Nanosyntax for the historical morphology of Tibetic verbs

Kitty Wenying Liu (University of Cambridge)

DM Meets Nano workshop
Brno, 7-9 July 2026

coming up –

- overview of empirical domain: verbal morphology in Tibetic languages
  - difficult to analyse decompositionally
- Nano is good here ! 👍
  - case study: √kab ‘to cover’
  - synchronic and diachronic domains
- my analysis has one loose end, so we weigh up some fixes and their implications

Tibetic languages

- spoken in the Himalayas + surrounding regions, and in diaspora
- shared typological traits in modern Tibetic languages:
  - morphologically / pragmatically ergative, evidentiality systems, reflexes of highly complex syllable structures
  - verbal morphology involves monosyllabic stems + periphrastic constructions
- historical attestation from 7th c. onwards
  - 7-9th c. = Old Tibetan ≈ Proto-Tibetic
  - 9-20th c. = Classical Tibetan – based on Old Tibetan, heavily standardised, heavily diglossic
- (cf. Romance family)

in this study:

- Classical Tibetan (CTib)
  - literary language from the last millennium
  - verb forms from Hill (2010)'s verb dictionary compendium, semantics from Zeisler (2004)

- Themchen Tibetan (ThCh)
  - very morphophonologically modern conservative variety
  - from the conservative Amdo Tibetan subgroup – northeast of Tibetan cultural region, in China's Qinghai Province
  - data from Haller (2004) – fieldwork from mid 1990s

CTib verbal morphology

	IMP	PST	PRS	FUT
'to be separated'	—	*bskal*	*bskal*	*bskal*
'to chew'	*dmur*	*dmur*	*dmur*	*dmur*
'to blame'	*^{n}phyas*	*^{n}phyas*	*^{n}phya*	*^{n}phya*
'to chew'	*rmes*	*rmes*	*rme*	*rme*
'to chop'	*tshogs*	*btsogs*	*^{n}tshog*	*gtsog*
'to read'	*klogs*	*bklags*	*bklog*	*klog*
'to pierce'	—	*btsugs*	*zug*	*zug*
'to supplicate'	*khrold*	*bkral(d)*	*dgral*	*^{n}grel*

ThCh verbal morphology

	IMP	PFV	IPV
‘to melt (intr)’	—	[ɕə]	[ɕə]
‘to argue’	[n̥kʰon]	[n̥kʰon]	[n̥kʰon]
‘to suck’	[ni]	[ni]	[nə]
‘to chew’	[rdol]	[rdal]	[rdal]
‘to send’	[ʂkər]	[ɸkər]	[ʂkər]
‘to read aloud’	[tʰon]	[ton]	[ndon]
‘to be caught on something’	[ŋgə]	[kwa]	—
‘to lift’	[tɕʰoχ]	[ptɕaχ]	[ɲdʑoχ]

observations from the data

- TM-based paradigm in CTib -> AM-based paradigm in ThCh
  - both CTib and ThCh have an IMP stem
  - CTib has three tense-based stems, ThCh has two aspect-based stems
  - CTib FUT stem was lost (Tournadre & Suzuki, 2023); in OTib and CTib it was the least frequently occurring of the four stems, and its form the least predictable given knowledge of one of the verb's other stems (Liu, 2024)

observations from the data

- CTib: very complex stem alternations, involving some or all of:
  - segmental affixes, vowel ablaut, onset voicing alternation, onset manner alternation
- ThCh: similarly complex stem alternations, mediated by having one fewer stem and some phonological simplification

- both CTib and ThCh have very complex inflectional classes

observations from the data

- CTib: very complex stem alternations, involving some or all of:
  - segmental affixes, vowel ablaut, onset voicing alternation, onset manner alternation
- ThCh: similarly complex stem alternations, mediated by having one fewer stem and some phonological simplification

- ~~both CTib and ThCh have very complex inflectional classes~~
  - difficult to sort through, cf. Liu, 2024

observations from the data

- CTib: very complex stem alternations, involving some or all of:
  - segmental affixes, vowel ablaut, onset voicing alternation, onset manner alternation
- ThCh: similarly complex stem alternations, mediated by having one fewer stem and some phonological simplification
- how to generate these stem alternation patterns?

