footnote). Other nine mora cadences exist, such as HLHLHL, but more work needs to be done in this area.

They will sow seed once a year. It will grow ten-fold.

(6) GDSf *-jsye* in *buljsye* ‘virtue’ is X in cadence 2 HXHL:

24.477 hāvā vātā | basta pa□jsamā | **buljsye** vīrā (B:5+6+7)

They are bound to profit, reverence, to virtue.

(7) IASf *-nye jsa* in *uysānye jsa* ‘self’ is HL in cadence 1 HLLHL:

24.265ab ttuśśā biśśā | dharma □ uysā|**nye jsa** paysānde . (B:5+6+7)

He recognized all the *dharmas* as empty of self.

### 3.2.3.5 /u̯o/

*u+o>vo* /u+ō>u̯o/: *uā*-decl ASf (1)

(1) \*-*Cvo* is not attested but is implied by the resolved *-Cuvo* in ASf *paṃja-satā-saluvo* ‘500-year-old’ < *-saluā*-.

### 3.2.3.6 /u̯a/

*u+a>va* /u+a>u̯a/: *ua*-decl NAPm (1), *uā*-decl NSf (2), *attāti*-nouns formed from *ua*-adjectives (3)

(1) NAPm *bāysva* < *bāysua*- ‘arm’ Suv[H] 4.11. No examples in meter.

(2) NSf *haṃdara-ysaṃthva* Sgh[17] 42.1. No examples in meter.

(3) *tcarsvattātā* < *tcarsua*- ‘bright’ + *-attāti*-, on its own constitutes a cadence 2

HXHL with *ṣva* filling the X position:

23.40cd cvī ttīśā | **tcarsvattātā** □ | tto ye ce yuḍu | tīndi paḍandu (5+7+5+7)

who can make his brilliance, splendour?

### 3.2.3.7 /u̯e/

*u+e>ve* /u+ē>u̯e/: *ua*-decl NAPm (1)

(1) NAPm *bāysve* < *bāysua*- ‘arm’ IOL Khot 166/2v4. This occurs in a fragment

identified by M. Leumann as from Z and is in meter A or B (1967:374). *-ysve* is X in cadence 2 HXHL at the end of a line:

pāda 87 ] ye dasau | bāysve dyai (A or B, ending in a 7-mora segment)

... you saw the ten arms (transcription and translation Skjærvø)

### 3.2.3.8 /uī/

*u+i>uī (vī) /u+i>uī/*: *ua*-decl GDSm (1)

(1) GDSm *aysmuī* < *aysmua*- 'mind' occurs 11x in Z with this spelling against once with the spelling *aysmvī* (Z 19.31). In all cases it is four moras. The syllable *-muī* is X in cadence 2 HXHL:

22.274cd ttarrna dukha | kṣūna barīndi □ | samu hävī | aysmuī ārru (A:

5+7+5+7)

they bear woes with thirst, with hunger. It is merely the fault of one's own mind.

### 3.2.3.9 Onglide Diphthongs with present stems

In principle diphthongs will arise through contraction where the conditions permit. In Hitch in-press-a I noticed two instances with present stems ending in a vowel. The first is certain from the meter of Z, while the second could be challenged.

*u+a>va /uā/*. 2Siv.a. *ākṣva* < *ākṣu*- 'begin' + *-a*, is HX in cadence 2 HXHL:

2.100cd cvä tä tta ratä | hälysdä pamā□tä | ma ne dīñi | **ākṣva** būṣṣa .

(A:5+7+5+7)

Thus if from this you now have pleasure, would that I not appear limited. Begin, distribute!<sup>63</sup>

*i+a>ya* /j̥a/. 2Siv.a<sup>64</sup> *parya* < *pari-* ‘order’ + *-a* is three moras, likely LH, and syllabified *pa·rya*. in a 5-mora segment:

23.52cd mamā **parya** | aysvī tanī□mä | nai handarä | tcāraṇä trāmu

(A:5+7+6+7)

Order me. I will make it. No one else is so capable of it.

### 3.2.4 Triphthongs

Old Khotanese had two triphthongs, /j̥au/ and the mirror image /uaj/. Both feature an onglide, a core [a], and an offglide. The diphthong /j̥au/ was described in Hitch in-press-b as the underlying form of the IAP suffix *-yau*. This underlying form and a series of ordered rules can explain the complex metrical and morphophonological behaviors of this morpheme. It is possible that /j̥au/ is created when the 2P enclitic pronoun *-u* suffixes to final *-e* in *ne* ‘not’ + *-u* > *nyau* Z 23.95, 24.474 (see 3.6.3).

The triphthong /uaj/ is found as the initial syllable in *uaira-* ‘suitable’ and *uaiṣṣa* ‘awake’. If the initial syllable were /waj/, one would expect the words to be spelled

---

<sup>63</sup> This translation is based on the identification of *dīñi* as 1Sop.a from *dä-* ‘appear’ (Hitch in-press-b, footnote ±33 on *dīñi*). Emmerick’s translation: “What pleasure is thus at hand for you I would not see limited. Begin, distribute!”

<sup>64</sup> Emmerick identified *parya* as 2Piv.a (SGS:73), but that would be *\*parīta~\*parīya* < *pari-* + *-āta* with *i+ā>ī* contraction, formally identical to 2Pp.a *parīya* Z 12.28.

with *\*vai-*. There is a spelling *vuaira-* (DKS:36a) but this looks suspiciously like copyist influence. *uaiṣṣa* is restricted to Z and is so spelled four times. The triphthong /*u̥ai̥*/ is also found as the result of the umlaut of *au*. There is /*au̥*/ to /*u̥ai̥*/ umlaut in the *i*-infinitive *kṣamvaittā* < *kṣamotta-* (< *\*kṣamautta-*) ‘ask forgiveness of’ and in LSf *avachvaido* < *avachauda-* ‘unobstructed’ + *-i* (identified by Maggi, Rk 23(2)). In Hitch 2015b:314 I proposed that the *au*-stem GDSm ending *-uai* shows a kind of metathesis, *au+i>uai* /*au̥+i>u̥ai>u̥ai̥*/, resulting in the triphthong /*u̥ai̥*/. We know that the result is purely vocalic and not *\*/wai̥/* from metrical evidence. For example, GDSm *saruai~sarvai* ‘lion’ is six of seven times clearly three moras thus syllabifying *sa·ruai* /*sa·u̥ai̥*/. Finally, the result of the enclisis of the 3S enclitic pronoun to final *-o* also seems to result in a triphthong. *puṣṣvai* Z 15.11 is from *puṣṣo* + *-i* ‘his ... completely’. It counts three moras in meter so syllabifies *pu·ṣṣvai* /*buṣṣu̥ai̥*/ (Hitch 2015b:314).

### 3.2.5 EWKh and AOKh overlong vowels

As known since the early work of E. Leumann and detailed in Hitch 2015b, a final *-e* or *-o* arising from contraction is always metrically two moras. These syllables also tend to be found in the metrical X positions. Final *-e* and *-o* not resulting from contraction in contrast are essentially always short. These ambiguous final *-e* and *-o* spellings play a pivotal role in the Emmerick and Maggi analysis discussed below. Here the synchronic derivation is thought to involve an abstract stage in which the contracted vowels carry information about their origin on the surface.

Coalescence initially, or underlyingly, produces phonetically over-long vowels [*ṝ*] and [*ō̄*]. As these reduce in length to that of plain long vowels, [*ṝ*] and [*ō̄*], original

final non-contracted [ǣ̄] and [ō̄] shift phonetically to short [ǣ̆] and [ō̆].<sup>65</sup> At no point is the phonemic unity of nonfinal /ē/=[ǣ̄] and final /ē/=[ǣ̆], or of nonfinal /ō/=[ō̄] and final /ō/=[ō̆] interrupted. For this reason these allophones are written *e* and *o*. The new, phonetically long final [ǣ̄] and [ō̄] arising from contraction carry the information that there is an underlying contraction. A morphophonemic representation of these contracted, underlyingly over-long vowels can be given as /ē̄/ and /ō̄/ with double macrons.

### 3.2.5.1 /ē̄/

*a+e>e* /*a+ē̄>ē̄*/: *aā*-decl NAPf (1) , *aa*-decl NAPm (2).

(1) NAPf *-me* in *bitame* ‘doubt’ is H in cadence 1 **HLLHL**:

6.23cd     harāta’ bita**me** bāsśā □ klaiśa | häma thu ysama|śśaṃḍya balysä  
(A:5+7+5+7)

Suppress doubts, all kleśas. You will become a Buddha on earth.

(2) NAPm *āste* < *āstaa*- ‘bone’ is four moras as in the hemistich below where *-te* is X in cadence 2 **HXHL**

20.46cd     ā**ste** baysgu | tca□brrīya | śśāre pharu . (C:7+5+5)  
there lie thickly scattered about many bones,

---

<sup>65</sup> This shift with regard to the back vowels is facilitated by the original absence of phonemic short /ō̆/ in the language. The picture with the front vowels is more complex. There was original short /ē̆/ written *ā* contrasting with long /ē̄/ written *e*. It is here assumed that the phonemes were phonetically sufficiently distinct to permit a phonetically short allophone of /ē̄/ to arise. While /ē̆/ was perhaps [ɛ], /ē̄/ was perhaps [ǣ̄] which could shorten to [ǣ̆] in final position and maintain distinctiveness from [ɛ].



### 3.2.4.2 /ō̄/

*a+o>o /a+ō>ō̄/*: *aā*-declension ASf (1)

(1) ASf -*do* in *baśdo* ‘sin’ is X in cadence 2 HXHL:

24.50cd    ne ne hāmāte | dātā □ samu rro | baśdo nāsa (B:5+6+7)

Law will not arise. You too will only perform evil.

## 3.2.6 Vowels in Canonical Old Khotanese

In texts in Canonical script we begin to see a reduction in the number of vocalic contrasts. This may begin with loss of distinction between some unstressed short vowels in final position. It seems likely that many if not all unstressed final short vowels in polysyllabic words are lost by the time of Late Khotanese. But the first loss in distinctiveness was probably that between final *-ā* and *-i* as shown by the frequent interchange between them in Canonical Old Khotanese. Somewhat parallel to that merger, and occurring at perhaps the same time is the merger of *ei* /aɛ/ and *ai* /ai/ into /ai/ which can then be written either *ai* or *ei*. For instance, the ending of the NSm of the *aa*-declension in the Archaic Śgs is consistently *-ei* (and the GDSm is consistently *-ai*; Śgs:xx), while in the Canonical Z the ending may be *-ai* or *-ei* (SGS:297–298). Compare *alysānei* Śgs 2.1r4 with *alysānai* Z 13.47 ‘prince’, *balysūñavūysei* Śgs 2.3r4 with *balysūñavūysai* Z 6.47 ‘bodhi’, and *ysamaśandei* Śgs 3.8r4 with *ysamaśśandai* Z 2.8 ‘earth’.

## 3.2.7 Non final /ō̄/ in *tcohora*

Within COKh, the spellings *o* and *au* become mostly, but not always, interchangeable, indicating a merger, probably with a monophthongization of /aʊ/ to /ō̄/. In non final position, EWKh and AOKh distinguish long /ō̄/=[ō̄] from the

diphthong /au̯/=[au̯]. There is no non final short \*/ǒ/=[ǒ]. The existence of that vowel has been the subject of dispute. Its presence or absence can affect the analysis of the mid vowels. As Emmerick and Maggi pointed out, Leumann noted non final short *o* only in the first syllable in *tcohora*- ‘four’ (1991:68). It is odd that a particular vowel would be attested in only one word. The issue is complicated by the fact that the word is also spelled *tcahorV*, *tcahaurV* and *tcohaurV* in COKh.

The metrical system of Z is once again helpful. We can examine the instances of ‘four’ to study the metrical treatment of the first syllable. If it is light, then in contains a short vowel, if heavy, a long vowel. A rather high percentage of cases of hemistichs containing ‘four’ display defective meter, no matter how the first syllable is counted. Then there are many instances where the *uysnora*-effect<sup>66</sup> neutralizes any observation about the mora count of the first syllable. But there are three places where *tcohorǂ* counts HHL, implying phonemic long /ō/ in the first syllable. In all three, *tcohorǂ* by itself constitutes a five mora segment in meter A 5+7+5+7:

10.28ab    **tcohora** | haṭṭhe paysendä □ | kṣei’mo pā|rāmato vīri .

He recognizes the four truths in the sixth *pāramitā*.

10.29cd    pātco vari | ṣṭāni paysendä □ | **tcohori** | handare haṭṭhe

then at once he recognizes four other truths.

13.16ab    **tcohora** | phārre abhiñi □ | jāna kṛs|nāyana yāva .

The four stages, the *abhijñās*, the *dhyānas* up to the *kṛtsnāyatanas*,

---

<sup>66</sup> A word which outside of a cadence counts HHL will, at the end of a seven mora cadence count as if LHL. The effect is named after the HHL word *uysnora*- ‘being’ which often appears in that position, and is the example first given by Leumann 1912:16 (effect detailed in Hitch 2014:15–17).

This seems rather strong evidence for the structure /t<sup>s</sup>ōhōr/ at the time of the composition of Z and for the absence of \*/ō/. The other spellings with *tca-* then may indicate an historical change in this word. Numerals are notoriously irregular, frequently changing in the context of counting. In this case, we seem to have a change from /t<sup>s</sup>ōhōrV/ to /t<sup>s</sup>ahōrV/, likely stress related. There seems to be evidence in Z supporting this analysis. There are two places where copyist influence might be detected. In the hemistich below there is *ce tsohorä* filing a five mora segment. Presumably the copyist's language now had a short vowel in the initial syllable so he added *ce* to the first segment, hypercorrecting what was now read as four moras, back to five. The addition also appears to have caused the following column boundary □ to shift one akṣara to the left away from the segment and word boundaries.

22.223ab **ce tcohorä** | ātama sīyā□ndi | ttä dyān|yau jsa ramīndä . (A:5+7+5+7)

Those who have studied the four *Āgamas* find pleasure in meditation

Similarly, below, *tcahora* was changed to *tcahorai* restoring the five count. Note that the second instance of the word, *tcahori*, has not been altered even though the first vowel is *a*:

22.160ab **tcahorai** | nyanā sarbīndi □| **tcahori** | hālā māsta . (A:5+7+5/4+7)

Four great treasures will rise up for him in the four directions

A final piece of evidence has to do with the distribution of the spellings among the stages of Old Khotanese. If there was a sound change in the first syllable of /ō/ to /a/, then we might expect a tendency to find *o* in older texts, and *a* in later. This seems to be the case. I find only one instance of the morpheme in AOKh where it

shows *o*: LPM *tcohorvaretcoholśuvo* ‘44’ Śgs 3.6r4. Further in Z the spelling *tco-* is substantially more frequent, 30×, than *tca-* 12× (Glossar). There are very few *tco-* spellings in Suv compared to the *tca-* spellings, and there appear to be no *tco-* spellings in Sgh. These distributions suggest that the composer of Z said /t<sup>s</sup>ō-/ while the translator of Sgh said /t<sup>s</sup>a-/.

### 3.2.7 Emmerick, and Emmerick and Maggi Vowels

The vowel inventories proposed here differ somewhat from those proposed elsewhere in the literature. The most coherent and developed proposal was perhaps by R.E. Emmerick in ‘The Vowel Phonemes of Khotanese’ (1979). He recognized that *ei* is distinct from *ai* in the Śgs (243), and so that the language, with *au*, has three diphthongs. For the monophthongs he presented a quadrilateral with an OKh inventory of seven short and three long vowels (248).

**Table 3.4:** Monophthongs in Emmerick 1979

<i>i</i> /i/	<i>ī</i> /ī/		<i>u</i> /u/	<i>ū</i> /ū/
<i>ä</i> /ɛ̃/				
<i>e, ai</i> /ɛ/		/ə/		/o/ <i>o, au</i>
<i>a</i> /a/ <i>ā</i> /ā/				

There is again no need to discuss the well understood *i*, *ī*, *u*, *ū*, *a*, *ā*. Emmerick recognized that *ä* could be a front vowel rather than the schwa [ə] of Leumann and Herzenberg, but agreed with the latter that unstressed *ä* could be [ə] (240, 242). About the front mid vowels his examination pointed, “to a pronunciation of *ä* as [ɛ̃] and of *e* as [ɛ]” (241). For the back mid vowel he stated that, “*au* seems to have been

monophthongised to *o* right at the beginning of our transmitted texts” (245). He would probably have allowed for *au*=/au/ in EWKh but perhaps not in AOKh. For the spellings *o*, *au*, he thought, “it is difficult to see any reason for assuming a pronunciation other than [o] for Old Khotanese” (245). On his quadrilateral diagram he included arrows pointing from both *ä* and *a* to /ə/, and under the chart he noted “/ə/ [sic] was an allophone of /a/ in unstressed syllables written *ä*” (248). Technically speaking, he proposed six short vowels /i, ɛ, ε, α, o, u/ and an allophone [ə] of both unstressed *ä* and *a*.

In 1989 Emmerick reported that, “Old Khotanese had the following vowel phonemes: /i ī ɛ ε a ā o u ū ə/” (209). That is, seven short vowels now including /ə/ and three long. In 1991 he published a study with Mauro Maggi, “Thoughts on Khotanese *e* and *o*,” in which they revised the inventory for mid vowels, adding long /ē/ and /ō/ largely on the basis of metrical evidence. They propose seeing minimal pairs in “*āstē* ‘bones’ beside *āstě* ‘he sits; *dasō* ‘ten’ beside *dasǒ*, acc. sg. from *dasā* ‘thread’; *hamō* ‘bowl’ beside *hamǒ* from the adjective *hama-* ‘same’” (67). They have followed Leumann in adding macrons and breves to the normalized transcription. The practice in these pages is to rigidly separate the three symbologies, orthographic, phonemic and phonetic. It is not immediately obvious if the macrons and breves on the normalized transcription are intended as phonemic or phonetic information.

In 2009 there appeared Emmerick’s last statement on vowels: “The vowel system of the older language appears to have had 11 phonemes” (382). His table,

reproduced below (with phonemes in italics and not with slanted brackets), shows five long and six short vowels:

**Table 3.5:** Monophthongs in Emmerick 2009

	Front	Central	Back
High	<i>ī, i</i>		<i>u, ū</i>
Mid	<i>ē, e</i>	<i>ə</i>	<i>o, ō</i>
Low		<i>a</i>	<i>ā</i>

There is no other comment in Emmerick 2009 related to Old Khotanese vowels. It appears that he has incorporated long /*ē*/ and /*ō*/ from Emmerick and Maggi 1991. It also appears that he has abandoned the front vowel interpretation for *ā*, and returned to schwa.

There are several differences between the inventory elaborated by Emmerick and Maggi and that proposed here. One is that here it is thought necessary to clearly define stages in the development of vowels within OKh whereas Emmerick and Maggi at times appear to treat all OKh materials as having the same inventory, although they are certainly aware that vowels evolve within OKh. For instance, in AOKh, *dasau* ‘ten’ would be spelled differently than ASf *daso* ‘thread’. Only after the monophthongization /*a*ṽ > *ō*/ sometime in the Canonical period could the spellings merge.

Emmerick and Maggi, who correctly note that the final vowels in *āste* ‘he sits’ and *āste* ‘bones’ have disparate metrical treatments, conclude that ‘he sits’ is phonemically /*āstē*/ and ‘bones’ /*āstē*/. They could similarly compare the metrical behavior of the final vowels in ASf *ūtco* ‘water’ and ASf *baśdo* ‘sin’ and conclude

these words are phonemically /ūt<sup>s</sup>ǒ/ and /baždō/. In contrast, it is here proposed that almost all instances of orthographic *e* and *o* belong to single phonemes /ē/ and /ō/ with phonetically long non final allophones [ǣ] and [ō], and phonetically short final allophones [ǣ̆] and [ō̆]. The relatively uncommon instances of final phonetic long [ǣ] and [ō] from contraction are highly marked. Actually, the phonetics are probably somewhat inaccurate. These contracted vowels may appear in the metrical X positions whereas non final *e* and *o* may not. It may be that what I have been giving as non final [ǣ] (=e=/ē/) and final [ǣ̆] (=e=/ē̆/) are in fact phonetically distinct. Emmerick and Maggi do not discuss the distribution of the mid phonemes in their analysis, but it may be that they would have to argue that short *e* and *o* occur only in final position, and long *e* and *o* occur only in non final position, or in final position in the case of contracted *e* and *o*. There may be no way to decide between the approaches other than the principle of simplicity and even that could likely be debated.

### **3.3.0 The Six Alternations of *ä***

In EWKh the vowels would have been written consistently. As the language evolved, spelling alternations arose where earlier vocalic contrasts had weakened. These alternations reveal information about the values of the vowel signs and about the development of the language.

The position here is that the vowel written *e* was the long correspondent /ē/ of the vowel written *ä* /ě/. For Emmerick and Maggi 1991 *ä* is a short vowel with no long partner, probably /ə/, and *e* is written for both long /ē/ and short /e/ (see 3.2.7).

It is clear that *ä* is short, front and non-low. Within OKh it occasionally alternates with *i*, in final position with *e*, and in unstressed medial position with *a*. These alternations are rare in Śgs but common in Z reflecting the evolution of the vowel system over time. It is possible to identify six alternations involving *ä*:

1. *ä* > *i* mostly in syllables lacking primary stress.
2. *ä* > *e* in final position with secondary stress.
3. *e* > *ä* in unstressed final position.
4. *a* > *ä* in unstressed medial position.
5. *ä* elided in unstressed medial position next to *r* /*r̥*/.
6. *ä* is elided in the IASm suffix *-äna*.

### 3.3.1 Alternations of *ä* in AOKh (Śgs)

The most visible alternation in Śgs is *ä* for unstressed *a* in the abstract noun suffix *-āmatā-* which occurs five times: LSf *e'māta* 3.1v1 < *ā'matā-* 'dwelling place'; LSf *vaṣṭemāta* 'absorption' 3.12v3,13r5 (cf. GDSf *vaṣṭemate* 2.4v2, 3.12v2,13v1, *vaṣṭemate jsa* IASf 3.13r2); GDSf *bvemāte* 'knowledge' IOL Khot 189/1 (Śgs new folio); IASf *bvemāte jsa* 'knowledge' IOL Khot 188/2v2 (Śgs new folio). Otherwise *-āmatā-* with *a* is very frequent. Emmerick noted, "*-emate* GDSf (never *\*-emāte* as in Z)" (xx) as one of the archaic features of Śgs (xx) but the new folios attest it twice (see above). There are some cases of *i~ä*: NSm *gyastā balysi* 3.7v3 ~ *gyastā balysä* 15x; NSf *śiratātā* 3.4v3 ~ GDSf *śāratete* 3.r3 'goodness'; NAPm [*hva*]ndä *ahvaṃndä*



‘men, non-men’ 1.1v5 ~ *hva’ndi* ‘men’ 3.13v3.<sup>67</sup> There may be one case of *e~ä* in *ne...nä* 3.6v3, for *\*ne...ne* emphatic negative (note Emmerick’s hesitation p.79).

These are either early examples of the vowel system evolving, or are emendations by later copyists. In the case of the *-i-* in *śiratātā* NSf, the emendation would appear likely as even in Z, *śśāratāti-* ‘goodness’ is spelled 15 times with *-ä-*, never with *-i-* (cf. Glossar: 505b).

In Śgs the distinction in the *a*-declension, NS *-ä* against GDS *-i* is almost invariable (xix). For example *balysa-* ‘Buddha’ is NSm *balysä* 15×, *-i* 1×, GDSm *-i* 14×, never *-ä* (Śgs 121b, 126a). In contrast, in Z, either *-ä* or *-i* can be used for either NS or GDS, e.g.: NSm *lovä* 9.20, *lovi* 2.21; NSm *balysä* 2.55,56 ~ *balysi* 2. 60, 81, 84, 90; GDSm *balysi* 2.73, 125 ~ *balysä* 2.63, 67, 71, 102; NSm *orsä* ‘desire’ 24.171, 25.113 ~ *orsi* 24.163, 172, 199; NSm *ggarä* ‘mountain’ 4.40 ~ *ggari* 22.254, GDSm *ggari* 13.11 ~ *ggarä* 2.84,118. The phrase *gyatsa- balysa-* is consistently *gyastä balysä* in the NSm 11 times, but is also *gyastä balysä* in the GDSm consistently three times, never *\*gyasti balysi* (cf. Glossar:426b).

### 3.3.2 Alternations of ä in COKh (Z)

#### 3.3.2.1 -ä~-e alternation in the 3S

In Śgs, for type A verbs, the 3Sp.a. ends in *-ätä* while the 3Sp.m. ends in *-äte* (xix), and the type B 3Sp.a. ending is *-’ttä*<sup>68</sup> while the middle is *-tte* (cf. *ibid.* xx). This pattern, active *-ä*/middle *-e*, partly breaks down in Z. The developments are

---

<sup>67</sup> The alternative stems *māsta-* ~ *mišta-* ‘great’, cf. SuvO 36r7 *māstäñe* ~ SuvO 36v3 *mištāñe* IASf, both occurring in Śgs, do not show synchronic *ä~i* alternation.

<sup>68</sup>Emmerick lists an alternative in *-i*, but as far as I can tell, only *-ä* is attested in Śgs.

complex and interesting and discussed in detail below (3.3.2.1.1-5). Briefly, in Z, type A active has *-äte* (*-ite*) more often than *-ätä*, while A middle always has *-äte*, never *-ätä*. And B active seems to end regularly in *-ä* (*-i*), rarely in *-e*, while B middle has *-e* ~ *-ä* alternation, rarely *-i*. That is, A active and B middle show alternation, while A middle and B active maintain the older form:

### 3.3.2.1.1 3Sp.a. A *-ätä*, *-äte*, *-ite* in Z

In SGS, that is, before his study of Śgs, Emmerick stated that “In the oldest Kh., *-äte* is more common than *-ätä*” and pointed out that in Z *byehäte* occurs nine times while *byehätä* only three; *māñäte*, *māñite* seven times but *māñätä* only at 24.4. Only with *kṣam-* are there slightly more cases of *-ätä*, 6 as opposed to 5 of *-äte* (SGS:192). Others from Z: *sarbite* 4×~*sarbäte* 5×~*sarbätä* 1× 2.43 ‘rises’; *sahyätä* 1× 2.138 ~ *sahyäte* 1× 11.45 ‘endures’; *bañäte* 11.36, 50~*bañite* 13.28, 24.44 (no *-tä*) ‘binds’.

### 3.3.2.1.2 3Sp.m. A *-äte* in Z

In SGS Emmerick commented “*-äte* (mid.) is never in Z spelled *-ätä*, and I have not noticed *-ätä* (mid.) elsewhere in O.Kh.” (198-199). But in the verb list under *padīm-* ‘to make’ (p.69), which must be middle as shown by the 1Sp.m. *padīme* Z 23.132, he gives the OKh alternates *padimätä* IOL Khot 185/1b3~*padimäte* Suv[Or] 5..9, 8.33, 12.64. In any case 3Sp.m *-ätä* is rare in COKh.

### 3.3.2.1.3 3Sp.a. B *-ttä* ~ *-tti* ~ *-tte* in Z

The most regular spelling appears to be with *-ä*. *-i* is much less common, and *-e* has been noticed just three times. *-ä~-i*: *grūstā* 3×~*grūsti* 2.21 ‘he calls’; *kaṃggīndi* 2.28 ‘he digs’; *uysgursti* 20.42 ‘it tears off (intr)’.

In discussing his types B and D Emmerick stated: “-*te* for -*tä*, -*ti* is rarely found in O.Kh. All examples found in O.Kh. are preceded by -*y*- or -*i*-, -*ä*-“ (SGS:194). Emmerick lists six examples of which three are no longer valid. *daiye* was a misreading for [*hvaṃ*] *däye* ‘of a man’ (corrected by Maggi Kvb:0.40). The two type D examples are now regarded as type A vowel stems (cf. Hitch in-press-b) which take the type A 3Sp.a suffix -*ätä* rather than the type B -*ttä*: *naltsaiye*, *naltseiye* < *naltsa-* + -*ätä* (cf. *naltseyä* Z 24.405) ‘goes out’. Type A frequently ends in -*e* as discussed above. Two of the three remaining examples, *nirminde*, *närmände* (cf. also *närmändä*, *nirmändä*), are from *närmän-* ‘create magically’ which is a loanword. The third example is *bitte* 2.138 ‘pierces’.

#### 3.3.2.1.4 3Sp.m. B -*tte* ~ -*ttä* ~ -*tti* in Z

Emmerick noted, “O.Kh. has -*te*, *tä* and rarely -*ti*” (SGS:199): *ditte* 1.37 ~ *dittä* 6.6 ‘appears’; *īste* 2.61 ~ *īstä* 12.68 ‘returns’; *kašte* 2.111 ~ *kaštä* 22.320 ‘falls’; *ysānde* 2.80 ~ *ysāndi* 2.85 ‘shines’; *hvīnde* 1.41 ~ *hvīndi* 11.4, 13.131 ‘is called’.

#### 3.3.2.1.5 Stress

This pattern may be explained by stress if we assume that normally the final stem syllable in OKh carries primary stress, and that stressed -*ä* tends to -*e* while unstressed -*e* tends to -*ä*. In type A endings the final vowel would have a secondary stress and tend to -*e*, i.e.  $CV^1Cätä^2 > CV^1Cäte^2$ . In this way, the 3Sp.m. would remain -*äte*, but the active -*ätä* would tend to -*äte* because of final stress. With type the B 3S endings, the final vowel would not carry secondary stress. Unstressed -*ä* in the active remains -*ä*, while unstressed -*e* in the middle tends to become -*ä*. These

developments are illustrated below where AOKh forms from Śgs are compared to forms from Z or Sgh. Raised <sup>1</sup> indicates primary and raised <sup>2</sup> secondary stress:

		Śgs	Z, Sgh
Type A 3S	p.a. -ätä > -äte	<i>bye<sup>1</sup>hätä<sup>2</sup></i>	<i>bye<sup>1</sup>häte<sup>2</sup></i> Z 10.26
		<i>kṣa<sup>1</sup>mätä<sup>2</sup></i>	<i>kṣa<sup>1</sup>mäte<sup>2</sup></i> Z 4.18
		<i>pane<sup>1</sup>mätä<sup>2</sup></i>	<i>pane<sup>1</sup>mäte<sup>2</sup></i> Sgh[1] <sup>69</sup> 36.4
		<i>pa<sup>1</sup>sätä<sup>2</sup></i>	<i>pa<sup>1</sup>śäte<sup>2</sup></i> Z 5.29
		<i>va<sup>1</sup>ṣtätä<sup>2</sup></i>	<i>va<sup>1</sup>ṣtäte<sup>2</sup></i> Z 24.387
		<i>hvā<sup>1</sup>ñätä<sup>2</sup></i>	<i>hvā<sup>1</sup>ñäte<sup>2</sup></i> Z 2.7
	p.m. -äte > -äte	<i>nyū<sup>1</sup>jäte<sup>2</sup></i>	<i>nyū<sup>1</sup>jäte<sup>2</sup></i> Z 24.194
		<i>pana<sup>1</sup>mäte<sup>2</sup></i>	<i>pana<sup>1</sup>mäte<sup>2</sup></i> Z 2.125, 240, 22.207
		<i>hā<sup>1</sup>mäte<sup>2</sup></i>	<i>hā<sup>1</sup>mäte<sup>2</sup></i> Z 2.129 + 110×
Type B 3S	p.a. - <sup>i</sup> ttä > - <sup>i</sup> ttä	<i>pī<sup>1</sup>ttä</i>	<i>pī<sup>1</sup>ttä</i> Z 11.58 + 9×, * <i>pītte</i> 0×
		<i>hī<sup>1</sup>ṣti</i>	<i>hī<sup>1</sup>ṣtä</i> Z 2.219 + 22×, * <i>hīste</i> 0×
		<i>hū<sup>1</sup>ṣtä</i>	<i>hū<sup>1</sup>ṣtä</i> z 9.17, 12.60, 88, * <i>hūste</i> 0×
	p.m. -tte > -ttä	<i>ī<sup>1</sup>ste</i>	<i>ī<sup>1</sup>stā</i> Z 12.68
		<i>hvi<sup>1</sup>nde</i>	<i>hvi<sup>1</sup>ndi</i> Z13.131, 11.4

Related evidence comes from the type A vowel stems. Here the contraction processes (Hitch 2015b, in-press-a) take place before stress assignment. For

<sup>69</sup> As ms. 1 of Sgh is an Archaic text, this form may be a counter example, or from a copyist.

instance, with the active verb *ākṣu-* ‘begin’ + *-ätä* > *ākṣūtä* > *ākṣū<sup>1</sup>tä* (Śgs, Z) the final *-tä* is unstressed as is shown by the fact that it can be lost in advanced OKh, i.e., *ākṣū<sup>1</sup>tä* > *ākṣū<sup>1</sup>* Suv[Or] 12.53. And with the irregular middle verb *jsā-* ‘go’ + *-äte* (?) > *jsāte* > *jsāte* (Śgs, Z) the final *-e* is unstressed and so can shorten to *-ä*, i.e., *jsātä* Sgh[23] 244[9]. The determining factor in these *ä~e* alternations is stress, not whether the verb is type A or B.

### 3.3.2.2 Other Endings Showing *-e* > *-ä*

Weakening of final *-e* to *-ä* can be seen in Z in several endings: 3Pp.m. *-āre*: *oysārä*, *hämārä* (also *-i*, e.g., *diyāri*, *yanāri* SGS 200); 1Sp.m. *-e*: *ggihä* 12.51 (also *-i*, *parehi* 24.493); *ā*-declension GDS *atme* 6.5 ~ *atmä* 4.43; IAS *biśśe jsa* 4.117 ~ *biśśä jsa* 22.323; NAP *avitsare* 23.150 ~ *avitsarä* 23.95 ‘apsaras’; *i*-declension GDS *-ie* *balysūste* 10.5 ~ *balysūstä* 13.62 ‘bodhi’; IAS *-ie (jsa)*: *vasuttattetä* 3.93 ‘purity’ cf. *mulśde jsa* 5.19 ‘compassion’; *i*-declension GDS *-ie* *balysūste* 13.62 ~ *balysūstä* 10.5 ‘bodhi’; IAS *ie (jsa)* *śśāratete jsa* ‘goodness’ 24.191 ~ *vasuttattetä* ‘purity’ 3.93. These weakenings are probably also stress related. In the examples in this paragraph there is probably primary stress in the last stem syllable except with the ending *-āre* where the long *ā* is presumably stressed.

### 3.3.2.3 Non final *-ä* ~ *-e*

Emmerick reported that the alternation of *ä* with *e* is rare in non-final position, and gave no examples (Emmerick 1979:240). It is difficult to find examples. Nouns

showing *ä* > *e* in stressed non final syllable: NSf *bätäva*<sup>70</sup> Z 20.61, 24.417 ~ *bäteva* Z 6.15 'lightening'; possibly IASm *dätäna* 'appearance' 11×~*dätēna* 7×.

### 3.3.2.4 Non final -*ä* ~ -*i*

Non-final *ä*~*i* variation in Z can perhaps occur rarely under stress. *däte* 'he saw' 3S pf.tr.m has -*ä*- 12 times and -*i*- once (14.52; Glossar:445b). Emmerick notes that *mästa*- 'great' occurs 226 times in Z with -*ä*-, twice with -*i*- (12.39, 13.128) and the frequent *vätä* 'towards' is likewise normally spelled with -*ä*- (Emmerick 1979:240). It is conceivable that these represent emendations by later copyists. But note *dätte* 10× ~ *ditte* 6× ~ *dittä* 6.6 ~ *dättä* 9.27, 3S p.m. of *dä*- 'appear' showing stem alternation of the stressed -*ä*/*i*-. In contrast, -*i*- appears for -*ä*- in unstressed syllables with a degree of frequency. In verb stems this is most obvious in historical prefixes (cf. SGS 229-244): *nijaṣ-* ~ *näjsäṣ-* 'show' (SGS:53); *nimandra-* ~ *nämaṃdra-* ~ *nämandra-* ~ *numandra-* 'invite' (SGS:54). *närmän-* ~ *nirmän-* ~ *nirmin-* (SGS:55; BHS *nirmin-* SGS 233); *niṣem-* ~ *näṣem-* (57); *nihujs-* ~ *nähujs-* 'sink down' (58); *nihvarr-* ~ *nähvarr-* 'long for' (58); *ttätsa-* ~ *ttitsa-* 'cross' (38); *päha-* ~ *piha* 'strike' (83); *bäyāśś-* ~ *biyāśś-* 'open' (97); *birays-* ~ *bärays-* 'extend, spread' (98); *västāta-* ~ *vistāta-* ppp. < *višt-* 'place, establish' (124).

---

<sup>70</sup>*bätävā*- 'lightening' is listed as *bätavā*- by Leumann (473) and *bätava* by Bailey (DKS 282b), both with -*a*- not -*ä*- in the second syllable. Skjærvø SuvII:313b has the correct stem. Of the OKh forms known to me, three have -*ä*-, two have -*e*- and just one has -*a*-: *bätava* NS Z 22.271, *bätäve* NAP Suv[Or] 12.31, *bätäva* NS Z 20.61, 24.417, *bäteva* NS Z 6.15, *bätevo* NAP + -*ū* post-positive conjunction 'and' (Z 3.125). The stem makes most sense with -*ä*- as the underlying form which becomes -*e*- under stress; -*a*- perhaps shows the omission of the diacritic by a copyist.

