Development of Tibetic verb paradigms: Diachronic and paradigm-based explanations

Kitty Wenying Liu & Samuel Andersson*

Abstract. This paper examines two issues in the historical morphology of Tibetic verbs from the perspective of paradigm morphology. The first issue is why the historical future stem was lost after the Old Tibetan period, which we answer using the measure of paradigm conditional entropy. We hypothesise that the loss of the future stem may be due to its lower informativity compared to other paradigm cells. The second issue is diachronic restrictions on syncretism patterns in modern Amdo varieties, where we find that analogical extension never gives rise to imperfective-imperative syncretism to the exclusion of the perfective. We link this to the literature on *ABA, and show that discussions of synchronic morphological paradigms benefit from considering diachronic explanations.

Keywords. historical linguistics, historical morphology, verb paradigms, conditional entropy, *ABA, Tibetic languages, Old Tibetan, Amdo Tibetan

- **1. Introduction**. This paper examines two issues in the historical morphology of Tibetic verbs: the loss of the future stem after the Old Tibetan period, and restrictions in verb stem syncretisms in the verbal paradigm of modern Amdo varieties. We approach both of these questions from a paradigm morphology perspective, which produces novel insights for explaining the attested patterns. After an overview of Tibetic verbal morphology in Section 2, we discuss the loss of the Old Tibetan future stem in Section 3. We make use of conditional entropy, a measure of paradigm predictability (Ackerman and Malouf 2013), and hypothesise that the future stem is lost due to its being the least morphologically informative stem in the Old Tibetan verbal paradigm. Losing the future, as opposed to any other tense-aspect-mood category, minimally disrupts the paradigm's predictive relationships. In Section 4, we discuss patterns of syncretism in modern Amdo Tibetic verbal paradigms. Verbs in modern Amdo varieties have up to three distinct stems, and one specific pattern of stem syncretism (the imperfective and imperative share a stem to the exclusion of the perfective) is much rarer than other patterns. We identify a diachronic explanation for this: imperfective-imperative syncretisms only arise as a result of phonological mergers, while other syncretisms can also arise due to paradigm levelling. We hypothesise that this diachronic restriction on paradigm levelling may be semantically motivated, in ways which recall *ABA restrictions. We conclude that some *ABA restrictions may be fruitfully explained by taking diachrony into account.
- **2. Verbal morphology in Tibetic.** This paper discusses the verbal morphology of Old Tibetan, and of two modern Tibetic varieties in the Amdo subgroup. An Old Tibetan verb has up to four

^{*} We would like to thank Marieke Meelen for supervising the original Masters dissertation (Liu, 2024) that the work in this paper is based on, and the Fieldwork/Documentation Working Group at Yale and our LSA audience for their questions and comments. Authors: Kitty Wenying Liu, University of Cambridge (wl399@cam.ac.uk), Samuel Andersson.

¹ The label 'Old Tibetan' is used loosely in this paper to refer to historical written Tibetan generally, of both the Old Tibetan (7-11th c.) and Classical Tibetan (12-19th c.) periods. Old Tibetan in this proper sense approximates the common ancestor of the modern Tibetic family, Old Tibetan word forms have transparent reflexes in modern Tibetic varieties. Classical Tibetan was used as a diglossic literary language, largely maintaining the linguistic features of

distinct monosyllabic stems, which populate a four-cell morphological paradigm whose cells express different tense-aspect-mood (TAM) information. The cells and corresponding stems are generally referred to as the present, past, future, and imperative, after their primary functions, although each stem can also express other TAM functions when occurring with morphosyntactic particles (Zeisler, 2004). Table 1 shows some examples of Old Tibetan verbs (Hill, 2010:xv-xxi), using the Wylie romanisation scheme.

	Present	Past	Future	Imperative
to revere	'khur	bkurd	bkur	khurd
to trickle	gtig	btigs	btig	gtigs
to hear	nyan	mnyand	mnyan	nyond
to scatter	'gyed	bkyes	dgye	khyes
to break (tr.)	dkrum	dkrums	dkrum	dkrums
to explain	'chags	bshags	bshags	'chags
to be surprised	had	had	had	

Table 1. Examples of Old Tibetan verbs.