- ~~both CTib and ThCh have very complex inflectional classes~~
  - difficult to sort through, cf. Liu, 2024

this study: √kab ‘to cover’

		IMP	PST	PRS	FUT
CTib	SR	*k^{h}obs*	*bkab(s)*	*^{n}geb(s)*	*dgab*
	UR	o-kab-s	b-kab-s	n-V-kab	d-V-kab
		↓	↓	↓	↓
		IMP	**PFV**	**IPV**	**—**
ThCh	SR	[k^{h}op]	[kwap]	[ŋgep]	—
	UR	o-kab	b-kab	n-V-kab	—

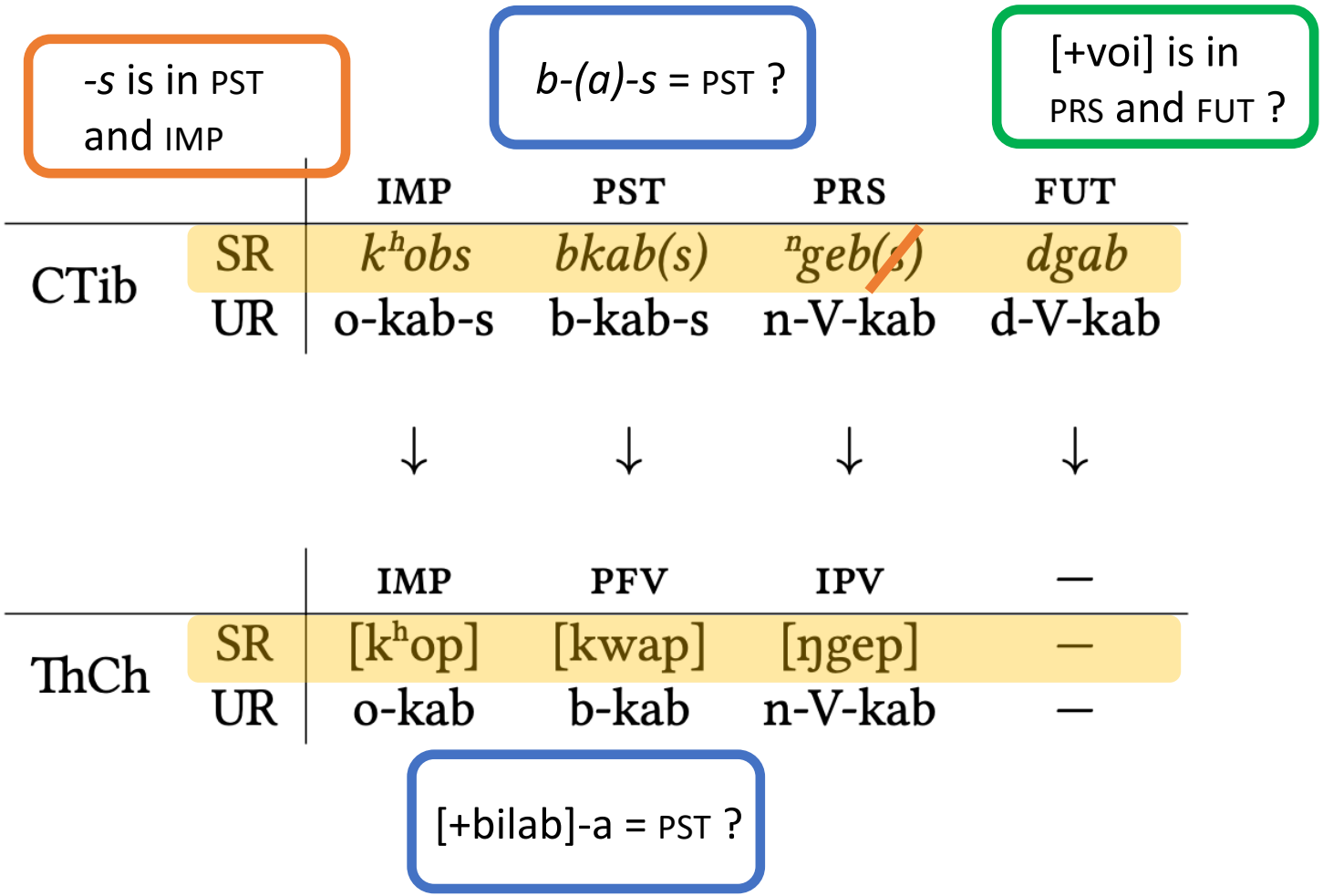
CTib verbs are difficult for decomposition