I notice one noun showing *ä* > *i* in unstressed non final syllable. *väśšeṣä* 9.5 ~ *viśšeṣä* 9.6 < Skt. *viśeṣa* ‘difference, distinction’. This is also possibly unstressed *i* > *ä*.

### 3.3.2.4.1 -*ä*- > -*i*- in 3Sp.m and 3Sp.a -*äte*, -*ätä*

About 3Sp.m type A -*äte* Emmerick noted ‘-ite is rare: *rrijite* Z 2. 77; 22. 165; 24. 658’ (SGS:199). *rrijäte* occurs at 4.51, 24.368 (Glossar:493a). About 3Sp. type A active -*ätä* he lists just one form with -*i*-, *māñite* ‘resembles’ 25.221, but also two with -*e*-, *peḍete* ‘clenches’ Z 6.13 and *paśsete* Sgh[24] 94.1 (SGS:192). In addition I notice *anuvarttite* (a./m.) 14.61 ‘conform to’, *paranirvite* (a.) 22.297 ‘attain complete nirvaṇa’; *pāraṃjite* (a./m.) ‘diminish’ H 142 NS 29 etc. 611v3 KT 5. 92; *bañäte* (a.) 11.36,50 ~ *bañite* 13.28, 24,444 ‘bind’; *sarbätä* (a.) 2.43 ~ *sarbite* 2.55+ ‘rise’.

Although 3Sp. type A endings with -*i*- are rare, they are much more common than those with -*e*-. The development here is *ä* > *i* in unstressed syllable.

### 3.3.2.5 *ä* > *i* in Final Position

As with *ä~e*, the *ä~i* alternation is most common in final position. The 3Pp.a. ending is most often -*indä* but occasionally -*indi*. Some examples: *kaṃggāndi* ‘they dig’ 2.28; *kūśindä* 7.2, 13.3 ~ *kūśindi* 11.64, 13.45 ‘they seek’; *ggei’śśindi* 6.4, 22.227 ~ *ggei’śśindä* 11.72 ‘they turn (tr)’; *tsīndi* 2.135 ~ *tsīndä* 2.56 ‘they go’; *parsīndi* 11× ~ *parsīndä* 4× ‘they escape’; *paṣṭindä* 24.484 ~ *paṣṭindi* 22.325 ‘they arise/set out’; *bañīndi* 22.324 ‘bind’; *bichānīndi* 24.413 ‘neigh’; *buvanīndä* 24.391 ~ *buvanīndi* 22.331 ‘be harmed, ruined’; *rrāysīndi* 2.46, 20.30 ‘cry out’ (of birds); *sarbīndä* 24.415 ~ *sarbīndi* 22.160 ‘rise’; *sājīndä* 24.646 ~ *sājīndi* 23.5 ‘learn’; *sāñīndi* 22.159 ‘raise’; *jsanīndä* 24.403 ~ *jsanīndi* 23.32 ‘strike; slay’; *narāmīndä* 3.99 ~ *narāmīndi* 22.173 ‘go out’ (of, from).

There are a variety of other *-ä* endings which sometimes show *-i* in Z. The 3Sp.a. B *-ittä*: *vasuštä* 5×~ *vasušti* 5×; 2S p.a. *-iä*: *yañä* 24.43 ~ *yañi* 5.42. 1S p.a. *-imä*: *hvarimi* 4.50. 1S opt.a. *-ä*: *parsi* 13.45. 3S opt.a. *-ä*: *kūši* 4.18. Present stem infinitive *-ä*: *hvāñi* (SGS:218). *i*-declension NS *tcarä* 21.16 ~ *tcari* 4.100; NAP *-iä* *hālštä* 4.60 ~ *hālšti* 5.74 ‘spear’; GDP *-iänu* *hālštänu* 24.422 ~ *hālštinu*. For 1Pp.a. *-ämä* Emmerick specifically notes: ‘*-ämi* has not been found’ (SGS:196).

### 3.3.2.6 *a > ä*

The vowel *a* can become *ä* in unstressed syllable. The best example may be with the abstract noun suffix *-āmatā-*, which, as shown above is written five times with *-ä-* in Śgs. There appear to be no cases of *-āmät-* in Z, but *-emäte* occurs 19×. The latter distribution suggests a mid front value for *ä* as the vowel is apparently assimilating to the vowels in adjacent syllables, i.e., /*ēmādē* > *ēmädē*/. The assimilation does not make sense if the value of *ä* is schwa.

### 3.3.2.7 *kāḍana > kāḍāna*

Śgs has *kāḍana* ‘on account of, for the sake of’ with medial *-a-* consistently 8 times (119-120, xx). For this reason it is probably a loan word (cf. Gdh *kridena* ‘to do with’; Baums and Glass ongoing s.v. *kidena*) rather than IAS to *kāḍa-* ‘done’ as suggested by Bailey (DKS:60a). But in Z, *kāḍana* is replaced by the consistent spelling *kāḍāna* with *-ä-* 28×, and the spelling *kāḍna* 2×. It is conceivable that this is a stress related development, i.e., that the weak stress had led the vowel to merge with /*ě*/. But it is also plausible that the word was reanalyzed to match the IAS *-āna*.



### 3.3.2.8 Elision of *ä* next to *r*

*ä* can be elided in unstressed syllable next to /r/: Śgs *bäsivārāṣā* 2× ~ *bäsivrāṣā* 4× (126-127); *bārūñāte* Suv[Or] 0.16~*brūñāte* Z 2.84 ‘shines’; *bārūñāre* IOL Khot 172/1a4~*brūñāre* Z 3.84 ‘they shine’; *ggarāna* IASm ‘mountain’ Z.14,28,41 ~ *ggarna* Z 2.86,128, 24.169,468; *hārāna* IASm ‘thing’ 1× in Z (12.18) ~ *hārna* 16× in Z; *harṣdi* 3Sp.a. Z 20.57 < \**harāṣdā* < *harāṣ-* ‘burst’, cf. 2S imper. *harāta*’ Z 6.23; *bārāh-* ~ *brāh-* ‘soar up’ (SGS:98).

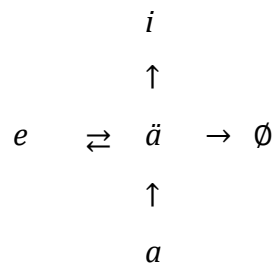
### 3.3.2.9 Elision of *ä* in IASm -*āna*

The *ä* in the IASm suffix -*āna* is sometimes elided. A thorough study has not been done but preliminarily elision seems to be possible after stem final sonorants, e.g., *padamna* ‘wind’ Z 2.43, *marāṇna* ‘death’ Z 6.30, *ṣṣamanna* ‘monk’ Z 2.19, *kṣīrna* ‘land’ Suv[Or] 6.2.34. Some of the examples in 3.3.2.8 also apply here.

### 3.3.2.10 Summary of *ä* alternations

The chart below summarizes the alternations discussed above.

**Table 3.6:** The alternations of *ä*



Under stress, *ä* can be written *e*. Similarly, unstressed *e* tends to be written *ä*. It seems plausible that *ä* and *e* are in a lax~tense relationship. The tense *e* would tend to appear under stress while the lax *ä* would tend to appear in unstressed position. This proposal is viable if *e* is long (tense) /ē/ and *ä* is short (lax) /ě/ as proposed

here. It is less tenable under the Emmerick and Maggi 1991 view that *e* ambiguously marks /ē/ and /e/ while *ä* is schwa.

The remaining alternations also suggest that *ä* was a short, front, mid vowel. In unstressed position, *ä* can be written *i*. This shows a tendency of the two short front vowels to merge when unstressed. When next to *r* /r/ or in the IASm suffix *-āna* after a sonorant, *ä* can be elided. This shows that *ä* is likely the least specified vowel in the system in terms of distinctive features, i.e., probably [-low, -high, -back, -long], which implies mid [-low, -high] and front [-back]. The *ä* is the default vowel, and in that sense resembles schwa in systems that have a phonemic schwa. Finally, the writing of unstressed *a* as *ä* shows that *ä* is likely mid, not high.

### 3.4 The Alternations of *o* and *au*

In EWKh it seems likely that the two vowel signs, *o* and *au*, were used consistently. A comparison of various *o* and *au* spellings in the AOKh Śgs and the COKh Z reveals interesting patterns. In most cases in Śgs there is no *o*~*au* alternation, either *o* or *au* is used. In Z, a more progressive stage, there is much more alternation but on the whole, the earlier pattern is preserved. Forms with *o* in Śgs have mainly *o* in Z and likewise forms with *au* in Śgs have mainly *au* in Z. Another interesting pattern in Z is the occasional appearance of final *-u* for *-o*. There also seems to be a difference in treatment between final and non-final position:

#### 3.4.1 EWKh non final *o*

Within the stems, *uysnora-*, *tcahora*, *hotā-*, and *hor-*, an *-o-* in Śgs is in Z mostly *-o-* but sometimes *-au-*. In meter, this vowel is always heavy, or two moras long. In

contrast, in the indeclinable words *kho*, *puşşo*, *buro*, *rro*, a final *-o* in Śgs is never written *\*-au* in Z but can sometimes be written *-u*. In meter, this vowel is always light, or one mora long. The conjunction *o* ‘or’ is a special case. Although indeclinable, it is always metrically heavy and never becomes *\*u* but is written *au* once in Z.

**Śgs non final *o* and *o* ‘or’**

**Z *o~au*** (metrically heavy *o*, *au*)

*uysnora-* ‘being’ (27×)

*uysnora-* (±85×) ~ *uysnaura-* (15×)

*hoto* ASf ‘power’

*hotā-* 27× ~ *hautā-* 4×

*horīndä* 3P pr.a. ‘give’,

*hor-* 8× ~ *haur* 4×:

*horāmatīnei* NSm ‘pert. to giving’

1S p.a. *horāmä* 11.18, *horimä* 11.75,

1P p.a. *horāmä* 24.517; 3S opt. a.

*hauri* 3.148; 2S imper. a. *haura*

12.10, *hora* 12.26, 24.463; 2P

imper.a. *hoḍa* 12.27, 24.414; pt. nec.

*haurāñi* 6.48, *horāñi* 12.28; inf.

*haurä* 13.71. Cf. also noun: *hora-*

(±30) ~ *haura-* (7×)

*tcohorvaretcoholśuvo*’ LPM ‘44’

*tco-/tca-*, *-hor-/haur-*, *-a/-ä/-i*, with *-*

*ho-* 36×, *-hau-* 6×; *tco-* 30×, *tca-* 12×

*o* conj. ‘or’ (12×)

*o* (±85×) ~ *au* (1×)

**Śgs final *o* in indeclinables**

**Z *o~u*** (metrically light *o*, *u*)

*kho* conj. ‘as; how; when; so that’ (19×)

*kho* very frequent, *khu* 1×

*puşşo* adv. ‘completely’

*puşşo* more than 100×, *puşşu* 1×

*buro* indefininte particle (10×)

*buro* (50+x) ~ *buru* (6×)

*rro* encl. adverb 'also, too' (6×)

*rro/ro* (±150×) ~ *rru/ru* (16×)

### 3.4.2 EWKh *au*

An *au* in Śgs is, in Z, written mostly *au* but sometimes *o*. There is no difference between declinable and indeclinable words, or between medial and final position.

#### Śgs *au*

#### Z *au~o*

*anau* 'without'

*anau* 4×

*thatau* 'quickly' (2×)

*thatau* 26× ~ *thato* 7×

*dasau* 'ten' (NAP 3×)

*dasau* 23× ~ *daso* 8×

*paḍauysānu* GDPm 'leading, chief'

*paḍauysa-* 8× ~ *paḍoysa-* 6×

*śau* NASm (4×)

*śśau* NASm 16× ~ *śśo* NASm 9×

*haudyau* IAP '7'

*haudyau* 8× ~ *hodyau* 1×<sup>71</sup>

### 3.4.3 EWKh *-o* and *-au* in Inflections

Leumann and Emmerick list some of the consistent or nearly consistent features in Śgs inflections as 'archaic orthography': *a*-decl NSm *-ä*/GDSm *-i*; 3Sp.a *-ätä*/3S p.m. *-äte*; *a*-decl IAS *-äna* never *-ina*; *aa*-decl NS *-ei*/GDS *-ai* (Śgs xix-xx). To this list can be added: *ā*-decl AS *-o* never *-au*; *aa*-decl AS *-au* never *-o*; all declensions IAP *-yau* never *-yo*; and probably some others but there are too few examples for firm conclusions (see 3.4.3.4-9).

---

<sup>71</sup>Numbers tend to be irregular. In Z, except for the IAP, forms with *hoda-* are more than twice as numerous (17×) as forms with *hauda-* (7×): NAP *hauda* 5× ~ *hoda* 13× '7', *haudama-* 1× ~ *hodama-* 3× '7th', *haudätä* 1× ~ *hodätä* 1× '70'.

### 3.4.3.1 ASf of the *ā*-declension, EWKh -o

In Śgs this is -o, never -au, 43×: *agāṣṭo* ‘inconceivable’, *gvārāṣkyo* ‘explanation’ (2×), *trāmāmato* ‘entry’, *tsūmato* ‘career’ (4×), *dyāmato* ‘appearance’, *nyasāmato* ‘despising’, *padamgyo* ‘description; manifestation’, *padāro* ‘maintenance’, *balysūñīgyo* ‘pert. to bodhi’, *bvāmatīgyo* ‘pert. to bodhi’, *rraṣṭo* ‘right’, *vaṣṭāmato* ‘absorption’ (9×), *śūraṅgamo* (6×) n. of meditation, *tto* ‘this’, *ṣavo* ‘night’, *ttuto* (6×) ~*ttuvo*~*tvo* ‘this’, *hastamo* ‘best’, *hoto* ‘power’, *hvānāmato* ‘speech, preaching’.

In Z the endings -a and -u are common (SGS:271-73). But the most common ending is still -o (ibid. 273) and about -au Emmerick noticed only two examples in OKh, both from Z, *b[ūṣ]ṣātau* ‘staircase’ 23.153 (beside *būṣṣāto*’ 23.166) and *Sīysau* 5.4 ‘Sītā’.

### 3.4.3.2 ASm of the *aa*-declension, EWKh -au

In Śgs this ending is consistently -au, never \*-o, 8×: *agūnau* ‘without characteristics’, *alysānau* ‘prince’ (2×), *kumṣatīnau* ‘pert. to sesame’, *ttuśau* ‘empty’, *nyāttara-kṣīrau* ‘pert. to a lesser land’, *spātainau* ‘consisting of flowers’, *hvanau* ‘doctrine; speech’,

In SGS, both -o and -au are described as ‘Common in O.Kh.’ (298). But it seems that -au still predominates, e.g., in Z: *ysamaśśandau* 9× ~ *ysamaśśando* 1× ASm ‘earth’; *hvanau* 5× ~ *hvano* 3×; *aggūnau* 2× (6.49,56); *ttuśśau* 2×; *āchānau* 2× ‘ill’; *āsto* 1× ‘bone; *kṣundau* 1× ~ *ksumdo* 1× ‘husband’.

### 3.4.3.3 IAPm and IAPf, EWKh -yau

In the AOKh Śgs this is consistently -yau, 30×: *agāṣṭyau* ‘inconceivable’, *anantanaryau* ‘deadly sin’, *avarrūṣkyau* ‘kleśa-free’, *uysnoryau* ‘being’, *kṣatryau* (2×)

‘umbrella’, *tcūryau* ‘four’, *tcei’maundyau* IAPm ‘able to see’, *daśyau* ‘ten’ (2×), *diśyau* ‘direction’, *pā’gañyau* *jsa* m. ‘power’, *pātāmyau* m. ‘confusion’, *pāškalyau* m. ‘thing analysed’, *pracyau* m. ‘cause’, *buljyau* (*aā*-decl.) f. ‘virtue, merit’, *ratanyau* ‘jewel’, *vicitriyau* ‘various’, *ttyau* ‘this’ (4×), *sañyau* ‘plan, method’, *salāvyau* ‘speech’, *handaryau* ‘other’, *harbiśyau* ‘all’, *hāryau* (3×), *haudyau* ‘7’,

In COKh an alternant *-yo* appears but in COKh *-yau* is still more common than *-yo*, e.g., in Z *dukhyo* 9× ~ *dukhyau* 51× (SGS:268)

#### 3.4.3.4 ASf of the *aā*-declension, EWKh -o

In Śgs this occurs twice with *-o*: *uysāno* ‘self’, *patāraho* ‘basis’.

In Z *-o* dominates over *-au*: *baśdo* 8× ~ *baśdau* 2× ‘sin’; *śśando* 7× ~ *śśandau* 3×; *bitamo* ‘doubt’ 2×; *brīyo* ‘love’ 3×; but *uysanau* 2× ~ *uysano* 1× (2.220). In Hitch 2015b I showed how this contracted *-o* is always metrically two moras (298–300). This may explain why it is far more commonly replaced by *au* than the metrically short *-o* of the ASf of the *ā*-declension.

#### 3.4.3.5 LPm and LPf, EWKh -uṣo

In Śgs the LP is most often spelled *-uvo*’, 14×: *avamātuvo*’ (2×) ‘immeasurable’, *kaṃthuvo*’ ‘city’, *kalpuvo*’ ‘eon’, *janavatuvo*’ ‘district’ (2×), *tcohorvaretcoholśuvo*’ LPm ‘44’, *parmihuvo*’ ‘settlement’, *buddha-kṣetruvo*’ ‘Buddha-field’ (Skt.), *yseruvo*’ < *ysāra-* ‘1000’, *lovadātuvo*’ ‘world-sphere’, *sītuvo*’ < ‘hundred’ *sata-* + *-iuvo*’, *harbiśuvo*’ ‘all’. Note also *āvuto*’ ‘village’ (< \**āvuvuvo*’ Hitch in-press-a), *parrāvo*’ (/baɪɹ̥ʷo/ \**parrātuvo*’) < *parrāta-* ppp. m. ‘elapsed (time)’, *mištāvo*’ *janavato*’ ‘in great districts’ (*mištāvo*’ has archaic *-āṣo*; Hitch in-press-a).

In Z, and elsewhere in OKh, the ending *-uvo'* can have a wide variety of spellings as catalogued in SGS (269-70, 279-80). It is never written with *-au* but there is one case of *-u'*: *Ttuṣātu'* 'Tuṣita' m. Z 24.182 (SGS:269)

#### 3.4.3.6 LSm and LSf, EWKh -o

The LS ending *-o*, never *-au* or *-u*, occurs 10× in Śgs: *āvuto* 'village', *kalpo* 'eon', *gā'to* 'group', *guvo'* 'ear', *janavato'* 'district', *naryo* 'hell' < *naria-*, *parmiho* 'settlement' m., *vairocano* n. of *kalpa*, *śūraṅgamo* (2×) n. of meditation.

Elsewhere in OKh there are no cases of *-au* in the LS. But in Z there are a few cases of *-u*: LSm *gūttāru*, *śrāvakayānu*, *samudru*, *sūtru*; LSf *Lmarīcu* (SGS:262-263, 277-78)

#### 3.4.3.7 3Pop, EWKh -īro

Śgs has two examples with *-o*: *vīro* 'be', *vamasīro* 'believe in'.

In Z, the 3P op.a or m can be either *-īro* or *-īru* (SGS:209, 210). But note the 2P op.a and m is almost exclusively *-īru*. Possibly the 2P was originally *-īru* while 3P was *-īro*, and that the weakening in the 3P of *-o* > *-u* was accelerated by the 2P.

#### 3.4.3.8 3Psj.m., EWKh -āro

In Śgs just *hāmāro* (2×) has this ending.

In SGS more examples of *-āro* are given than *-āru*, from Z and elsewhere (203, 205-6).

#### 3.4.3.9 3Siv.m type A, EWKh -āto

Śgs *hāmāto* which occurs 3× (once in a new folio, IOL Khot 190/3v2) is apparently the only example of the ending.

### 3.4.4 Exceptions

It is useful to note the exceptions to this pattern in Śgs, that is, the few cases of *o~au* alternations. They may be partly explained as influence from later copyists. There is *hotana-* ‘powerful’ (3×) ~ *hauvana-* (1×). The *-v-* for *-t-* shows later influence. The morpheme *lova-* ‘loka’ has *-o-* 5× and is spelled *lauva-* 1×. That this is a late addition is suggested by the consistent spelling in Z *lova-* 14×. The ppp *byauda-* is 3× with *-au-*, in *byaudāndä* 3P pf.tr. and *hubyaudä* (2×) NSm ‘well found’, and it is once with *-o-* *byode* 3S pf.tr.m. In Z there is *byauda-* 2× ~ *byoda-* 1×. Note also in Z the 3S of *byau-* is *byaude* 19× ~ *byode* 18×. Finally the *au-* declension appears to be showing *-o* in some places already in Śgs. With the NSm of *hamau-* ‘bowl’ there is *-au* three times, *darra-hamau* 3.14v1 ‘having a broken vessel’, *hamau* 3.13v4,5, and *-o* two times, *hatcasta-hamo* 3.14r4 ‘having a broken vessel’, *hamo* 3.14r2. There is also NAPm *hamo* once. With *pabastago* 3.2r5 ~ *pabastatau* 3.6r2 ‘continuously’ the *-g-* probably confirms Late influence.

### 3.4.5 Canonical Old Khotanese *o~au*

#### 3.4.5.1 *o* and length in monosyllables

Emmerick and Maggi 1991 list monosyllabic words to which Leumann in his edition of Z had assigned long *ō* or short *o* to according to metrical evidence. The short *o* words *kho* and *rro* are discussed in 3.2.1. They have short /o/ in Z as shown by the spellings *khu* and *rru*. The third listed short *o* word, *jsa*, occurs just 5 times. According to the Leumanns it is 3× ‘partikel “eben, da”’, and 2× ‘nebenform zu -jsa’ (Glossar:429b). It is not clear how these relate to the entry *jsau*, *jsa* ‘indeed’ in DKS (116b). The long *ō* words would mostly have been written *au* in EWKh. The only one



attested in Śgs, *śau*, is so written four times. In Z these words show *o~au* alternation as listed below. *o* ‘or’ and *ko* ‘if’ had /ō/ in EWKh show alternation with *au* in Z.

<b>Long ō in Z</b>		<b>Other Z Spellings</b>
<i>o</i>	‘or’	<i>o~au</i> (see 3.2.1)
<i>ko</i>	‘if’	<i>ko~kau</i>
<i>co</i>	‘who ... you all’	<i>kyau~cau~co</i> < <i>ce</i> + <i>-ū</i>
<i>tto</i>	‘thus ... you all’	<i>tto~ttau</i> < <i>tta</i> + <i>-ū</i>
<i>tto</i> ASf	‘this’	<i>tto</i> (1×) < <i>ttuto</i>
<i>tto</i>	‘thus’	<i>tto</i> (4×) [strengthened <i>tta</i> , like <i>ka&gt;ko?</i> ]
<i>tso</i> 2S imp.	‘come’	<i>tso~tsau~tsu</i> (irregular verb)
<i>dro</i>	‘hair’	stem <i>drau-</i> . 2.136 <i>dro-mase</i> ‘the size of a hair’ ~ <i>drau-mujsya</i> 6.4 LS ‘hair pore’
<i>no</i>	‘nine’	<i>no~nau</i> , also <i>notä~nautä</i> ‘90’, <i>nautama-</i> ‘90th’
<i>no</i>	‘ship’	stem <i>nau-</i> . 2× in Z <i>no</i> NS
<i>no</i>	‘not ... you all’	<i>no</i> 1×, <i>nau</i> 1×, but also <i>nyo</i> 1×, <i>nyau</i> 2×. < <i>ne</i> + <i>-ū</i>
<i>po’</i> LP	‘feet’	stem <i>paa-</i> . <i>po’</i> 3× ~ <i>pau’</i> 10× ~ <i>pät’o</i> 2× ~ <i>pvo’</i> 3× (/pa- + <i>-ūō</i> / > /pa <sub>u</sub> o/ <i>pau’</i> > /pōo/ <i>po’</i> OR /pā- + <i>-ūō</i> / > /pewō/ <i>päto’</i> or > /pūō/ <i>pvo’</i> )
<i>rro</i> AS	‘plain’	stem <i>rraa-</i> : <i>rro</i> 3× ~ <i>rrau</i> 3×
<i>śśo</i>	‘one’	NSm <i>śśau</i> 9× ~ <i>śśo</i> 5×; ASm <i>śśau</i> 7× ~ <i>śśo</i> 4×
<i>ho</i>	‘voice’	stem <i>hau-</i> : NS <i>ho</i> 1×, AS <i>ho</i> 1×; IAS <i>hauna</i> 3× ~ <i>hona</i> 9×.

### 3.4.5.2 ko

The issues around *ka*, *ko*, *kau*, *kai* are complex. It appears that there is a separate word *ko* which is not identical to *ka + -ū*. *ko* appears twice in the Śgs where one might expect a more conservative *\*kau* if the form were *< ka + -ū*. Emmerick defines it as “if, strengthened form of *ka*” (Śgs 120). It is curious that in Z *ko/kau* is used only with the subjunctive or optative (exc. 11.54), while *ka* may also be used with the indicative (Glossar:407a). By itself this might indicate that *ko* was syntactically different from *ka*, and so might be a separate word. But it would be premature to infer this since *kai*, which is a contraction of *ka + -ī* is also used only with the subjunctive and the optative (ibid.). Still, it is quite plausible that there were originally two different forms, a simple word *ko /kō/* and a contraction *kau (ka + -ū) /kaū/* which merged formally in */kō/ ko, kau* with merger of */ō/* and */aū/*. Simple */kō/* is perhaps visible in two spellings in:

2.25cd     cu va ne **ko** | ṣāte biśśu bu□tta | tte klaiśa | biśśu jita āro (A:5+7+5+7)

How much less should he know all, should his *kleśas* be utterly removed.

2.4cd     biśye ysama|śśandai satvā □ | cu va ne **kau** | hastari āya (A:5+7+5+7)

[a] being in the whole world ... how much less if there should be one better!

Contracted *ko* is perhaps found in

5.106cd     nāvuñi mā | jīvātā balysa □ | tterā harāte | **ko** rro dātāmā .

(A:5+7+5+7)

Not without merit is my life, Buddha, since so much has been left that I have seen you.

### 3.5.0 Umlaut Evidence

There are two morphophonemic processes within OKh which reveal much about the vowels. Both involve suffixing. Umlaut may affect a stem vowel when an umlaut triggering suffix is attached, e.g. *arra-* ‘arm’ + *-<sup>i</sup>a* LS > *erra*. Contraction occurs when a vowel final stem meets a vowel initial suffix (detailed in Hitch 2015b and in-press-a).

The umlaut p.s.p.a.rocess is complex. The umlaut potential in a suffix, traditionally shown by *’*, is drawn toward the stressed vowel in the stem. An intervening consonant can have one of three effects. (1) It can absorb the umlaut potential, e.g., *drūja-* ‘lie’ + *’e jsa* > *drūje jsa* IASf. (2) It can realize the umlaut potential, e.g., *uryāna-* ‘garden’ + *’a* > *uryāñna* LSm. (3) It can let the umlaut potential pass on towards the stressed vowel, e.g., *kanthā-* ‘city’ + *’e* > *kīnthe* GDSf. In the examples below, the stems all end in consonants of type (3), those which are transparent or neutral to umlaut, in order to show the effects of umlaut on the vowels.

#### 3.5.1 Umlautable Vowels

/a/	a ----- /ī/	ī	<i>tcari-</i> + <i>’u</i> > <i>tcīru</i> ASf ‘face’; <i>pat-</i> + <i>’ttā</i> > <i>pītā</i> 3S p.a ‘fall’
/ā/	ā ----- /ē/	e	<i>bātā-</i> + <i>’<sup>i</sup>e jsa</i> > <i>bete jsa</i> IASf ‘wind’; <i>bāḍa-</i> + <i>’<sup>i</sup>a</i> > <i>beḍa</i> LSm ‘time’
/ū/	ū ----- /ui/	vī, uī	<i>mūrā-</i> + <i>’<sup>i</sup>e jsa</i> > <i>mvīre jsa</i> IASf ‘coin’; <i>rrūd-</i> + <i>’<sup>i</sup>ttā</i> > <i>rrvītā</i> 3Sp.a ‘grow (intrans.)’; <i>ūra-</i> + <i>’<sup>i</sup>a</i> > <i>uīra</i> LSm ‘womb’

/ō/	<i>o, au --/ue/</i>	<i>ve, e</i>	<i>hotā- + <sup>i</sup>e jsa &gt; hvete jsa</i> IASf ‘power’ Śgs 132b; <i>hor- + <sup>i</sup>ttā &gt; heḍā</i> 3Sp.a ‘give’ (19×); <i>puror- + <sup>i</sup>ttā &gt; LKh pīreḍa</i> ‘to take away’
/au/	<i>au, o --/uai/</i>	<i>vai</i>	<i>kṣamotta- + <sup>i</sup>e &gt; kṣamvaittā</i> infinitive ‘ask forgiveness of’; Lsf <i>avachvaido</i> Rk 23(2) < <i>avachauda-</i> ‘unobstructed’ + <sup>i</sup> o.

### 3.5.2 Not Umlautable Vowels

/ī/	<i>ī</i>	<i>haṃbīr- + <sup>i</sup>ttā &gt; haṃbīḍā</i> 3Sp.a ‘to be filled’; <i>bīr- + <sup>i</sup>ttā &gt; bīḍā</i> 3Sp.a ‘to throw, sow’; <i>ksīra- + <sup>i</sup>a &gt; kṣīra</i> LSm ‘land’.
/ē/	<i>ē</i>	<i>brem- + <sup>i</sup>ā &gt; bremā</i> 2S p.a. ‘to weep’; <i>byev- + <sup>i</sup>o &gt; byevo</i> 1S opt.a. ‘to obtain (tr.)’
/i/	<i>i</i>	<i>naṣkīrr- + <sup>i</sup>ttā &gt; naṣkīrrḍā</i> 3Sp.a LKh ‘to cut’ (OKh * <sup>i</sup> ā- not <sup>i</sup> - ?); <i>jīn- + <sup>i</sup>ttā &gt; jīnḍā</i> 3Sp.a ‘to destroy’; <i>bī- + <sup>i</sup>ttā &gt; bītte</i> 3Sp.a ‘pierce’.
/ě/	<i>ā</i>	<i>ggān- + <sup>i</sup>ttā &gt; ggānḍā</i> 3Sp.a ‘to buy’; <i>patāḍ<sup>72</sup> + <sup>i</sup>ttā &gt; patālttā</i> 3Sp.a ‘to cut off’; <i>dāta- + <sup>i</sup>e &gt; dāte</i> infinitive ‘to see’; <i>nā- + <sup>i</sup>ttā &gt; nāttā</i> 3Sp.a ‘to sit down’.
/u/	<i>u</i>	<i>purr- + <sup>i</sup>ttā &gt; purḍā</i> 3Sp.a ‘to overcome’; <i>huṣṣ- + <sup>i</sup>ttā &gt; huṣṣā</i> 3Sp.a ‘to grow (intr.)’; <i>yuda- + <sup>i</sup>e &gt; yuḍe</i> infinitive ‘to make, do’; <i>paju- + <sup>i</sup>ttā &gt; pajuttā</i> 3Sp.a ‘to conceal’
/o/	<i>o, au</i>	<i>hotā- + <sup>i</sup>e jsa &gt; hote jsa</i> IASf ‘power’; <i>hauda- + <sup>i</sup>e &gt; hauḍe</i> infinitive ‘to give’; <i>haṣkaunda- + <sup>i</sup>e &gt; haṣkaunde</i> infinitive ‘to make’.

---

<sup>72</sup>SGS *patālt-*

### 3.5.3 Umlaut phonetics

A series of observations may be made about the five vowel umlaut processes. Regardless of the length of the source vowel, the target is always long. The source vowel is always non front /a, ā, ū, ō, au/ while the target is either a front monophthong /ī, ē/ or a diphthong or triphthong with a front final component /ui, ue, uaj/. The derivations  $a > \bar{i}$  and  $\bar{a} > e$  may reveal something about the phonetics of short /a/. In systems with a distinction between short and long vowels, or between lax and tense vowels, the long or tense vowels will be phonetically at the periphery of the vowel quadrilateral, while the short or tense partners will be phonetically more towards the interior. For instance, the tense vowel in English *beet* is phonetically in the high front corner, while the lax partner in *bit* is towards the interior. Khotanese long  $\bar{a}$  will be at the lower edge of the quadrilateral while short  $a$  will be phonetically somewhat higher. This is perhaps proven by umlaut where the source and target vowels are, in parallel fashion, separated by phonetic height. As source  $a$  is higher than source  $\bar{a}$ , target  $\bar{i}$  is higher than target  $e$ . The phonetic distance between  $a$  and  $\bar{a}$  will be more limited if there is a mid central schwa, than if there is not. Sanskrit has no phonemic schwa and so the pronunciation of short  $a$  may be schwa-like. These observations affect the quality of  $\bar{a}$ . If the vowel is mid front [ɛ] as argued here, there is room for more distance between  $a$  and  $\bar{a}$  than if  $\bar{a}$  is [ə] as argued by Emmerick and Maggi.

The umlaut  $\bar{u} > \bar{v}$  is elegantly parallel to  $o > ve$ . Phonemically these are / $\bar{u} > \bar{u}i$ / and / $\bar{o} > \bar{u}e$ /. The source vowels are phonetically round and back while the target

diphthongs have a /ɤ/ on-glide and a target second component which is front but which matches the height of the source.

The umlaut of the diphthong *au* > *vai* has much in common with the preceding two. Phonemically this is /aɤ > ɤai/. The target has a /ɤ/ on-glide and the remaining components /ai/ match the starting and finishing heights of the source /aɤ/. This analysis supports the existence of triphthongs in Old Khotanese. In Hitch 2015b:314 I proposed that *uaira-* (in Z *uaira-* 6×, *ueirä* 23.8) ‘suitable’ and *uaiṣṣa* ‘awake’ begin with a triphthong. If they do not, one might expect, for instance, a spelling \**vaira-* to alternate with *uaira-* but the only OKh variant is *vuairä* Suv[Or] 12.17 (DKS:36a). *uaiṣṣa* is restricted to Z and is so spelled five times. It is also useful to reckon with a mirror image triphthong /jau/ in the synchronic derivation of the IAP morpheme -*yau*.

### 3.6.0 Enclisis

The contraction processes related to enclisis can reveal information about vowel phonetics. The 3S enclitic pronoun *-i*, the 2Sp.a enclitic verb ‘be’ *-ī*, and the emphatic *-ī* attach only to words otherwise ending in a phonetically short vowel. After phonetically long vowels, the form of the 3S enclitic is *yä* and there is no contraction.

#### 3.6.1 3S enclitic pronoun *-i*

The 3S accusative and genitive-dative enclitic pronoun for all genders *-i* has the following behavior when attached to vowels (cf. *Lessons* 11.1-3; Glossar s.v. *yä*, *i*):

*-a + -i* > *-ai* (*-ei*)      *tta + -i* > *ttai* Z 2.53 ‘thus ... him’; *ka + -i* > *kai* Z 3.66 ‘if ... to one’, *kei* Z 12.29 ‘if ... to him’; *ṣṣāvyau jsai* Z 12.71 ‘from the

Śrāvakas ... from him'; *ttyau jsei* Z 24.441 'from these ... him'; *mai* Z 5.108 'not ... him'; *parrījai* Z 5.108 'rescue him'; *prraṇāhānai* Z 12.4 'his vows'; *cāndākei* Z 14.78 'great ... his'.

-ä + -i > -ī      *ttitä + -i > ttīti* Z 22.327 'then of him'; *atä + ī > atī* Z 24.650 'very ... for him'; *ārrä + -i > ārrī* Z 12.95 'fault ... of him'; *tterä + -i tterī* Z 3.147 'as many ... to him'; *kyerī* Z 23.3 'how many ... to him'; *nāstī* Z 9.15 'there is no ... of it'; *biśśī* Z 2,68 'all his'; *harbiśśī* Z 5.18 'all ... for one'; *bendī* Z 13.28 'on it'; *māñandī* Z 2.4 'resembling him'; *vāti* Z 5.27 'in his ...'; *vīri* Z 22.172 'for ... of him'.

-u + -i > -uī, -vī      *ku + -i > kuī* Z 2.85, *kvī* Z 2.66 'when ... (of) him'; *pharu + -i > pharuī* Z 13.22, *pharvī* Z 22.247 'very ... (by) him'; *cu + -i > cvī* Z 1.41 'what his'; *aśśu + -i > aśvī* Z 5.31 'horse ... him'; *thvī* Z 2.30 'you ... him'; *biśvī* Z 17.24 'everything ... to him'; *harbiśvī* Z 16.5 'all ... it'; *aysvī* Z 2.52 'I ... it'; *ttaṃdvī* Z 22.148 'to this degree ... it'; *trāmvī* Z 2.59 'so ... to him'; *samvī* Z 5.16 'only ... of it'; *dīruī* Z 23.113 'bad to him'; *kāḍaruī* Z 24.53 'sword on him'.