As shown in Table 1, stems of the same verb can differ by a combination of prefixation, suffixation, changes to the stem vowel, and consonant voicing alternations. This is an instance of gestalt exponence as discussed by Matthews (1991:170-179) and Blevins et al (2019), where the meaning of each paradigm cell cannot be localised to a specific exponent. For example, b- or -s do not individually expone the past stem, since b- also occurs in the future stem and -s in the imperative stem, but the past is exponed when b- and -s co-occur. The extent of gestalt exponence makes it difficult to analyse Old Tibetan verbal inflection through concatenation of individually meaningful morphs, and the present paper instead uses a paradigm-based approach.

Some apparent irregularities in Table 1 can be explained. The verb 'to be surprised' lacks an imperative stem because it denotes a categorically involuntary action (cf. Bialek, 2020). The future of 'to explain' has an -s, which is rare for future forms, perhaps because of morphological extension from the -s found in all other paradigm cells of this verb. Some forms are also affected by historical sound changes (e.g. g- and d- occur in complementary distribution, as do -s and -d after consonants; for 'to hear', *bnyand > mnyand through assimilation; for 'to explain', *'shags > 'chags through fortition; Hill, 2019:9ff.). However, although some of the irregularities can be explained, Old Tibetan verbal morphology nevertheless forms a complex synchronic system, parts of which were opaque and non-productive for all except its earliest users. We return to the idea of which realisations of Old Tibetan paradigm cells are synchronically predictable in Section 3.

_

Old Tibetan (including verbal morphology) but with some lexical changes and orthographic standardisations (see Tournadre & Suzuki, 2023:190ff. for discussion). Since Classical Tibetan is attested in more texts, and in many dedicated reference grammars and dictionaries, much of the current knowledge of Old Tibetan morphosyntax builds on that of Classical Tibetan, making it difficult to extricate the two without further specific research. Therefore, in this paper, we do not specifically distinguish between verbal morphology in Old vs Classical Tibetan. While this is not ideal, differences between Old and Classical Tibetan pale in comparison to the extent of linguistic difference between historical written Tibetan and any modern spoken variety. We refer to the conglomerated Old-and-Classical-period historical language as 'Old Tibetan' in our discussions of verbal morphology in this paper, because the morphological system originates in the Old Tibetan system, and was likely already a fossilised archaism at the start of the Classical Tibetan period (Bialek, 2022:6).

The development of the verbal paradigm after the Old Tibetan period was affected by language-wide phonological changes and by a restructuring of the verbal paradigm's primary functions. All Tibetic varieties underwent syllable structure reduction after the Old Tibetan period, leading to phonological mergers between distinct Old Tibetan syllables, to different extents across the modern Tibetic language family. Many more innovative varieties lost alternations between present-past-future-imperative stems altogether, and the aforementioned mergers impacted the varieties which still retain stem alternation, such as the Amdo subgroup discussed in Section 4. In modern varieties with stem alternation, the four-cell Old Tibetan paradigm primarily distinguished by tense and mood is replaced by a three-cell paradigm which is primarily distinguished by aspect and mood. The Old Tibetan present, past, and imperative stems developed into the modern imperfective, perfective, and imperative stems respectively (Zeisler, 2004; Bielmeier et al, 2018). The future stem is lost in all modern varieties (Tournadre & Suzuki, 2023:367), a puzzle which we attempt to shed light on in Section 3 below.

3. Loss of the Old Tibetan future. As stated above, no modern Tibetic variety retains a distinct reflex for the Old Tibetan future stem. We hypothesise that this may be due to the comparative lack of informativity of the Old Tibetan future stem compared to the three other stems, in a sense which we make formally precise immediately below.

We operationalise morphological informativity as conditional entropy, following Ackerman and Malouf (2013; see also Ruan, 2024). Given knowledge of a morphological paradigm (i.e. all morphosyntactic distinctions and their realisations across all morphological classes), the conditional entropy of paradigm cell A given paradigm cell B answers the question: how difficult is it to predict a word's realisation for A, given knowledge of the realisation for B? The ecological validity of conditional entropy as a measure of informativity can be increased by incorporating information about the frequency of the paradigm cells and declension classes. We present calculations without frequency information here, and refer the interested reader to Liu (2024) for additional discussion. We base our calculations based on the paradigms in Table 2, and Table 3 shows the conditional entropy of these paradigms. Conditional entropy is calculated using the scripts from Andersson (2025), which are freely available online and which can be used to calculate conditional entropy for any language, with or without frequency information.