- working definition:

decomposition = being able to identify sub-word form-function mappings (‘morphemes’)

this study: √kab ‘to cover’

		IMP	PST	PRS	FUT
CTib	SR	*k^{h}obs*	*bkab(s)*	*^{n}geb(s)*	*dgab*
	UR	o-kab-s	b-kab-s	n-V-kab	d-V-kab
		↓	↓	↓	↓
		IMP	**PFV**	**IPV**	**—**
ThCh	SR	[k^{h}op]	[kwap]	[ŋgep]	—
	UR	o-kab	b-kab	n-V-kab	—


-s is in PST and IMP
b-(a)-s = PST ?
[+voi] is in PRS and FUT ?
IMP
PST
PRS
FUT
CTib
SR
kʰobs
bkab(s)
ⁿgeb(s)
dgab
UR
o-kab-s
b-kab-s
n-V-kab
d-V-kab
IMP
PFV
IPV
—
ThCh
SR
[kʰop]
[kwap]
[ŋgep]
—
UR
o-kab
b-kab
n-V-kab
—
[+bilab]-a = PST ?

CTib verbs are difficult for decomposition

- working definition:

  decomposition = being able to identify sub-word form-function mappings (‘morphemes’)

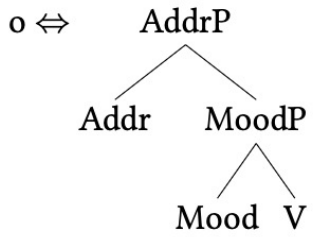
- this is what paradigm-based approaches call ‘gestalt exponence’ (Blevins, 2016; also Matthews 1991:180ff, Anderson 1992:55-56)
  - traditionally a ‘gotcha’ to demonstrate the limits of decomposition

CTib verbs are difficult for decomposition

- difficulty: we can’t satisfyingly map form to function
- solution: use a framework where morphemes contain an intervening layer of representation between form and function

CTib verbs are difficult for decomposition

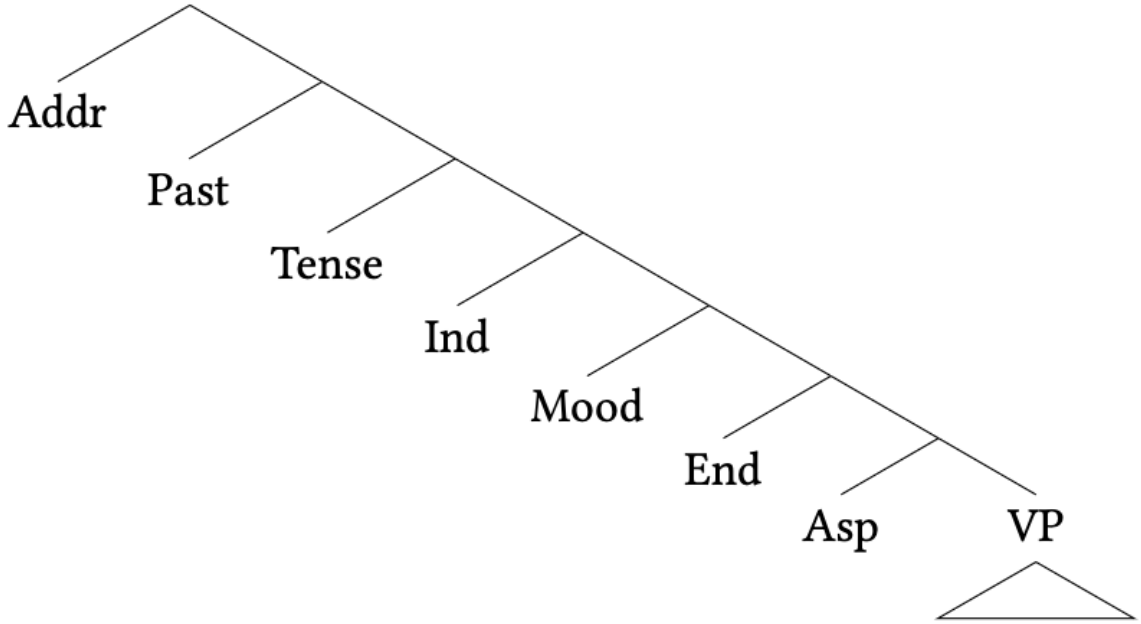
- difficulty: we can't satisfyingly map form to function
- solution: use a framework where morphemes contain an intervening layer of representation between form and function
- Nano is such a theory
  - the trees of LIs act as this intervening layer
  - FSEQ and lexicalisation algorithm can help to motivate decompositions