-e + i > -ai, -ei      *kye, ce + -i > kyai* Z 14.5, *cai* Z 4.34 'who ... it'; *ne + ī > nei* Z 12.44, *nai* Z 2.55 'not ... him'; *kīnthe + -i > kīnthai* Z 22.138 'of the city ... of him'; *kādai* Z 2.63 'very ... him'; *cūḍai* Z 21.20 'why ... it'; *dātai* Z 4.8 'saw him'; *vajsiṣṭai ... paysāndai*

Z 23.168 'he saw her ... he recognized her'; *parrai* Z 22.146  
'its fetlocks'; *ysarai* Z 24.404 'thousand ... his'.

-o + i > -vei, -vai    *kho + -i > khvei* Z 14.78 'as ... his', *khvai* Z 2.127 'how  
(-vī, -ai)    ... him' (*khvī* Z 23.39 'how ... of him' < \**khu + -i*); *puṣṣo + -i >*  
*puṣṣvai* Z 15.11 'completely ... his'; *ggaṃdyo + -i >*  
*\*ggaṃd(i)yvai* (\**CCyV* disallowed?) > *ggaṃdyai* Z 2.100 'a  
gong ... his'; *stuno + -i > stunai* Z 22.159 'pillar ... to him'.

The patterns allow some observations. The processes *a+i>ai*, *ä+i>ī* and *u+i>uī*, *vī*,  
are identical with nominal contractions. That the last is phonemically /u+i>uī/ may  
be demonstrated with the meter of Z. The words *pharvī*, *aśvī*, *biśvī*, and *aysvī* all  
count three moras, weigh LH, and syllabify /*(C)V·Cui*/.<sup>73</sup> The clearest example may  
be in the hemistich below where *biśvī* is LH in cadence 3 **LHLHL**:

17.24ab    **biśvī** haysāndu □ | mara paḍā | āsta pharu (C:7+5+5)

everything is known to him. Once there were many here

In the next three hemistichs, *pharvī*, *aśvī*, and *aysvī* all count three moras in a five  
mora segment:

22.247ab    **pharvī** nuva | hvaṃ'ndā aysura □ | nāta yakṣa | dīvate gyasta

(A:5+7+5+7)

Behind him there will be many men, Asuras, Nāgas, Yakṣas, deities, gods,

5.31ab    **aśvī** uys|malstu hataiṣṭāndā □ | maulu | ysarrno baste . (A1:5+9+3+7)

They gave him a well-groomed horse. He bound on a golden diadem,

2.52ab    **aysvī** vā | ttīyā hamī □ hīmā | thatau | maṃdru pātcu (A1:5+9+3+6)

<sup>73</sup> *harbiśvī* Z 16.5 should be HLH /*haabi·ṣui*/. However, it is the final word in a type A  
hemistich and makes no sense metrically or semantically.



then I will quickly change it afterwards by a spell

The process  $e+i>ei$ , if known to Leumann, might have been what prompted his transcription. But the older form is probably  $e+i>ai$ , with the  $ei$  spellings possible after the diphthongs merged. This process also suggests that  $/\bar{e}/$  is phonetically lower than mid, possibly even low. While the phonemic process  $/\bar{e}+i>aj/$  involves disassimilation (mid+high>low+high), phonetically it may not  $[\text{æ}+i>aj]$ .

The last process was perhaps originally  $o+i>vai$  (khvai 9×, khvei 7×). It is similar to  $e+i>ai$  but retains a rounded element.  $khv\bar{i}$  may reflect a stage where  $kho$  is  $khu$ , and so  $u+i>v\bar{i}$ .

The one instance of  $-\bar{u} + -i$  in *aysmū-t-ī* Z 24.196 'her thought' is clearly Late Khotanese. From the glossary in SuvII:245b there are two cases of  $-v\bar{i}$  from the LKh manuscript P and three of  $-t\bar{i}$  from manuscript Q which Skjærvø describes as "slightly archaizing MKh [Middle Khotanese]" (SuvII:64) but has heavy LKh influence such as *hime* for 3Sp.m *hämätä* 'be' Suv[Q] 6.2.65

### 3.6.2 3S enclitic pronoun -yā

In complementary distribution with  $-i$  is the enclitic  $-yā$  of the same meaning which occurs almost always after long vowels:

$-e + yā > -e-yā$       *rre-yā* Z 5.44 'king ... his'; *kīnthe +ī > kīnthe-yā* (*Lessons* 11.3).

$-o + yā > -o-yā$       *o-yā* Z 4.114 'or ... one'; *pāto'-yā* Z 12.9 'at his feet' (cf. *pau'* *ye* Z 5.105 'id.');

$-\bar{a} + yā > -\bar{a}-yā$       *śśaṃdā-yā* Z 2.125 'the ground ... by him'; *bitamā-yi* Z 24.230 'doubt in him'; *hā-yā* Z 5.46 'up at him'.

-ī + yä >	-ī-yä	<i>salī-yi</i> Z 14.30 ‘years ... his’; <i>uvī-yä</i> Z 3.30 ‘one’s senses’;
-ū + yä >	-ū-yä	I find no examples. The behavior is listed in <i>Lessons</i> 11.3 without examples.
-ai + yä >	-ai-yä	<i>ṣṣai-yä</i> Z 5.27 ‘even for him’
-au+ yä >	-au-yä	<i>hävvyau-yä</i> Z 14.56 ‘through (our) own ... him’; <i>haudyau-yä</i> Z 22.135 ‘with seven ... it’s’; <i>hvatanaui</i> Z 23.2 ‘into Khotanese ... it’; <i>āljsanyau-yä</i> Z 24.243 ‘with songs ... him’.
-ei + yä >	*-ei-yä	(no examples known to me).

After short vowels in Z I notice the 3S -yä only three times: *śś<ä>ru-yä* 24.655 and *biśśu-yä* 24.657 on the same broken leaf; and, *jāṅga-tä* Z 13.131 ‘a disappearance of him’. There is also one exception in Śgs: *kho-yi* 3.2v2.

### 3.6.3 Another parallel, *e+u>yau* and *o+i>vai*

The 2P enclitic pronoun -u suffixes to short *a*, *u*, *ä* in predictable fashion:

<i>a+u&gt;au</i>	NAP <i>hambadau</i> ‘fulfilled ... your’ Z 5.23
<i>u+u&gt;ū</i>	<i>aysū</i> ‘I ... you’ Z 23.18
<i>ä+u&gt;ū</i>	1Spf.tr <i>hvataimū</i> ‘I have told you’ 22.332

But when it comes to phonetically short final -e, one outcome appears to be *e+u>yau*. This is best seen with *ne* ‘not’ + -u > *nyau* Z 23.95, 24.474, *nyo* 24.263. The other outcome *e+u>au* is seen clearly in *käḍau jsa* ‘very ... by you’ Z 22.233. The construction *kyau* ‘that ... your’ Z 3.15 +4×, *cau* ‘what ... your’ Z 15.10 +1× might be either outcome as there may not be an effective way to write /kʸj̥äʷ/.

The enclisis *e+u>yau* appears to be parallel to the *o+i>vai* illustrated above with *khvai* and *puṣsvai*. Phonemically these are /ē+u>j̥äʷ/ and /ō+i>ʷaj̥/. In both cases, a

mid vowel from one side of the quadrilateral meets a high vowel from the other, and the result is a triphthong with the first element or on-glide matching its source vowel in palatality and rounding, and its last element or off-glide matching the enclitic in palatality and rounding. This symmetry itself ought to lend credence to the idea of triphthongs existing in Old Khotanese.

## Chapter 4: Stem and Suffix Metanalysis

### 4.1 Morphological Reanalysis

The simplest kind of morphological reanalysis is when discrete words with no internal morphological structure are assigned a structure by speakers. This is common with words of foreign origin. The English word *hamburger* was originally borrowed as an adjective as in *Hamburger sandwich*, referring to a German city. Partly because of the word *ham*, and perhaps partly because of shortening, speakers separated out the syllables *burger* as an independent noun, and this became the basis of nouns for related types of prepared food, such as *cheeseburger*, *chickenburger*, or *soy burger*. There is even *beefburger*, which demonstrates loss of the knowledge that a *Hamburger sandwich* was always made from *beef*. *Cranberry* (< Low German *kranebeere* ‘crane-berry’) has given rise to *cran-* as an English morpheme as in *cranapple* or *cranraspberry* juice. The name of the Pacific island group *Bikini* became the name of a two piece bathing suit, which was reanalysed into *bi·kini* allowing formation of *monokini*, a term for a one piece suit.

Inuktitut *inuk* ‘person’ has plural *inuit* ‘people’ which is used as an adjective in English, as in *inuit art*, *inuit carving*. From that use, English (and French) speakers have created a false plural, *inuits*. A reverse process is the currently acceptable use of *data*, originally plural of *datum*, as a singular mass noun. This can also be pluralized to *datas*.

The type of morphophonological reanalysis involving the shift in a morpheme boundary is called metanalysis. The usual example for this phenomenon is two

related shifts between Middle and Modern English. Middle English *a naddre, a napron, a noumpere, a nauger* became Modern *an adder, an apron, an umpire, an auger*. The reverse happened when Middle *an ewt, an ekename* became Modern *a newt, a nickname*. French *licorne* ‘unicorn’ is derived *unicorne > une icorne > l’icorne > licorne*. Arabic *nāranj* was borrowed into Old French and more or less evolved as *norenge > une norenge > une orange* and then borrowed into English.

A shift in a morpheme boundary can create a new stem shape which propagates throughout a paradigm. The Persian present stem *hast-* ‘be’ was developed through the reanalysis of the historical third singular as the zero or base form.

**Table 4.1:** Persian metanalysis

	Proto-Iranian	Avestan	Middle Persian <sup>74</sup>	Persian
1S	* <i>as-mi</i>	<i>ah-mi</i>	<i>hēm</i>	<i>hast-am</i>
2S	* <i>as-i</i>	<i>a-hi</i>	<i>hē</i>	<i>hast-i</i>
3S	* <i>as-ti</i>	<i>as-ti</i>	<i>ast</i>	<i>hast</i>

In Middle Persian, this common present has been remodeled and most of the paradigm begins with *h-* (cf. also 1P *hōm*, 2P *hēd*, 3P *hēnd*). The 3S *ast* is no longer synchronically analyzable as stem+suffix. In the evolution to New Persian, the *h-* spreads from the other forms in the paradigm, *ast > hast*, and the new 3S shape *hast* becomes the new stem *hast-*. The new stem shape is also reflected in Persian 1P *hast-im*, 2P *hast-id*, and 3P *hast-and*. In this evolution the morpheme boundary shifted to the right from Proto-Iranian \**s·t* to Persian *st-*.

---

<sup>74</sup> M. Maggi brought the importance of Middle Persian in the process to my attention (Middle Persian forms from Durkin-Meisterernst 2014:233-234 § 475).

Another metanalyzed stem shape is reflected in the Persian present stem *kon-* ‘make, do’. From a root *\*kar-*, Old Iranian had a *nu-*present stem seen in Avestan 3S *kərənaoiti* (cf. Vedic *kṛnóti*). This stem was remodeled in Old Persian as seen in 3S *kunautiy* <ku-u-n-u-ṭ-iy>. Middle Persian has the present stem *kwn-*. In the historical process, the initial *\*n* of the suffix was reinterpreted as the final consonant in the stem.<sup>75</sup>

A shift in a morpheme boundary can also reshape a suffix. G. E. Dunkel provides an example from Sanskrit: “Shift of morpheme boundary created the Sanskrit *-iṣ-*aorist: *tāri-ṣ-am* > *tār-iṣ-am*; thence e.g. *śaṃs-iṣ-am*” (1987:26).

The Old Khotanese metanalysis presented here involves the development of new shapes for both a range of present stems and a set of suffixes. A series of historically consonant final stems has been reanalyzed as vowel final, while simultaneously a series of historically *\*/t/-*initial suffixes was reanalyzed as */tt/-*initial. The morpheme boundary shifted to the left from *\*/Vt·tV/* to */V·ttV/*. This reanalysis has been previously overlooked because the data are complex. It is necessary to prove or explain several interlocking or interdependent points. Many of these have been detailed in the five articles summarized above (see 1.0).

## 4.2 The two sides of the morpheme boundary

The proof that the synchronic analysis of *butte* is stem *bu-* + suffix *-tte* has two components. It is necessary to show that the stem *bu-* behaves like a vowel final stem with no need to presuppose an abstract stem final *-d-*. Further, from the point

---

<sup>75</sup> This example was suggested by M. Maggi. Cf. Mayrhofer 1992:307–308, Hoffman and Forssman 2004:213–214.

of view of the synchronic description, there are advantages to a shape *bu-* and disadvantages to a shape *\*bud-*. At the same time it is necessary to prove that the metanalyzed suffixes now begin with double /tt/. The two components are interlocked and it is challenging to describe one without involving the other and to avoid circularity. The current approach has several steps:

- Identifying all the stems that exhibit conjugated forms containing *VttV*.
- Identifying all of the non *VttV* forms, both conjugational and derivational, based on these stems. These exhibit contraction of stem final vowel with suffix initial vowel.
- Comparing the contractions of the *VttV* exhibiting stems with other contractions exhibited by the language and showing there is no synchronic trace of a historical final consonant with the contractions of the *VttV* exhibiting stems.
- Giving evidence that the suffixes in the *VttV* conjugations begin with double /tt/.
- Giving evidence from other suffixes that shows that the *VttV* exhibiting stems are vowel final.
- Describing the historical developments which led to the metanalysis.

Before outlining those steps, it is useful to review the position of Skjærvø relating to the *VttV* exhibiting stems from a decade ago.

### **4.3 Skjærvø 2004**

The stems involved in the metanalysis have been listed by most previous scholarship as ending in *-d-* or *-h-*. For instance, the Leumanns (Glossar), Emmerick

(SGS) and Bailey (DKS) list *bu-* as *bud-* ‘know’ and *hva-* as *hvah-* ‘strike’. This is largely for historical reasons. But scholars have also seen traces of a final *-d-* or *-h-* in some conjugated forms such as in the *-v-* in 3Pp.m *buvāre* ‘they know’ or the *-ai-* in 3Sp.a *hvaittä* ‘he strikes’. Skjærvø is the exception. He recognizes that there is more to be understood about these stems.

I have departed from Emmerick’s synchronic-diachronic verb stem system in not positing stem forms that do not appear in actual forms. Instead, verbal stems are usually given as they appear in the verbal nouns in *-āmatā-* (*-matā-*) or the 2P imper. Thus, for instance, no stems end in *h* or *d* (which become *y* or nothing in Khotanese depending on the phonetic context), and Emmerick’s *hvah-* (< \**h<sub>u</sub>ahaja-*), *sad-* (< \**sadaja-*), and *bud-* are given as *hvay-*, *say-*, and *buv-*. The 3S pres. ind. forms can be “generated” from these stems by rules such as *y + t > itt* (*’itt*) (*hvaittä*, *saittä*) and *v + t > (u)tt* (*butte*). (SuvII:229).

Skjærvø’s approach is a step in the right direction. It marks a break in tradition by attempting to give synchronically underlying stem forms and suggesting rules to derive the actually attested forms. He proposes stems *hvay-* for *hva-* (SGS *hvah-*), *say-* for *se-* (SGS *sad-*) and *buv-* for *bu-* (SGS *bud-*). In his glossary he also gives a range of stems ending in *-y-* including *guhay-*, *pva’y-*, *bva’y-* for *gguha-*, *puṣa-* and *buṣa-* (SGS *gguhād-*, *puva’d-* and *buva’d-*).

A historical stem final *\*d* is certainly reflected in stem final *y* in a particular set of stems. These are all type A verbs in *-āy-* or *-ūy-*. The examples below show conjugated forms with *-y-*. All Proto-Iranian roots and prefixes are from EDIV.

*t(tu)vāy-* < \**ati-H<sub>u</sub>ad-* ‘convey across’ 3Sp.a *tvāyātā* Z 6.19, 3Pp.a *ttuvāyīndā* Z

1.187

*t(t)rāy-* < \**ati-(H)raHd-* ‘rescue, deliver’ 3Sp.a *trāyāte* Z 12.4

*nāṣāy-* < \**niž-HaH(a)d-* ‘establish; appoint’ 3Sp *nātā’yātā* IOL Khot 155/3r3



*pajāy-* < \**pati-ǰad-* ‘ask for, beg’, 3Pp.m *pajāyāre* Z 3.86 (type A assumed)

*pātāy-* < \**pati-HaH(a)d-* ‘speak’ 3Sp.a *pātāyātā* IOL Khot 155/3.r3

*bāy-* < \**Hud-* ‘lead’, 3Sp.a *bāyātā* Suv[Or] 6.3.32

*rrūy-* < \**raud-* ‘lose’ 3Sp.a *rrūyāte* Z 2.29, 3Pp.a *rrūyīndā* Z 12.40

In my analysis, other than type A in *-āy-* or *-ūy-* descending from roots in final \**d*, there are no inherited stems in Old Khotanese ending in *-y*.<sup>76</sup> In other words, there are none from Emmerick’s other classes, B, C, or D.

Skjærvø also proposes there was a final synchronic *-y-* in two stems which other scholars have regarded as ending in *-h-*. *hvay-* for *hva-* (SGS *hvah-*) has been mentioned above. In the glossary to Suv he also lists *hamiy-* (SuvII:362a) for *hamā-* (SGS *hamäh-*) ‘change (intr)’. There are stems ending synchronically in *-h-* in Old

---

<sup>76</sup> Previous scholarship gives two type A verbs in *-āy-* or *-ūy-* which do not reflect a historical stem final \**d*. One, now meaning ‘appear’ (semantics Emmerick St.I:24–25) is listed in SGS as *āy-* ‘be reflected’ A, in DKS s.v. *āy-* ‘see’, and in Glossar as *āy-* ‘sich spiegeln’. The other is listed in SGS as *pūy-* ‘look’, in DKS as *pūy-* ‘observe, watch over’, and in Glossar as *pūy-* ‘beobachten’. Both stems derive from OIr \**daiH-* ‘look, see’ with prefixes (EDIV:48–49). It is possible that the stem shapes are vowel final *āyā-* and *pūyā-* which would give diphthongization /*ě+ǰ>ǰV*/ or semivocalization /*ě+ā>yā*/ in the conjugated forms:

*āyā-* < \**ā-daiH-* ‘appear’

3Sp.m *āyāte* Z 4.100, 101, 6.52, 9.4, 13.99, 24.489 • /*āyǰǰēdē* < *āyě-* + *-ēdē*/

3Pp.m *āyāre* Z 3.90, 7.45 • /*āyyārē* < *āyě-* + *-ārē*/

*pūyā-* < \**pati-* or \**apa-daiH-* ‘observe’

3Sp.m *pūyāte* Z 22.249 • /*pūyǰǰēdē* < *pūyě-* + *-ēdē*/

*aṃcā-*p.ptc NAPf *pūyaṃce* Z 20.20 • /*pūyǰāncē* < *pūyě-* + *-ančē*/

Cheung also suggests the stem shape *āyā-*, without explanation, but retains *pūy-* (EDIV:49). The nouns *āyāna-* ‘mirror’ and *āyānaa-* ‘example’ (discussed below) might also reflect *āyā-*. Double *yy* may be written as in IAPm *indriyyau* Suv[Or] 5.9 but may be simplified as in IAPm *indriyo* Z 14.75 where the meter shows an underlying /*yy*/ (Hitch 2016a).

Khotanese, but like the stems in *-y-*, these are restricted to type A<sup>77</sup>, and there are none in Emmerick's B, C, or D.

There is no compelling evidence to see a Pre-Khotanese stem final *\*d* or *\*h* becoming synchronic *y* in any stem. Skjærvø's approach to these stems marks progress in the field but we need to go a little further.

#### 4.4 Conjugated present stems showing *VttV*

Two of the stems attested with *VttV*, *pat-* 'fall' and *hot-* 'be able' were not affected by the metanalysis. The *VttV* forms known are 3Sp.a *pīttā*, 3Sp.m *hotte*, 2Pp.a *patta*, 3Siv.m *hautto*, 3Sij.m *hautta*. These stems retain the stem final *-t-* when a vowel initial suffix is attached as for example in *pat-* + *-īnda* > 3Pp.a *patīndā* Z 4.59 'they fall' or *hot-* + *-āre* > 3Pp.m *hotāre* Z 13.36 'they are able'. They are also the only *VttV* stems reflecting a Proto-Iranian final *\*-t-*, that is, *\*pat-* and *\*fra-uat-*. The other stems with *VttV*, as shown below, reflect a Proto-Iranian final *\*d*, *\*y*, or a laryngeal. The stems *pat-* and *hot-* do not need further examination.

The chart below lists all of the other attested forms of present stems showing orthographic *tt* between vowels. Only one spelling is listed for each attested stem and suffix combination. These combinations are known for 3Sp.a, 3Sp.m, 3Siv.a, 2Piv.a and 3Sij.m. They are predicted for 2Pp.a, 3Siv.m and 3Sij.a. The leftmost column organizes inflected forms by suffix as well as by Emmerick's distinction

---

<sup>77</sup> I notice the following synchronic stems in *-h-* in Old Khotanese: *ārīh-* 'share', *ārūh-* 'move, shake', *eh-* 'reach', *ggañih-* 'moisten', *jeh-* 'cleanse', *dr(r)āh-* 'fly (up)', *patārah-* 'support oneself (on)', *pareh-* 'restrain oneself', *parmīh-* 'damage, destroy', *bārāh-* 'soar up', *byeh-* 'obtain', *ysah-* 'cease', *ysānah-* 'bathe', *yseh-* 'make cease', *rūh-* 'attack', *hambrīh-* 'share', *hamīh-* 'change (tr)'.

between type B and C conjugations. The second column shows the synchronic stem shapes proposed here. The remaining columns list the stem shapes given by Emmerick (SGS), Bailey (DKS), the Leumanns (Glossar) and Skjærvø (SuvII).

**Table 4.2:** All known *VttV* categories with vowel stems

	Here	SGS	DKS	Glossar	Suv
<b>3Sp.a -<i>ttä</i> B</b>					
<i>nättä</i>	<i>nä</i> - ‘sit down’	<i>näd-</i>	<i>näša’d-</i>	<i>nä-ta’d-</i>	<i>näy-</i>
<i>pajättä</i>	<i>pajā-</i> ‘ask for’	<i>pajäd-</i>	<i>pajad-</i>	<i>pa-jät-</i>	—
<i>bitte</i>	<i>bi-</i> ‘shoot’	<i>bid-</i>	<i>bid-</i>	<i>bid-</i>	—
<i>bei’ttä</i>	<i>bäšä</i> - <sup>78</sup> ‘release’	—	( <i>bei’ttä</i> )	( <i>b[iš]eittä</i> )	
<i>rrvittä</i>	<i>rru-</i> ‘grow (intr)’	<i>rrūd-</i>	<i>rruv-, rrv-</i>	<i>rrud-</i>	—
<i>haṃbruittä</i>	<i>haṃbru-</i> ‘grow together’	<i>haṃbrūd-</i>	( <i>haṃbruittä</i> )	<i>haṃbrud-</i>	—
<b>3Sp.a -<i>ttä</i> C</b>					
<i>khaittä</i>	<i>kha-</i> ‘wound’	<i>khad-</i>	( <i>khasta-</i> )	<i>khad-</i>	—
<i>ttähvaittä</i>	<i>ttähva-</i> ‘cross (a river)’	<i>ttähvah-</i>	<i>ttähvah-</i>	<i>ttä-hvad-</i>	—
<i>*[pä]hai[ttä]</i>	<i>päha-</i> ‘strike’	<i>pähad-</i>	( <i>khasta-</i> )	<i>pähad-</i>	—
<i>pvai’ttä</i>	<i>puša-</i> ‘fear’	<i>puva’d-</i>	<i>puva’d-, pva’d-</i>	<i>puva’d-, pvad-, puva’y-puv’-</i>	
<i>baittä</i>	<i>ba-</i> ‘be bound’	<i>bad-</i>	<i>bad-</i>	<i>bad-</i>	—
<i>bvai’ttä</i>	<i>buša-</i> ‘mount’	<i>buva’d-</i>	<i>buva’d-</i>	<i>buva’d-</i>	<i>bva’y-</i>
<i>maitti</i>	<i>ma-</i> ‘be intoxicated’	<i>mad-</i>	<i>mad-</i>		<i>may-</i>
<i>saittä</i>	<i>se-</i> ‘appear’	<i>sad-</i>		<i>sad-, sai-</i>	<i>say-</i>
<i>hvaittä</i>	<i>hva-</i> ‘strike’	<i>hvah-</i>	<i>hvah-</i>	<i>hvad-</i>	<i>*hvay-</i>
<b>3Sp.m -<i>tte</i> B</b>					
<i>dätte</i>	<i>dä-</i> ‘appear’	<i>did-</i>	( <i>dai:-di-</i> )	<i>did- (däd-), diy-</i>	<i>diy-</i>
<i>nuvatte</i>	<i>nuva-</i> ‘lie down’	<i>nuvad-</i>	<i>nuvad-</i>	<i>nu-vat-</i>	—
<i>nütte</i>	<i>nü-</i> ‘lie down’	<i>nuvad-</i>	<i>nuvad-</i>	<i>nu-vat-</i>	—
<i>pajuttä</i>	<i>paju-</i> ‘cover’	<i>pajud-</i>	<i>pajud-</i>	<i>pa-jut-</i>	—
<i>butte</i>	<i>bu-</i> ‘know’	<i>bud-</i>	<i>buv-, bv-</i>	<i>bud-, buv-</i>	<i>buv-, bv-</i>
<i>bütte</i>	<i>bü-</i> ‘be fragrant’	<i>būd-</i>	( <i>bütte</i> )	<i>būd-</i>	—
<i>hamatte</i>	<i>hama-</i> ‘sober up’	<i>hamad-</i> 1967:46	( <i>hamatte</i> )	—	( <i>ha</i> ) <i>may-</i>
<i>hamättä</i>	<i>hamä-</i> ‘change (intr)’	<i>hamäh-</i>	<i>hamih-, hamy-</i>	<i>ha-mäh-</i>	<i>hamiy-</i>
<b>3Siv.a -<i>ttu</i> C</b>					
<i>pva’ttu</i>	<i>puša-</i> (see above)				
<b>2Piv.a -<i>tta</i> B</b>					
<i>nätta</i>	<i>nä-</i> (see above)				
<i>haṃbitta</i>	<i>haṃbi-</i> ‘pierce’	<i>haṃbid-</i>	( <i>haṃbitta</i> )	<i>haṃ-bid-</i>	—
<b>2Piv.a -<i>tta</i> C</b>					
<i>pähatta</i>	<i>päha-</i> (see above)				
<i>puva’tta</i>	<i>puša-</i> (see above)				
<b>3Sij.m -<i>tta</i> B</b>					
<i>butta</i>	<i>bu-</i> (see above)				

<sup>78</sup> The proposed stem shape *bäšä-* ‘release’ is discussed in 4.6.6 below.

## 4.5 The vowel contractions exhibited by the *VttV* stems

The stems in Table 2 are organized below by final vowel, *-a-*, *-i-*, *-u-*, *-ū-*, *-ä-*, *-e-*. For each final vowel all known contractions are listed with all known examples. Then for each contraction evidence is given from elsewhere in the language, from both the nominal and verbal systems, showing that the process is usual.

### 4.5.1 Contractions with stems ending in *-a-*

Some *VttV* stems ending in *-a-* are not attested with a vowel initial suffix: *buṣa-*, *ma-*, *hva-*, *nuva-* ('lie down'), *hama-*. These are not examined below. This leaves *kha-*, *ttāhva-*, *pāha-*, *puṣa-*, *ba-* as *VttV* stems in final *-a-* which exhibit contraction behavior. To this group we may add for examination other likely type B stems in *-a-* which may have the same morphophonology even though a *VttV* form is not attested: *gguha-* 'wound' (SGS *gguhā-*), *nuva-* 'make a noise' (SGS *\*nuvad-*), *nei'hva-* 'cross over' (SGS *nei'hvah-*), and *hatca-* 'be broken' (SGS *hatcy-*). After listing below the known inflected forms for each contraction of these nine stems, some evidence for the same contractions is given from the type A stems in *-a-* presented in TypeA, and from nominals ending in /a/ presented in Hitch 2015b and Hitch 2016a.

#### 4.5.1.1 *a+a>ā* in *VttV*

3Ssj.a *khā* Z 13.81 < *kha-* 'wound' • /k<sup>h</sup>ā < k<sup>h</sup>a- + -a/

2Siv.a *pvā*' Sgh[17] 31.9, Sgh[17] 100.2 < *puṣa-* 'fear' • /buā < buṣā < buṣa- + -a/

#### type A stems in *-a-* evidence

Not attested

## nominal evidence

NSf of *aā*-declension nominals

*brīyā* 'love' Z 13.58 < *brīyāā*- 'love' • /βrīyā < βrīya- + -a/

VSm of *aa*-declension

*Maitrā* Z 3.12 < *Maitraa*- 'Maitreya' • /majrā < majtra- + -a/

NAPm of *aa*-declension

*āchā* Z 3.144 < *āchaa*- 'disease' • /āč<sup>h</sup>ā < āč<sup>h</sup>a- + -a/

### 4.5.1.2 *a+i>ai* in *VttV*

1Sp.a: -a- + -īmā > -aimā

*gguhaimā* Z 2.199 < *gguha*- 'wound' • /guhajmē < guha- + -īmē/

3Pp.a: -a- + īndā > -aindā

*ttāhvaindā* Z 13.20, 21, 24.277 < *ttāhva*- 'cross (a river)' • /tēh<sup>w</sup>ajndē < tēh<sup>w</sup>a- + -īndē/

*nuvaindā* Z 2.46, *nvaindā* Z 20.30 < *nuva*- 'make a noise' • /nuwajndē < nuwa- + -īndē/

*nei'hvaindā* IOL Khot 190/6a-b.v3 (Śgs) *nei'hva*- 'cross over' • /naēh<sup>w</sup>ajndē < nalh<sup>w</sup>a- + -īndē/

*pvai'ndā* Sgh[25] 241.2, 253[49], SI P 53.11+ v2, *pvai'ndi* Z 11.10, 20.19, K5 641.a2, *pvai'ndā* Z 22.118 < *puṣa*- 'fear' • /buaṅndē < buṣajndē < buṣa- + -īndē/

*baindā* Z 3.107, 22.273 < *ba*- 'be bound' • /βajndē < βa- + -īndē/

## type A stems in -a- evidence

1Sp.a: -a- + -īmā > -aimā

*nimandrainā* Z 2.50 < *nāmandra-* ‘invite’ • /*němandra*ĩmě < *němandra-* + -  
ĩmě/

*p[ai]m[ä]* Suv[Or] 11.14 < *pa-* ‘protect’ • /*ba*ĩmě < *ba-* + *ĩmě*/

*buysaimā* Z 20.23, 24 < *buysa-* ‘extinguish’ • /*βuzai*mā < *βuza-* + -*ĩmě*/

3Pp.a: -a- + ĩndä > -aindä

*nāmandraindä* Z 12.61 < *nāmandra-* ‘invite’ • /*němandra*ĩndě < *němandra-* + -  
ĩndě/

*paindä* Z 16.5 < *pa-* ‘protect’ • /*ba*ĩndě < *ba-* + -*ĩndě*/

3Pop.a: -a- + ĩru > -airu

*buysairu* Z 4.63 < *buysa-* ‘extinguish’ • /*βuzai*ru < *βuza-* + -*ĩru*/

## nominal evidence

denominal adjective suffixes -ĩña- (KhSuf:129b, 131b) and -ĩnaa- (KhSuf:135)

NAPm *muḍaiña* Z 2.48 ‘of the dead’ < *muḍaa-* ‘corpse’ + -*ĩña-* + -*a* • /*muḍai*ña  
< *muḍa-* + -*ĩñ-* + -*a*/

NSm *baśdainei* Sgh[1,17] 36.3-4 ‘of sin’ < *baśdaā-* ‘sin’ + -*ĩnaa-* + -*ä* •  
/βaždaiñě < βažda- + -*ĩna-* + -*ě*/

### 4.5.1.3 a+ä>ei>ai in VttV

3Sop.a *gguhei* Sgh[1] 36.3, *gguhαι* Sgh[17] 36.3 < *gguha-* ‘wound’ • /*guha*<sup>e</sup> <  
*guha-* + -*ě*/

3Sop.a *pihei* Sgh[1] 39.1, *pāhai* Sgh[17] 39.1, Sgh[11] 211.4 < *pāha-* ‘strike’ •

/bēha<sup>e</sup> < bēha- + -ě/

3Sop.m *hatcai* Śgs 3.13v3 < *hatca-* ‘break’ • /hat<sup>s</sup>aḡ < hat<sup>s</sup>a- + -ě/<sup>79</sup>

### type A stems in -a- evidence

3Sp.a: -a- + -ätä > \*-eitā, -aitā, -aiyā

*ttātsaiyi* Z 13.27 < *ttātsa-* ‘cross (a river)’ • /tēt<sup>s</sup>aiḡē < tēt<sup>s</sup>aḡē < tēt<sup>s</sup>aḡē <

tēt<sup>s</sup>a- + -ēdē/

*nāmaṃdraiyā* Z 24.465 < *nāmandra-* ‘invite’ • /nēmandraiḡē < nēmandraḡē <

nēmandraḡē < nēmandra- + -ēdē/

*paṭhaiyā* Z 2.175 < *paṭha-* ‘burn’ • /baṭ<sup>h</sup>aiḡē < baṭ<sup>h</sup>aḡē < baṭ<sup>h</sup>aḡē < baṭ<sup>h</sup>a- + -

ēdē/

3Pp.a: -a- + -ändā > -eindā

*nuveindi* Z 24.420 < *nuva-* ‘make a noise’ • /nuwaḡendē < nuwa- + -ēndē/

3Sp.m: -a- + -äte > \*-eite, -eiye, -eiyā, -aite, -aiye

*naltseye* Z 22.202, *naltsaiye* Z 24.252, *naltseyā* Z 24.405 < *naltsa-* ‘go out’ •

/nalt<sup>sh</sup>aiḡē < nalt<sup>sh</sup>aḡē < nalt<sup>sh</sup>aḡē < nalt<sup>sh</sup>a- + -ēdē/

*spaiye* Or. 12637/71.1, *spaiye* Z 2.201, 20.64, IOL Khot 155/3, *spaiyā* Z 19.18 <

*spa-* ‘self satisfy’ • /sbaiḡē < sbeiḡē < sbaḡē < sba- + -ēdē/

3Sop.a: -a- + -ā > \*-ei, -ai

*buysai* Z 20.24 < *buysa-* ‘extinguish’ • /βuzaḡ < βuza- + -ě/

---

<sup>79</sup> M. Maggi suspects a copyist error here for *hatcai*<*ya*> at the end of a clause and before a punctuation mark. \**hatcaiya* would reflect *hatca-* + *-īya*, the regular 3Sop.m ending.



2Piv.a: *-a- + -äta > \*-eita, -aiya*

*paiya* Z 22.277 < *pa-* ‘protect’ • /*bai̯ʔa* < *bae̯ʔa* < *bae̯da* < *ba-* + *-ëda*/

3Sij.m: *-a- + -äta > \*-eita, -eiya -aiya*

*vatseiya* Neb 50.30 < *vatsa-* ‘descend’ • /*wat<sup>s</sup>ai̯ʔa* < *wat<sup>s</sup>ae̯ʔa* < *wat<sup>s</sup>ae̯da* < *wat<sup>s</sup>a-* + *-ëda*/

ä-infinitive: *-a- + -ä > \*-ei, -ai*

*buysai* ‘extinguish’ Sgh[10] 205.1 < *buysa-* ‘extinguish’ • /*βuzaj̯* < *βuzae̯* < *βuza-* + *-ë*/

#### nominal evidence

*aa*-decl NSm *a+ä>ei>ai*

*alysānei* ‘prince’ Šgs 2.1r4 < *alysānaa-* + *-ä* • /*alzānae̯* < *alzāna-* + *-ë*/

*pīsai* ‘teacher’ Z 11.42 < *pīsaa-* + *-āna* • /*bīsaj̯* < *bīsae̯* < *bīsa-* + *-ë*/

*aa*-decl IASm *a+āna>eina>aina*

*vineina* Z 13.106 < *vinaa-* ‘vinaya’ • /*winae̯na* < *wina-* + *-ëna*/

*rräjseina* ‘sharp’ Sgh[1] 39.1 < *rräjsaa-* + *-āna* • /*rëd<sup>z</sup>ae̯na* < *rëd<sup>z</sup>a-* + *-ëna*/

*āstaina* ‘bone’ Z 5.8 < *āstaa-* + *-āna* • /*āstaj̯na* < *āstae̯na* < *āsta-* + *-ëna*/

#### 4.5.1.4 *a+ā>ā* in *VttV*

*āmāta*-noun

NSf *pähāmata* Z 7.22 < *pāha-* ‘strike’ • /*bëhāmada* < *bëha-* + *-āmad-* + *-a*/

*āka*-agent noun

NAPm *pihāka* '(tree) choppers'<sup>80</sup> Sgh[11] 210.4 < *pāha*- 'strike' • /*bēhāka* <

*bēha*- + *-āk*- + *-a*/

*āñā*-p.nec:

NSm *puvā'ñi* Z 18.35 < *puṣa*- 'fear' • /*buāñē* < *buzāñē* < *buṣa*- + *-āñ*- + *-ē*/

ASm *pvā'ñu* Z 6.36 < *puṣa*- 'fear' • /*buāñu* < *buzāñu* < *buṣa*- + *-āñ*- + *-u*/

### type A stems in *-a*- evidence

1Pp.a *a+āmā>āmā*

*pāmā* Suv[Or] 6.1.39, [Or]11.9 (*-mī*, + *-ī*) < *pa*- 'protect' • /*bāmē* < *ba*- + *āmē*/.