In Table 2, Σ refers to the stem. Vowels in parentheses denote non-concatenative changes to the stem vowel. Subscript V means that the stem-initial consonant is voiced, while subscript U means that the stem-initial consonant is voiceless. g- represents both g-/d- allophones (see Section 2 for discussion). -d and -s are kept distinct because they are only allophones when occurring after another consonant. n- represents orthographic <q> (Wylie romanisation '-), thought by some to have represented prenasalisation (e.g. Jacques, 2012b).

CLASSES	Present	Past	Future	Imperative
1	n-Σ	b-Σ-s	b-Σ	Σ(o)-s
2	n-Σ-d	b-Σ-s	b-Σ	Σ(o)-s
3	Σ-d	b-Σ-s	b-Σ	Σ(o)-s
4	g-Σ(o)	b-Σ-s	b-Σ	Σ(o)-s
5	g- Σ(o)	b-Σ	g-Σ	Σ(o)-s
6	n - $\Sigma_{ m v}$	b - Σ_{u}	g - $\Sigma_{ m v}$	$\Sigma_{\rm u}({\rm o})$ -s
7	$n-\Sigma_v(e)-d$	b - Σ_{u}	g - $\Sigma_{ m v}$	$\Sigma_{\rm u}({\rm o})$ -s
8	n - $\Sigma_{ m v}$	b-Σ _u -s	g - Σ_{v}	$\Sigma_{\rm u}({\rm o})$ -s
9	$n-\Sigma_v(e)-d$	b-Σ _u -s	g - $\Sigma_{ m v}$	$\Sigma_{\rm u}({\rm o})$ -s
10	n-Σ	Σ-s	n-Σ	Σ-s
invar 1	n-Σ	n-Σ	n-Σ	n-Σ
invar 2	b-Σ	b-Σ	b-Σ	b-Σ
invar 3	g-Σ	g-Σ	g-Σ	g-Σ
invar 4	Σ	Σ	Σ	Σ

Table 2. Old Tibetan verb paradigm (Coblin, 1976; Hill, 2010:xv-xxi; Jacques, 2012a; Liu, 2024)

H(column row)	Present	Past	Future	Imperative	E(row)
Present		0.768	0.340	0.340	0.482
Past	1.000		0.143	0.143	0.429
Future	1.258	0.829		0.544	0.877
Imperative	0.972	0.544	0.258		0.591
E(column)	1.077	0.714	0.247	0.342	Avg: 0.595

Table 3. Conditional entropy of the Old Tibetan verb paradigm from Table 2. Each cell represents H(column|row), the difficulty of guessing the column paradigm cell given knowledge of the row paradigm cell. E(X) is the average of the values in each row/column, Avg is the average conditional entropy of the whole table.

Table 3 shows that the future stem is the least morphologically informative stem in the paradigm. It is both the easiest to predict when a verb's other stems are known, H(Future|E(col)) = 0.247, and the most difficult to use to predict other stems, H(E(row)|Future) = 0.877. Losing the future therefore minimally disrupts the paradigm's predictive relationships, a factor we hypothesise is important in the diachronic development of morphological paradigms. This echoes Ruan (2024), who argues that conditional entropy may predict the direction of paradigm levelling in Germanic verb paradigms. Despite being relatively easy to predict, knowing the future stem of a verb is relatively unhelpful for learning the other forms in the paradigm (cf. Albright, 2008). More informative stems were either retained diachronically, as is the case with the imperative, or else repurposed, as with the present and past which were reanalysed as imperfective and perfective respectively (see Section 2). Based on our results from Tibetic, we believe it may be fruitful to examine other patterns of morphological change through the lens of conditional entropy. In Section 4 below, we turn from the study of Old Tibetan paradigm changes over time to the synchronic study of modern Tibetic paradigms after the diachronic changes have taken place.