/o-/ in ThCh

FSEQ

- after Starke & Cortiula (2021), Cortiula (2023), Demonie & De Clercq (2023), Wiland (2025)



lexicalisation algorithm

- Caha's 2023 algorithm (from De Clercq et al eds. 2025), CLBs after Starke (2018) and Wiland (2025)

    1. move nearest labelled non-remnant
    2. pied-pipe (recursive)
    3. open a new workspace for a CLB

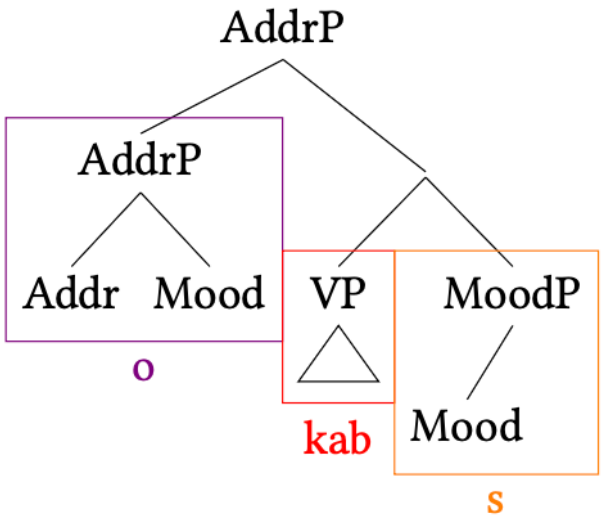
decomposition for CTib (top) and ThCh (bottom)

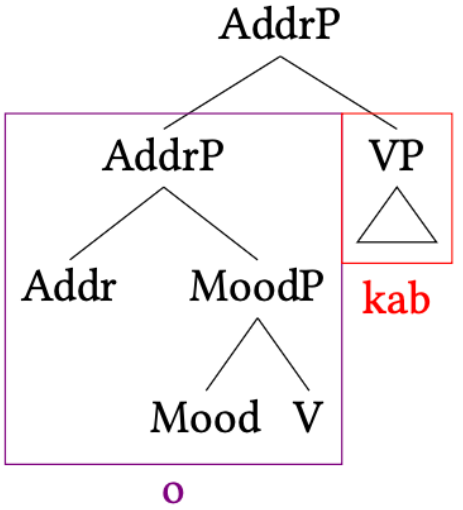
	SR	UR	stem	VP	Asp	End	Mood	Ind	Tense	Past	Addr
a)	*k^{h}obs*	o-kab-s	IMP	kab			-s				o-
b)	*bkab(s)*	b-kab-s	PST	kab			-s	b-			
c)	*^{n}geb(s)*	n-V-kab	PRS	kab			V-	n-			
d)	*dgab*	d-V-kab	FUT	kab			V-		d-		

	SR	UR	stem	VP	Asp	End	Mood	Ind	Tense	Past	Addr
a)	[k^{h}op]	o-kab	IMP	kab			o-				o-
b)	[kwap]	b-kab	PFV	kab			b-				
c)	[ŋgep]	n-V-kab	IPV	kab			V-	n-			