3Pp.m *a+āre>āre*

*vatsāre* IOL Kh 159/7r2 < *vatsa*- 'descend' • /*wat<sup>s</sup>ārē* < *wat<sup>s</sup>a*- + *-ārē*/

3Psj.a *a+āro>āro*

*numāṇdrāro* IOL Khot 147/3v2 Neb:107–109 < *nāmandra*- 'invite' •

/*nēmandrāro* < *nēmandra*- + *-āru*/

p.nec *a+āñ>āñ*

NSm *paṭhāñā* Suv[Or] 8.23 < *paṭha*- 'burn' • /*baṭ<sup>h</sup>āñē* < *baṭ<sup>h</sup>a*- + *-āñē*/

---

<sup>80</sup> Previous scholarship regards *pihāka* as an attributive adjective or participle. For instance, Degener translates once as participle, *rrājsā pihāka paḍe* "scharfe schneidende Äxte" (KhSuf:43–44), and once as an adjective, *rrājsā pihāka paḍe ājumāta* "bringt scharfe Äxte zum Schneiden!" (KhSuf:49). Perhaps most recently, Canevascini gives "fetch sharp, felling axes" (Sgh:86, 152). But there are many cases of the appositional use of the *āka*-agent noun ("Sehr oft stehen nom. ag. auf *-āka*- in Apposition," KhSuf:43). This passage can also be interpreted as appositional use: 'fetch sharp (tree) choppers, axes.'

## nominal evidence

GDPm *aā*-decl *a+ānu>ānu*

*āchānu* Z 24.222 < *āchaa*- ‘disease’ • /*āč<sup>h</sup>ānu* < *āč<sup>h</sup>a*- + -*ānu*/

*ṣṣāvānu* Z 3.120 < *ṣṣāvaa*- ‘śrāvaka’ • /*ṣṣāvānu* < *ṣṣāwa*- + -*ānu*/

*balysūñāvūysānu* Śgs 3.6v1 < *balysūñāvūysaa*- ‘Bodhisattva’ •

/βalzūñawūzānu < βalzūñawūza- + -*ānu*/

### 4.5.1.5 *a+ū>∅+ū*<sup>81</sup> in *VttV*

*ūna*-noun

IASm *pihyūnāna*<sup>82</sup> < \**pāhūnāna* ‘with a blow’ Sgh[1] 39.1 < *pāha*- ‘strike’ •

/bēhyūnēna < \*bēhūnēna < *bēha*- + *ūn*- + -*ēna*/

## type A stems in *-a*- evidence

ASf *numadrūṇo* Z 2.60, *nimaṃdrūṇo* Z 22.97 < *nāmandra*- ‘invite’ •

/nēmandrūnō < *nēmandra*- +-*ūn*- + -*ō*/

### 4.5.1.6 *a+∅>a*

2Sij.a<sup>83</sup> *puva*’ Z 2.102, 24.263 < *puṣa*- ‘fear’ • /*bua* < *buṣa* < *buṣa*- + -*∅*/

---

<sup>81</sup> There is very little evidence for contractions of the type *V+ū*. As *a+i>ai*, one might expect *a+ū>au*. Instead, on the basis of the little evidence given here, it would appear that suffix initial long /ū/ deletes a stem final /a/ in the same way that the oblique plural suffixes delete a stem final /a/ on nominal stems (Hitch 2016a).

<sup>82</sup> This sole example appears to show LKh influence, cf. LKh *pva’styūna*- ‘fearsome’ < *pva’sta*- ‘feared’ (KhSuf:170–171)

<sup>83</sup> A 2Sij.a category is not generally recognized. E. and M. Leumann listed *puva*’ as imperative, but suggested, “(inj.?)” (Glossar:465a). It is attested twice with prohibitive *ma* where injunctive might be expected. Since all OKh words must end in a vowel, the consonant final stems take the 2Siv.a *-a* and 2Siv.m *-u* suffixes. A proposed injunctive zero

## 4.5.2 Contractions with stems ending in *-i-*

There are two *VttV* stems in *-i-*, *bi-* ‘shoot’ and *ham̐bi-* ‘pierce’. There are no examples of *ham̐bi-* with vowel initial suffixes. The stem *bi-* shows two contractions. Three other verb stems in *-i-* show the same contractions as *bi-*, *ji-* ‘cease’, *pari-* ‘order’, and *hāvi-*~*hīvi-* ‘appropriate’. These three were discussed in Hitch 2016a in the context of diphthongization and semivocalization.

### 4.5.2.1 *i+i>ī*

3Pp.a *bīndi* Z 24.408 < *bi-* ‘shoot’ • /βīndě < βi- + -īndě/

#### other *-i-* present stems

1Sp.a *parīmā* Z 12.51, 23.33 < *pari-* ‘order’ • /barīmě < bari- + -īmě/

3Pp.a *parīndā* Neb 158.29 (SI P3.5r1) < *pari-* ‘order’ • /barīndě < bari- + -īndě/

2Pp.m<sup>84</sup> *hīvirā* Suv[Or] 6.1.64, 6.2.44 (Skjærvø 2Pp.m) < *hīvi-* ‘appropriate’ •

/hīwīrě < hīwi- + -īru/

---

suffix *-∅* would only be visible with a vowel stem. The 2Siv.a *pvā*’ is a regular formation (4.5.1.1).

<sup>84</sup> In SGS, the 2P middle ending *-īru/-īrā* is considered to be optative, probably for historical reasons (203). At the same time Emmerick noted, “No certain example of the 2 pl. mid. has been found” (SGS:200). Skjærvø treats some cases of *-īru/-īrā* as 2Pp.m in the Glossary to 2004b. The suffix is not included in his list of conjugations (SuvII:230). That *-īru/-īrā* has spread from 2Pop.m to 2Pp.m is probably proven by a passage from Suv[Or] 6.2.44: *u trāyāta nu u hīvirā nu. u parhalāta nu. u anārra nā padīmīru. u tsāṣtatetu śśāratetu nā yanda.* ‘And you rescue them and appropriate them and protect them and make them guiltless and give them peace and well-being.’ The stems *trāy-*, *parhal-*, and *yan-* take the type A *-āta* or type B *-tta* active endings while the stems *hāvi-* and *padīm-* take the middle ending *-īru/-īrā*.

## nominal evidence

-īnaa- (m.), -īṃg<sup>y</sup>ā- (f.) denominal adjective

NSm *jaḍīnai* Z 4.92 ‘of ignorance’ < *jaḍia*- ‘ignorance’ • /jaḍīnai < g<sup>y</sup>ḍi- + -īna- + ě/

ASf *gyaḍīju* Z 2.173 ‘of ignorance’ < *jaḍia*- ‘ignorance’ • /jaḍīju < jaḍi + -īṃg<sup>y</sup>- + -o/

NSf *jaḍīgya* Z 7.45 ‘of ignorance’ < *jaḍia*- ‘ignorance’ • /jaḍīja < jaḍi + -īṃg<sup>y</sup>- + -a/

-īya- denominal adjective

NSm *narīyā* Z 4.62 ‘of hell’ < *naria*- ‘hell’ • /na.īyē < na.ī- + -īy- + -ě/

NSf *narīya* Z 4.113 ‘of hell’ < *naria*- ‘hell’ • /na.īya < na.ī- + -īy- + -a/

### 4.5.2.2 *i+ā>ī*

3Sop.a *bī* Z 21.15 • /βī < βi- + -ě/

### other -i- present stems

3Sp.m *jīte* IOL Khot 8/1v3, *jītā* Z 24.482, *jīye* Z 1.38 + 12x, Sgh[22] 43[6],

Suv[Or] 12.43, Kvb 0.41, *jīyā* Z 6.31, 13.127, 15.8, *jīyi* P45.2b4, 3a4 < *ji*- ‘cease’ • /jīṛē < jīdē < ji- + -ědē/

3Sp.m *hāvīye* Z 5.79, *hivīye* Z 3.138, 6.46, *hāvīyā* Z 3.138, 9.17, *hivīyā* Z 7.37 <

*hāvi*- ‘appropriate’ • /hēwīṛē < hēwīdē < hēwi- + -ědē/

2Sp.a *parī* Z 22.112, 23.114 < *pari*- ‘order’ • /barī < bari- + -<sup>i</sup>ě/

3Sp.a *parītā* Z 24.450, *parīyi* Z 24.442 < *pari*- ‘order’ • /barīṛē < barīdē < bari- + -ědē/

2Pp.a *parīya* SI P50r4, Z 12.28, *parīy[a]* IOL Kh 151/5b3 < *pari-* ‘order’ • /barīʔa

< barīda < bari- + -ēda/

### nominal evidence

*ia*-decl

N<sub>Sm</sub> *gyaḍi* Z 3.63 < *gyaḍia-* ‘ignorance’ • /g<sup>y</sup>aḍi < g<sup>y</sup>aḍi- + -ě/

I<sub>ASm</sub> *gyaḍīna* Z 5.56 < *gyaḍia-* ‘ignorance’ • /g<sup>y</sup>aḍīna < g<sup>y</sup>aḍi- + -ēna/

L<sub>Sm</sub> *narī*<sup>85</sup> Sgh[24] 203.4 < *naria-* ‘hell’ • /na.i < na.i- + -ě/

#### 4.5.3 Contractions with *VttV* stems ending in *-u-*

The *VttV* stems ending in *-u-* are *rru-* ‘grow (intr)’, *haṃbru-* ‘grow together’, *paju-* ‘cover’, and *bu-* ‘know’. As *bu-* is well attested, just a selection of forms from that verb from each category is given. The other *-u-* present stems compared here include three type A stems *ākṣu-* ‘begin’, *ju-* ‘live’, and *du-* ‘beat’ (SGS *ākṣū-*, *juv-*, \**dav-*) which were discussed in Hitch 2016a. The probably type B *ju-* ‘fight’ (SGS *juv-*) was also discussed in Hitch 2016a. To this group may be added the stem *hāru-* ‘grow (intr)’ (SGS *hārūd-*) as it is etymologically related to both *rru-* ‘grow (intr)’ and *haṃbru-* ‘grow together’.

---

<sup>85</sup> This analysis of *narī* was pointed out by an anonymous reviewer to Hitch 2015b. The L<sub>Sm</sub> ending *-ā* was common with the *a*-stems in Old Khotanese (SGS:262). Another instance with an *ia*-stem is with the adjective *anāstania-* ‘beginningless’ in the phrase *anāstanī saṃtsāri* ‘in the beginningless cycle of existence’ SI A 5r3 (SDTV3:239; Late OKh) where both adjective and noun show L<sub>Sm</sub> *-ā* (*-i*).

#### 4.5.3.1 *u+a>va>uva*

(diphthongization /ua/ and diphthong resolution /Cuwa/)

2Sp.m *bva* Sgh[8] 214[2], *buva* Suv[Or] 5.33, Z 3.9, 24.52 < *bu*- ‘know’ • /βuwa  
< βua < βu- + -a/

#### other -u- present stems

2Siv.a. *ākṣva* Z 2.100 < *ākṣu*- ‘begin’ • /aṭṣua < aṭṣu- + -a/

#### nominal evidence

NAPm of *ua*-declension nominals

*hajva* Z 23.35, *hajuva* Z 22.321 < *hajua*- ‘wise’ • /hajuwa < hajua < haju- + -a/

*bāysva* Suv[H] 4.11 < *bāysua*- ‘arm’ • /βāzua < βāzu- + -a/

*briya-ṣṣuva* Z 20.21 < *briya-ṣṣua*- ‘lovingly renowned’ (Hitch 2015b) • /briya-

ṣuwa < βriya-ṣua < βri- + -a- + ṣu- + -a/

*pajsama-ṣuva* Śgs 3.10r4 < *pajsama-ṣṣua*- ‘of honorable repute’ (Hitch 2015b)

• /bad<sup>z</sup>ama-ṣuwa < bad<sup>z</sup>ama-ṣua < bad<sup>z</sup>am- + -a- + ṣu- + -a/

NSf of *uā*-adjectives

*tcarṣuva* Z 22.155 < *tcarṣua*- ‘brilliant’ • /t<sup>s</sup>aṣuwa < t<sup>s</sup>aṣua < t<sup>s</sup>aṣu- + -a/

*haṃdara-ysaṃthva* Sgh[17] 42.1, Sgh[9] 42.2, *haṃdara-ysaṃthuva* Sgh[9]

42.1 < *haṃdara-ysaṃthua*- ‘of another birth’ • /handaṃa-zant<sup>h</sup>uwa <

handaṃa-zant<sup>h</sup>ua < handaṃ- + -a- + zant<sup>h</sup>u- + -a/

-attāti-de-adjectival nouns

NSf *hajvattātā* Z 11.42 < *hajvattāti*- ‘wisdom’ < *hajua*- ‘wise’ • /hajuatādē <

haju- + -atād- + -ē/

NSf *tcarṣvattātā* Z 23.40 < *tcarṣvattāti*- ‘splendor’ < *tcarṣua*- ‘brilliant’ •

/t<sup>s</sup>aɿṣuatādē < t<sup>s</sup>aɿṣu- + -atād- + -ě/

IASf *bārrā-ṣvattete* Suv[Or] 6.3.19 < *bārrā-ṣvattāti*- ‘manly fame, machismo’

(Hitch 2015b:317–318) < *bārrā-ṣṣua*- ‘of manly renown’ • / βērā-ṣuatēdē

< βērā-ṣuatād- +<sup>i</sup>ē < βērā-ṣu- + -atād- + -<sup>i</sup>ē/

#### 4.5.3.2 *u+ā>vā>uvā*

(semivocalization /wā/ and glide resolution /Cwā>Cuwā/ (Hitch 2016a))

1Pp.m *bvāmane* Sgh[17] 253[52], *buvā[mane]* Sgh[7] 199[4] < *bu*- ‘know’ •

/βuwāmanē < βwāmanē < βu- + -āmanē/

3Pp.m *bvāre* Z 2.67 + 33x, *buvāre* Z 2.121+ 7x, and many more < *bu*- ‘know’ •

/βuwārē < βwārē < βu- + -ārē/

1Ssj.m *buvāne* Z 2.216, Sgh[10] 213[2], *buvāni* Z 2.206, *bvāne* Z 5.113, 11.77,

12.1, 24.259, IOL Khot 30/8v3 < *bu*- ‘know’ • /βuwānē < βwānē < βu- + -  
ānē/

2Ssj.m *bvā* IOL Khot 189/5v3 (Śgs) < *bu*- ‘know’ • /βwā < βu- + -ā/

3Psj.m *bvāro* Z 11.7, Suv[SI<sup>1</sup>] 8.13, IOL Khot 30/8r2, *bvāru* Z 2.105, 11.76 <

*bu*- ‘know’ • /βwāro < βu- + -āro/

3Psj.m *bvānde* Z 2.244, 23.372 < *bu*- ‘know’ • /βwāndē < βu- + -āndē/

*āka*-agent noun ‘realizer’

NSm *bvākā* Z 6.50, IOL Khot 16/3a4, b1, *buvāki* IOL Khot 147/5r3 (Neb:111) <

*bu*- ‘know’ • /βuwākē < βwākē < βu- + -āk- + -ě/



*āñā*-p.nec

NSm *bvāñi* Z 11.23, *bvāñä* SI P3.5v1, IOL Khot 29/8b3 (\*Z), 151/2v3

(Neb:153) < *bu-* 'know' • /βwāñě < βu- + āñ- + -ě/

NSf *bvāña* Z 10.8, SI P4.8v4, 11r5, 12v2, 14r5, 18r1, *bvā[na]* IOL Khot 24/5r2

< *bu-* 'know' • /βwāña < βu- + āñ- + -a/

*āmātā*-noun

NSf *bvāmata* Z 1.31 +, Śgs 3.6v2, *buvāmata* Suv[Or] 5.31, 33 < *bu-* 'know' •

/βuwāmada < βwāmada < βu- + -āmad- + -a/

NAPf *bvāmate* Śgs 2.7v4, Z 6.20 < *bu-* 'know' • /βwāmadē < βu- + -āmad- + -ē/

### **other -u- present stems**

1Pp.a. *ākṣuvāmā* Z 23.43 < *ākṣu-* 'begin' • /aṭṣuwāmě < aṭṣwāmě < aṭṣu- + -

āmě/

3Pp.m *juvāre* Z 1.33, *juvāri* Sgh[25] 241.3, *juvāre* Z 2.46, 20.30, 24.407, 499 < *ju-*

'fight' • /juwārē < jwārē < ju- + -ārē/

*āñā*-p.nec

NASn *juvāñu* IOL Khot 8/1v3 (not in KhSuf) < *ju-* 'fight' • /juwāñu < (jwāñu <)

ju- + -āñ- + -u/

*āna*-p.ptc.m

NSm *juv[ā]nā* Sgh[10] 206.7 < *ju-* 'fight' • /jwāñě < ju- + -ān- + -ě/

NAPm *juvānā* Sgh[10] 206.7 (agreeing with NAPm *hva'ndā* 'men') < *ju-* 'fight' •

/jwāñě < ju- + -ān- + -ě/

*āmata*-noun

Nsf *hāruvāmata* Suv[Or] 10.45 < *hāru*- ‘grow (intr)’ • /hāɹuwāmada <

hāɹwāmada < hāɹu- + -āmad- + -a/

### nominal evidence

GDPm of *ua*-nominals

*ysānvānu* Z 23.162 < *ysānu*a- ‘knee’ • /zānwānu < zānu- + -ānu/

*hajvānu* Z 10.4, Vim 3.39.3, Sgh[22] 43[7], *hajuvānu* Suv[Or]8.54, Sgh[10]

52[1] < *hajua*- ‘wise’ • /haǰwānu < haǰu- + -ānu/

*kṣīruvānu* Suv[Or] 6.1.38 2x, 12.57, 63 < *kṣīru*a- ‘of the land’ • /ṭṣīɹuwānu <

ṭṣīɹwānu < ṭṣīɹu- + -ānu/

*hāruvānu* Z 24.398 < *hāru*a- ‘merchant’ • /hāɹuwānu < hāɹwānu < hāɹu- + -

ānu/

*hālysdamjsuvānu* Suv[Or] 6.2.75 < *hālysdamjsua*- ‘present’ • /hēɹzdand<sup>z</sup>uwānu

< hēɹzdand<sup>z</sup>wānu < hēɹzdand<sup>z</sup>u- + -ānu/

#### 4.5.3.3 *u+i>vī>uvī*

(semivocalization /wī/ and glide resolution /Cwī>Cuwī/ (Hitch 2016a))

3Sop.m *buvīya* Z 13.62 < *bu*- ‘know’ • /βuwīʔa < βwīʔa < < βwīda < βu- + -īda/

3Pop.m *buvīru* Z 23.93, *buvīri* Z 22.317, *bvīru* Z 11.71 < *bu*- ‘know’ • /βuwīru <

βwīru < βu- + -īru/

3Pp.a *rrvīndä* Z 18.25 < *rru*- ‘grow (intr)’ • /rrwīndě < rru- + -īndě/

### other -u- present stems

3Pp.a. *ākṣvīndä* Z 15.111, 24.407, *ākṣvīndi* Z 23.133, *ākṣuvīndä* Z 24.499,

*ākṣ[v]īndä*<sup>86</sup> Suv[Or] 10.46 < *ākṣu-* ‘begin’ • /aṭṣuwīndě < aṭṣwīndě < aṭṣu- + īndě/

3Sopt.a. *ākṣvī* (later for *-īyā* < *-ītā*) SI P 65.3r2 • /aṭṣwī < aṭṣwīřě < aṭṣwīdě <

aṭṣu- + -īdě/

3Pp.a. *jvīndä* Neb 176.9, 11 (= SI P 4.9v7 2x) < *ju-* ‘live’ • /jwīndě < ju- + -īndě/

1Ssj.a. *jvīñi* Sgh[1] 39.2 < *ju-* ‘live’ • /jwīñi < ju- + -īñi/

3Pop. *duvīre*<sup>87</sup> Sgh[17] 243[32] < *du-* ‘beat’ • /duwīro < ḍwīro < du- + -īro/

3Pp.a. *hāruvīndä* Suv[Or] 10.20 < *hāru-* ‘grow (intr)’ • /hāruwīndě < hārūwīndě

< hāru- + -īndě/

### nominal evidence

There is no OKh nominal evidence but LKh *aysmvīnaa-* ‘mental’ < *aysmu-*

‘mind’ + *-īnaa-* (KhSuf:139a) shows the predicted OKh form \*/azmwīna-

< azmu- + -īna-/.

---

<sup>86</sup> So read by Emmerick and Vorob’ëva-Desjatovskaja 1995, SI M 13.7r7 (1995:196). Skjærvø read “*ākṣuīndä*” Or. 59r7 (SuvII:31). The plate 142d (Emmerick and Vorob’ëva-Desjatovskaja 1993) is not legible at this point. In Old Khotanese, as far as I can tell, the spelling *uī* is used only for the diphthong /uī/ while *vī* may be used ambiguously for either the diphthong or the sequence /wī/ (Hitch 2015b:311).

<sup>87</sup> Canevascini noted, “The ending *\*-īre* of *duvīre* points to a scribal mistake for *-īro*, 3 pl. opt. of *duv-*” (Sgh:157 §243.2).

#### 4.5.3.4 *u+e>ve>uve*

(diphthongization to /uē/ and diphthong resolution to /Cuwē/)

1Sp.m *bve* Z 2.66, 23.34, 49, Sgh[10] 68.4, Sgh[6] 99[2], *buve* Z 6.45, *bu<sup>e</sup>ve* IOL

Khot 153/4v1 (Neb:140) < *bu-* ‘know’ • /βuwē < βuē < βu- + ē/

#### **other -u- present stems**

No examples.

#### **nominal evidence**

NAPm *bāysve* IOL Khot 166/2v4, *bāysuve* Z 21.27 < *bāysua-* ‘arm’ • /βazuwē <

βāzuē < βāzu- + -ē/

NAPf *tcarṣṣuve* Z 22.210 < *tcarṣṣua-* ‘brilliant’ • /t<sup>s</sup>aṣuwē < t<sup>s</sup>aṣuē < t<sup>s</sup>aṣu- +

-ē/

#### ***u+ū>ū***

*ūna*-noun ‘covering’

NSm *pagyūni* Z 22.138 < *paju-* ‘cover’ • /baḷūni < baḷu- + -ūn- + -ē/

#### **nominal evidence**

No examples. There are examples of *u+u>ū*, e.g., ASm *aysmū* Z 5.30 < *aysmua-*

‘mind’ + -u, *ysānū* Z 22.149 < *ysānua-* ‘knee’ + -u, *kṣīrū* Z 24.390 < *kṣīrua-*

‘of the land’ + -u.

#### 4.5.4 **Contractions with *VttV* stems ending in -ū-**

Just two *VttV* exhibiting stems end in -ū-, *nū-* ‘lie down’ and *bū-* ‘be fragrant’. As far as I can determine, there are no other stems in Old Khotanese, verbal or nominal, that unambiguously show final -ū-. Both *nū-* and *bū-* have a 3Sp.m ending in -ūtte, 3Sp.m *nūtte* Z 22.129 /nūttē < nū- + -ttē/ and 3Sp.m *būtte* Z 3.57, 85, 20.3, Or.

12637/57.2a3 /βūt̄t̄ē < βū- + -ttē/. However, each appears to have a different 3Pp.m formation, cf. 3Pp.m *nūyāre* Z 24.168, and 3Pp.m *bvāre* Z 3.118~*buvāre* Z 3.35. The spellings *bvāre*, *buvāre* are ambiguous, meaning either ‘they know’ < *bu-* or ‘they are fragrant’ < *bū-*. The Old Iranian etyma are homophonous with OIr *\*baud-* meaning both ‘feel, sense’ and ‘smell (intr)’ (EDIV:14–16). The long *ū* in *bū-* ‘be fragrant’ may be a late development to keep the verb distinct from *bu-*, in which case *bvāre*~*buvāre* represent older *\*bu-* ‘be fragrant’<sup>88</sup> and there is semivocalization of short *u*>*v*. It may be preferable to transcribe *bū̃-*. In contrast, 3Pp.m *nūyāre* Z 24.168 shows *y*-epenthesis between long vowels. Until now it has been generally accepted that *nū-* is a contracted form of *nuva-* ‘lie down’. I hesitate to accept that view without parallel cases of historical contraction, but offer no better explanation.

#### 4.5.4.1 $\bar{u}+\bar{a}>\bar{v}\bar{a}>uv\bar{a}$ or $u+\bar{a}>\bar{v}\bar{a}>uv\bar{a}$

3Pp.m *bvāre* Z 3.118, *buvāre* Z 3.35 < *bū-* ‘be fragrant’ • /βuwārē < βwārē < βu- + -ārē/

#### 4.5.4.2 $\bar{u}+\bar{a}>\bar{u}y\bar{a}$

3Pp.m *nūyāre* Z 24.168 < *nū-* ‘lie down’ • /nūyārē < nū- + -āre/

---

<sup>88</sup> Although its relationship to Proto-Iranian *\*baud-* is unclear (cf. Emmerick’s discussion, SGS:332) the noun *buša-* ‘perfume’ also has short *-u-* reflected in NAPn *buvi’* Z 3.123 /βuě < βuzě < βuz- + -ě/, NAPm *bva’ñi* Śgs 2.43 /βuañě < βuzañě < βuz- + -añā/, NSm *bū’* Suv[Or] 8.23, [Q] 11.a2 /βuě < βuzě < βuz- + -ě/, IASm *bū’ṇa* Z 2.42+4× /βuěṇa < βuzěṇa < βuz- + -ěna/.

#### 4.5.5 Contractions with *VttV* stems ending in *-ä-*

The *VttV* exhibiting stems are *dä-* ‘appear’, *nä-* ‘sit down’, *pajā-* ‘ask for’, and *hamä-* change. There are perhaps no other Old Khotanese verbs ending in *-ä-*.<sup>89</sup> No nominal stem ending in *-ä-* has ever been suggested by anyone. There are thus no exact morphophonological parallels. Approximate parallels are available from stems in *-i-*. The vowels *ä* and *i*, phonetically [ɛ] and [ɪ], merge sometime within Old Khotanese and then the spellings interchange. The contractions *ä+i>i* and/or *ä+i>ī* are evidenced by two enclitics.

##### 4.5.5.1 *ä+a>ya*

2Sp.m *dya* Śgs 2.8v2 • /δj̄a < δě- + -a/

2Siv.a *nya* Z 2.188, 5.52 • /nj̄a < ně- + -a/

*andaa-p.ptc*

ASm *pajyandau* ‘beggar’ Z 11.19 main ms. and variant 8 < *pajā-* ‘ask for’ •

/baḷyandaḷ < baḷě- + -anda- + -u/

##### present stems in *-i-*

2Siv.a *parya*<sup>90</sup> Z 23.52 < *pari-* ‘order’ • /barj̄a < bari- + -a/

##### 4.5.5.2 *ä+i>ī*

3Sop.a *pajj̄yi* Z 11.19 • /baḷj̄ī < baḷj̄idě < baḷě- + -idě /

3Pp.a *pajj̄indä* Z 2.190 • /baḷj̄indě < baḷě- + -indä/

---

<sup>89</sup> The possible existence of the shapes *āyā-* ‘appear’ (SGS *āy-*) and *pūyā-* ‘observe’ is discussed in a footnote to 4.3 above. Both would be type A, so not exhibiting *VttV*.

<sup>90</sup> Emmerick identified *parya* as 2Piv.a (SGS:73), but that would be *\*parīta~\*parīya* < *pari-* + *-āta* with *i+ä>ī* contraction, formally identical to 2Pp.a *parīya* Z 12.28.

3Pp.a *nīndä* Z 3.60 • /nīndě < ně- + -īndě/

3Pop.m *dīru* Z 11.73, 21.25 • /δīru < δě- + -īru/

2Ssj.a *dīñi* Z 2.100 • /δīñi < δě- + -īñě/

### **ä + enclitic -i or -ī**

Two enclitics, both traditionally listed as *-ī* with long *ī*, contract with word final *-ä* producing *-ī*. As illustrated in TypeA, the 3S enclitic pronoun appears to have the underlying shape *-i* with short *i*. For the 2Sp.a enclitic ‘be’ there is no evidence concerning length so the shape *-ī* is tentatively retained. These contractions show either *ä+i>ī* or *ä+ī>ī*.

3S enclitic pronoun *-i*

NSm *kamalä* ‘head’ + *-i* > *kamalī* ‘its head’ Z 22.145

NAPf *tcei’mañä* ‘eye’ + *-i* > *tcei’mañī* ‘his eyes’ Z 22.169

2Sp.a enclitic ‘be’ *-ī*

NSm *kṣāmovä* ‘forgiving’ + *-ī* > *tterä kṣamovī* ‘you are so forgiving’ Z 2.177

NSm *balysä* ‘Buddha’ + *-ī* > *thu balysī* ‘you are a Buddha’ Z 22.258

### **present stems in -i- showing *i+ī>ī***

(see above)

#### **4.5.5.3 ä+ä>yā>iyā~äyā**

3Pp.m *dyāre* Z 3.113, 137 +, Sgh[15] 69.1, *däyāre* Z 5.12, 23.13, *diyāre* Z 4.100,

Sgh[10] 162.1, *dyārä* Z 17.3, *diyāri* Z 4.37 • /δëyā.ē < dyā.ē < δě- + -ā.ē/

3Psj.m *diyāru* Z 4.91 • /δiyā.ru < dyā.ru < δě- + -ā.ru/

3Pp.m *hamyāre* Z 6.17, Suv[L] 16.36, *hamāyāre* Z 24.2 • /haměyā.ē < hamyā.ē  
< hamě- + -ā.ē/

3Psj.m *hamyāro* Z 2.140 • /hamyā.ɔ < hamě- + -ā.ɔ/

*āna*-p.ptc.m, for *dyāna*- 'bright'; good-looking'

NSm *dyānā* Suv[Or] 6.3.42, *d*]yānā Suv[C] 3.100 • /dyāně < dē- + -ān- + -ě/

NSf? *dāyāna* Z 9.27 • /dēyāna < dyāna < dē- + -ān- + -a/

NAPm *dyāna* Suv[Or] 6.3.16, *dhyāna* Suv[Or] 6.3.28 • /dyāna < dē- + -ān- + -a/

+ *-ātara*-comparative

NAPm *dyānātara* Suv[Or] 10.22 'brighter; better looking' • /dyānēda.ɔ < dē- +  
-ān- + -ēda.ɔ + -a/

*āmatā*-noun 'appearance', homophonous with *dyāmatā*- 'seeing' from *de*- 'see'.

NSf *dyāmata* Z 24.483 • /dyāmada < dā- + -āmad- + -a/

ASf *dyāmato* Śgs 2.3r4 • /dyāmadō < dē- + -āmad- + -o/

*āñā*-p.nec

NASn *nyāñu* Suv[Or,C] 6.3.17, *niyāñu* Z 2.221, 3.30, *ñānu* SI P65.3v1 • /niyāñu  
< nyāñu < nē- + -āñ- + -u/

**present stems in -i- showing *i+ā>yā>iyā***

(more evidence for *i+ā>yā>iyā* in Hitch 2016a)

3Pp.a *iyāre* Z 3.144, Śgs 3.13r3~*jiyāre* Z 6.1 < *ji*- 'cease' • /jiyāre < jyāre < ji- +  
-āre/

3Pp.a *kṣyāre* Z 21.23~*kṣiyāre* Z 2.43 < *kṣi*- 'be troubled' • /ṭṣiyāre < ṭṣyāre  
< ṭṣi- + -āre/



1Psj.m *hīvyāmane* Suv[Or] 6.1.39 < *hīvi-* ‘appropriate’ • /hīvyāmanē < hīwi- +  
-āmanē/

participle of necessity *-āñā-*: *i+āñ>yañ*

NAPm *paryāñā* ‘should be ordered’ Suv[Or] 6.3.23 < *pari-* ‘order’ • /ba.ɣyāñā <  
ba.ɣi- + -āñ- + -a/

abstract noun *-āmatā-*: *i+āmat>yāmat>iyāmat*

NAPf *āriyāmate* ‘*saṃrañjanīya-dharmas*’ Z 6.20 < *āri-* ‘mix, meet (intr)’ + -  
*āmatā-* + -e • /ā.ɣiyāmadē < ā.ɣyāmadē < ā.ɣi- + -āmad- + -ē/

**nominal evidence showing *i+ā>yā>iyā***

GDPf *aśyānu* Z 22.310, Suv[Or] 6.1.63, 10.33 < *aśiā-* ‘nun’ + *-ānu* • /ažyānu <  
aži- + -ānu/

GDPm *paḍāṃjsiyānu* Suv[Or] 2.2, *paḍāṃjsiyānu* Śgs 3.7r5 < *paḍāṃjsia-*  
‘former’ + *-ānu* • /baḍānd<sup>z</sup>iyānu < baḍānd<sup>z</sup>yānu < baḍānd<sup>z</sup>i- + -ānu/

#### 4.5.6 Contractions with *VttV* stems ending in *-e-*

The only *VttV* exhibiting stem ending in *-e-* is *se-* ‘appear’, the morphophonology of which was partly detailed in TypeA. Also described there are the type A verbs *uysde-* ‘look up (at)’, *de-* ‘see’, *bāysde-* ‘observe’ and *yse-* ‘be born’, as well as one further stem *āve-* ‘accumulate’ which may be either A or B. No nominal stems end synchronically in *-e-*. Stems in *-e-* behave partly like stems in *-a-* and partly like stems in *-i-* or *-ā-*. On the one hand *e+i>ai* is like *a+i>ai*, and on the other hand *e+ā>yā>āyā~iyā* is like *i+ā>yā>iyā* or *ā+ā>yā>āyā*.

#### 4.5.6.1 *e+i>ai*

1Sp.a *saimä* Z 13.60 < *se-* ‘appear’ • /sai̯mĕ < *se-* + -i̯mĕ/

3Pp.a *saindä* Z 3.116 (+ 20x in Z), IOL Khot 7/7.r4, 141/6.r2, r3, r4 (2x), v1,  
29/8.a4, 20/15.v2, 189/5v1 (*Śgs*), *seindä* Z 3.108, *saindi* Z 2.227 (+ 6x in  
Z) < *se-* ‘appear’ • /sai̯ndĕ < *se-* + -i̯ndĕ/

3Sop.a *saiyä* Z 7.26, 10.35, *saiyi* Z 13.107 < *se-* ‘appear’ • /sai̯ĭĕ < sai̯dĕ < *se-* + -  
i̯dĕ/

#### other -e- final present stems

3Sop.a *āvaiyā* Z 18.3 < *ave-* ‘accumulate’ • /āwai̯ĭĕ < āwai̯dĕ < āwē- + -i̯dĕ/

1Sp.a *daimä* ‘I see’ Z 6.7 • /dai̯mĕ < δē- + +i̯mĕ/

3Pp.a *uysdaindä* ‘they look up (at)’ Z variant 12 to 20.20 • /uzdai̯ndĕ < uzδē- +  
-i̯ndĕ/

3Pp.a *daindä* ‘they see’ *Śgs* 3.4r2, *deindä* Z 4.91, 9.11, 21.32 • /dai̯ndĕ < δē- +  
-i̯ndĕ/

3Pp.a *bāysdaindä* ‘they observe’ Z 3.17 • /βāzdai̯ndĕ < βāzδē- + -i̯ndĕ/

1Ssj.a<sup>91</sup> *daiñä* ‘may I see’ SI P62.1a6 • /dai̯ñĕ < δē- + -i̯ñĕ/

3Pop.a *daira* ‘they would see’ Z 5.72 • /dai̯ra < dai̯ru < δē- + -i̯ru/.<sup>92</sup>

3Pop.a *ysairo* ‘they would be born’ Z 8.25 • /zai̯ru < zē- + -i̯ru/

---

<sup>91</sup> The interpretation of *daiñi* as 1Ssj.m with the active ending rather than the usual middle is discussed in Hitch in-press-b footnote ± 28.

<sup>92</sup> Cf. SGS:209.