4. Patterns of syncretism in modern Amdo verbs. In this section, we discuss verb stem alternation patterns in two Amdo varieties, Themchen (data from Haller, 2004) and Labrang (data from Liu, 2024). The Labrang dataset is phonologically more innovative, with a more reduced syllable structure compared to the Themchen dataset, but we found no significant differences between the morphological structure of the two varieties (Liu, 2024).

Verbs in both varieties may have an invariant stem across the paradigm, a full paradigm of three distinct stem alternants, or two cells sharing one stem to the exclusion of the third cell, as exemplified in Table 4. We refer to these patterns by using capital letter variables for each stem.

Pattern	Themchen		Labrang		
AAA	/nkhon/, /nkhon/, /nkhon/ 'to ar- 305		/lox/, /lox/, /lox/ 'to return'		
	gue'				
ABC	/ŋgəp/, /kwap/, /khop/ 'to cover'	77	/dzə/, /tsi/, /tshi/ 'to ask'	25	
AAB	/ptəl/, /ptəl/, /thəl/ 'to approach'	100	/nan/, /nan/, /non/ 'to listen'	26	
ABB	/nə/, /ni/, /ni/ 'to suck'	40	/zu/, /zwi/, /zwi/ 'to avoid, flee	10	
			from'		
ABA	/rdzəç/, /bdzəç/, /rdzəç/ 'to run'	63	/lox/, /lax/, /lox/ 'to read silently'	3	

Table 4. Different patterns of syncretism in the Themchen and Labrang verb paradigms, with the type frequency of each syncretism pattern in each variety. Stems of the same verb are listed in the order imperfective, perfective, and imperative.

The syncretic patterns arise due to a combination of inheritance, phonological mergers, and paradigm extention. Interestingly, all three of these pathways create AAA, AAB, and ABB patterns, while ABA patterns only arise through inheritance and phonological mergers, never through analogical extension from one cell to another. Table 5 shows modern Amdo syncretisms that are inherited from Old Tibetan.

	Translation	PRS / IPFV	PST / PFV	FUT	IMP
OTib		sprod	sprad	sprad	sprod
Themchen	to hand over	/şpol/	/şpal/	_	/şpol/
OTib		gcar	gcar	gcar	gcor
Themchen	to hit	/çtear/	/çtcar/	_	/çtcor/
OTib		bgo	bgos	bgo	bgos
Labrang	to divide	/go/	/gi/	_	/gi/
OTib		gtan	gtan	gtan	_
Labrang	to open (e.g. a	/tan/	/tan/	_	_
	door)				

Table 5. Syncretisms in modern Amdo verbs that are inherited from Old Tibetan.

Table 6 shows modern Amdo syncretisms that arise due to historical sound change, as a result of syllable structure simplification.

	Translation	PRS / IPFV	PST / PFV	FUT	IMP
OTib		slog	bslogs	bslog	slogs
Labrang	to come back	/lox/	/lox/		/lox/
OTib		ldad	bldad	bldad	ldod
Labrang	to lick	/da/	/da/	_	/do/
OTib		ʻphur	phur	dphur	phurd
Themchen	to massage	/mphər/	/hər/	_	/hər/
OTib		rgyug	brgyugs	brgyug	rgyugs
Themchen	to run	/rdzəç/	/bdzəç/	_	/rdzəç/

Table 6. Syncretisms in modern Amdo verbs that arise due to phonological changes that merged distinct Old Tibetan verb stems.

Finally, Table 7 shows modern Amdo syncretisms that arise due to analogical extension.

	Trans-	PRS / IPFV	PST / PFV	FUT	IMP
	lation				
OTib		ldad	bldad	bldad	ldod
Themchen	to chew	/rdal/	/rdal/ */brdal/		/rdol/
OTib		sbyong	sbyangs	sbyang	sbyongs
Labrang	to learn	/dzaŋ/ */dzoŋ/	/dzaŋ/		/dzoŋ/
OTib		'khur	bkurd	khur	khur
Themchen	to carry	/ŋkʰər/	/kʰər/ */pkər/		/kʰər/
OTib		'go	'gos	'go	gos
Themchen	to	/ŋgo/	/ŋgi/		/ŋgi/ */gi/
	climb				
OTib		dgod	bgad	bgad	dgod
Labrang	to laugh	/ga/ */go/	/ga/	_	_

Table 7. Syncretisms in modern Amdo verbs that arise due to analogical extension. Bold text marks the reflex that has been extended into another cell, and in the cell whose realisation has been altered by the extension, * marks the expected reflex for that cell based on regular sound change.