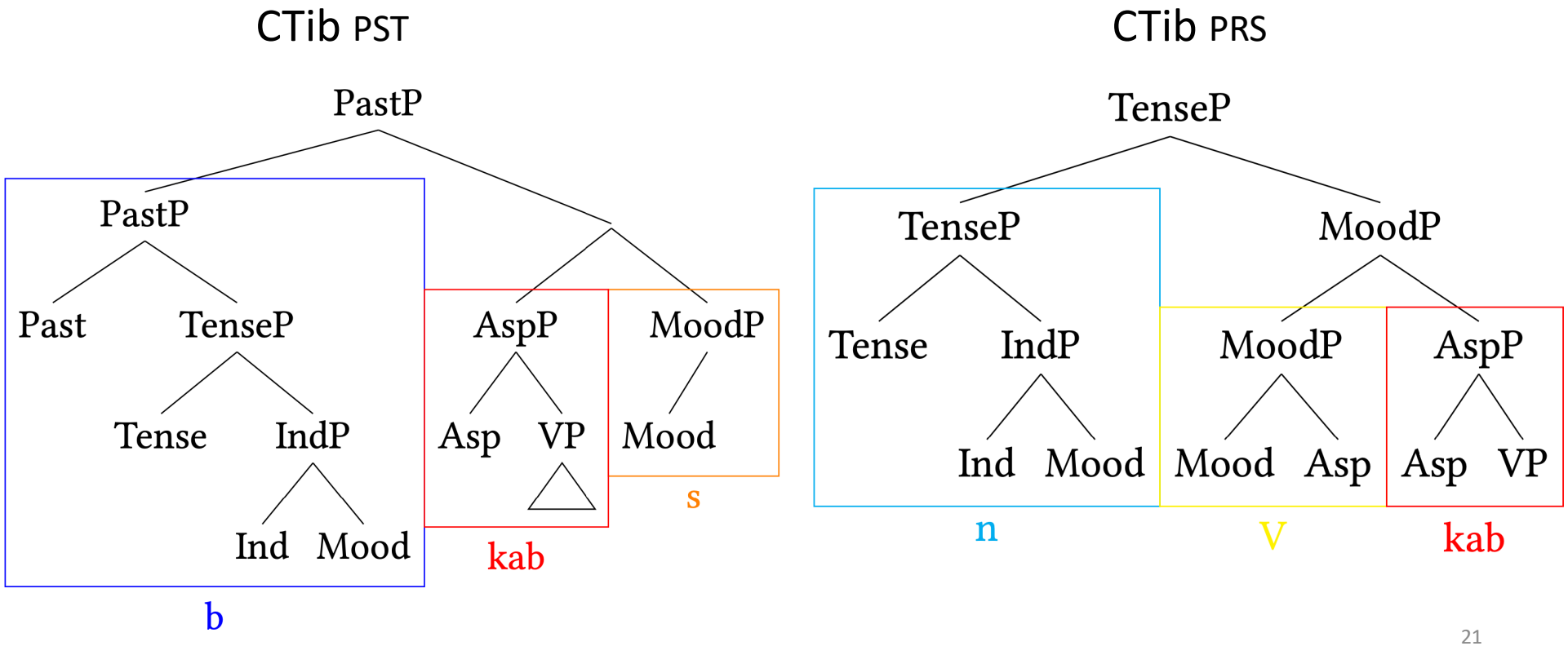
IMP in CTib and ThCh: loss of /-s/

CTib /o-kab-s/ > ThCh /o-kab/



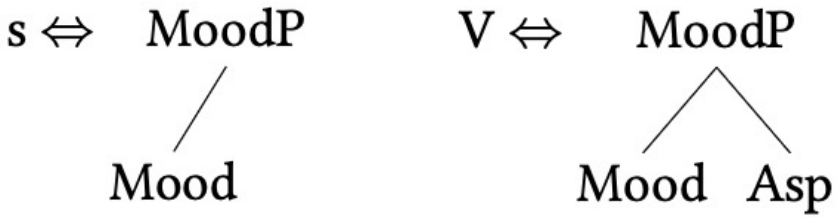


CTib /-s/ and /V-/: suffixal vs prefixal Mood



advantages of this Nano analysis

- we get a decompositional analysis, but preserve the intuition that form and function are not transparently associated in this system
- theory-internal requirements (FSEQ, algorithm) help to supply features for exponents
- lexicalisation requirements can model allomorphy – CTib /s-/ and /V-/ as allomorphs at Mood



diachronic account

	SR	UR	stem	VP	Asp	End	Mood	Ind	Tense	Past	Addr
a)	*kʰobs*	o-kab-s	IMP	kab			-s				o-
b)	*bkab(s)*	b-kab-s	PST	kab			-s	b-			
c)	*ⁿgeb(s)*	n-V-kab	PRS	kab			V-	n-			
d)	*dgab*	d-V-kab	FUT	kab			V-		d-		