#### 4.5.6.2 *e+ā>yā*

*āmatā*-noun

NSf *syāmata* Z 4.85, 5.77, 9.7, IOL Khot 163/1.r4 (Rk 94.12) < *se-* ‘appear’ •

/syāmata < *sē-* + *-āmat-* + *-a*/

IASf *syemāte* Z 9.10 2x, SI M 38.v3, *syemate* Z 2.210 < *se-* ‘appear’ • /syēmātē <

*sē-* + *-āmat-* + *-iē*/

#### other *-e-* final present stems

1Pp.a *dyāmā* Z 4.54, *dāyāmā* Z 14.56 ‘we see’ • /*ḍēyāmē* < *ḍyāmē* < *ḍē-* + -

*āmē*/

1Pp.m *ysyāmane* Sgh[17] 253[50] ‘we are born’ • /*zyāmanē* < *zē-* + *-āmanē*/

3Pp.m *ysyāre* Z 3.62, *ysiyāre* Z 13.59 ‘they are born’ • /*zēyārē* < *zyārē* < *zē-* + -

*ārē*/

3Piv.a *dyāṃdu* Z 14.58 ‘let them see’ • /*ḍyāndu* < *ḍē-* + *-āndu*/

1Psj.a *dyāma* Sgh[7] 198.4 ‘may we see’ • /*ḍyāma* < *ḍē-* + *-āma*/

3Ssj.m *ysyāte* SI P 65.2+ r3 ‘may he be born’ • /*zyādē* < *zē-* + *-ādē*/

*āka*-agent noun: *-āka-*

NSm *dyākā* Z 5.72, *dāyākā* 8.13 ‘seer’ • /*ḍēyākā* < *ḍyākā* < *ḍē-* + *-āk-* + *-ē*/

NAPm *bāysdy[ā]ka* ‘watchers’ Z 24.642 • /*βāzḍyāka* < *βāzḍē-* + *-āk-* + *-a*/

*āñā*-p.nec: *-āñā*

NSm *āvyāñi* Z 12.23 < *avē-* ‘accumulate’ • /*āvyāñē* < *āwē-* + *-āñ-* + *-ē*/

NAPf *uysdyāñe* Suv[Or 36r4; not in ed.] 6.3.22, *uysdiyāñe* Suv[Or] 6.3.21 ‘must

be seen’ • /*uzḍiyāñē* ) *uzdyāñē* < *uzḍē-* + *-āñ-* + *-ē*/

NAPm *dyāña* Z 24.441, *diyāña* Z 2.153, *däyāña* Z 2.210 ‘must be seen’ • /*dyāña*

< *δē-* + *-āñ-* + *-a/*

*āna*-p.ptc.m: *-āna-*

NAPm [*uysdy*] *āna* Z 23.151 ‘looking up’ • /*uzδē-* + *-ān-* + *-a/*

*ānaa*-p.ptc.a: *-ānaa-*

NSm *dyānai* Rk 94.7 ‘visible’ • /*dyānaḡ* < *δē-* + *-āna-* + *-ě/*

NSm *adyānei* Rk 94.7 ‘invisible’ • /*aḡyānaḡ* < *a-* + *δē-* + *-āna-* + *-ě/*

*āmatā*-noun: *-āmatā-*

NSf *dyāmata* IOL Khot 156/7 r3 ‘seeing’ • /*dyāmad-* < *δē-* + *-āmad-/*

ASf *ysyāmato* Sgh[8] 226.3 ‘birth’ • /*zyāmadō* < *zē-* + *-āmad-* + *-ō/*

GDSf *ysyemate* ‘birth’ Z 24.222 • /*zyēmadē* < *zyāmad-* + *-<sup>i</sup>ē* < *zē-* + *-āmad-* + *-<sup>i</sup>ē/*

IASf *bāysdyemāte jsa* IOL Khot 149/1 v6, *bāysdiyemetä jsa* IOL Khot 148/3.r6

• /*βazḡyēmadē d<sup>z</sup>a* < *βazḡyēmadē d<sup>z</sup>a* < *βazḡyāmad-* + *-<sup>i</sup>ē d<sup>z</sup>a* < *βāzda-* + *āmad-* + *-<sup>i</sup>ē d<sup>z</sup>a/*

IASf *ysyemate* ‘birth’ Z 24.165, *y[sy]emāte jsa* Sgh[25] 241.5 • /*zyēmadē* <

*zyāmad-* + *-<sup>i</sup>ē* < *zē-* + *-āmad-* + *-<sup>i</sup>ē/*

*āmatīṅgyā*-adjective: *-āmatīṅgyā-*

IASf *adyematīje* Z 6.7 ‘unseeing’ • /*aḡyēmadījē* < *aḡyāmadīṅg<sup>y</sup>-* + *-<sup>i</sup>ē* < *a-* + *δē-*

+ *-āmad-* + *-īṅg<sup>y</sup>-* + *-<sup>i</sup>ē/*

NAPm *ysyāmatīgya* Z 24.238 ‘birth’ • /zyāmatīgya- < zē- + -āmad- + -īṅg<sup>y</sup>- + -

a/

āñ-causative: + -āñ

3Sp.m *ysyāñāte* Suv[Or] 6.2.72 ‘he produces’ • /zyāñēdē < zē- + -āñ- + ědē/

#### 4.5.7 Summary of the contractions exhibited by the *VttV* stems

When the final vowel of a *VttV* exhibiting stem meets a suffix initial vowel, the resulting contraction shows no trace of a synchronic stem final consonant. This is especially clear with  $a+a>\bar{a}$  seen in 3Ssj.a *khā* and 2Siv.a *pvā*. But it should also be clear when one compares the contractions exhibited by other vowel final stems, both nominal and verbal. An orthographic *y* or *v* resulting from some contractions does not reflect an underlying synchronic stem final consonant. Rather it reflects a synchronic process of either diphthongization or semivocalization well attested with vowel final nominals and other vowel final present stems. Other than the conjugated forms showing *VttV*, there is no evidence in Old Khotanese that the *VttV* stems once ended in a consonant.

#### 4.6 *tt*-initial suffixes contain double /tt/

In “*tt* in Old Khotanese” (Hitch 2015a) the metrical structure of the *Book of Zambasta* was used to demonstrate that orthographic *tt* between vowels is ambiguous, standing sometimes for single /t/ and sometimes for double /tt/. Essentially all examples of orthographic *VttV* fall into four categories: 1) Indian loans, 2) conjugated present stems, 3) past stems, and 4) forms containing the de-adjectival \**tara*-comparative, \**tama*-superlative and \**tāti*-noun suffixes. The first

two categories, Indian loans and conjugated present stems have double /tt/ while the other categories have single.

Not mentioned in that study is the related observation that orthographic - $\check{V}dV$  in native words appears to exhibit parallel ambiguity. There are no examples in Indian loans but past stems and the de-adjectival suffixes have underlying single /d/ while conjugated present stems apparently have double /dd/. In all three categories, the orthographic  $VdV$  reflects historical  $*VrtV$ . The past stems are historically derived from present stem ending in  $*-Vr-$  with a  $*-ta-$  suffix (KhSuf:224):  $-Vda- < *-Vrta-$ . The de-adjectival forms reflect  $*-Vr- + *-tara-$ ,  $*tama-$  or  $*-tāti-$ . The conjugated presents reflect  $*-Vr- + *-tV$ . Some examples of each category will now be compared with the meter of Z to show the ambiguity. For a form to be testable, it must have a short vowel before the  $d$ , that is,  $\check{V}dV$ . If the syllable containing the short  $\check{V}$  counts heavy in the meter, this shows that there is underlying /dd/. In those cases, orthographic  $\check{V}dV$  syllabifies / $\check{V}d.dV$ /. The metrical analytical apparatus was defined in Hitch 2014 with additions in 2015b, and summarized above in 1.5.3.

#### 4.6.1 Past stems in - $Vda-$

Of the 39 past stems ending in  $-Vda-$  listed by Emmerick in SGS (172) I notice just ten occurring with short vowel in Z:  $\check{a}ph\check{a}r-:\check{a}ph\check{a}da-$ ,  $\check{a}ph\check{i}da-$  ‘be disturbed’;  $\check{a}spar-:\check{a}spu\check{d}a-$  ‘tread’; 1.  $bar-:bu\check{d}a-$  active ‘carry’; 2.  $bar-:bu\check{d}a-$  middle ‘ride’;  $m\check{a}r-:mu\check{d}a-$  ‘die’;  $yan-:y\check{a}da-$ ,  $yu\check{d}a-$  ‘make, do’;  $haph\check{a}r-:haph\check{a}da-$ ,  $hapha\check{d}a-$  ‘be distracted’;  $hamb\check{i}r-:hamb\check{a}da-$  ‘be filled’,  $hay\check{a}r-:hay\check{a}da-$  ‘rejoice, be happy’;  $hvar-:hv\check{a}da-$  ‘consume’. The evidence is overwhelming that the forms contain underlyingly single

/d/. An orthographic sequence -*ṣḍṣ* in past stems always counts LL. An example from each of the nine qualifying past stems follows.

3Ppf.intr.m *āphāḍa* is **HLL** in cadence 1 **HLLHL**:

24.163cd ttāna cu klaiś|yau jsa . □| hamata | **āphāḍa** satva (B:5+6+7)<sup>93</sup>

because beings themselves were confused by kleśas.

3Spf.tr.m *āspuḍe* is **HLL** in cadence 1 **HLLHL**:

23.169cd samu balysi | **āspuḍe** śśando □| varnātai | sthvirā subhūtā

(A:5+7+5+7)

As soon as he Buddha trod upon the earth, the Elder Subhūti honoured him.

The ppp *buḍa-* occurs four times (5.35, 23.138, 141, 164) always as the 3Ppf.tr *buḍāndā*, *buḍāndi* and always at the end of a line in cadence 1 **HLLHL**, where the *uysnora*-effect<sup>94</sup> makes the evidence ambiguous, e.g.:

5.35cd haḍā kāḍe | vasute niṣo□ndā | ttāmu hā | śśāya **buḍāndi** .

(A:5+7+5+7)

The day was very clear, calm. Then the Śākyas rode off.

---

<sup>93</sup> In this hemistich the metrical apparatus may be defined as an example. 24.163=chapter and line number. cd=graphic column spread over the four columns, a, b, c, d, of the main manuscript. |=metrical segment boundary. □=graphic column boundary. (B:5+6+7)=type B meter with segments counting 5, 6, and 7 moras.

<sup>94</sup> Name given in Hitch 2014:15–17 for the metrical principle through which a normally heavy syllable may be counted as light in the antepenultimate position in a seven mora segment.

3Spf.tr.m *muḍā* counts **LL** in cadence 4 **LLLLHL**:<sup>95</sup>

24.509 arahandā | sūradā □ vara | **muḍā** śśāte pā[tc]o (B:5+6+7)

The Arhat Sūrata lies dead there too.

3Ppf.tr *yādāndā* occurs as **LHL** in cadence 3 **LHLHL**:

3.5cd **yādāndā** balysā | pha□ru padya | pajsama kāḍe . (C:7+5+5)

In very many ways they did honours to the Buddha.

2Spf.tr.m *yuḍai* is three moras long, LH, in a five mora segment:

5.41cd brīyai<sup>96</sup> | spāṣṭe tta kāṣṭe □ | śśuru **yuḍai** | bryandama pūra

(A:5+7+5+7)

He looked upon him with affection. Thus he thought: 'You have done well, most beloved son.'

3Ppf.intr.m *haphāḍa* counts **LLL** in cadence 4 **LLLLHL**:

4.76ab vittarkyau | **haphāḍa** uysnaura □ | vikalpyo jsa | saindā cu nāstā

(A:5+7+6+7)

Beings are distracted by *vitarkas*. Through *vikalpas*, (things) appear which do not exist.

The NSf adjective *haṃbaḍa* counts **HLL** in cadence 1 **HLLHL**:

---

<sup>95</sup> The related noun *muḍaa-* 'corpse' occurs 16 times in Z: NAPm *muḍā* 2.48, 93, 20.29, 34, 62, 24.420, GDP *muḍāni* 2.53, ASm *muḍau* 4.9, 24.453, GDSf *muḍye* 8.25. Another noun *muḍa-* 'corpse' occurs twice: NSm *muḍā* 24.450, ASm *muḍu* 13.135. Both nouns appear once with the denominal adjective suffix *-īñā-* (KhSuf:131b): NAPf *muḍīñi* 'of the dead' 2.44, NAPm *muḍaiñā* 'of the dead' 2.48. In all of these forms, *mu-* counts L showing single /ḍ/.

<sup>96</sup> For original five mora *\*brītyai* (+ *-i*) as suggested by E. and M. Leumann (Glossar:480b) and pointed out to me by Mauro Maggi p.c. The same emendation is due at 5.46 where *brrīyai* also occupies a five mora segment. The IAS is formed in the same way as the GDS *brītāye* Z 22.200, *brītaā-* + *-i* with umlaut, diphthongization and diphthong resolution /βṛīṛēyē < βṛīṛīe < βṛīṛi-ē < βṛīṛa- + *-i*ē/. Unresolved IASf *brītye jsa* counts five moras HHL at Z 13.59, 20.8, 12.



16.42ab purvatī | dīvi khu □ purra | **haṃbaḍa** trāmā (B:5+6+7)

The Pūrvavideha-dvīpa is like the full moon

The <sup>i</sup>e-infinitive form *hayāḍe* counts as three moras in a six mora segment:

14.84cd kho ni kṣamāte | **hayā**□**ḍe** tta bī|nāñu pyuvā're (B:5+6+7)

as it pleases them to rejoice, so do they hear the music.

#### 4.6.2 *tara*-comparatives and *tāti*-nouns

There are no qualifying *ṽḍV tama*-superlatives but there is a discussion of *buḍḍama*- 'most' below. There is one qualifying *tāti*-noun, *āyiḍāti*- 'joy' < \**āyāra*- 'joyful' (Degener \**āyiḍa*- KhSuf:279a) cf. *hayār*-:*hayāḍa*- 'rejoice'. Below, the first two syllables of ASf *āyiḍetu* (*āyiḍāti*- + <sup>i</sup>-u) count as three moras in a six mora segment:

12.75ab ka bodhi|satvā **ā**□**yiḍetu** nājsaṣḍe . (B:4+6+7)

If a Bodhisattva exhibits merriment,

There are five qualifying *tara*-comparatives. The syllable *-va-* below in *iśvaḍara*- 'more powerful' < Skt *īśvara*- 'Lord' is **L** in cadence 4 **LLLLHL**:

23.96cd samu ku vajsī|tī'rā handaru □| gyastu ce **iś**|**vaḍarā** umyau jsa

(A:5+7+6+7)

When you merely behold another god who is more powerful than you,

ASm *pṇadaḍaru* < *pandaḍara*- 'more foolish' < *pandara*- 'foolish' fills a five mora segment showing the count is **HLLL**:

2.134ab hanāna vajsā|ta'ndai ṣṭāna □| **pṇadaḍaru** | karaṇu yāḍaimā

(A:5+7+5+7)

Though a man with sight, I did a feat more foolish than that of a blind man.

ASm *buḍaru* < *buḍara-* ‘more, greater’ < *\*bura-*<sup>97</sup> ‘much’ counts **LLL** in cadence 4

**LLLLHL:**

13.126ab *riddhā-pāt|yau jsa kalpu jsī|na □ u kyeri | buḍaru kṣamīyā .*

(A:5+7+5+7)

Through the *ṛddhipādas*, one has life for a *kalpa* and for however more (*kalpas*) one should wish.

Adverbial *ysittaḍaru* < *ysittaḍara-* ‘very soon’ < *ysittara-* ‘soon’ (cf. DKS:350b ‘reduced, short’) fills a five mora segment showing that the syllable *-tta-* counts **L** in

**HLLL:**

2.70cd **ysittaḍaru** | *hā usahyāmā □ | vā haṃgrī|sādu biśśīnda .* (A:5+7+5+7)

Very shortly we will deign to go hence. Let them assemble here from all sides.

The first three syllables of NAPm *hatāḍarāṃjsya* (+ *-āṃjsia-* adjective suffix) < *hatāḍara-* ‘former’ < *hatāra-* ‘former’ cf. *hatāro*, *hatāru* ‘formerly, once’, count as **LLL** in a five mora segment:

9.28ab *parrījāte | satva dukhyo jsa □ | kho rro hatāḍa|rāṃjsya balysa*

(A:5+7+5+7)

He rescues beings from woes just like the former Buddhas.

The evidence of the *tara-*comparatives and *tāti-*noun which contain the shape *ṽḍV* in *Z* show without doubt that these words all contain single /ḍ/.

---

<sup>97</sup> Related to the postpositive particle of indefiniteness *buro*. From the cognates Av *būri-* ‘much’, Skt *bhūri-* ‘much, many’, Mayerhofer 1996:268–269 offers Indo-Iranian *\*b<sup>h</sup>uHri-* > Proto-Iranian *\*buHri-* which would give Old Iranian *\*būri-* and then Khotanese *\*bura-* by shortening of *ū* in disyllables before *m, n, r* (cf. Emmerick 1989:210). The absence of an adjective *\*bura-* in Old Khotanese suggests that *buḍara-* is a fossilized inherited form from Pre-Khotanese rather than synchronically derived.

### 4.6.3 Present stems in *-r-* conjugated with *tt*-initial suffixes

Emmerick listed sixty-six present stems in *-r-* in SGS:168. Of these, apparently only one, *bar-* ‘ride’ has forms which can be examined in Z for evidence of suffixes beginning with double */tt/*. Many simply do not occur in Z, or do not occur with a suitable suffix. We cannot examine type A stems because they do not have *VttV* conjugated forms. Type B stems with long vowel before the *-r-* like *hambār-*, 3Sp.a *hambīḍā* Z 2.194, ‘be filled’ will count that syllable as heavy just from the vowel length. Umlautable type B stems with long vowel like *bār-* ‘rain’, 3Sp.a *beḍā* Z 22.121, or *haur-* ‘give’, 3Sp.a *heḍā* Z 2.30, have a heavy stem syllable both before and after umlaut. Similarly excluded are 3Sp active forms of type B stems with short *-a-* because this is umlauted to long *-ī-* as in 3Sp.a *ttuvīḍā* Z 22.156 < *ttuvar-* ‘bring’, or 3Sp.a *bīḍā* Z 18.15 < *bar-* (active) ‘carry’. Finally, of the remaining candidate stems only *bar-* (middle) ‘ride’ has forms with short vowel plus *-ḍ-*. In SGS Emmerick listed *haphāḍe* 4.72 and *hayāḍā* 24.215 as present indicative forms from *haphār-* ‘be distracted’ and *hayār-* ‘rejoice’ respectively. If these were present forms they would qualify for examination here but they are perfects.<sup>98</sup>

A first hint that there might be something interesting with this category is the unique form 3Sp.m *baḍḍe* Z 13.149 ‘he rides’. Elsewhere this is always spelled *baḍe*. The double *ḍḍ* suggests a consonant cluster. The spelling was usually avoided,

---

<sup>98</sup> 4.72 *haphāḍe hūnāna* ‘he is distracted by a dream’ is parallel with 4.76 *vittarkyau haphāḍa uysnaura* ‘beings are distracted by *vitarkas*’ four lines later. They are both perfect forms. 24.215 *hayāḍā* ‘one rejoiced’ is parallel in tense with the perfect *panata* ‘she rose up’ earlier in the line. Another argument for perfect status, which may appear circular at this point, is that they are both LLL in the meter, which prohibits reading an underlying double */ḍḍ/*.

perhaps partly because the radical *ḍa* is tall and placing one under the other risks the akṣara descending into the line of writing below.

*baḍ(ḍ)e* 'he rides' is spelled once with double *ḍḍ* and five times with single *ḍ* in Z. It is clear from the metrical position of *baḍe*, *baḍḍe* in Z that the first syllable is heavy, and the word is underlyingly */βaḍḍē/*. In five of six instances the meter unquestionably confirms the weight HL:

*baḍḍe* is **HL** in cadence 2, **HXHL**:

13.149ab trāmu bal|ysūṣṭa mahāyā ni | kho rraha-|bārai **baḍḍe** (A1:5+9+3+7)

Such for bodhi is the Mahāyāna as a charioteer rides.

*baḍe* is **HL** in cadence 2, **HXHL**:

22.150cd hamye bruī<sup>99</sup> | kvī rre bvaittä . | kari jambu|tīvī **baḍe** . (A:5+7+5+7)

On the same morning as the king mounts it, he will ride right around Jambudvīpa.

*baḍe* is **HL** in cadence 1<sup>A</sup>, **HLLHHL**:

2.95ab rre biṃbā|ysārā rro hā **baḍe** | haṃtsa | hārvyau ttīyā . (A1:5+9+3+7)

King Bimbisāra too rides off then with the merchants.

*baḍe* is **HL** in a nine mora cadence ending HL, **HLLLLHL**:

13.26cd prattyeka|-yāni kho aśśā **baḍe** | u nir|vānā kho kaṃtha .

(A1:5+9+3+7)

The Pratyekayāna is as the horse is ridden; and Nirvāṇa is like the city.

*baḍe* is HL in cadence 1, **HLLHL**:

24.120ab hāna dāya | brīk[ā<sup>100</sup>] □ aḍari | **baḍe** kuberā (B:5+6+7)

Look up, beloved. There rides another Kubera.

There is one example where *baḍe* might be considered to be two moras long in order to correctly make up the middle, normally six mora, type B segment:

24.249ab uryānu | **baḍe** handaru | kālu padāya (B:5+7+7)

He rides to a park at another time. On the way ...

---

<sup>99</sup> As shown in Hitch 2015b:310–313, the spelling *uī* or *vī* may represent a diphthong /ui/ which counts as two moras. *bruī* counts for two moras. *kāḍaruī* 24.53 counts as 5 and *pharuī* 13.22, 24.229 counts as three. The latter is spelled *pharvī* once 22.247 and counts 3. Similarly, *hastarvī* 2.234 counts 5 and *śśārvī* 2.169 counts 3. Presumably, *bruī* could also be written *\*brvī*.

<sup>100</sup> Emmerick followed the E. and M. Leumann restoration *brīk*[o]. M. Maggi (p.c.) points out that the VS of the *ā*-declension is now attested as *-ā* in *dīvatā* < *dīvatā*- ‘goddess’ at Suv[Or] 5.16, Suv[SI<sup>3</sup>] 17.31,32.

However, it is reasonably certain that *baḍe* is HL here too. The most common cadence in a type B meter six mora segment is B<sub>1</sub>: HL + word boundary + 3 additional syllables. Of a sample of 153 hemistichs, 68% had B<sub>1</sub>. In addition, 87.5% begin HL, and 98% begin H (Hitch 2014:7–8). It is the second part of the segment which is irregular.

The metrical evidence plainly shows that *baḍ(ḍ)e* consists of three moras, HL. There is double *ḍḍ* written one time because there was phonemic /ḍḍ/ in this word. The only explanation for this requires that the 3Sp.m synchronic suffix /-ttē/ has double /tt/. When added to *bar-* /βαῤ-/ , the sequence /-ῤtt-/ becomes /-ḍḍ-/. The new cluster retains the retroflexion and voicing of /-ῤ-/ as well as the geminate stop value of /-tt-/. The derivation /βαḍḍē < βαῤ- + -ttē/ is natural. The metrical behavior of *ba(ḍ)ḍe* provides independent evidence that the historical present system conjugational suffixes beginning with single \*t have been reanalyzed as beginning with double /tt/. As shown above, other suffixes beginning historically with single \*t, in the past stems, *tāti*-nouns and *tara*-comparatives, have not been subject to metanalysis. In those cases, orthographic *VḍV* reflects a single consonant, just as does *VttV* in the same categories (Hitch 2014a).

#### 4.6.4 The odd case of *buḍḍamu*

There is a spelling which might be considered counter evidence to the claim made above. This involves a *tama*-superlative, *buḍḍama*- ‘most’, historically < \**bura*- ‘much’ + \*-*tama*-. It occurs one time in Z where the meter confirms double /ḍḍ/. Adverbial *buḍḍamu* counts HLL in cadence 1 **HLLHL**:

4.74ab cu ye brātā | jsānye kei'tā □ | ttu ye hūña | **buḍḍamu** daiyā  
(A:5+7+5+7)

One mostly sees in a dream that which one thinks while awake,

If not for the double orthographic *ḍḍ*, one would expect the meter to be defective here. Since there is no known adjective *\*bura-* in Khotanese, it is possible that this superlative form is inherited. Most superlatives are synchronically derived from adjectives but several are not such as *ustama-* 'future', *miṣṭama-*<sup>101</sup> 'greatest', *hatama-* 'foremost' and *hastama-* 'best'. It is tempting to regard the double *ḍḍ* as an effort to force the meter to work but I know of no other instances of such a practice. It is also odd that the better attested corresponding comparative *buḍaru* 'more' clearly has single /ḍ/ in ten<sup>102</sup> out of eleven cases in Z. In 13.126 below *buḍaru* is LLL in cadence 4 **LLLLHL**, and in 23.369 it occurs twice, first as three moras in a five mora segment and then it constitutes a three mora segment:

13.126ab riddhā-pāt|yau jsa kalpu jsīna □ | u kyeri | **buḍaru** kṣamīyā .  
(A1:5+9+3+7)

Through the *ṛddhipādas*, one has life for a *kalpa* and for however many more ... one should wish.

---

<sup>101</sup> *miṣṭama-* 'greatest' has an unclear historical relationship with *māsta-* 'great' (Skjærvø, St.I:104 < *\*masha(h)-tama-*, not repeated in SuvII:326b), but cannot be synchronically derived from it.

<sup>102</sup> *buḍaru* is three moras in Z 1.45, 4.91, 6.18, 13.126, 14.97, 15.12, 18.38, 23.367, 23.369 2×. In 13.88 could be interpreted as four moras in order to make up a usually five mora segment:

13.88cd śśo haḍā | sūtro tta hvīnde □ | ṣā **buḍaru** | baśdo nāste (A:5+7+?+7)  
for one day ... So it is said in the *sūtra*: 'The latter does the greater evil.'

23.369ab pharu **buḍaru** | balysä samāhāna | **buḍaru** | mästa vimūha

(A1:5+9+3+7)

The meditations of the Buddha are many more, the great *vimokṣas* more

At best, *buḍamu* constitutes questionable counter evidence.

#### 4.6.5 -*l*tt- and -*ṃ*tt-

The stems showing -*l*tt- and -*ṃ*tt- in those conjugations where where vowel final stems exhibit *VttV*, also convey information about the /tt/ initial suffixes. There are seven stems which show -*l*tt- in the categories 3Sp.a, 3Sp.m, 2Piv.m:

*ḍ*+tt > *l*tt:

*ggaḍ*- < \**gart*- ‘lie about’, 3Sp.m *ggaltte* Z 24.450, cf. 3Ppm *ggaḍāre* Z 2.44.

\**hamggaḍ*- < \**ham-gart*- ‘result, develop’ 3Sp.m *hamggaltte* Z 4.47, 5.84, cf.

3Spf.intr.f *hamggälsta* Z 5.27.

*baḍ*- < \**u*art- ‘move; writhe’ 3Sp.m *baltte* Z 20.57, cf. 3Pp.m *baḍāri* Z 24.503.

*nyūḍ*- < \**ni-u*art- (SGS) ‘rush down’ 3Sp.m *nyūltte* Z 17.14, 24.419, cf. 3Pp.m

*nyūḍāre* Z 17.14.

\**patāḍ*-<sup>103</sup> < \**pati-kart*- ‘cut off’, 3Sp *patälttä* Z 12.49, *patä’ttä* Sgh[1] 37.3, cf.

ppp ASn *patälstu* Z 5.78, Npm *patälste* Z 24.

*l*+tt > *l*tt:

*saṃkhäl*- < \*?-*xard*- ‘be tainted’, 3Sp.m *saṃkhiltte* Z 24.227, *saṃkhilttä* Z

22.259, cf. 3Sop.a *saṃkhalitā* Sgh[17] 37.3, *saṃkhalī* Sgh[1] 37.3, *āmata-*

---

<sup>103</sup> Earlier scholarship gives the stem as *patält*- (Glossar:456b, SGS:67, DKS:206a, Sgh:269b).



noun NSf *saṃkhālāmata* ‘defiling’ Vim 3.39.3.9, ppp + *a*-privative NSm  
*asaṃkhilstā* Z 6.39 ‘unstained’.

*pātāl*-<sup>104</sup> < \**pati*-Hard- ‘prosper’ B, 2Piv.m *pātā’lta* IOL Khot 149/1v1, cf. ppp

*pātālsta*- IOL Khot 148/r6, r7, 148/3v7, *p[ä]tā’lsta*- IOL Khot 148/1r2.

There is one stem, *nuvaṃth*- ‘be removed’, which shows *-ṃtt*- in one of these conjugations: 3Sp.m *nuvaṃtte* Z 20.14, 23.30. This word certainly has double /tt/ even if the cluster were to derive from a suffix with single /t/, i.e., /nuwanttē < nuwant<sup>h</sup>t(t)ē < nuwant<sup>h</sup>- + -t(t)ē/. It is more difficult to tell if the *-lta*- forms contain double /tt/. The historical derivation might suggest single /t/, e.g., \*/-rt- + -tV > -ḍ- + -tV > -ltV-/. There is no obvious reason to claim double /tt/ but arguments in favor of that are presented below in the section on cluster reduction.

#### 4.6.6 The 3Sp.a suffix *-ittā* and vowel final stems.

A major morphophonological pattern that must be explained is the behavior of the 3Sp.a *-ittā* with vowel stems. The suffix contains what is elsewhere the potential for umlaut, traditionally symbolized by superscript <sup>i</sup>. With consonant final stems the potential will be realized as palatalization of one of the stem final consonants or as umlaut of the stem vowel, according to the rules given in Hitch 1990. Some examples:

---

<sup>104</sup> This is the stem form given by Emmerick (SGS:84). The Leumanns gave the stem as *pātāly*- (Glossar:463a s.v. *pātāy*-) as did Bailey (DKS:237a) probably because of the *ly* in 3Pp.m *pātālyāre* Z 23.94, IOL Khot 148/1r2, r6, r7, *pyālyāre* IOL Khot 148/3v7. With all three instances of *pātālyāre* in IOL Khot 148/1r2, there is subscript *pyā* under *pā* indicating the later pronunciation. It may be preferable to regard the *ly* as Late Khotanese influence. The etymology, the ppp, and the fact that no other stems end in *ly* suggests this stem ends in *l*.

*vasuštä* < *vasus-* ‘become pure’ • /wasuště < wasuštĕ < wasus- + <sup>i</sup>ttě/ (Z 3.68)

*nimalśdä* < *nimalys-* ‘rub down’ • /nimalždě < nimalžddě < nimalz- + <sup>i</sup>ttě/ (Z 22.147)

*padaśdä* < *padajs-* ‘burn (intr)’ • /baḍaždě > baḍažddě > baḍad<sup>z</sup>- + <sup>i</sup>ttě/ (Z 11.36)

*paysendä* < *paysān-* ‘recognize’ • /bazēndě < bazēnddĕ < bazān- + <sup>i</sup>ttě/ (Z 4.97)

*jsīndä* < *jsan-* ‘strike’ • /d<sup>z</sup>īndĕ < d<sup>z</sup>īnddĕ < d<sup>z</sup>an + <sup>i</sup>ttě/ (Z 24.418)

*beḍä* < *bār-* ‘rain’ • /βēḍḍĕ < βāɹ- + <sup>i</sup>ttě/ (Z 22.121)

Note also that when the corresponding type B middle suffix, 3Sp.m *-tte*, with no umlaut potential, is attached to a stem final palatal *-ñ-*, the nasal assimilates to the dental and a potential for umlaut is released and drawn towards the stressed vowel (Hitch 1990:12.2.2):

*hvīnde* < *hvañ-* ‘be called’ • /h<sup>w</sup>īndĕ < h<sup>w</sup>īnddĕ < h<sup>w</sup>añ- + <sup>i</sup>ttĕ/ (Z 5.13)

*pajsīnde* < *pajsañ-* ‘be struck’ • /bad<sup>z</sup>īndĕ < bad<sup>z</sup>īnddĕ < bad<sup>z</sup>añ- + <sup>i</sup>ttĕ/ (Z 5.86)

The behavior of the suffix *-ittä* with vowel final stems is different. In this case, what for convenience may still be called the umlaut potential behaves as if it is /i/ and contracts with the vowel into a diphthong. Four contractions are attested. Two stems show *u+<sup>i</sup>uī~vī*:

*rrvītä* Z 15.4 < *rru-* ‘grow (intr)’ • /ruittĕ < ru- + <sup>i</sup>ttĕ/

*hambruītä* Z 22.128 < *hambru-* ‘grow together’ • /hambruittĕ < hambru- + <sup>i</sup>ttĕ/

Eight stems show *a+<sup>i</sup>ai*:

*khaittä* IOL Khot 28/5 < *kha-* ‘wound’ • /k<sup>h</sup>aittĕ < k<sup>h</sup>a- + <sup>i</sup>ttĕ/

*ttähvaittä* Z 13.21 < *ttähva-* ‘cross (a river)’ • /tēh<sup>w</sup>aṭtē < tēh<sup>w</sup>a- + -<sup>i</sup>ttē/

\*[pā]hai[ttä] Sgh[17] 31.6 < *pāha-* ‘strike’ • /bēh<sup>a</sup>ṭtē < bēha- + -<sup>i</sup>ttē/

*pvai’ttä* Z 2.101 < *puṣa-* ‘fear’ • /bu<sup>a</sup>ṭtē < buṣ<sup>a</sup>ṭtē < buṣa- + -<sup>i</sup>ttē/

*baittä* Z 5.55 < *ba-* ‘be bound’ • /β<sup>a</sup>ṭtē < βa- + -<sup>i</sup>ttē/

*bvai’ttä* Z 14.92 < *buṣa-* ‘mount’ • /βu<sup>a</sup>ṭtē < βuṣ<sup>a</sup>ṭtē < βuṣa- + -<sup>i</sup>ttē/

*maitti* IOL Kh 148/2r4 < *ma-* ‘be intoxicated’ • /maṭtē < ma- + -<sup>i</sup>ttē/

*hvaittä* Z 2.16 < *hva-* ‘strike’ • /h<sup>w</sup>aṭtē < h<sup>w</sup>a- + -<sup>i</sup>ttē/

One stem shows *e<sup>i</sup>>ai*:

*saittä* Z 1.35 < *se-* ‘appear’ • /saṭtē < se- + -<sup>i</sup>ttä/

Three stems show umlaut absorption *ä<sup>i</sup>>ä*, in the same way as consonant stems in -*äC*- (Hitch 1990:191 10.2):

*nättä* Z 13.28 < *nä-* ‘sit down’ • /nēttē < nē- + -<sup>i</sup>ttē/

*pajätta* Z 12.42 < *pajā-* ‘ask for’ • /baṭtē < bajē- + -<sup>i</sup>ttē/

*bei’ttä* Z 5.18 < *bäṣa-* ‘release’ • /βēttē < βēṣtē < βēṣ- + -<sup>i</sup>ttē/<sup>105</sup>

Note that the stems ending in -*a-* or -*e-* are members of Emmerick’s type C. With the current analysis, there is no need to retain that categorization. Instead, those

---

<sup>105</sup> A similar orthographic and phonological situation is shown by the type A stems *käṣ-* ‘think’ (SGS *kät’-*) and *häṣ-* ‘send’ (SGS *hei’*) when followed by suffixes beginning in -*ä-*:

3Sp.a *kei’ttä* Z 2.5 + 10×, *kei’yä* Z 19.85 • /kēṭtē < kēṭdē < kēṣdē < kēṣ- + -ēdē/

2Pp.a *kei’ta* Z 20.67, *kä’ta* Z 23.97 • /kēṭda < kēṣda < kēṣ- + -ēda/

3Sop.a (-*ä*) *kei’t* Z 13.31 • /kēṭ < kēṣ < kēṣ- + -ē/

2Piv.a *kei’ta* Z 22.277 • /kēṭda < kēṣda < kēṣ- + -ēda/

3Sp.a *hei’ttä* Z 24.431, *hä’ttä* Sgh [10] 207.2 • /hēṭdē < hēṣdē < hēṣ- + -ēdē/

The nouns *bäṣa-* ‘poison’ and *näṣa-* ‘nectar’ also show /ēṭ/ written *ei’*:

NSm *bei’t* Z 3.76 • /βēṭ < βēṣ < βēṣ- + -ē/

IASm *bei’ña* Z 11.55 • /βēṭña < βēṣña < βēṣ- + -ēna/

IASm *nei’ña* Z 3.85 • /nēṭña < nēṣña < nēṣ- + -ēna/

stems may be more precisely described as type B stems ending in *-a-* or *-e-*. This view is parallel to that expressed in TypeA where it was argued that Emmerick's type D stems may be described as type A stems ending in *-a-* or *-e-*. The four categories of SGS, A, B, C, D, may be reduced to two, A, B, in the synchronic analysis.