As shown in the tables above, all patterns of syncretism can arise through all three mechanisms, with the exception that syncretism between the imperfective and imperative to the exclusion of the perfective never arise through analogical extension. This may be because the imperfective and imperative stems do not share semantic properties to the exclusion of the perfective (Zeisler, 2004). The imperfective and perfective are both aspectual categories, to the exclusion of the imperative. We suggest that the perfective and imperative also share aspectual information to the exclusion of the imperfective, both describing bounded events in the near future or past, unlike the ongoing imperfective. This type of reasoning, where patterns of analogical extension are diachronically constrained by semantic similarity, comes from work such as Cristofaro (2010) on semantic maps. Echoes of this can also be found synchronically in so-called overlapping decompositions of feature spaces in relation to *ABA restrictions (see Bobaljik & Sauerland, 2018).

Our findings bring a diachronic perspective to discussions on the *ABA constraint in three-cell morphological paradigms. *ABA was first formulated in work on adjective-comparative-superlative paradigms, where patterns like *good*, *better*, *best* (ABB) or Latin *bonus*, *melior*, *optimus* (ABC) are permitted, but not ABA patterns of syncretism like *good*, *better*, *goodest* (Bobaljik, 2012). *ABA has also been extended to other domains where paradigm cells do not necessarily have an intuitively obvious ordering, but where syncretism restrictions still appear to hold (e.g. Caha, 2017; Wiese, 2004; Wiese, 2008).

In a study of Germanic verb paradigms, Andersson (2018) argues that all patterns of syncretism, including ABA, are attested synchronically, but that conjugation class frequencies in Proto-Germanic may offer a diachronic explanation for why ABA patterns of syncretism are rare. This is similar to the Tibetic situation above. Imperfective-imperative syncretisms are attested synchronically in both Themchen and Labrang, but diachronically there are fewer pathways which lead to these types of syncretism, arguably for the semantic reasons outlined above. Tibetic therefore suggests that *ABA-like restrictions, where some patterns of syncretism are rare or unattested relative to others, may be fruitfully examined from a diachronic perspective.

5. Conclusions. In this paper we have examined two issues in Tibetic verbal morphology: the loss of the Old Tibetan future across modern Tibetic varieties, as well as patterns of syncretism between paradigm cells in modern Amdo Tibetic. We have argued that patterns of conditional entropy in Old Tibetan verb paradigms may predict why the future was lost rather than any other paradigm cell: losing the future minimally disrupts the paradigm's predictive relationships. Examining the modern varieties, after the loss of the future, we have identified a diachronic restriction on syncretism: while all logically possible patterns of syncretism are attested synchronically in Amdo Tibetic, imperfective-imperative syncretism to the exclusion of the perfective never arises through analogical extension. We have offered a possible semantic explanation for this fact. Some patterns of syncretism may therefore be rarer than others not because of any synchronic constraints but because of diachronic restrictions on how they came to be. We have connected this to previous work on *ABA, and suggested that exploring *ABA from a diachronic perspective may be productive.

References

Ackerman, Farrell & Robert Malouf. 2013. Morphological organization: The Low Conditional Entropy Conjecture. *Language* 89(3). 429–464. https://doi.org/10.1353/lan.2013.0054

Albright, Adam. 2008. Explaining universal tendencies and language particulars in analogical change. In Jeff Good (ed.), *Language Universals in Language Change*. https://doi.org/10.1093/acprof:oso/9780199298495.003.0007

Andersson, Samuel. 2018. (*)ABA in Germanic verbs. *Glossa* 3(1) https://doi.org/10.5334/gjgl.733

Andersson, Samuel. 2025. morphological-entropy. https://github.com/SJAndersson/morphologi-cal-entropy [Accessed 2025-04-29]

Bialek, Joanna. 2020. Old Tibetan verb morphology and semantics: An attempt at a reconstruction. *Himalayan Linguistics* 19(1). https://doi.org/10.5070/h919145017

Bialek, Joanna. 2022. A textbook in Classical Tibetan. Abingdon: Routledge.