	SR	UR	stem	VP	Asp	End	Mood	Ind	Tense	Past	Addr
a)	[kʰop]	o-kab	IMP	kab			o-				o-
b)	[kwap]	b-kab	PFV	kab			b-				
c)	[ŋgep]	n-V-kab	IPV	kab			V-	n-			

diachronic account

- phonological changes erode /-s/ --> /o-/ and /b-/ now need to lexicalise Mood, to replace the CTib function of /-s/
- the use of the CTib PST stem in perfective contexts leads to speakers adding an End feature to FSEQ – now there are two aspect features, so we can get an aspectual contrast through containment
- this new aspectual distinction is more salient than the previous tense distinction, leading to the loss of the tense domain (Tense and Past)
- the CTib FUT stem is lost during this, due to functional pressures as mentioned earlier and/or the phonological erosion of /d-/

an issue with the CTib derivation

- /-s/ and /V-/ are both supposed to realise Mood, and their distribution is not conditioned by a featural natural class
  - i.e. they are morphomic
  - in the current formulation, /-s/ always blocks /V-/

	SR	UR	stem	VP	Asp	Mood	Ind	Tense	Past	Addr
a)	*kʰobs*	o-kab-s	IMP	kab		-s				o-
b)	*bkab(s)*	b-kab-s	PST	kab		-s		b-		
c)	*ⁿgeb(s)*	n-V-kab	PRS	kab		V-		n-		
d)	*dgab*	d-V-kab	FUT	kab		V-		d-		

both /-s/ and /V-/ need to be generated at or near Mood

- /-s/ must lexicalise Mood, bc Mood is the only shared feature between IMP and PST, the two stems that have /-s/
- /V-/ should be generated adjacent to the root, so that in inflectional classes that don't have the voicing alternation, the features lexicalised by /V-/ can be taken over by the root

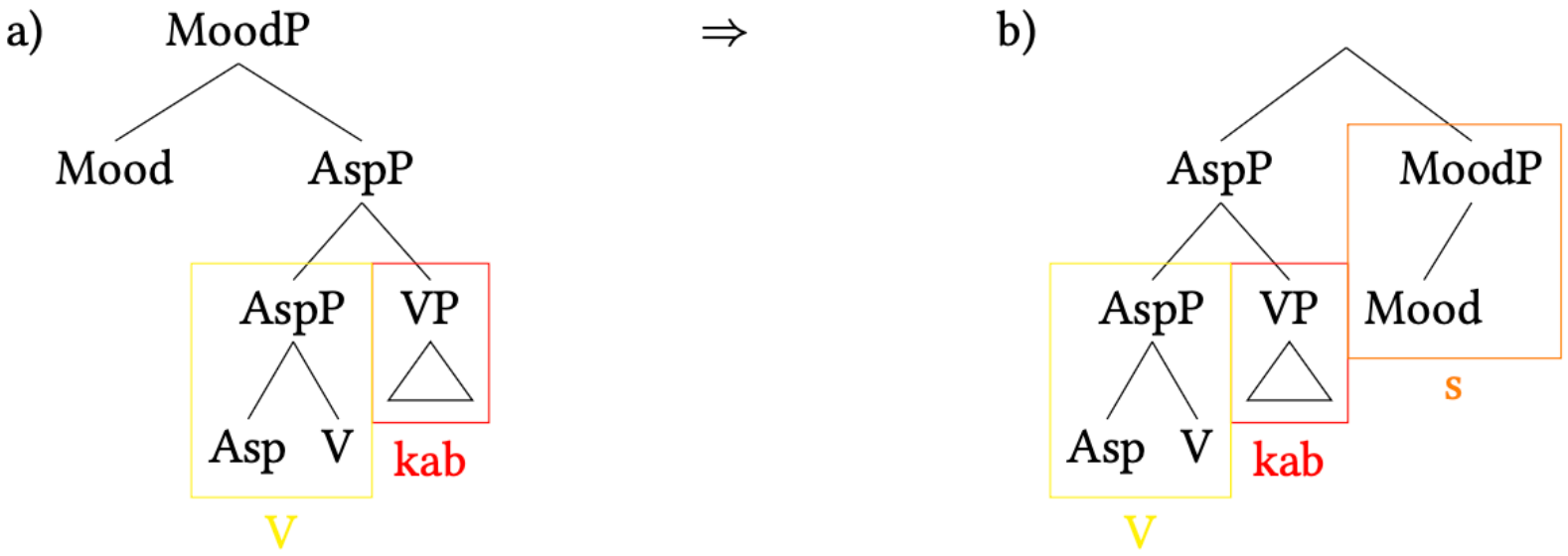
	SR	UR	stem	VP	Asp	Mood	Ind	Tense	Past	Addr
a)	*kʰobs*	o-kab-s	IMP	kab		-s				o-
b)	*bkab(s)*	b-kab-s	PST	kab		-s		b-		
c)	*ⁿgeb(s)*	n-V-kab	PRS	kab		V-		n-		
d)	*dgab*	d-V-kab	FUT	kab		V-		d-		