The contraction  $a + \overset{i}{\curvearrowright} ai$  as seen in the 3Sp.a of type B stems in *-a-* requires a more rigid definition. As shown in Hitch 2015b (Hitch 2015b), the  $-i$  will also umlaut the stem final /a/ of *aa-* and *aā-* declension nominals. The umlauted vowel and the suffix vowel then immediately contract to a diphthong. The umlauted vowel is never visible on the surface, so its exact quality is difficult to determine. It is tentatively given as short /i/. From the *aā-* declension three suffixes have umlaut potential, GDSf  $-ie$ , IASf  $-ie$  (*jsa*) and LSf  $-ia$ . E.g.:

GDSf *buljsye* Z 24.477 < *buljsaā-* 'virtue' +  $-ie$  • /βuld<sup>z</sup>iē < βuld<sup>z</sup>i-ē < βuld<sup>z</sup>a- +  $-iē$ /

GDSf *baśdye* Z 24.428 < *baśdaā-* 'sin' +  $-ie$  • /βažd<sup>i</sup>ē < βaždi-ē < βažda- +  $-iē$ /

IASf *uysānye jsa* Z 24.265 < *uysānaā-* 'self' +  $-ie jsa$  • /uzāniē d<sup>z</sup>a < uzani-ē d<sup>z</sup>a < uzāna- +  $-iē d^z a$ /

IASf *śśādye jsa* Z 13.145 < *śśādaā-* 'goodness' +  $-ie jsa$  • /šēd<sup>i</sup>ē d<sup>z</sup>a < šēdi-ē d<sup>z</sup>a < šēda- +  $-iē d^z a$ /

LSf *brītya* Z 2.169 < *brītaā-* 'love' +  $-ia$  • /βiīd<sup>i</sup>a < βiīdi-a < βiīda- +  $-ia$ /

LSf *śśaṃdya* Z 5.29 < *śśaṃdaā-* 'earth' +  $-ia$  • /šand<sup>i</sup>a < šandi-a < šanda- +  $-ia$ /

From the *aa-* declension one suffix, LSm  $-ia$ , conveys umlaut. E.g.:

LSm *drau-mūjsya* Z 6.4ab < *drau-mūjsaa-* 'hair pore' • /drau-mūd<sup>z</sup>ia < drau-mūd<sup>z</sup>i-a < drau-mūd<sup>z</sup>a- +  $-ia$ /

L<sub>Sm</sub> *hvanya* Śgs 3.6v1 < *hvanaa*- ‘exposition’ (corrected by Skjærvø, IOL Khot 11/3v1 note d) • /h<sup>w</sup>an̄ja < h<sup>w</sup>ani-a < h<sup>w</sup>ana- + -<sup>i</sup>a/

Thus there are three distinct effects of the umlaut potential on /a/.

1. As well known for around a century, with consonant final stems, either nominal or verbal, an /a/ in the last syllable of the stem will umlaut to /ī/ when the intervening consonants are transparent to umlaut, e.g., ASf *tcīru* Z 21.22 < *tcari*- ‘face’ + -<sup>i</sup>u, infinitive *khīṣṭe* < *khaṣṭa*- ‘drink’ + -<sup>i</sup>e.
2. As shown in Hitch 2015b and illustrated above, with the nominals, a stem final /a/ will umlaut to /i/ and feed a contraction with the suffix initial vowel producing a diphthong, e.g., GDSf *buljsye* Z 24.477 < *buljsaā*- ‘virtue’ + -<sup>i</sup>e
3. As argued here, with the 3Sp.a suffix -<sup>i</sup>ttā, a type B present stem final /a/ will form the diphthong /aj/ with the umlaut potential, e.g., *khaittā* < *kha*- ‘wound’ + -<sup>i</sup>ttā

These effects occur in differing phonological environments. The critical difference between 2 and 3 seems to be that in 2 the umlaut potential is associated with a vowel initial suffix while in 3 the umlaut potential is associated with a consonant initial suffix. The three environments may be formulated as follows:

1. /-aC(C)- + -<sup>i</sup> > -īC(C)-/
2. /-a- + -<sup>i</sup>V > -i-V > -īV/
3. /-a- + -<sup>i</sup>C- > -ajC/

It is probably not relevant that 2 is attested with nominal stems and 3 with verbal.

The pattern in 2 could in theory be tested with present stems in -a- and the three

conjugational suffixes with the shape  $-iV$ : 2Sp.a  $-i\ddot{a}$  ( $-i$ ), 1Sop.a  $-i\ddot{o}$ , 3Sop.a  $-i\ddot{a}$  (SGS:192, 206–207). For example, one would predict 2Sp.a  $*p\ddot{a}hy\ddot{a}$  <  $p\ddot{a}ha-$  +  $-i\ddot{a}$  ‘you strike’ or 3Sop.a  $*khy\ddot{a}$  <  $kha-$  +  $-i\ddot{a}$  ‘may you wound’. However, as far as I can tell there are no conjugated forms of these types in the corpus.

Only one type B present stem in long  $/\bar{a}/$  is known,  $b\ddot{a}\ddot{s}\bar{a}$ - ‘release’, which is attested once. 3Sp.a  $bei'tt\ddot{a}$  counts **LHL** in cadence 1 **HLLHL** in a hemistich from Z, reflecting phonemic  $/\beta\ddot{e}a\ddot{i}t\ddot{t}\ddot{e}/$ , pre-Khotanese  $*b\ddot{a}\ddot{s}a\ddot{i}t\ddot{t}\ddot{a}$ :

Z 5.18 samu kho haṃ|būvu **bei'ttā** .□|harbiśī | āchai jīye . (A:5+7+5+7)

Just as when one releases<sup>106</sup> a fester all pain disappears for one,

With nominal and verbal stems ending in a consonant, a long  $-\bar{a}$ - will umlaut to  $-e-$ , e.g.  $bete\ jsa$  Z 4.3 <  $b\bar{a}t\bar{a}$ - ‘wind’, LSm  $eha$  Z 20.41 <  $\bar{a}ha$ - ‘mouth’, 3Sp.a  $beḍ\bar{a}$  Z 22.121 <  $b\bar{a}r$ - ‘rain’. There are nominals ending in long  $/\bar{a}/$ , from the  $\bar{a}a$ -declension. Here, an umlauting suffix, LS  $-i\ddot{a}$ , does not modify the stem vowel, e.g., LSm  $nit\bar{a}ya$  Z 5.75 <  $n\bar{a}t\bar{a}a$ - ‘river’, LSm  $haḍ\bar{a}ya$  Z 2.134 <  $haḍ\bar{a}a$ - ‘day’. The principle which appears to be operating here is that a  $/y/$  is inserted between the stem long vowel and the suffix vowel. This may also be observed with the GDSm suffix  $-i$ , e.g. GDSm  $n\bar{a}t\bar{a}yi$  Z 13.51,  $haḍ\bar{a}yi$  Z 11.17. The umlaut potential is absorbed by the  $/y/$  after epenthesis. With the one type B present stem in  $-\bar{a}$ -,  $b\ddot{a}\ddot{s}\bar{a}$ -, the umlaut potential in the suffix  $-i\ddot{t}\ddot{t}\ddot{a}$  contracts with the suffix vowel to  $ai$ . There are thus also three environments involving  $-\bar{a}$ - and  $-i\ddot{a}$ :

1.  $/-\bar{a}C(C)- + -i\ddot{a} > -eC(C)-/$
2.  $/-\bar{a}- + -i\ddot{V} > -\bar{a}-y-i\ddot{V} > -\bar{a}yV/$

---

<sup>106</sup> Emmerick ‘treats’, Bailey ‘opens’.

3. /-ā- + -<sup>i</sup>C- > -aḷC/

#### 4.7 Other suffixes with the *VttV* exhibiting stems

Among the conjugational suffixes, the only ones beginning with a consonant are those beginning with /tt/. There appear to be derivational suffixes which synchronically seem to begin with a consonant and are attested with vowel final stems including the *VttV* stems. These are the deverbal noun suffix *-(a)na-* or *-(a)naa-*, and possibly the causative suffix *-ñ-*. In addition, the causative suffix *-āñ-* provides evidence of stem shape.

##### 4.7.1 The deverbal noun suffix *-na-* and vowel stems

The *(a)na-* and *(a)naa-* deverbal noun suffixes are traditionally listed as *-ana-* and *-anaa-* because an *-a-* appears with consonant final stems, as in *āljsana-* ‘song’ < *āljs-* ‘sing’, LKh *pasūjsana-* ‘fuel’ < *pasūjs-* ‘burn’ (KhSuf:25 §3.A.9). There is no suffix initial *-a-* with the vowel stems<sup>107</sup>. There may be four or five vowel final stems showing these suffixes.

---

<sup>107</sup> Degener gives the traditional view: “Im allgemeinen tritt *-ana-* unmittelbar an den Präs.st., z. B. *āljsana-* von *āljs-*, *pasūjsana-* von *pasūjs-*. Endet der Präs.st. auf *-d-*, so geht dies intervokalisches verloren, z. B. *pagyūna-* von *pajud-*, *puva’ṇa-* von *puva’d-*” (KhSuf:24 §3.A.9). *pagyūna-* ‘covering’ < *paju-* ‘conceal’ does not contain *-(a)na-* as Degener (KhSuf:26) but rather the denominal verb suffix *-ūna-* as in *hayirūṇa-* ‘joy’ < *hayār-* ‘rejoice’, *pīrūna-* (LKh) ‘painting’ < *pīr-* ‘write, paint’ (KhSuf:167–169). Some of the other forms listed by Degener under *-ana-* may not belong there. *āphārana-* ‘disturbance’ is not synchronically directly derived from *āphār-* ‘be disturbed’ but is influenced by *āphārgyā-* ‘disturbance’ (SuvII:242a). *ātīna-* ‘mirror’ and *āyīnaa-* ‘example’ appear to show *-īnaa-* but that is otherwise a denominal adjective suffix (KhSuf:133–152). There is no other evidence for a stem *\*āyī-* which could be joined with *-(a)na-*. M. Maggi p.c. offers, “Possibly Kh. *ātīna-* ‘mirror’ and *āyīnaa-* ‘example’ are inherited from Ir. *\*ā-daiH-ana(-ka)-* (cf. Avestan *daēnā-*), while *āyāna-* ‘mirror’ and *āyānaa-* ‘example’ are analogical with the metanalysed verb.”

**āyā-** ‘appear’ (semantics St.I:24–25 Emmerick), SGS, DKS *āy-*, EDIV:48–49 *āyā-*, < \**ā-daiH-* (or \**ā-daiH-ja-* if *dä-* ‘appear’ is *daiH-ja-*; see footnote 38)

*āyāna-* ‘mirror’ • /*āyēn-* < *āyē-* + -n- /:

L<sub>Sm</sub> *āyāña* Z 4.100, 9.27, 24.489

I<sub>ASm</sub> *āyārna* (for expected \**āyānāna*) Z 24.491 either HLLL or HHL

*āyānaa-* ‘example’ • /*āyēna-* < *āyē-* + -na- /:

G<sub>D</sub>Sm *āyānai* Z 19.89

N<sub>Sm</sub> *āyānai* Z 19.93

**puša-** ‘fear’ SGS *puva’d*, SuvII *puvay’-* (s.v. *puva’ña-*, *pvay-*), DKS *puvqđ-*, *pvqđ-*. Note that synchronically suffixation takes place before the loss of intervocalic -*ṣ-* as shown by the spread of retroflexion to the -*ṇ-*. Just a selection of forms is given.

*puva’ña-* (*pva’ña-*, *pvaṇa-*, *puvaṇa-*) ‘fear’ • /*buaṇ-* < *buzāṇ-* < *buza-* + -n- /

I<sub>AS</sub> *puva’ñāna* SI M20.2b1, *pva’ñāna* Z 12.131, *pvaṇāna* Suv[] 3.6

N<sub>AP</sub> *puva’ña* Sgh[] 55.2(2x), 55.3, 79.8, 199[1], Z 5.89, 22.218; *pva’ña* Z 4.114,

Sgh[] 213[3], 253[43], 253[53]

**ba-** ‘be bound’ SGS *bad-*, DKS *bad-* s.v. *bañ-*, *baittä*.

*bana*-<sup>108</sup> ‘bond’ • /*βan-* < *βa-* + -n- /

N<sub>Sm</sub> *bani* Z 20.14, *banä* Vim 3.93.3v3

L<sub>Sm</sub> *bano* Sgh[] 213[4]

---

<sup>108</sup> In KhSuf:9 about *bana-* Degener wrote, “Zu *bañ-* binden” (KhSuf:9). She was indicating a relationship or association with *bañ-*, not a derivation for which her “idea was rather \**band-a-*” (email 8 November 2012). Here, historically *bana-* < \**band-a-*, but after the *VttV* metanalysis and the development of the stem shape *ba-*, the noun was also metanalyzed [*ban*][*a-*] > [*ba*][*na-*].



NAPm *bana* IOL Khot 22/6 a1

IAPm *banyau* Sgh[] 212.1

*banaa*-<sup>109</sup> ‘prisoner’ • /βana- < βa- + -na-/

NAPm *banā* Z 5.100, \**ban*[ā] Sgh 115.2

**paba-** ‘be bound’ is attested in Old Khotanese only in the (*a*)*na*-deverbal noun *pabana-* which is translated ‘continuum, connexion, nexus’ by Bailey (DKS:212a), ‘connexion’ by Emmerick (SGS:71 and Z 5.81), and ‘rebirth, nexus’ by Emmerick (Śgs:124b). The present stem is also reflected in the causative stem *pabañ-* ‘bind’ which is attested only in LKh 3Pp.a *pabañīda* (SGS:71).

*pabana-* ‘nexus, connection’ • /baβan- < baβa- + -n-/.

NSm *pabanä* Z 5.81, Śgs 3.7r5\*, 7v3, KT3 53.a1, IOL Khot 147/3.v5

ASm *pabanu* Śgs 3.7v3, 14v2, Suv[] 10.63

IAPm *pabanyau* Suv[] 18.91

**prahau-** ‘put on (clothing)’, SGS, DKS *prahauy-*, SuvII *prohauy-*. The present stem occurs three times, always in the participle of necessity NAPm *prahauyāñā* Suv[Or] 1.21, 6.3.16, *prohauyāñā* Suv[Or] 6.3.28. The stem final -y- perceived by earlier scholarship may be epenthetic. Compare *skau-* ‘touch’ where a -y- is also synchronically inserted before long vowels.<sup>110</sup> Besides the ppp *prahausta-*, the only

---

<sup>109</sup> Degener has *bana-* + -aa- (KhSuf:18).

<sup>110</sup> With -y- before long vowel: 1Ssj.a *skauy[īñā]* Suv[C] 3.97, 3Ssj.m *skauyāte* Z 22.148, NSf ? *skauyāñā* Or. 12637/36a1. Without: 3Sp.m *skaute* Z 22.144 (*skau-* + -āte), 3Sop.a *skvaiya* Z 13.56 (*skau-* + -<sup>i</sup>a). The evidence for *ysau-* and epenthetic -y- is from LKh: 3Sp.m *ysautte* SI 153r3, *āñā-p.nec ysauyāñā* KT1 1.100r2. Only one form of *vanau-* ‘become inactive’ is attested, 3Sp.m *vanautai* Z 4.72, 7.25 (*vanau-* + -āte + -ī). *sau-* is found only in LKh and only

other form possibly from this verb is the noun *prahaṇa*- ‘clothing’. The noun occurs more than 25× in Old Khotanese, always with retroflex *-ṇ-*. In Late Khotanese it seems to frequently if not always have dental *-n-*. Bailey gives “Kuci Sanskrit *parhūṇa*,”<sup>111</sup> without gloss, but apparently suggesting the possibility of a loanword in Khotanese. T. Burrow suggested (1934:514) and H. Lüders agreed (1936:35) that Niya Gāndhārī *prahoni*, which occurs in just one document (318), was connected to Khotanese *prahaṇa*-. If a loan into Khotanese, the present stem *prahau-* could be a back formation. The speakers may have reinterpreted a loanword simplex [*prahaṇa-*] as [*prahau*][*na-*].

(a)*na*-noun *prahaṇa*- ‘clothing’ • /praha<sup>u</sup>na- < praha<sup>u</sup>- + -na-/

N<sub>Sm</sub> *prrahaṇä* Z 2.63, IOL Khot 20/5r1, *prahaṇä* IOL Khot 150/1v7

(Neb:118), *prahoṇä* Z 11.36

A<sub>Sm</sub> *prahaṇu* Kvb 0.56, 216.1, Suv[Or] 13.17, 18.96, *prrahoṇu* Sgh[17]

243[31], *prahoṇu* Sgh[1] 35.1, IOL Khot 147/3r2

A<sub>Sm</sub> *prahoṇä* Sgh[10] 35.1

I<sub>ASm</sub> *prrahaṇäna* Suv 8.36, *prrahoṇäna* Z 21.32, *prahaṇna* Suv 8.3

---

with the *āṇa*-p.nec *sauyāṇa*- ‘must be rubbed, ground’. There is *y*-insertion also in *nūyāre* < *nū-* + *-āre* discussed above. Also it may be found in the GDS of the *āa*-declension, e.g. *nātāyi* Z 13.51 ‘river’, *haḍāyi* Z 11.17 ‘day’, and perhaps in LS *nitāya* Z 5.75, *haḍāya* Z 2.134, which, if correct means that the *au*-declension LS also has it, e.g. *amatauya* IOL Khot 155/5v1 (\*Z) ‘disaster’, *hamauya* KV 54.1 ‘goblet’ (St.III:165 Maggi). There is also *y*-insertion when the 3S enclitic pronoun *-ä* attaches to a word ending in a long vowel. Traditionally this is written as a separate word *yä*. Compare also L<sub>Sm</sub> *nātātuṣo*’ Sgh[1] 36.3 ‘river’ with inserted non-etymological *t* after the long vowel.

<sup>111</sup> He also lists a second shape, *parhyaṇa* and refers to “P. Bagchi, Deux Lexiques Sanskrit-Chinois I 325: II 1267” (DKS:255a), which I have not seen.

NAPm *prahaṇe* Suv[Or] 1.21, 15.30, *prahone* Z 3.44, 55, 87, 4.36, 5.34, Suv[Or]

6.1.2, Sgh[17] 253[54], 253[59]

IAPm *prahaṇyau* Suv[Or] 1.10, *prahonyau* Sgh[3] 125.3, 131.2

#### 4.7.2 The causative suffixes *-ñ-* and *-āñ-* with vowel stems

There appear to be several synchronic causative suffixes in Old Khotanese. A full discussion is reserved for elsewhere. With several present stems ending in *-a-*, there may be a suffix *-ñ-* which gives causative meaning. These stems reflect OIr *\*bad-* and *\*ščad-* which in Khotanese have been metanalyzed to *(-)ba-* and *-tca-*.

**pabañ-** ‘bind’ < *paba-* ‘be bound’. *paba-* is attested in the *(a)na*-noun *pabana-* discussed above. The causative stem *pabañ-* is attested only in LKh.

LKh 3Pp.a *pabañīda* (SGS:71)

**bañ-** ‘bind’ < *ba-* ‘be bound’

3Sp.a *bañāte* Z 11.36, 50, *bañite* Z 13.28, 24.444

3Pp.a *bañīndi* Z 22.324

**\*vatcañ-** ‘cripple’ < *\*vatca-* (< OIr *\*ava-ščad-*) suggested by the ppp *vatcasta-* which is attested once (SGS:117 compares *hatcañ-*).

**hatcañ-** ‘break (tr)’ < *hatca-* ‘be broken’ (< OIr *\*fra-ščad-*) attested in OKh as 3Sop.m *hatcai* Śgs 3.13v3 • /hat<sup>s</sup>a<sup>e</sup> < hat<sup>s</sup>a- + -ě/.

3Sp.a *hatcañāte* Z 6.34

3Pp.a *hatcañīndi* Z 22.198

**hambañ-** ‘compose (writings)’ < *hamba-* ‘be bound together’ which seems to be known only in LKh 3Sp.a *hambette* (SGS:142, DKS:461b, 463b) reflecting OKh \*/hamba<sup>1</sup>ttě < hamba- + -<sup>1</sup>ttě/.

3Pp.a *hambañindä* Z 24.385

The causatives (-)bañ- and -tcañ- historically reflect OIr \*-band-*a*ja- and \*-sčand-*a*ja-. It may be argued that they are inherited simplexes. The pattern seen in *ba-* : *bañ-* :: *hatca-* : *hatcañ-* led to reanalysis, [*bañ*] > [*ba*][*ñ*] and [*hatcañ*] > [*hatca*][*ñ*], which created the -ñ- synchronic causative suffix.

While there may be doubt about the synchronic status of -ñ-, there should be no doubt about -āñ-. The suffix is well attested with stems ending in a consonant. Some Old Khotanese examples:

*paṣṭāñ-* ‘satisfy; promote’ < *paṣṭ-* ‘arise, set out’ (SGS:77)

*handajāñ-* ‘ripen’ < *handaj-* ‘be ripened’ (SGS:140)

*huṣāñ-* (SGS *hvā’ñ-*) ‘make dry’ < *huṣ-* ‘become dry’ (SGS:153, 157)

When attached to a stem ending in -a-, there is coalescence  $a + \bar{a} > \bar{a}$ ,

*puṣāñ-* (SGS *pvā’ñ-*) ‘frighten’ < *puṣa-* (SGS *puva’d-*) ‘be afraid’ (SGS:89, 85)

*bitcāñ-* ‘break up’ LKh (SGS:96) implies \**bitca-* < \**vi-sčad-* (cf. DKS:283a)

With other vowels there is semivocalization:

*ysyāñ-* ‘cause to be born’ (SGS:114) < *yse-* (SGS:114 *ysai-*) ‘be born’ • /zyāñ- < zē- + -āñ-/

*dyāñ-* ‘make appear, reveal’ (SGS:48) < *dā-* (SGS:45 *did-*) ‘appear’ • /ḍyāñ- < ḍē- + -āñ-/

The causatives *puṣāñ-* (*pvā'ñ-*) 'frighten' and *dyāñ-* 'make appear' are derived from the metanalyzed *VttV* stems *puṣa-* 'be afraid' and *dā-* 'appear'. The causative shapes show no synchronic trace of historical stem final *\*-d-*.

#### 4.8 The historical evolution of the metanalysis

The metanalysis was facilitated by the convergence of some historical processes which led to the existence of variants of the types *CiyV~CyV*, *CäyV~CyV*, and *CuvV~CvV*. The easiest to observe is the rise of variants within stems. Below are given the examples *ttuvāy-~tvāy-* and *päyaurā-~pyaurā-*. Historically the longer forms precede, but synchronically they are treated as stylistic variants, used in Z where convenient for the meter. The easiest process to understand is the development of the secondary declensions, mostly from loss of intervocalic *-k-*. The endings *\*-CikV* and *\*-CukV* developed to *-CiyV* and *-CuvV*, and then to *-CyV* and *-CvV*. For instance, *\*narikV* > *\*nariyV* > *naryV* for *naria* - 'hell' and *\*bāzukV* > *bāysuvV* > *bāysvV* for *bāysua* - 'arm'. A second process relates to verb stems which end in *-u-* and *-i-*. With *ākṣu-* 'begin' the development is something like *\*ā-xšaip/b-V* > *ākṣuvV* > *ākṣvV* and with *āri-* 'be mixed' it is perhaps *\*ā-raiθ-V* > *\*āriδV-* > *āriyV* > *āryV*. The last two historical processes are largely parallel. In both, the longer variants *-CiyV*, *-CäyV*, and *-CuvV* developed first, as they did with the stem internal process first listed. But, at least in the last two cases, the shorter variant became regarded as synchronically more basic, and the longer as more derived. That is, historically *-Ciyv* > *-CyV* and *-CuvV* > *-CvV* but in the synchronic derivation, *-CyV* > *-CiyV* and *-CvV* > *-CuvV*. The presence of these variants facilitated the metanalysis, encouraging, for example, *buvāre* (< *\*budāre*) to become *bvāre* and inducing the historical fact of a

stem final consonant (\*-v- < \*-d-) to be forgotten. When the stem became *bu-* the suffix in *butte* became *-tte*.

#### 4.8.1 Glide resolution

A possible factor leading to the metanalysis involved the historical processes which created variants of the types *CyV~CiyV* (*CyV~CäyV*) and *CvV~CuvV*. These variations matched a possibly pre-existing pattern of glide resolution (Hitch 2014:17–23) and were reinterpreted according to that pattern.

The process of glide resolution permitted the composer of the *Book of Zambasta* to use variants of stems with resolved and unresolved glides where needed to make up the meter. A shape *CvV* may resolve to *CuvV*, and a shape *Cyv* may resolve to *CiyV* or *CäyV*. The resolved forms within stems are a mora longer. The two mora stem *ttvāy-* ‘convey’ may be resolved to three mora *ttuvāy-*. Below, 3Sp.a *tvāyātā* with unresolved stem is four moras HLL in a five mora segment:

6.19bd    kho ṣā dhā|raṇā kṣaya-nāḍa □ | ce **tvāyātā** | harbiśśā vāma .

(A:5+7+5+7)

as this *Kṣayanātā-dhāraṇī*, which conveys across all seas.

3Ssj.a *ttuvāya* with resolved stem is **LHL** in cadence 1 HLLHL:

24.239    ṣā muho nve | bāgyo □ ttāru | tcalco **ttuvāya** (B:5+6+7)

He would convey us across to the further shore instead of<sup>112</sup> a ship.

---

<sup>112</sup> On *nve bāgyo* ‘instead of a ship’ see Skjærvø, St.III:112.

In the two hemistichs below, the variants NSf *pyaura*~*päyaura* ‘cloud’ have differing metrical values. In the first hemistich below, *pyaura* is **HL** in cadence 1 **HLLHL**, while in the second hemistich below, *päyaura* is **LHL** in cadence 1 **HLLHL**.

4.107cd kho ju anā|bhoggāna mästu □| ātāsi | ggaljāte **pyaura** (A:5+7+5+7)

As without effort a cloud thunders loudly in the sky,

24.475ab ttye haḍai | sarbāte □ kāḍe | mästä **päyaura** (B:5+6+7)

On this day, a very big cloud rises up.

The composer of Z chose to resolve the clusters in *tvāy*- and *pyaurā*- for metrical reasons.<sup>113</sup>

Variations in stems of the type CrV~CVrV may also be related. In 3.3.2.8 *brūñ*-~*bārūñ*- ‘shine’, and *brāh*-~*bārāh*- ‘soar up’. Another prominent example is the stem for ‘particle, sand-grain’ which features a range of spellings, of which some vary according to this principle, e.g., *grvīca*~*ggurvīca*-, *gruīca*~*gguruīca*- (cf. Glossar:419a).

## 4.8.2 Secondary declensions

The synchronic contraction behavior of the secondary declensions has been detailed in Hitch 2015b and Hitch 2016a. Diachronically the declensions arose through the loss of intervocalic consonants. Many of the words in question earlier had a *\*-kǎ-* suffix. Emmerick described the development this way: “The intervocalic

---

<sup>113</sup> The stem variations *syatā*- ~ *sāyatā*- ‘sand’ and *nyaṇḍa*- ~ *niyaṇḍa*- ‘nigrantha’ are also likely due to glide resolution but no certain metrical contrasts are available. ASm *byanu* Z 3.67 ‘obstacle’ counts LLL showing copyist replacement of original *biyanu* which is attested in that spelling at Z 22.309 where it counts LLL.

\*-k- was voiced, disappeared, and the resulting hiatus was closed by vowel contraction” (SGS:295).

#### 4.8.2.1 \*-ikV-

The lenition likely proceeded in stages. With original \*-ikV- the stages may have been \*-ikV- > \*-igV- > \*-iyV- > -iyV-<sup>114</sup> > \*-iV- which would then give one of three synchronic realizations depending on the quality of the suffix vowel V: 1 coalescence to  $\bar{i}/\bar{i}/$ , 2 diphthongization to  $yV/\bar{i}V/$ , or 3 semivocalization to  $yV/yV/$ . Following those realizations, further possible synchronic developments are 4 diphthong resolution  $CyV > CiyV \sim C\ddot{a}yV / C\bar{i}V > CiyV \sim C\ddot{e}yV/$ , and 5 glide resolution  $CyV > CiyV \sim C\ddot{a}yV / CyV > CiyV \sim C\ddot{e}yV/$ . The difference between the last two is visible in the meter of Z where 4 shows no mora increase while 5 shows a one mora increase as with *pyaurā- > päyaurā-* and *ttvāy- > ttuvāy-* above. It is important to recognize the five distinct morphophonological outcomes of \*-ikV- in Old Khotanese. Many of the examples of \*-ikV- come from Prakrit and these may have been borrowed with the \*k at an advanced stage of lenition.

##### 1. coalescence

*i+ä>ī*: NSm *Anāhapiṇḍī* Z 22.209 proper name cf. Skt *Anāthapiṇḍika*;

*i+i>ī*: GDS *gyaḍī* Z 2.133 < *jaḍia-* ‘folly’ cf. Skt *jaḍa* ‘stupid’.

##### 2. diphthongization

*i+e>ye*: NAPf *ūvāysye* Suv[Or] 6.1.40 ‘lay woman’ cf. Skt *ūpāsikā*;

*i+a>ya*: LSm (-<sup>i</sup>a) *narya* Z 3.72 < *naria-* ‘hell’ cf. Skt *narikā*;

---

<sup>114</sup> This stage is deliberately not starred. The reason is given below.



*i+o>yo*: LSm (-<sup>i</sup>o) *naryo* Z 13.72.

3. semivocalization GDP *ūvāysyānu* Suv[Or] 6.1.63 cf. Skt *ūpāsikā*;

4. diphthong resolution NAPm *anyattīrthiya* Z 2.1 'heretic' cf. Skt *anyatīrthika*;

5. glide resolution GDP *paḍāṃjsiyānu* Śgs 37r5 'former' reflects the \*-ika- > -ia-  
adjective suffix (cf. Konow 1932:61, KhSuf:75–76).

#### 4.8.2.2 \*-ukV-

With original \*-ukV- the stages may have been \*-ukV- > \*-ugV- > \*-uyV- > -uwV-<sup>115</sup> > \*-uV- which would then in parallel fashion give one of three synchronic realizations depending on the quality of the suffix vowel *V*: 1 coalescence to *ū* /*ū̄*/, or 2 diphthongization to *vV* /*ū̄V*/, or 3 semivocalization *vV* /*wV*/. Here also two further developments can occur: 4 diphthong resolution *CvV*>*CuvV* /*Cū̄V*>*CuwV*/, and 5 glide resolution *CvV*>*CuvV* /*CvV*>*CuwV*/. The difference between the last two is again visible in the meter of Z where 4 shows no mora increase while 5 shows a one mora increase.

##### 1. coalescence

*u+ä>ū*: NSm *bāysū* Z 2.139 < *bāysua*- 'arm' < \**bāzuka*-, cf. Av *bāzu*- 'arm'.

*u+u>ū*: ASm *ysānū* Z 22.149 < *ysānua*- 'knee' < \**zānuka*-, cf. Av *zānu*- 'knee'.

##### 2. diphthongization

*u+a>va*: NAPm *baysva* Suv[H] 4.11 < *bāysua*- 'arm' ;

*u+e>ve*: NAPm *bāysve* IOL Khot 166/2v4 < *bāysua*- 'arm' ;

---

<sup>115</sup> This -uwV- stage is deliberately not starred. The reason is given below.

*u+i>vī~uī*: GDS *busvī* IOL Khot 160/1v4 < *busua*- ‘firewood’ < \**abi-sauka*- cf.

Av *ātrə-saoka*- ‘fuel’ (DKS:301b).

*u+o>vo*: (underlying diphthong resolved:) ASf *ṣṣuvo* Z 22.333 < *ṣṣuā*- ‘report’

\**śraukā* cf. Av *sra vah*-, Manichean Sogdian *srwq* ‘speech’ (DKS:412a).

3. semivocalization GDPm *ysānvānu* Z 23.162 < *ysānua*- ‘knee’ < \**zānuka*-, cf. Av *zānu*- ‘knee’.

4. diphthong resolution NAPm *bāysuve* Z 21.27 < *bāysua*- ‘arm’ (as above).

5. glide resolution GDPm *kṣīruvānu* SI M13.5r4 < *kṣīrua*- ‘of the country’ reflects the \**-uka*- > *-ua*- adjective suffix (cf. Konow 1932:61).

Both the \**-ikV*- and \**-ukV*- evolutions had a late stage with an intervocalic semivowel, *-iyV*- and *-uvV*-. It is plausible that these shapes persisted as alternates to the regular contracted forms and then were reinterpreted as resolved variants.

#### 4.8.2.3 \**-akV*-

The development of the *aa*- and *aā*- declensions was somewhat different. In the case of the non-umlauting suffixes, the lenition was perhaps \**-akV*- > \**-agV*- > \**-ayV*- > \**-ahV*- > \**-aV*- which would then give one of two processes depending on the quality of the suffix vowel *V*: 1 coalescence *a+a>ā* /*a+a>ā̄*/, or *a+e>e* /*a+ē>ē̄*/, or *a+o>o* /*a+ō>ō̄*/; or 2 diphthongization to /*aV̄*/ (*ai* /*aī*/, *ei* /*aē*/, *au* /*aū*/).

1. coalescence

*a+a>ā*: NAPm *āchā* Z 3.144 < *āchaa*- ‘disease’;

*a+e>e*: NAPf *bitame* Z 6.23 < *bātamaā*- ‘doubt’ < \**ui-tamH-a-kā*- (cf.

EDIV:376);

$a+o>o$ : ASf *baśdo* Z 13.87 ‘sin’ < *baśdaā-* < *\*bazdi-ǰā-kā-*, cf. Av *bazda-* ‘ill’

(SuvII:311b).

## 2. diphthongization

$a+ä>ei(>ai)$ : NSm *āchei* Z 13.108 < *āchaa-* ‘disease’ < ?<sup>116</sup>;

$a+i:ai$ : IAS *āstaina* Z 5.8 < *āstaa-* ‘bone’ < *\*asta-ka*, cf. Av *ast-* (SuvII:245a).

In the case of the umlauting suffixes the lenition is harder to formulate. We need to address the evolution of umlaut at the same time. Corresponding to the *-iyV-* and *-uvV-* stages above the stage here is conceptualized as *\*-aĥV-*. The voiced laryngeal [ĥ] would be transparent to umlaut, cf. the transparency of the voiceless laryngeal in LSm *eha* Z 20.41 < *āha-* ‘mouth’. So it is conceivable that at this stage of the evolution there was *\*-īĥV-* which would further evolve as *\*-īĥV-* > *\*-īV-* > *-iyV-* > *-yV-* or perhaps *\*-īĥV-* > *\*-īV-* > *-yV-* > *-iyV-*. These umlauted and resolved forms also contributed to the metanalysis as they added more variants of the type *CyV~CiyV* (*CäyV*).

## 6. umlaut and diphthongization

$a+<sup>i</sup>e>ye$ : GDSf *baśdye* Z 24.428 < *baśdaā-* ‘sin’;

$a+<sup>i</sup>a>ya$ : LSf *śśaṃdya* Z 5.29 < *śśaṃdaā-* ‘earth’.

## 4. diphthong resolution GDSf *baśdiye* Z 13.81.

### 4.8.3 Historically vowel final present stems

The next set of processes to be considered are the contractions featured by present stems which are historically vowel final. These are in many cases identical

---

<sup>116</sup> Etymologies are offered by Bailey (DKS:16b) and Skjærvø (SuvII:241a). Konow tentatively compared Av *aχti* ‘pain’ (1932:113a).

to the contractions exhibited by the nominals, and in all cases phonologically consistent with them. The present stems ending in short /i/, /u/ and /a/, and not derived through metanalysis,<sup>117</sup> present the same five shapes and associated processes as nominal stems ending in the same vowels discussed in the preceding section.

/i/ *pari-* ‘order’, *hāvi-* ‘appropriate’, *ji-* ‘cease’, *kṣi-* ‘be troubled’

1. coalescence 3Sp.a (-i- + -ätä) *jīte* IOL Khot 8/1v3, *paritā* Z 24.450
2. diphthongization 2Siv.a *parya* Z 23.52
3. semivocalization 3Pp.a *vyāre* Z 3.144, *kṣyār[e]* Z 21.23, *hāvvyāre* Z 4.68, p.nec  
NAPm *paryāñña* Suv[Or] 6.3.23
4. diphthong resolution (cf. *diya* below)
5. glide resolution 3Pp.a *jiyāre* Z 6.1, *kṣiyāre* Z 2.43

/u/ *ju-* active ‘live’, *ju-* middle ‘fight’, *ākṣu-* ‘begin’

1. coalescence 3Sp.a (-u- + -ätä) *ākṣūtā* Z 4.15, *jūtā* Z 11.47 ‘live’
2. diphthongization 2Siv.a *ākṣva*
3. semivocalization  
*u+ā>vā*: 3Pp.m *vyāre* Z 1.33 ‘fight’  
*u+i>vī*: 3Pp.a *ākṣvīndā* Z 15.111, 1Ssj.a *juvīñi* Sgh[1] 39.2 ‘live’
4. diphthong resolution (cf. *buva* below)

---

<sup>117</sup> A possible exception is *ju-* middle ‘fight’. No *VttV* conjugation is attested, but deriving from \**Hīaud-* ‘fight’, cf. Av *yūδ-* ‘fight’ (EDIV:176), and with a past stem *justa-*, a 3Sp.m \**jutte* might be predicted.