Bielmeier, Roland & Driem, Georg L. van & Volkart, Marianne. 2018. *Comparative Dictionary of Tibetan Dialects (CDTD). Volume 2: Verbs*. Berlin: De Gruyter. https://doi.org/10.1515/9783110554076

Blevins, James P, Farell Ackerman, Robert Malouf. 2019. Word and Paradigm morphology. In Jenny Audring & Francesca Masini (eds.), *Oxford Handbook of Morphology*, 264–284. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199668984.013.22

Bobalijk, Jonathan D. 2012. *Universals in comparative morphology: Suppletion, superlatives and the structure of words*. Cambridge, MA: MIT Press. https://doi.org/10.1515/lingty-2016-0038

- Bobaljik, Jonathan D & Sauerland, Uli. 2018. ABA and the combinatorics of morphological features. *Glossa* 3(1). https://doi.org/10.5334/gjgl.345
- Caha, Pavel. 2017. How (not) to derive a *ABA: The case of Blansitt's generalisation. *Glossa* 2(1). https://doi.org/10.5334/gjgl.348
- Carstairs-McCarthy, Andrew. 1987. Allomorphy in inflexion. London: Croom Helm.
- Coblin, W South. 1976. Notes on Tibetan verbal morphology. *T'oung Pao* 62(1-3). 45–70. https://www.jstor.org/stable/4528050
- Cristofaro, Sonia. 2010. Semantic maps and mental representation. *Linguistic Discovery* 8(1). https://doi.org/10.1349/ps1.1537-0852.a.345
- Haller, Felix. 2004. *Dialekt und Erzählungen von Themchen [Dialect and folktales from Themchen]*. Bonn: VGH Wissenschaftsverlag.
- Hill, Nathan W. 2010. *A lexicon of Tibetan verb stems as reported by the grammatical tradition*. https://github.com/tibetan-nlp/lexicon-of-tibetan-verb-stems
- Jacques, Guillaume. 2012a. An internal reconstruction of Tibetan stem alternations. *Transactions of the Philological Society* 110(2). 212–224. https://doi.org/10.1111/j.1467-968x.2012.01293.x
- Jacques, Guillaume. 2012b. A new transcription system for Old and Classical Tibetan. *Linguistics of the Tibeto-Burman area*. John Benjamins Publishing Company 35(2). https://crlao.ehess.fr/docannexe/file/1780/ltba2012_libre.pdf
- Liu, Kitty Wenying. 2024. *Verb paradigms in Tibetic: Morphophonology in diachronic perspective*. University of Cambridge. (MPhil dissertation.) https://doi.org/10.17605/osf.io/z9sr2
- Matthews, Peter H. 1991. *Morphology: An introduction to the theory of word-structure*. 2nd edn. Cambridge: Cambridge University Press.
- Ruan, Junyu. 2024. *Predictability changes in morphologization: A case study of German Umlaut*. The Ohio State University. (PhD dissertation.) https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send?accession=osu1713366453818817&disposition=inline
- Tournadre, Nicolas & Suzuki, Hiroyuki. 2023. *The Tibetic Languages: An introduction to the family of languages derived from Old Tibetan*. Villejeuf: LACITO-Publications. https://lacito.cnrs.fr/the-tibetic-languages/
- Wiese, Bernd. 2004. Unterspezifizierte Stammparadigmen: zur Systematik des Verbablauts im Gegenwartsdeutschen. *30. GGS-Tagung*. Mannheim: IDS Mannheim. https://nbn-resolving.org/urn:nbn:de:hebis:30-1127248
- Wiese, Bernd. 2008. Form and function of verbal ablaut in contemporary standard German. In Robin Sackmann (ed.), *Explorations in integrational linguistics: Four essays on German, French, and Guaraní*, 97–151. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.285.03wie
- Zeisler, Bettina. 2004. *Relative tense and aspectual values in Tibetan languages*. Berlin: De Gruyter. https://doi.org/10.1515/9783110908183