now: 5 attempts to save my analysis, because I think the gains here outweigh this issue

(in order of increasing successfulness)

1: if /-s/ blocks /V-/, can we generate /V-/ in the previous cycle?

- no: we will always generate /V-kab-s/



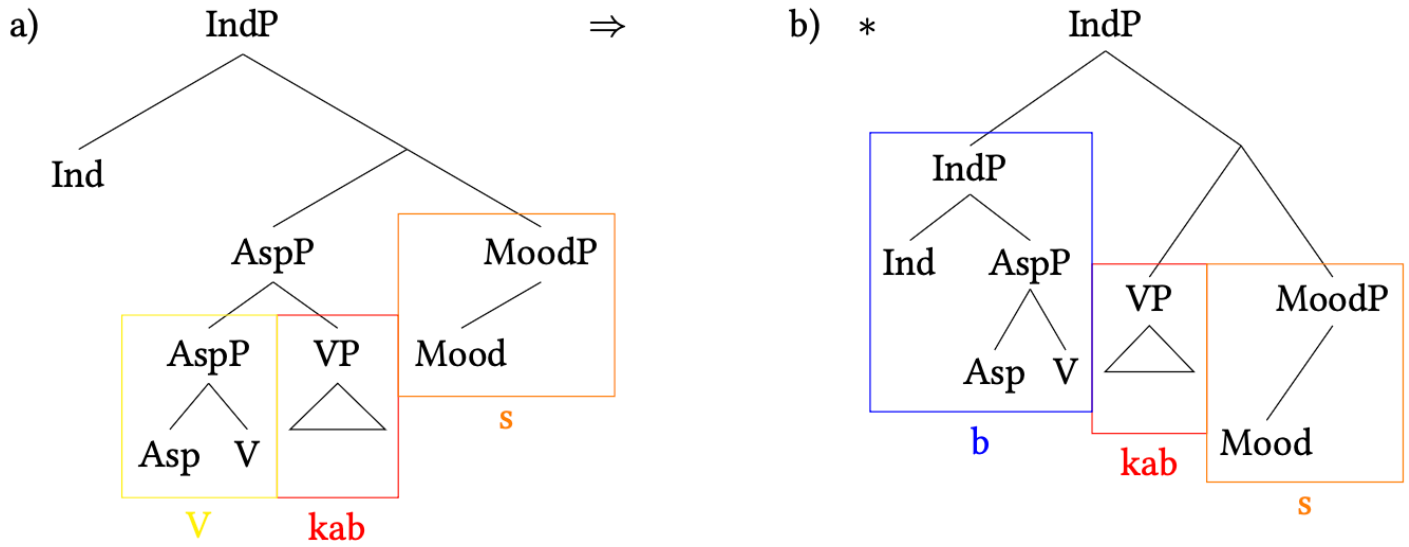
2a: can one of /s-/ or /V-/ overwrite the other one?

- no: in the current algorithm, affix$_j$ can't overwrite (i.e. lexicalise the features of) a pre-existing affix$_i$ if affix$_i$ is on one side of the root and affix$_j$ needs to be on the other side of the root
- this also means that we can't have a derivation where we derive both /-s/ and /V-/, and overwrite /-s/ with /n-/ or /d-/

	SR	UR	stem	VP	Asp	Mood	Ind	Tense	Past	Addr
a)	*kʰobs*	o-kab-s	IMP	kab		-s				o-
b)	*bkab(s)*	b-kab-s	PST	kab		-s	b-			
c)	*ⁿgeb(s)*	n-V-kab	PRS	kab		V-	n-			
d)	*dgab*	d-V-kab	FUT	kab		V-		d-		

2b: so can /b-/ overwrite /V-/?