5. glide resolution 1Pp.a *ākṣuvāma* Z 23.43, 3Pp.m *juvāre* Z 2.46, 3Pp.a

*ākṣuvindä* Z 24.499

/a/ *nāmandra-* 'invite', *buysa-* 'extinguish', *pa-* 'protect', *ttätsa-* 'cross', *paṭha-*

'burn', *naltsa-* 'go out'

1. coalescence  $a+\bar{a}>\bar{a}$ : 1Pp.a (-*āmä*) *pāmä* Suv[Or] 6.1.39, 3Pp.m (-*āre*) *vatsāre*

IOL Kh 159/7r2, p.nec (-*āñā-*) NSm *paṭhāñä* Suv[Or] 8.23

2. diphthongization

$a+\bar{i}>ai$ : 1Sp.a (-*īmä*) *nimandrainä* Z 2.50, *buysaimä* Z 20.23;

$a+\bar{a}>ei>ai$  3Sp.a (-*ätä*) *pathaiyä* Z 2.175, *nāmaṃdraiyä* Z 24.465, *ttätsaiyi* Z

13.27; 3Sp.m *naltseye* Z 22.202;  $\bar{a}$ -infinitive *buysai* Sgh[10] 205.1.

$a+u>(au)>o$  2Siv.m (-*u*) *naltso* Z 5.30; 2Siv.a (-*u*) *buyso* Z 5.50.

In addition, as shown in TypeA, there is a set of type A verbs in *-e-*. In contraction with *-ī-* they behave like stems in *-a-* ( $e+\bar{i}>ai$ ) but in contraction with *-a-* or *-ā-* they behave like stems in *-i-* ( $e+a>ya$ ,  $e+\bar{a}>yā$ ).

/e/ *de-* 'see', *uysde-* 'look up (at)', *bāysde-* 'observe', *yse-* 'be born' (SGS: *dai-*,

*uysdai-*, *bāysdai-*, *ysai-*)

1. coalescence (does not occur)

2. diphthongization

$e+\bar{i}>ai$ : 1Sp.a *daimä* Z 6.7; 3Pp.a *uysdaindä* Z variant 12 to 20.20, *bāysdaindä* Z

3.17, *daindä* Śgs 3.4r2; 1Ssj.a (-*ñä*) *daiñä* SI P62.1a6

$e+\bar{a}>ei>ai$  (nine categories attested; see above): 3Sp.a (-*ätä*) *daitä* Z 24.416;

3Sp.m (-*äte*) *ysaite* Z 2.212; 2Pp.a (-*äta*) *bāysdaiya* Z 23.8; 3Sop.a (-*ä*); *dai*

Z 2.226

*e+a>ya*: 2Siv.a *uysdya* SI A 18v4, *dya* Z 24.324

*e+e>ye*: 2Ssj.a \**dye* (attested only in the resolved form *diye* SI M14.1v4)

### 3. semivocalization

*e+ā>yā* (thirteen categories attested; see above): 3Pp.m *ysyāre* Z 3.62; *āka*-agent noun NSm *dyākā* Z 5.72; *āña*-p.nec NAPf *uysdyāñe* Suv[Or] 6.3.22, NAPm *dyāña* Z 24.441; *āmatā*-noun IASf *bāysdyemāte jsa* ‘observation’ IOL Khot 149/1 v6.

4. diphthong resolution 2Siv.a *uysdāya* Z 2.84, *dāya* Z 24.120; 2Ssj.a *diye* SI M14.1v4

5. glide resolution 3Pp.m *ysiyāre* Z 13.59; *āka*-agent noun *dāyākā* Z 8.13; *āña*-p.nec NAPf *uysdiyāñe* Suv[Or] 6.3.21, NAPm *dāyāña* Z 2.210; *āmatā*-noun *bāysdiyemetä jsa* IOL Khot 148/3.r6.

#### 4.8.4 Hist. cons. final, but synchronically vowel final present stems

The next set of items to compare involves the present stems which feature *VttV* in certain conjugations. If we set aside those *VttV* forms, the remaining conjugations exhibit exactly the same contractions as stems, nominal and verbal, ending in a vowel detailed above. Stem final *-i-* and *-ä-* show identical behavior and are treated together.

/i, ě/ *bi-* ‘pierce’, *dä-* appear, *nä-* ‘sit down’, *pajā-* ‘ask for’, *hamä-* ‘change (intr)’,

(SGS: *bid-*, *did-*, *näd-*, *pajäd-*, *hamäh-*)

#### 1. coalescence

$\ddot{a}+\bar{i}>\bar{i}$ : 3Pp.a *nīndä* Z 3.60; 1Ssj.a *dīñi* Z 2.100; 3Pop.m *dīru* Z 11.73; 3Sop.a  
*pajīyi* Z 11.19;

$i+\bar{i}>\bar{i}$ : 3Pp.a *bīndi* Z 24.408;

$i+\ddot{a}>\bar{i}$ : 3Sop.a *bī* Z 21.15.

## 2. diphthongization

$\ddot{a}+a>ya$ : 2Sp.m *dya* Śgs 2.8v2; 2Siv.a *nya* Z 2.188.

## 3. semivocalization

$\ddot{a}+\bar{a}>y\bar{a}$ : 3Pp.m *dyāre* Z 3.113, *hamyāre* Z 6.17.

## 4. diphthong resolution (no example)

5. glide resolution 3Pp.m *dāyāre* Z 5.12, 23.13, *diyāre* Z 4.100, *hamāyāre* Z 24.2.

/u/ *hāru*- ‘grow (intr)’, *paju*- ‘conceal’, *bu*- ‘know’, *rru*- ‘grow (intr)’ (SGS: *hārūd*-,  
*pajud*-, *bud*-, *rrūd*-)

## 1. coalescence:

$u+\bar{u}>\bar{u}$  *ūna*-noun (KhSuf:167–169) NSm *pagyūni* Z 22.138 ‘covering’.

## 2. diphthongization:

$u+e>ve$ : 1Sp.m *bve* Z 2.66

$u+a>va$ : 2Sp.m *bva* Sgh[8] 214[2]

## 3. semivocalization:

$u+\bar{a}>v\bar{a}$ : 3Pp.m *bvāre* Z 2.67; 2Ssj.m *bvā* IOL Khot 189/5v3 (Śgs); *āmata*-noun  
NSf *bvāmata* Z 1.31, IASf *bvemate jsa* Z 3.140

$u+\bar{i}>v\bar{i}$ : 3Pp.a *rrvīndä* Z 18.25.

4. diphthong resolution: 1Sp.m *buve* Z 6.45; 2Sp.m *buva* Z 3.9.

5. glide resolution 3Pp.a *hāruvīndä* Suv[Or] 10.20; 3Sp.m *buvāre* Z 2.121;

*āmata*-noun NSf *hāruvāmata* Suv[Or] 10.45, NSf *buvāmata* Suv[Or] 5.31,

IASf *buvemate jsa* Suv[Or] 5.31.

/a/ *kha*- ‘wound’, *pāha*- ‘strike’, *puṣa*- ‘fear’, *ba*- ‘be bound’, *ttāhva*- ‘cross’ (SGS:

*khad*-, *pāhad*-, *puva’d*-, *bad*-, *ttāhvah*-)

1. coalescence

*a+a>ā*: 3Ssj.a *khā* Z 13.81; 2Siv.a *pvā*’ Sgh[17] 31.9

*a+ā>ā*: *āmāta*-noun NSf *pāhāmata* Z 7.22 ‘striking’; *āka*-agent noun NAPm

*pihāka* ‘(tree) choppers’<sup>118</sup> Sgh[11] 210.4; *āñā*-p.nec NSm *puvā’ñi* Z 18.35.

2. diphthongization

*a+i>ai*: 3Pp.a *ttāhvaindä* Z 13.20, *puvai’ndi* Z 11.10, *baindä* Z 3.107.

*a+ā>ei>ai*: 3Sop.a *pihei* Sgh[1] 39.1, *pāhai* Sgh[17] 39.1.

There are also type B stems in *-e*-.

/ē/ *se*- ‘appear’, *āve*- ‘obtain’ (SGS: *sad*-, *āvad*-)

2. diphthongization

*e+i>ai*: 1Sp.a *saimä* Z 13.60; 3Pp.a *saindä* Z 3.116; 3Sop.a *saiyä* Z 7.26, *āvaiyä* Z

18.3.

3. semivocalization

*e+ā>yā*: *āmatā*-noun NSf *syāmata* Z 4.85; p.nec NSm *āvyāñi* Z 12.23.

If not for the existence of the *-tt-* forms with these stems, it is possible that earlier scholarship might have considered the stems to be vowel final. In none of the forms listed above is there a *-d-* or *-h-* reflecting the historical stem final consonant. The *-y-*

---

<sup>118</sup> For *pihāka* see footnote 80.



and -v- resulting from diphthongization and semivocalization are synchronically derived from the stem final vowel.

#### 4.8.5 Parallel Historical Evolution

The information above may be consolidated to show that the evolution of the secondary declensions and that of the *VttV* stems is parallel. In both cases a stem final consonant is lost before a vowel. The information is organized by stem vowel in pre-Khotanese. Within each section the nominal evidence is followed by that of the non-*VttV* stems. These did not have historical final consonants but feature the same contractions in Khotanese as the nominals and the *VttV* stems. The *VttV* evidence concludes each section. A selection of stems is examined. The Old Iranian forms are from Cheung (EDIV) or Skjærvø (SuvII) with a few new proposals. Those scholars use divergent transcriptional systems which produces some inconsistencies.

##### 4.8.5.1 \**uC+V* > *uw+V* > *u+V*

###### nominals

\**bāzuka-* (cf. Av *bāzu-*) > *bāysua-* ‘arm’: NSm *bāysū* Z 2.139 (*u+ä>ū*), NAP

*bāysva* Suv[H] 4.11 (*u+a>va /u+a>ṽa/*)

###### non-*VttV* stems

\**juṽ-*<sup>119</sup> (cf. Av *juua-*) > *ju-* ‘live’: 3Sp.a *jūtā* Z 11.47 (*u+ä>ū*), 3Pp.a *juvīndā* SI P

4.9v7 (*u+ī>vī /u+ī>wī/*)

\**pati-dauH-* (EDIV:67) > *padu-* ‘be made smoky’: 3Sp.m *padūte* Z 6.39 (*u+ä>ū*)

---

<sup>119</sup> This is a late eastern Old Iranian root seen also in Bactrian *ζoo-*, *ζo-* ‘live’ (Sims-Williams 2007:213a), Manichean Sogdian *ju-*, Yaghnobi *žu-* ‘live’, Yazghulami *žaw-:žod* ‘revive after an illness’, Pashto *žwāk*, *žwand* ‘life’ (EDIV:223).



## VttV stems

\**pāti-gaud-* (EDIV:114) > \**pajud-* > *paju-* ‘cover’: *ūna-* noun ‘covering’ NSm

*pagyūni* Z 22.138 (*u+ū>ū*)

\**Hraud-* (EDIV:193) > \**rrud-* > *rru-* ‘grow (intr)’: 3Pp.a *rrvīndā* Z 18.25 (*u+i>vī*

*/u+i>wī/*)

\**baud-* (EDIV:14) > \**bud-* > *bu-* ‘know’: 2Sp.m *bva* Sgh[8] 214[2] (*u+a>va*

*/u+a>uā/*); 3Pp.m *bvāre* Z 2.67 (*u+a>va /u+ā>wā/*), 3Pop.m *bvīru* Z 11.71

(*u+i>vī /u+i>wī/*)

### 4.8.5.2 \**iC+V > iy+V > i+V* and \**āC+V > äy+V > ä+V*

#### nominals

\**pānti-ka-* (SuvII:297a) > *pādia-* ‘manner’: IAS *pādīna* Z 13.29 (*i+ā>ī*), NAPm

*padya* Z 10.9 (*i+a>ya /i+a>ja/*)

#### non-VttV stems

\**jaiH-* (EDIV:223) ‘perish’ > *ji-* ‘cease’: 3Sp.a *jīte* IOL Khot 8/1v3 (*i+ā>ī*)

\**pari-jaH*<sup>120</sup> > *pari-* ‘order; deign’: 3Sp.a *parītā* Z 24.450 (*i+ā>ī*)

#### VttV stems

\**uaid-* (EDIV:410) > \**bid-* > *bi-* ‘pierce’: 3Sp.a *bīndi* Z 24.408 (*i+i>ī*), 3Sop.a *bī* Z

21.15 (*i+ā>ī*)

\**did*<sup>121</sup> > \**dād-* > *dā-* ‘appear’: 3Pop.m *dīru* Z 11.73, 2Sp.m *dya* Śgs 2.8v2

---

<sup>120</sup> This is a new suggestion for the etymology of *pari-*. OIr \**jaH-* ‘demand, request’ has suitable semantics but this suggestion still does not explain the shape of the ppp *parsta-* (cf. SGS:73).

\**ni-had-* (EDIV:125) \**nād-* > *nä-* ‘sit down’: 3Pp.a *nīndä* Z 3.60 (*ä+i>ī*); 2Siv.a *nya*

Z 2.188 (*ä+a>ya /ë+a>ja*/)

\**pati-ĵad-* (EDIV:221) > \**pajād-* > *pajā-* ‘ask for’: 3Sop.a *pajīyi* Z 11.19 (*ä+i>ī*),

\**fra-maiθH-* (EDIV:260) > \**hamäh-* > *hamä-* ‘change (intr.)’: 3Pp.m *hamyāre* Z

6.17 (*ä+ā>yā /ë+ā>yā*/)

#### 4.8.5.3 \**eC+V* > \**ey+V* > *e+V*

##### **nominals**

no examples

##### **non-VttV stems**

\**zāĵa-* cf. Av *zaiia-* (SuvII:331) > *yse-* ‘be born’, 3Sp.m *ysaite* Z 2.212, 3Pp.m

*ysyāre* Z 3.62

\**daiH-* (EDIV:48) > *de-* ‘see’: 3Pp.a *daindä* Śgs 3.4r2 (*e+i>ai*), 2Siv.a *dya* Z

24.324 (*e+a>ya /ë+a>ja*/)

---

<sup>121</sup> The etymology of this verb is difficult. Skjærvø: “< \**did-ĵa-* \**dista-*, secondary passive and pp. of \**dida-*, cf. Av. *diδā-* < *dāy*<sup>2</sup> ‘to see’ (for *dista-* < \**dida-*, cf. Skt *datta-* < *dadā-*)” (SuvII:280b s.v. *diy-*). Emmerick: “< \**di-da-* ... Them. of \**di-dā-* (like Av. *daθa-* them. of *daθā-* redupl. pres. of *dā-*), cf. Av. *diδā-* redupl. pres. of <sup>2</sup>*dā(y)-*” (SGS:45). The metanalysis makes another proposal possible. Perhaps the stem never ended historically in \**d*, but was derived initially < \**daiH-ĵa-* (\**daiH-* ‘see’ EDIV:48), semantically ‘see’ > ‘be seen’ > ‘appear’. Then to keep *dā-* ‘appear’ distinct from *de-* ‘see’, the former was transferred to type B, where it displays /tt/ initial suffixes, and a new ppp *dista-* was created. The two stems are still indistinct in some cases, such as with the *āmatā-*-noun *dyāmatā-* which can mean either ‘appearance’ (< *dā-* ‘appear’) or ‘seeing’ (< *de-* ‘see’). This derivation helps explain the morphophonological relationship between *dā-* ‘appear’ and two proposed stem shapes *āyā-* ‘appear’ and *pūyā-* ‘observe’ (see above footnote 76, and below).

## VttV stems

\**saiH-* (< Indo-European \**skeiə*<sub>2</sub> ‘shine’; Yakubovich 2002:548) > *se-* and \**sad-*  
*aia-* (cf. Av *sad/δaiia-* SuvII:352b) > \**sed-* > *se-* ‘appear, seem’: 1Sp.a *saimä*  
Z 13.60 (*e+i>ai*), *āmatā-*noun NSf *syāmata* Z 4.85 (*e+ā>yā /ē+ā>yā/*)

### 4.8.5.4 \**ǎC+V* > \**aĥ+V* > *a+V*

#### nominals

\**asta-ka* (cf. Av *ast-*; SuvII:245a) > *āstaa-* ‘bone’: IAS *āstaina* Z 5.8 (*a+ǎ>ai*).  
\**ui-tamH-a-kā-* (cf. EDIV:376) > *bātamaā-* ‘doubt’: NAPf *bitame* Z 6.23 (*a+e>e*)  
\**pād(a)-* (Av, Old Persian *pād(a)-*; SuvII:302a) > \**pāa*<sup>122</sup> > *paa-* ‘foot’

#### non-VttV stems

\**paH* (EDIV: 288) > *pa-* ‘protect’: 1Pp.a *pāmā* Suv[Or] 6.1.39 (*a+ā>ā*), *paindä* Z  
16.5 (*a+i>ai*)  
\**pari-θāu-* (cf. \**pari-θāuqia-* SuvII:294b) > *paṭha-* ‘burn up (tr)’: *āñā-*p.nec NSm  
*paṭhāñā* Suv[Or] 8.23 (*a+ā>ā*), 3Sp.a *paṭhaiyā* Suv[Or] 6.4.22 (*a+ǎ>ai*)  
\**aua-čiau* (EDIV:41) > *vatsa-* ‘descend’: 3Pp.m *vatsāre* IOL Kh 159/7r2 (*a+ā>ā*)  
\**niš-čiau* (EDIV:41; \**nir-čāuqia-* SuvII:286a) > *naltsa-* ‘go out’ 2Siv.m *naltso* Z  
5.30 (*a+u>o*)

---

<sup>122</sup> In SGS Emmerick set up monosyllabic and polysyllabic *āa*-declensions (305–308). His monosyllabic category declines exactly like the *aa*-declension, so synchronically the three nouns in question may be listed as *daa-* ‘fire’, *paa-* ‘foot’ and *rtaa-* ‘plain’. Further proof that *paa-* ends in short /a/ is shown by the form IAP *pyau* SI M13.14v3. The oblique plural suffixes delete the stem final /a/ in the *aa*-declension, e.g., GDPm *āchānu* Z 24.222 < *āchaa-* ‘disease’, IAP *stāryau* Z 23.22 < *stāraa-* ‘star’, LP *ggāṭhuvo* Z 22.315 < *ggāṭhaa-* ‘householder’. The glide in *pyau* may be resolved as in *pāyau* jsa Z 22.148, not as \**pāyau* or \**payau*.

\**ui-zau* (EDIV:471–472; \**ui-zāuajā*- SuvII:162, 6.2.34) > *buysa*- ‘extinguish’:

2Siv.a *buyso* Z 5.50 (*a+u>o*)

### **VttV stems**

\**ati-hvah* (EDIV:142) > *ttähva*-: ‘cross; penetrate’: 3Pp.a *ttähvaindä* Z 13.20

(*a+i>ai*)

\**xad*- (EDIV 439) > *kha*- ‘wound’: 3Ssj.a *khā* Z 13.81 (*a+a>ā*)

\**pati-xad*- (EDIV:439) > *pāha*- ‘strike’: *āmatā*-noun NSf *pāhāmata* Z 7.22

‘striking’ (*a+ā>ā*)

\**api-šad* (cf. \**api-šadaja* SuvII:308b s.v. *pvay*-) > *puša*- ‘fear’: 2Siv.a *pvā*’ Sgh[17]

31.9, 100.2 (*a+a>ā*), *āñā*-p.nec NSm *puvā’ñi* Z 18.35 (*a+ā>ā*), 3Pp.a

*puvai’ndi* Z 11.10 (*a+i>ai*)

\**bad*- (cf. \**badja*- SuvII:312a s.v. *basta*-) > *ba*- ‘be bound’: 3Pp.a *baindä* Z 3.107,

22.273 (*a+i>ai*)

### **4.8.6 Summary of the parallel evolution**

The information sketched above shows that the evolution of the *VttV* stems from consonant final to vowel final is exactly parallel with the evolution of the secondary declensions. In both cases, a stem final consonant is lost when followed by a vowel initial suffix. The vocalic contractions exhibited by both, as well as by the historically vowel final stems, are identical.

Most of the secondary declension nominals derive from loss of intervocalic \**k*. At least one example of loss of intervocalic \**d* is known. From OIr \**pāda*- there is *paa*- ‘foot’. This provides the closest parallel with the *VttV* stems. That is, \**pād+V* > \**pā+V* > *pa+V* ‘foot’ is similar to \**xad+V* > *kha+V* ‘wound’ and several others (see above).

The intervocalic consonants were completely lost from both nominals and present stems. No trace remains in Khotanese. The traces seen by other scholars are all morphophonological effects explained by the synchronic grammar. This was demonstrated for the nominals in Hitch 2015b and Hitch 2016a, and can be observed for the present stems in the morphophonological comparisons above.

In the spelling alternations 2Sp.m *bva* Sgh[8] 214[2] ~ *buva* Suv[Or] 5.33 ‘know’, and NASn *nyāñu* Suv[Or, C] 6.3.17 ~ *niyāñu* Z 2.221 < *nä-* ‘sit down’, it has generally been thought that the *v* and *y* reflect the historical *\*d*. Two things show that this is not the case. On the one hand, the verbs which are historically vowel final show exactly the same morphophonological pattern. The *y* in 3Pp.m *jiyāre* Śgs 3.13r3 ~ *jiyāre* Z 3.29 < *ji-* ‘cease’ < *\*jaiH-* does not derive from a historical final consonant and neither does the *y* in 2Siv.a *dya* Śgs 3.4r2 < *de-* ‘see’ < *\*daiH-*. The *v* in 3Pp.a *jiyāre* SI P 4.9v7 < *ju-* ‘live’, or in 2Siv.a *ākṣva* Z 2.100 < *ākṣu-* ‘begin’ does not show a stem final consonant as shown by 3Sp.a *jūtā* Z 11.47 and 3Sp.a *ākṣūtā* Z 4.15 both featuring coalescence *u+ä>ū /-u- + -ědě > -ūdě/*. On the other hand, there is no corresponding consonant, no *y*, *v* or anything else, when the stem final vowel is *a*. No forms of *paa-* ‘foot’ contain a consonant which could by any stretch be interpreted as reflecting a final *\*d*. Likewise, with the exception of the *VttV* conjugations, no form of *puṣa-* ‘fear’ shows a consonant which could reflect a final *\*d*. Those segments are gone without a trace, at least in the synchronic analysis.

#### 4.9 Skjærvø’s Proposal

Skjærvø proposed that stems ending historically in *\*-ad-* evolved to stems ending in *\*-ay-* in Khotanese. For his stems *hvay-* ‘strike’ and *say-* ‘appear’ he gives this

formula for generating the 3Sp.a: “ $y + t > itt$  (*itt*) (*hvaittä, saittä*)” (SuvII:229). Those are inconvenient examples in this context since neither clearly ended historically in \*-ad-. While some scholars do propose that the stem *hva-* ‘strike’ ended historically in \*-d-, others see it ending in \*-h-<sup>123</sup>, and the stem *se-* ‘appear’ has a different stem vowel, at least by the time of pre-Khotanese. But we may be able to apply his rule to his other \*-ay- stems. For instance, his stem *puva’y-* ‘fear’ (here *puṣa-*, SGS *puva’d-*) would synchronically derive under his formula as  $puva’y + tä > 3Sp.a\ puvai’ttä$ . That formula looks superficially promising but it will not work with any other conjugations featuring *VttV*. For instance,  $puva’y + tu$  would result in *\*puvai’ttu* not the attested 3Siv.a *pva’ttu* Z 2.101, and  $puva’y- + ta$  would give *\*puvai’tta* not the attested 2Piv.a *puva’tta* Z 24.474.

Skjærvø’s stem shape and formula fail with vowel initial suffixes as well, where there is no trace of a -y- on the surface. The addition of the 2Siv.a suffix -a does not result in *\*puva’ya* but *pvā’* Sgh[17] 31.9 with contraction  $a+a>\bar{a}$  /buā < buṣā < buṣa- + -a/. Similarly, the addition of the *āñā-p.nec* suffix does not produce *\*puva’yāñā-* but *pvā’ñā-* as in NSm *pvā’ñi* Z 18.35 showing  $a+\bar{a}>\bar{a}$  • /buāñě < buṣāñě < buṣa- + -āñ- + -ě/. Finally, Skjærvø’s stem shape and formula do not explain how 3Sp.a *puvai’ttä*, 3Siv.a *pva’ttu* and 2Piv.a *puva’tta* developed the underlying double /tt/ clearly proven to exist from metrical position in Z (Hitch 2015a). The stem *puṣa-*

---

<sup>123</sup> Konow *hvah-* (1932:145), then *hvada-* (1949:1320; E. and M. Leumann *hvad-* (Glossar:528b), Palmira Cipriano *\*hvad-* (1998:210–235, especially 224–227), Bailey *hvah-* (DKS:508b); Emmerick *hvah-* (SGS:156); Skjærvø *hvay-* (SuvII:373a). Cf. also footnote 18 above.



'fear', as all others historically ending in *\*-ad-*, ends in the vowel *-a-* in Old Khotanese.

#### 4.10 The metanalysis

The stems which show *VttV* in certain conjugations have here had their conjugations divided into two classes, those showing *VttV* (3Sp.a, 3Sp.m, 3Siv.a, 2Piv.a, 3Sij.m are attested while 2Sp.a, 3Siv.m, 3Sij.a are predicted) and those showing contraction of stem and suffix vowel (all others). As the latter class lost the historical stem final *\*d* or *\*h*, the *VttV* exhibiting class underwent the metanalysis. For example, as the stem became vowel final *bu-* where the suffixes were vowel initial, that is,  $/[\beta u w][V] > [\beta u][V]/$ , in the conjugations where the suffix historically began with *\*t* the morpheme boundary shifted to the left, that is,  $/[\beta u t][tV] > [\beta u][ttV]/$ . As 2Sp.m *buva* became analyzed as *bu-* + *-a*, 3Sp.m *butte* became analyzed as *bu-* + *-tte*. Suffixes historically beginning with single *\*t* now began with double */tt/*. Similar examples of rebracketing are 3Sp.a  $[nät][tä] > [nä][ttä]$  'sit down', 3Siv.a  $[pva't][tu] > [pva'] [ttu]$  'fear', 2Piv.a  $[hambit][ta] > [hambi] [tta]$  'pierce', and 2Piv.a  $[pähat][ta] > [päha] [tta]$  'strike'.

The 3Sp.a forms showing the umlaut potential expressed on the surface as a diphthong like *rrvittä* 'grow (intr)' or *baittä* 'be bound' cannot be formulated with rebracketing without resorting to underlying forms. Underlyingly,  $/[rut][i'tě] > [ru][i'ttě]/$  and  $/[\beta at][i'tě] > [\beta a][i'ttě]/$ . The metanalysis in these forms is in principle the same. The umlaut potential is not attached to a segment as the transcriptions might suggest, but to a suffix. The potential remains with the suffix as the suffix onset is adjusted. Actually, 3Sp.a *nättä*, which does not show the presence of the

umlaut potential on the surface, underlyingly must show the same structure,  
 /*[nět][<sup>i</sup>tě]* > *[ně][<sup>i</sup>ttě]*/.

As shown above, the 3Sp.m *baḍ(ḍ)e* < *bar-* ‘ride’, contains double /ḍḍ/ which is not predicted by historical origin. Instead, the geminate has a synchronic explanation. After the metanalysis the suffix *-tte* began with initial double /tt/. This combined with stem final *-r-* producing double *-ḍḍ-* /βaḍḍē < βaḍ- + -ttē/.

#### 4.10.1 Cluster reduction (R)CCC > (R)CC

With stems ending in a consonant other than *-r-*, there is a process of consonant cluster reduction in the synchronic derivation where there is a suffix beginning with /tt/. In the derivation a range of other morphophonological processes may also take place. It is necessary to describe these in order to clearly describe the reduction.

Where there is a single stem final voiceless consonant, the resulting cluster is voiceless:

##### **VC+ttV > VCtV where C is voiceless**

*ch* 3Sp.m *gvašte* Neb 75.40 < *gwach-* ‘be digested’ • /gwaštē < gwač<sup>h</sup>- + -ttē/

*ts* 3Sp.a *patāste* Z 2.190 < *patāts-* ‘give up, abandon’

• /badēstē < badēt<sup>s</sup>- + -ttē/

*šš* 3Sp.m *panašte* Z10.2 < *panašš-* ‘perish’ • /banaštē < banaš- + -ttē/

*šš* 3Sp.a *huṣṭä* Z 11.50 < *huṣṣ-* ‘grow (intr)’ • /huṣṭē < huṣ- + -<sup>i</sup>ttē/

*s* 3Sp.a *namaštä* Z 12.55 < *namas-* ‘worship’ • /namaštē < namas- + -<sup>i</sup>ttē/

*t* 3Sp.a *pītä* Z 2.29 < *pat-* ‘fall’ • /bītē < bat- + -<sup>i</sup>ttē/

Where there is a single stem final voiced consonant, but not *ḍ* or *l*, the resulting cluster is voiced as the suffix assimilates to the stem:

**VC+ttV > VCdV where C is voiced but C is not *ǰ* or *l***<sup>124</sup>

(There is also deaffrication *j>ś* /*ǰ>ž*/ and *js>ys* /*dʒ>z*/ in the first three examples.)

- j* 3Sp.m *uskuśde* Z 11.50 < *uskuǰ*- ‘rise up (against)’ • /uskuždē < uskuǰ- + -ttē/  
*js* 3Sp.m *vataysde* Z 17.12 < *vatajs*- ‘flow down’ • /wadazdē < wadadʒ- + -ttē/  
*js* 3Sp.a *padaśdä* Z 11.36 < *padajs*- ‘burn (tr)’ • /baðaždē < baðadʒ- + -<sup>i</sup>ttē/  
*ś* 3Sp.m *pyūśde* Z 14.73 < *pyūś*- ‘hear’ • /byūždē < byūǰ- + -ttē/  
*ys* 3Sp.m *haraysde* Z 2.132 < *harajs*- ‘extend’ • /haɹazdē < haraz- + -ttē/  
*ñ* 3Sp.m *jsīnde* Z 19.10 < *jsañ*- ‘be struck; slain’ • /dʒīndē < jsañ- + -ttē/  
*n* 3Sp.a *ggändä* Z 22.209 < *ggän*- ‘buy’ • /gëndē < gën- + -<sup>i</sup>ttē/  
*rr* 3Sp.a *purdä* Z 24.400 < *purrr*- ‘overcome’ • /burdē < bur- + -<sup>i</sup>ttē/

Where the single stem final voiced consonant is *ǰ* or *l*, the addition of *tt* results in *litt*. This pattern was detailed in 4.6.5 above. There is no voicing assimilation, and there is orthographic *tt*:

---

<sup>124</sup> 3Sp.a *sūštä* Z 20.12, 24.323 is an exception if it derives from *sūjs*- ‘burn’. Two 3Sp.m forms reflecting OIr \**xand*- ‘laugh’ may also be exceptions. These are 3Sp.m *naṣkhamttä* Z 12.75, IOL Khot 9/5v3 ‘laugh’, and 3Sp.m *bihante* Z 5.26 ‘smile’. All scholars present these stems as *naṣkhan*- and *bihan*-. With all other type B stems ending in *n* there is voicing of a /*tt*/ initial suffix as in 3Sp.a *jsīndi* Z 7.24 < *jsan*- ‘strike; slay’, 3Sp.a *paysendä* Z 4.97 < *paysān*- ‘recognize’, or 3Sp.a *jändä* Z 3.28 < *jān*- ‘destroy’. It is unlikely that the stem shapes given by earlier scholars are incorrect. The related stem *khan*- ‘laugh’ has the final *n* confirmed by 3Pp.a *khanīndi* Z 20.20. These descendants of \**xand*- are the only stems in *n* which do not voice the suffix, and they are also the only stems with a ppp in *-ṃtta*-, i.e., *bihamttā*- and LKh *khatta*- < \**khamttā*- (SGS:99, 25–26, 174). As M. Maggi (p.c.) points out, *khan*- and its relatives are the only attested type B verbs from roots ending in \**-nd*-. The clusters *-ṃtt*- and *-nt*- have a historical explanation < \**-nd+t*- which differs from *-nd*- < \**-n+t*-.

**VC+ttV > VCttV where C is *ḡ* or *l***

*ḡ* 3Sp.m *ggaltte* Z 24.450 < *ggaḡ-* ‘lie about’ • /galttē < gaḡ- + -ttē/

*l* 3Sp.m *saṃkhiltte* Z 24.227 < *saṃkhāl-* ‘be tainted’ • /sank<sup>h</sup>ilttē < sank<sup>h</sup>ēl- +  
-ttē/

There are few present stems which end in two consonants and which are attested with type B /tt/ initial suffixes. These stems end in just four clusters, *ls*, *lys*, *rs* and *ṃth*. The first segment, a resonant (*R*), /l, ɹ, n/ is voiced while the second /s, t<sup>h</sup>, z/ may be either voiced or voiceless. The /tt/ of the suffix assimilates to the voicing of the second segment:

**RCtt > RCt where C is voiceless**

*ls* 3Sp.a *pulstā* Z 3.75 < *puls-* ‘ask’ • /bulštē < buls- + -<sup>i</sup>ttē/

*ls* 3Sp.m *ggeiste* (*\*ggalste*<sup>125</sup>) Z 4.98 < *ggei’ls-* ‘revolve; return’  
• /gaēste < galstē < gals- + -ttē/

*rs* 3Sp.a *parstā* Z 2.68 < *pars-* ‘escape’ • /barštē < bars- + -<sup>i</sup>ttē/

*rs* 3Sp.a *harstā* Z 24.428 < *hars-* ‘be left, remain’ • /barštē < bars- + -<sup>i</sup>ttē/

*rs* 3Sp.a *uysgursti* Z 20.42 < *uysgurs-* ‘tear off’ • /uzgurštē < uzgurs- + -<sup>i</sup>ttē/

---

<sup>125</sup> The OKh instances of a stem *\*ggals-* (inchoative < *\*gart-*; SGS:31), also 3Pp.m *ggei’lsarā* Z 150, *ggei’lsāre* Z 2.164, *ggei’sāre* Z 20.16, *ggeisāre* Z 20.56, may show LKh influence in the spelling. Compare OKh *palsārā* ‘crowns, garlands’ with LKh *pe’sārā* (DKS:221b), OKh *balsa-* ‘monument’ with LKh *be’sa-* (DKS:272b), OKh *balysa-* ‘Buddha’ with LKh *bai’ysa-*, *be’ysa-* (JS:480a). In all those instances, the loss of *l* may be marked with the apostrophe and the vowel spelling may change. It is odd that there is no *\*ggals-* spelling. There is also no *\*ggalśś-* spelling for the causative *ggei’śś-* ‘make revolve’. If Bailey is correct to relate the noun *\*ggālsāra-* seen in LS *ggālserai* Z 22.149 ‘on his neck’ to *ggei’ls-* (DKS:90b), then the older stem may be *\*ggāls-*.

*m̄th* 3Sp.m *nvam̄tte* Z 20.14 < *nuvam̄th*- ‘be removed’

- /*nwanttē* < *nuwant<sup>h</sup>*- + *-ttē*/

***RCtt* > *RCd* where *C* is voiced**

*lys* 3Sp.a *nimalśdä* Z 22.147 < *nimalys*- ‘rub down’

- /*nimalždě* < *nimalz*- + *-<sup>i</sup>ttě*/

*lys* 3Sp.m *nuvalysde* Z 17.18 < *nuvalys*- ‘flow down’

- /*nuwalzdē* < *nuwalz*- + *-ttē*/

The stems listed above reveal at least six processes.

1) The stem umlaut seen with 3Sp.a forms like *namaštä* < *namas*-, *pīttä* < *pat*-, and *padaśdä* < *padajs*- is detailed in Hitch 1990.

2) The deaffrication of the stem final as seen in *gwach*- > *gvašte* /č<sup>h</sup> > š/, *patäts*- > *patäste* /t<sup>s</sup> > s/, *uskuj*- > *uskuśde* /j > ž/, *vatajs*- > *vataysde* Z 17.12 /d<sup>z</sup> > z/ is straightforward.

3) The rightward point of articulation assimilation as shown by *huṣṭä* < *huṣṣ*- /ṣt > ṣt/, and *pyūṣḍe* < *pyūṣ*- /zd > zd/ is also straightforward, if not purely orthographic.

4) Voicing assimilation of the suffix to the stem final consonant as outlined above is phonologically straightforward but it has curious apparent exceptions as listed in 5).

5) When the stem final consonant is *ḍ* or *l*, the /tt/ initial does not voice. This process was outlined in 4.6.5 above and is detailed further below in this section.

6) The most challenging process to define in these derivations is consonant cluster reduction. Most of the 3S forms in the list above show a historically predicted number of consonants in the cluster, as the historical suffixes contained single \**t*.

The metanalysis and the reformulation of the synchronic suffixes to /tt/ initial required a reanalysis of the established forms. One course of evolution might have been to have two kinds of suffixes, those beginning with double /tt/ to be used with vowel final stems, and those beginning with single /t/ to be used with consonant final stems. The double /ḏḏ/ in *baḏ(ḏ)e* < *bar-* discussed above suggests this was not the case. Instead, the language appears to have made use of consonant cluster reduction. We may see cluster reduction in LSm *aysmya* < *aysmua-* ‘mind’ /*azmja* < *azmwja* < *azmja* < *azmūi-a* < *azmu-* + <sup>i</sup>a/ (Hitch 2015b:315–317).<sup>126</sup> The rule with the /tt/ suffixes may be as follows: Where the stem has a single final consonant other than *r* (discussed above), the underlying three consonant cluster *Ctt* or *Cdd* is reduced to two, *Ct* or *Cd*. Where the stem has two final consonants the underlying four consonant cluster, which begins with a resonant *R*, is reduced to three, *RCtt* > *RCt* and *RCdd* > *RCd*. These statements may be combined as *(R)CCC* > *(R)CC*.

For the consonant cluster rule to be correct, then *saṃkhiltte* < *saṃkhāl-* and *ggaltte* < *ggaḏ-* would not undergo reduction. That is, the conjugated forms contain double /tt/. This structure cannot be proven in the meter of Z since the *l* already makes position. There may be other evidence. Some statistics regarding orthographic clusters containing *tt* in Z were reported in Hitch 2015a:684–685.

There are ten cases of *VlttV* with double *tt* (all from stems in *-l-* or *-ḏ-*) and none of

---

<sup>126</sup> With most *ia*-stems, the resolved forms showing *-CiyV-* are less frequent than the unresolved forms with *-Cya-*. The forms of *indria-* ‘sense’ exhibit the reverse distribution (Hitch 2015b:3). The *indriyV-* shapes may be more common than the *indryV-* shapes because the latter may contain an undesirably heavy consonant cluster. The resolved forms may reflect another kind of cluster reduction, *RCCC* > *RCCVC*.

\**VltV* with single *t*. At the same time, the common phonological clusters /*st/* and /*št/* are always written with single *t* as *st* and *št* and never as \**stt* or \**št*. This suggests that *VlttV* contains double /*tt/* otherwise single *t* would be written. A shape \**VltV* would non-ambiguously represent single *t*, since in clusters /*d/* is written *d*. There is a superficially parallel distribution with *r* in *Z*, which might be regarded as counter evidence: there are twenty-nine examples of *VrttV* (all in Indian loans<sup>127</sup>) and none of \**VrtV*. But preconsontantal *r* in the Brāhmī is written with a superscript diacritic. Single orthographic *t* in this case risks reading as [d]. Further, the cluster /*tɿ/*, which is also restricted to Indian loans, may be written either *VtrV* (over 100×) or *VttrV* (45×). The distribution of *VrttV* is not relevant to interpreting *VlttV* which likely has double /*tt/*.

The form 3Sp.m *nvam̐tte* exhibits a unique combination of processes. The stem *nuvam̐th-* is clearly known from 3Pp.m *nvam̐thārā* Z 24.423. The 3Sp.m certainly has double /*tt/*. Even before the metanalysis it would have contained double /*tt/* as stem final /*t<sup>h</sup>/* joined the suffix initial single /*t/*. After the metanalysis it is synchronically derived /*nuwanttē* < *nuwant<sup>h</sup>ttē* < *nuwant<sup>h</sup>- + -ttē/*. The RCCC > RCC cluster reduction rule applies and the aspiration on the first segment is lost.

#### 4.10.2 Possible role of *se-* ‘appear’ in the metanalysis

As mentioned above, the stem *se-* may reflect two distinct historical stems with similar shape and meaning, pre-Khotanese *se-* and \**sed-*. The \**d* final stem produced *VttV* forms with double /*tt/* and may have produced \**seyV-* forms before vowels. It

---

<sup>127</sup> All examples are Indian loans containing (-)*varttV* such as NSm *cakkravarttā* Z 24.199 ‘*cakravartin*’ or 1Sp.a *varttimā* Z 24.492 ‘practise’.

would have been difficult to keep these latter distinct from the resolved forms of *se-* (*\*säyV-* < *syV-* < *se-V*; cf. NSf *syāmata* Z 4.85 ‘appearance’). The two stems may have merged with conjugated *VttV* forms coming from *\*sed-* and the others from both *se-* and *\*sed-*. In this way, this common verb may have encouraged the metanalysis. The merger and metanalysis is sketched in Table 3. We do not have attested any *VttV* forms other than the 3Sp.a *saittā* which features the umlaut potential. A 2Pp.m form *\*setta*, for example, would establish the stem vowel and confirm the history.

**Table 4.3:** *se-*, *\*sed-* and the metanalysis

Pre-Khotanese		Old Khotanese	
[se][āmatā-]	>	[se][āmatā-]	<i>syāmatā-</i>
[se][īmā]	>	[se][īmā]	<i>saimā</i>
[*sed][ <sup>i</sup> tā]	>	[se][ <sup>i</sup> ttā]	<i>saittā</i>
[*sed][*ta]	>	[se][tta]	<i>*setta</i>

#### 4.10.3 Fossils

It is possible that the synchronic variations *Cv~Cuv* and *Cy~Ciy* in both nominals and present stems are fossils echoing an earlier stage of the language. For instance in the evolution from OIr *\*baud-* ‘feel, sense’ to Old Khotanese *bu-* ‘know’ there may have been a pre-Khotanese phonetic evolution before vowel initial suffixes something like *\*[baud-V- > βuδ-V- > βuw-V- > βu-V-]*. Similarly, the evolution from *\*fra-maiθH* to *hamä-* ‘change (intr.)’ before vowels may have been something like *\*[fra-maiθH-V- > haměĥ-V- > haměy-V- > hamě-V-]* and that from *\*ni-had-* to *nä-* ‘sit down’ like *\*[ni-had-V- > nęd-V- > něy-V- > ně-V-]*. For a time, the last two stages in each evolution may have coexisted before the patterns were reanalyzed as reflecting



synchronic glide or diphthong resolution. That is, just before the metanalysis, *buva* and *bva* coexisted with *buva* being older. Similarly, *hamäyāre* was older than *hamyāre*, and *niyāñu* was older than *nyāñu*. After the metanalysis, the forms *bva*, *hamyāre* and *nyāñu* became the basic synchronic forms and the others became regarded as the derived forms. The existing phonetic material was given a new morphophonological interpretation, perhaps influenced by glide resolution illustrated above by *pyaurā-~päyaurā-* and *tvāy-~ttuvāy-*.

The process would have been exactly parallel with the nominals. For instance, *bāysuve* may have historically preceded *bāysve* but in Old Khotanese the basic synchronic form is *bāysve* and the derived is *bāysuve*.

All final *-u-*, *-i-*, *-ä-* and *-e-* stems show this pattern. The final *-a-* stems do not. They do not exhibit the same kind of fossils. There is no *\*pāhayāmata* beside *pāhāmata*, reflecting pre-Khotanese *\*pāhad- + -āmatā-*. Instead, there is a break in the historical morphophonology. Possibly the metanalysis occurred first with the *-u-*, *-i-*, *-ä-* and *-e-* stems. After the suffixes in the *VttV* forms in these stems were reinterpreted as being /tt/ initial, the new suffix shapes naturally spread to the *\*-ad-* and *\*-ah-* stems. For instance, even though 3Sp.m *hamatte* Z 2.167 ‘sober up’ historically derives from *\*fra-mad-tai*, when the 3Sp.m suffix was reanalyzed to /ttē/, the stem became synchronically *hama-*. Reanalyses of the type [hamat][te] > [hama][tte] followed those of the types [but][te] > [bu][tte], [bit][te] > [bi][tte], and [dät][tä] > [dä][ttä].

Such an explanation is helpful in understanding the evolution of the *\*-ah-* stems. There are two type A stems ending in *-ah-* in Old Khotanese, *patārah-* ‘support

oneself (on)' and *ysah-* 'cease'. Conjugated forms like 3Sp.m *patārahāte* Suv[Or] 10.40 and 3Sp *ysahāte* Z 24.423 demonstrate that the usual evolution was pre-Khotanese *\*ahV* > Old Khotanese *ahV*. That is, *\*h* was retained in *ahV*. With the type B verbs historically in *-ah-* like *ttāhva-* 'cross (a river)', *nei'hva-* 'cross over' or *hva-* 'strike', only the metanalysis can explain their evolution to stems in *-a-*. There is no historical phonological explanation.

#### 4.11 Summary

The Old Khotanese metanalysis described in these pages affected a significant number of present stems including several common ones such as *bu-* 'know'. It involved a morphophonological restructuring that was both major and subtle. It was major because it impacted many stems and suffixes. It was subtle because it has escaped detection until now. Identification of the metanalysis allows the explanation of all forms of the affected present stems in a systematic way with morphophonological rules which are either attested elsewhere in the grammar or are logical extensions of attested rules. It allows the description of the language to be simplified in several ways:

- There is no need to recognize a separate type C present stem as all stems in that category of Emmerick's may be synchronically analyzed as type B vowel stems. Together with the redefinition of Emmerick's type D as type A vowel stems (TypeA) this redefinition permits the present stem types to be reduced to just A and B.
- All type B verbs take the same set of suffixes regardless of whether they end in a consonant or a vowel. The same was shown for type A verbs in (TypeA)

- A single set of vocalic contraction rules defines all contractions seen with nominals and present stems ending in vowels.

## Abbreviations

	metrical segment boundary
□	space between orthographic columns
1S, 1P	first person singular, first person plural
2S, 2P	second person singular, second person plural
3S, 3P	third person singular, third person plural
a	1) active verb; 2) first column of text in the main manuscript of Z
A	1) meter type A; 2) SGS present stem type A
A1	subtype of meter type A
ab	spread of hemistich across first and second columns of text in the main manuscript of Z
ac	spread of hemistich across first and third columns of text in the main manuscript of Z
AOKh	Archaic Old Khotanese
AS	Accusative Singular
Av	Avestan
B	1) meter type B; 2) SGS present stem type B
bd	spread of hemistich across second and fourth columns of text in the main manuscript of Z
BHS	Buddhist Hybrid Sanskrit
C	1) any consonant; 2) meter type C; 3) SGS present stem type C
C1	subtype of meter type C
cd	spread of hemistich across third and fourth columns of text in the main manuscript of Z
COKh	Canonical Old Khotanese
D	SGS present stem type D
DKS	<i>Dictionary of Khotan Saka</i> , Bailey 1979
EDIV	<i>Etymological Dictionary of the Iranian Verb</i> , Cheung 2007
EMC	Early Middle Chinese
EOT	Earlier Old Turkic
EWKh	Earliest Written Khotanese
f	feminine noun
Gdh	Gāndhārī (Prakrit)
GDP	Genitive-Dative Plural
GDS	Genitive-Dative Singular
Glossar	<i>Glossar</i> to E. and M Leumann 1933–36, pp. 385–530
H	a heavy metrical syllable counting two moras
IAP	Instrumental-Ablative Plural
IAS	Instrumental-Ablative Singular
ij	injunctive
intr	intransitive
IOL Khot	India Office Library Khotanese; manuscript designation in Skjærvø 2002
iv	imperative

KhSuf	<i>Khotanische Suffixe</i> , Degener 1989
KT1–KT3	Khotanese Texts I–III, Bailey 1969
L	a light metrical syllable counting one mora
LKh	Late Khotanese
LKh	Late Khotanese
LMC	Late Middle Chinese
LOT	Later Old Turkic
LOT	Later Old Turkic
LP	Locative Plural
LS	Locative Singular
m	1) middle verb; 2) masculine noun
n	neuter noun
NAP	nominative-accusative plural
Neb	<i>Nebenstücke</i> , E. Leumann 1920
NS	nominative singular
OIr	Old Iranian
OKh	Old Khotanese
op.	optative
Or.	Oriental; manuscript designation in Skjærvø 2002
OT	Old Turkic
p	present indicative
p.nec	participle of necessity
ppp	past passive participle
Rk	<i>Ratnakūṭa</i> , §23–25 = Maggi 2015; §94–157 = Skjærvø 2003
Sgh	<i>The Khotanese Saṅghāṭasūtra</i> , Canevascini 1993; square brackets [ ] contain manuscript numbers
SGS	<i>Saka Grammatical Studies</i> , Emmerick 1968a
Śgs	<i>The Khotanese Śūraṅgamasamādhisūtra</i> , Emmerick 1970
SI P, SI M	Serindia Petrovsky, Serindia Malov; manuscript designations in Emmerick and Vorob’eva-Desjatovskaja 1995
sj	subjunctive
Skt	Sanskrit
St.I–III	<i>Studies in the Vocabulary of Khotanese, I–III</i> , Emmerick and Skjærvø 1982, 1987, 1997
Suv	<i>The Khotanese Suvarṇabhāsottamasūtra</i> , text and translation, Skjærvø 2004a; square brackets [ ] contain manuscript abbreviations
SuvI	<i>The Khotanese Suvarṇabhāsottamasūtra</i> , vol. I, v-lxxix, Skjærvø 2004a
SuvII	<i>The Khotanese Suvarṇabhāsottamasūtra</i> , vol. II, Skjærvø 2004b
Tib	Tibetan
tr	transitive
V	any vowel, short, long or diphthong
ṽ	any short vowel
Ṿ	any long vowel
Vim	<i>Vimalakīrtinirdeśasūtra</i> , Skjærvø 1986

VP	vocative plural
VS	vocative singular
X	metrical syllable often counting as if two light syllables but not equivalent to a heavy syllable; possibly only from contracted vowels and the IAP suffix <i>-yau</i>
Z	<i>The Book of Zambasta</i> , Emmerick 1968b

## Bibliography

- Asmussen, Jes Peter. 1961. *The Khotanese Bhadracaryādeśanā. Text, Translation, and Glossary, Together with the Buddhist Sanskrit Original*. Vol. 2. Historisk-Filosofiske Meddelelser Udgivet Af Det Kongelige Danske Videnskabernes Selskab 39. København: Ejnar Munksgaard.
- Bailey, H.W. 1938. "Vajra-Prajñā-Pāramitā." *Zeitschrift Der Deutschen Morgenländischen Gesellschaft* 92: 579–93.
- . 1944. "A Turkish-Khotanese Vocabulary." *Bulletin of the School of Oriental and African Studies* 11 (2): 290–96.
- . 1946. "Gāndhārī." *Bulletin of the School of Oriental and African Studies* 11: 764–97.
- . 1958. "Languages of the Saka." In *Handbuch Der Orientalistik. 1te Abt. Der Nahe Und Der Mittlere Osten, 4ter Bd. Iranistik, 1ter Absch. Linguistik.*, 131–54. Leiden: E. J. Brill.
- . 1961. *Indo-Scythian Studies being Khotanese Texts Volume IV*. Cambridge: Cambridge Univ. Press.
- . 1963. *Indo-Scythian Studies being Khotanese Texts Volume V*. Cambridge: Cambridge Univ. Press.
- . 1966. "The Sudhana Poem of Rddhiprabhava." *Bulletin of the School of Oriental and African Studies* 29: 506–32.
- . 1969. *Indo-Scythian Studies being Khotanese Texts I-III*. 2nd ed. Cambridge: Cambridge Univ. Press.
- . 1979. *Dictionary of Khotan Saka*. Cambridge: Cambridge Univ. Press.
- . 1981. *Khotanese Buddhist texts*. Rev. ed. (orig. 1951). Cambridge: Cambridge Univ. Press.
- Baums, Stefan, and Andrew Glass. ongoing. *A Dictionary of Gāndhārī*. Munich and Seattle. [https://www.gandhari.org/n\\_dictionary.php](https://www.gandhari.org/n_dictionary.php).
- Brasington, Ron. 2012. *A NetLogo Interface to Updated UPSID*. <http://www.personal.rdg.ac.uk/~llsling1/upsid.netlogo/UPSID.updated.interface.html>.
- Brough, John. 1962. *The Gāndhārī Dharmapada*. London: Oxford Univ. Press.

- Burrow, Thomas. 1934. "Iranian Words in the Kharoṣṭhi Documents from Chinese Turkestan." *Bulletin of the School of Oriental and African Studies* VII (3): 510–16.
- . 1937. *The Language of the Kharoṣṭhī Documents from Chinese Turkestan*. Cambridge: Cambridge Univ. Press.
- Canevascini, Giotto. 1993. *The Khotanese Saṅghāṭasūtra, A critical edition*. Beiträge Zur Iranistik 14. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Cheung, Johnny. 2007. *Etymological Dictionary of the Iranian Verb*. Leiden Indo-European Etymological Dictionary Series 2. Leiden: Brill.
- Cipriano, Palmira. 1998. *La Labiovelare Iranica, Dalle Sue Origini Indoreuropee Agli Sviluppi Attuali*. Biblioteca Di Ricerche Linguistiche E Filologiche 48. Viterbo/Rome: Università della Tuscia/Università "La Sapienza."
- Clauson, Gerard. 1962. *Turkish and Mongolian Studies*. London: The Royal Asiatic Society of Great Britain and Ireland.
- . 1972. *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*. Oxford: Clarendon Press.
- . 1973. "The Turkish-Khotanese Vocabulary Re-Edited." *Islâm Tetkikleri Enstitüsü Dergisi* V (1–4): 37–45.
- Dankoff, Robert, and James Kelly, eds. 1984. *Maḥmūd El-Kāšyarī, Compendium of the Turkic Dialects. Part 1, 1982. Part II, 1984. Part III, 1985*. 3 vols. Sources of Oriental Languages and Literature 7. Duxbury, Mass.
- Degener, Almuth. 1989. *Khotanische Suffixe*. Alt- und Neu-Indische Studien, herausgegeben vom Institut für Kultur und Geschichte Indiens und Tibets an der Universität Hamburg 39. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Dresden, Mark J. 1955. *The Jātakastava or Praise of the Buddha's Former Births, Indo-Scythian (Khotanese) text, English translation, grammatical notes, and glossaries*. Transactions of the American Philosophical Society. New series., 45, part 5. Philadelphia: American Philosophical Society.
- Dunkel, G. E. 1987. "A Typology of Metanalysis in Indo-European." In *Studies in Memory of Warren Cowgill (1929-1985). Papers from the Fourth East Coast Indo-European Conference, Cornell University, June 6-9, 1985*, edited by Calvert Watkins, 7–37. Untersuchungen Zur Indogemanischen Sprach- Und Kulturwissenschaft, n.F., 3. Berlin; New York: W. de Gruyter.
- Durkin-Meisterernst, Desmond. 2014. *Grammatik des Westmitteliranischen: (Parthisch und Mittelpersisch)*. Sitzungsberichte / Österreichische Akademie der



Wissenschaften, Philosophisch-Historische Klasse 850. Wien: Verl. der Österr. Akad. der Wiss.

Edel'man, D.I. 1987. "Šugnano-Rušanskaja Jazykovaja Gruppa." In Rastorgueva 1987, 236–347.

Edel'man, D.I., and A.L. Grjunberg. 1987. "Afganskij Jazyk." In Rastorgueva 1987, 6–154.

Edgerton, Franklin. 1985. *Buddhist Hybrid Sanskrit Grammar and Dictionary. Volume 1 Grammar. Volume II: Dictionary.* (First edition, New Haven, 1953). 2 vols. Delhi: Motilal Banarsidass reprint.

Emmerick, R.E. 1967. "The Ten New Folios of Khotanese." *Asia Major* N.S. xiii (1–2).

———. 1968a. *Saka Grammatical Studies.* London Oriental Series 21. London: Oxford Univ. Press.

———. 1968b. *The Book of Zambasta, A Khotanese Poem on Buddhism.* London Oriental Series 21. London: Oxford Univ. Press.

———. 1970. *The Khotanese Śūraṅgamasamādhisūtra.* London Oriental Series 23. London: Oxford Univ. Press.

———. 1979. "The Vowel Phonemes of Khotanese." In *Studies in Diachronic, Synchronic, and Typological Linguistics. Festschrift for Oswald Szemerényi*, edited by Bela Brogyanyi. Amsterdam Studies in the Theory and History of Linguistic Science, Series IV, Current Issues in Linguistic Theory, Vol. II. Amsterdam: John Benjamins.

———. 1981. "The Consonant Phonemes of Khotanese." In *Monumentum Georg Morgenstierne*, edited by Gilbert Lazard, 1:185–209. Acta Iranica 21. Leiden: E. J. Brill.

———. 1987. "The Transition from Old to Late Khotanese." In *Transition periods in Iranian history: actes du Symposium de Fribourg-en-Brisgau (22-24 mai 1985)*, by Societas Iranologica Europaea. Leuven: E. Peeters.

———. 1989. "Khotanese and Tumshuqese." In *Compendium Linguarum Iranicarum*, edited by Rüdiger Schmitt, 204–29. Wiesbaden: Dr. Ludwig Reichert Verlag.

———. 1991. "Iranian." In *Indo-European Numerals*, edited by Jadranka Gvazdanović, 289–347. Trends in Linguistics, Studies and Monographs 57. Berlin/New York: Mouton de Gruyter.

- . 1992a. *A Guide to the Literature of Khotan*. 2nd ed. Studia Philologica Buddhica, Occasional Paper Series, III. Tokyo: The International Institute for Buddhist Studies.
- . 1992b. “The Dunhuang MS Ch 00120: Its Importance for Reconstructing the Phonological System of Khotanese.” In *Turfan and Tun-Huang, the Texts: Encounter of Civilizations on the Silk Route*, edited by Alfredo Cadonna. *Orientalia Venetiana* 4. Florence: Leo S. Olschki.
- . 1995. “On the St. Petersburg Folios of the Khotanese Suvarṇarbhāsottamasūtra.” In *Au Carrefour Des Religions. Mélanges Offerts à Philippe Gignoux*, edited by Rika Gyselen. *Res Orientales*, VII. Bures-sur-Yvette: Groupe pour l’Étude de la Civilisation du Moyen-Orient.
- . 1998. “Khotanese *ei*.” In *Proceedings of the Third European Conference of Iranian Studies. Part 1. Old and Middle Iranian Studies.*, 93–97. *Beiträge Zur Iranistik* 17. Wiesbaden: Reichert.
- . 2009. “Khotanese and Tumshuqese.” In *The Iranian Languages*, edited by Gernot Windfuhr. London and New York: Routledge.
- Emmerick, R.E., and Mauro Maggi. 1991. “Thoughts on Khotanese *e* and *o*.” In *Corolla Iranica. Papers in Honour of Prof. Dr. David Neil MacKenzie on the Occasion of His 65th Birthday on the Eighth of April 1991*, edited by R.E. Emmerick and Dieter Weber, 67–73. Frankfurt am Main: Peter Lang.
- Emmerick, R.E., and Edwin G. Pulleyblank. 1993. *A Chinese Text in Central Asian Brahmi Script. New Evidence for the Pronunciation of Late Middle Chinese and Khotanese*. Vol. LXIX. Serie Orientale Roma. Rome: Istituto Italiano per il Medio ed Estremo Oriente.
- Emmerick, R.E., and András Róna-Tas. 1992. “The Turkish-Khotanese Wordlist Revisited.” *Central Asiatic Journal* 36 (3–4): 199–241.
- Emmerick, R.E., and P.O. Skjærvø. 1982. *Studies in the Vocabulary of Khotanese I*. Veröffentlichungen Der Iranischen Kommission 12. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- . 1987. *Studies in the Vocabulary of Khotanese II*. Veröffentlichungen Der Iranischen Kommission 17. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- . 1997. *Studies in the Vocabulary of Khotanese III*. Veröffentlichungen Der Kommission Für Iranistik, Nr. 27. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

- Emmerick, R.E., and Margarita I. Vorob'eva-Desjatovskaja. 1995. *Saka Documents, Text Volume III: The St. Petersburg Collections*. Corpus Inscriptionum Iranicarum, Part II Inscriptions of the Seleucid and Parthian Periods and of Eastern Iran and Central Asia, Vol. V Saka. London: School of Oriental and African Studies.
- Erdal, Marcel. 1979. "The Chronological Classification of Old Turkic Texts." *Central Asiatic Journal* 23: 151–75.
- Gamkrelidze, Thomas V. 1978. "On the Correlation of Stops and Fricatives in a Phonological System." In *Universals of Human Language. Volume 2. Phonology*, edited by Joseph H. Greenberg. Stanford: Stanford University Press.
- Gercenberg, L. G. [Л.Г. Герценберг]. 1965. *Xotano-Sakskij Jazyk [Хотано-Сакский Язык]*. Moscow: Nauka.
- Grjunberg, A.L. 1987. "Mundžanskij Jazyk." In Rastorgueva 1987, 155–235.
- Harmatta, J. 1969. "The Bactrian Well-Inscription from Kara Tepe." In *Buddijskie Peščery Kara Tepe v Starom Termeze*, 82–127. Moscow: Nauka.
- Hitch, Doug. 1984. "Kharoṣṭhī Influences on the Saka Brāhmī Scripts." In *Middle Iranian Studies*, edited by W. Skalmowski and A. van Tongerloo, 187–202. *Orientalia Lovaniensia Analecta* 16. Leuven: Peters.
- . 1987. "Old Khotanese Consonantal Palatality, Aspiration and Retroflexion." Paper presented November 16, 1987, at the Middle Eastern Studies Association meeting in Baltimore.
- . 1989. "Old Turkic 'b' and 'p.'" In *Religious and Lay Symbolism in the Altaic World and Other Papers. Proceedings of the 27th Meeting of the Permanent International Altaistic Conference, Walberberg, Federal Republic of Germany, June 12th to 17th, 1984*, edited by Klaus Sagaster, 130–46. Wiesbaden: Harrassowitz.
- . 1990. "Old Khotanese Synchronic Umlaut." *Indo-Iranian Journal* 33: 177–98.
- . 2014. "Meter in the Old Khotanese Book of Zambasta." *Ars Metrica* 2014/11: 1–43. <http://ars-metrica.germ-ling.uni-bamberg.de/ars-metrica-201411/>.
- . 2015a. "Tt in Old Khotanese." *Journal of the American Oriental Society* 135 (4): 663–87.
- . 2015b. "Contracted Diphthongs in Old Khotanese." *Indo-Iranian Journal* 58: 293–394.

- . in-press-a. “Contracted Semivowels in Old Khotanese.” *Indo-Iranian Journal*
- . in-press-b. “Old Khotanese Type A Stems in -a- and -E-.” *Bulletin of the School of Oriental and African Studies*
- Hoffmann, Karl, and Bernhard Forssman. 2004. *Avestische Laut- und Flexionslehre*. 2., durchgesehene u. erw. Aufl. Innsbrucker Beiträge zur Sprachwissenschaft 115. Innsbruck: Institut für Sprachen und Literaturen der Universität Innsbruck.
- Hovdhaugen, Even. 1971. “Turkish Words in Khotanese Texts, A Linguistic Analysis.” *Norsk Tidsskrift for Sprogvidenskap* XXIV: 772–823.
- Konow, Sten. 1932. *Saka Studies*. Oslo Etnografiske Museum. Bulletin 5. Oslo: [Glückstadt and Hamburg, Printed by J.J. Augustin].
- . 1935. “Ein Neuer Saka-Dialekt.” *Sitzungsberichte Der Preussischen Akademie Der Wissenschaften Phil.-Hist. Kl. XX*: 772–823.
- . 1941. *Khotansakische Grammatik mit Bibliographie, Lesestücken und Wörterverzeichnis. Mit einer Schrifttafel*. Leipzig: Harrassowitz.
- . 1948. “Khotanese āriyāmata.” *Acta Orientalia (Copenhagen)* 20: 77–80.
- . 1949. “Primer of Khotanese Saka. Grammatical Sketch, Chrestomathy, Vocabulary, Bibliography.” Edited by G. Morgenstierne. *Norsk Tidsskrift For Sprogvidenskap* 15: 1–136.
- Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- Leumann, Ernst. 1912. *Zur nordarischen Sprache und Literatur ; Vorbemerkungen und vier Aufsätze mit Glossar*. Strassburg: Deutsche Morgenländische Gesellschaft. 1966 reprint, Nendeln, Liechtenstein: Kraus.
- . 1919. *Maitreya-Samiti, Das Zukunftsideal Der Buddhisten; Die Nordarische Schilderung in Text Und Übersetzung, Nebst Sieben Andern Schilderungen in Text Oder Übersetzung, Mit Einer Begründung Der Indogermanischen Metrik*. Strassburg: K. J. Trübner. <http://catalog.hathitrust.org/Record/100322037>.
- . 1920. *Buddhistische Literatur: nordarisch und deutsch. I. Teil: Nebenstücke*. Abhandlungen für die Kunde des Morgenlandes 15.2. Leipzig: Brockhaus. 1966 reprint, Nendeln, Liechtenstein: Kraus.
- Leumann, Ernst, and Manu Leumann. x1933-36. *Das nordarische (sakische) Leergedicht des Buddhismus. Text und Übersetzung von Ernst Leumann. Aus dem Nachlass herausgegeben von Manu Leumann*. Abhandlungen für die Kunde des

- Morgenlandes 20.1–3. Leipzig: Deutsche Morgenländische Gesellschaft. 1966 reprint, Nendeln, Liechtenstein: Kraus.
- Leumann, Manu. 1967. "Neue Fragmente Des Altkhotanischen 'Lehrgedichts.'" *Zeitschrift Der Deutschen Morgenländischen Gesellschaft* 117: 366–75.
- Lindau, Mona, Kjell Norlin, and Jan-Olof Svantesson. 1985. "Cross-Linguistic Differences in Diphthongs." *UCLA Working Papers in Phonetics* 61: 40–44.
- Lüders, Heinrich. 1936. "Textilien Im Alten Turkistan." *Abhandlungen Der Preussischen Akademie Der Wissenschaften Jahrgang 1936* (3): 3–38.
- Maddieson, Ian. 2009. *Patterns of Sounds*. Cambridge Studies in Speech Science and Communication. Cambridge University Press.  
<https://books.google.ca/books?id=iYBdPwAACAAJ>.
- Maddieson, Ian, and Kristin Precoda. ongoing. *The UCLA Phonological Segment Inventory Database*.  
<http://www.linguistics.ucla.edu/faciliti/sales/software.htm#upsid>: Last accessed 16 March 2016.
- Maggi, Mauro. 1995a. "Late Khotanese Numerals from 1 to 36 in Tibetan Transcription." In *Proceedings of the Second European Conference of Iranian Studies*, 425–43. Rome: Istituto Italiano per il Medio ed Estremo Oriente.
- . 1995b. *The Khotanese Karmavibhanga*. Serie Orientale Roma 74. Rome: Istituto Italiano per il Medio ed Estremo Oriente.
- . 2004. "The Manuscript T III S 16: Its Importance for the History of Khotanese Literature." In *Turfan Revisited: The First Century of Research into the Arts and Cultures of the Silk Road*, edited by D. Durkin-Meisterernst et al., 184–90 and plate 36 p. 457. Berlin: Dietrich Reimer.
- . 2009. "Annotations on the Book of Zambasta, I." In *Literarische Stoffe Und Ihre Gestaltung in Mitteliranischer Zeit*, edited by Desmond Durkin-Meisterernst, Christiane Reck, and Dieter Weber, 153–71 and Tafel I. Beiträge Zur Iranistik 31. Wiesbaden: Riechert.
- . 2015. "A Folio of the Ratnakūṭa (Kāśyapaparivarta) in Khotanese." *Fa Gu Foxue Xuebao 法鼓佛學學報 = Dharma Drum Journal of Buddhist Studies* 17: 101–43.
- . in-press. "Annotations on the Book of Zambasta, IV: Ronald E. Emmerick's Notes." Draft of 11 August 2013.
- Mayrhofer, Manfred. 1996. *Etymologisches Wörterbuch des Altindoarischen*. Vol. I. Indogermanische Bibliothek. II. Reihe, Wörterbücher. Heidelberg: C. Winter.

- Melikishvili, I. 1970. "The conditions of markedness for the features voicing, voicelessness, labiality, and velarity. (in Georgian)." *Matsne* 4: 137–58.
- . 1972. "Otnoshenije Markirovannosti v Fonologii (uslovija Markirovannosti v Klasse Shumnykh Fonem)." *Avtoreferat dissertatsii na soiskanije uchenoj stepeni kandidata filologicheskikh nauk., Tbilisi.*
- Nadelyaev (Наделяев), V.M. (В.М.), D.M. (Д.М.) Nasilov (Насилов), E.R. (Э.Р.) Tenishev (Тенишев), and A.M. (А.М.) Shcherbak (Щербак), eds. 1969. *Drevnetyurkskij Slovar' (Древнетюркский Словарь)*. Leningrad (Ленинград): Izdatel'stvo "Nauka" (Издательство «Наука»).
- Nadzhip, E.N. 1971. *Modern Uigur*. Languages of Asia and Africa. Moscow: Central Department of oriental Literature.
- Paxalina, T.N. 1987a. "Vaxanskij Jazyk." In *Rastorgueva 1987*, 408–73.
- . 1987. "Iškašimskij Jazyk." In *Rastorgueva 1987*, 474–536.
- Pulleyblank, Edwin G. 1984. *Middle Chinese, A Study in Historical Phonology*. Vancouver: University of British Columbia Press.
- . 1986. "Chinese in Khotanese Brahmi. Presented at the 32nd International Congress on Asian and North African Studies, Hamburg, 25-30 August, 1986."
- Rastorgueva, V.S., ed. 1987. *Novoiranskije Jazyki: Vostočnaja Gruppya*. Moscow: Nauka.
- Reetz, Henning. undated. *Simple Interface to the UPSID Database*. (see Maddieson and Precoda). <http://web.phonetik.uni-frankfurt.de/upsid.html>: Last accessed 16 March 2016.
- Sander, Lore. 1968. *Paläographisches Zu Den Sanskrithandschriften Der Berliner Turfansammlung*. Verzeichnis Der Orientalischen Handschriften in Deutschland, Supplementband 8. Wiesbaden: Franz Steiner Verlag.
- Sims-Williams, Nicholas. 1990. "Chotano-Sogdica II: Aspects of the Development of Nominal Morphology in Khotanese and Sogdian." In *Proceedings of the First European Conference of Iranian Studies, held in Turin, September 7th-11th, 1987, by the Societas Iranologica Europaea. Part 1. Old and Middle Iranian Studies*, edited by Gherado Gnoli and Antonio Panaino, 275–96. Istituto Italiano per il Medio ed Estremo Oriente.
- . 2007. *Bactrian Documents from Northern Afghanistan II: Letters and Buddhist Texts*. Vol. III. Studies in the Khalili Collections, Volume III. London: Nour Foundation.

- Skjærvø, P.O. 2004b. *This Most Excellent Shine of Gold, King of Kings of Sutras, The Khotanese Suvarṇabhāsottamasūtra. Volume II, Manuscripts, Commentary, Glossary, Indexes*. Sources of Oriental Languages and Literatures 61. Central Asian Sources VI. Cambridge, Mass.: The Department of Near Eastern Languages and Civilizations, Harvard University.
- . 2004a. *This Most Excellent Shine of Gold, King of Kings of Sutras, The Khotanese Suvarṇabhāsottamasūtra. Volume I, The Khotanese Text with English Translation and the Complete Sanskrit Text*. Sources of Oriental Languages and Literatures, 60. Central Asian Sources V. Cambridge, Mass.: The Department of Near Eastern Languages and Civilizations, Harvard University.
- . 1986. "Khotanese Fragments of the Vimalakīrtinirdeśasūtra." In *Kalyāṇamitrārāgaṇam, Essays in Honour of Nils Simonsson*, 229–60. Oslo: Norwegian Univ. Press.
- . 2002. *Khotanese Manuscripts From Chinese Turkestan in The British Library, A Complete Catalogue with Texts and Translations. With Contributions by Ursula Sims-Williams*. Corpus Inscriptionum Iranicarum, Part II, Vol. V: Saka, Texts VI. London: The British Library.
- . 2003. "Fragments of the Ratnakūṭa-sūtra (Kāśyapaparivarta) in Khotanese." In *Religious themes and texts of pre-Islamic Iran and Central Asia. Studies in honour of Professor Gherardo Gnoli on the occasion of his 65th birthday of 6th December 2002*, edited by Carlo G. Cereti, Mauro Maggi, and Elio Provasi. Beiträge Zur Iranistik 24. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Tedesco, Paul. 1926. "Ostiranische Nominalflexion." *Zeitschrift Für Indologie Und Iranistik* 4.
- Tudkeao, Chanwit. 2010. "Zentralasiatische Versionen Des Ratnaketuparivarta." Unpublished PhD dissertation, Munich: Ludwig-Maximilians-Universität München.
- von Gabain, Annemarie. 1974. *Alttürkische Grammatik*. 3rd ed. Porta Linguarum Orientalium, XXIII. Leipzig: Harrassowitz.
- von Hinüber, Oskar. 1981. "Die Paiśāci Und Die Entstehung Der Sakischen Orthographie." In *Studien Zum Jainismus Und Buddhismus. Gedenkschrift Für L. Asldorf*. Wiesbaden: Steiner.
- . 2001. *Das ältere Mittelindisch Im Überblick*. Österreichische Akademie Der Wissenschaften, Philosophisch-Historische Klasse Sitzungsberichte, 467. Band. = Veröffentlichungen Der Kommission Für Sprachen Und Kulturen Südasiens, Heft 20. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.

Yakubovich, Ilya. 2002. "Nugae Sogdicae." *Bulletin of the School of Oriental and African Studies* 65 (3): 544-49.