- no:
  - to derive both /-s/ and /V-/, we derive /V-/ first at Asp and then /-s/ at Mood
  - this means /-s/ traps /V-/, so /V-/'s features are now inaccessible



3: incorporate backtracking to the algorithm, after CLB-formation

- this works for FUT /d-V-kab/:
  - derive /kab-s/, merge Tense, backtrack to Mood and derive /V-kab/

- doesn't work for PRS /n-V-kab/:
  - PRS has exactly the same features as PST until Past, meaning that it will derive identically as PST, with nothing to trigger backtracking

	SR	UR	stem	VP	Asp	Mood	Ind	Tense	Past	Addr
a)	*kʰobs*	o-kab-s	IMP	kab		-s				o-
b)	*bkab(s)*	b-kab-s	PST	kab		-s		b-		
c)	*ⁿgeb(s)*	n-V-kab	PRS	kab		V-		n-		
d)	*dgab*	d-V-kab	FUT	kab		V-		d-		

4: postulate null feature

- if we could add one feature that conditions the split /-s/ vs /V-/ stems, all of these issues would melt away
- I don't have a good candidate for a semantically meaningful feature, so we can add a null one solely to condition this allomorphy
  - this could be powerful if it can also derive other morphomic behaviours in the wider system

	SR	UR	stem	VP	Asp	?	Mood	Ind	Tense	Past	Addr
a)	*kʰobs*	o-kab-s	IMP	kab			-s				o-
b)	*bkab(s)*	b-kab-s	PST	kab			-s		b-		
c)	*ⁿgeb(s)*	n-V-kab	PRS	kab		V-			n-		
d)	*dgab*	d-V-kab	FUT	kab		V-				d-	

5: suppletion

- e.g. use a pointer that turns /kab-s/ into /gab/ in the presence of /n-/ or /d-/
- I'm hesitant to use suppletion for a process that is phonologically regular, but maybe it's admissible here because there are only 50 verbs with voicing alternation in CTib (Hill, 2010)

	SR	UR	stem	**VP**	**Asp**	**Mood**	**Ind**	**Tense**	**Past**	**Addr**
a)	*kʰobs*	o-kab-s	IMP	kab		-s				o-
b)	*bkab(s)*	b-kab-s	PST	kab		-s	b-			
c)	*ⁿgeb(s)*	n-V-kab	PRS	kab		V-	n-			
d)	*dgab*	d-V-kab	FUT	kab		V-		d-		

summary

- Nano has lots of potential for providing decompositional accounts of phenomena that don’t seem transparently decompositional

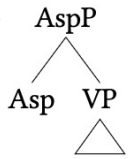
- how should we model morphomic phenomena in NS?

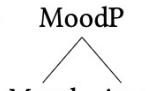
thank you / děkuji / ཐུགས་རྗེ་ཆེ།

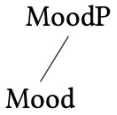
wl399@cam.ac.uk

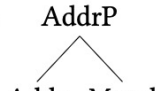
Appendix: LIs

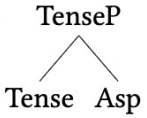
LIs for CTib derivation

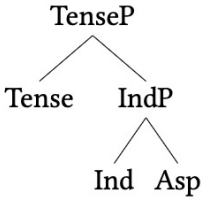
(1) √ ⇔ 

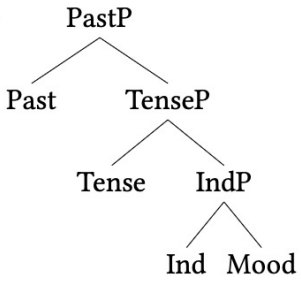
(2) V ⇔ 

(3) s ⇔ 

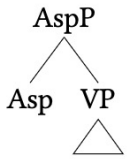
(4) o ⇔ 

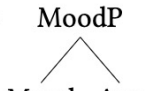
(5) d ⇔ 

(6) n ⇔ 

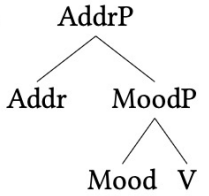
(7) b ⇔ 

LIs for ThCh derivation

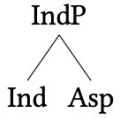
(8) √ ⇔ 

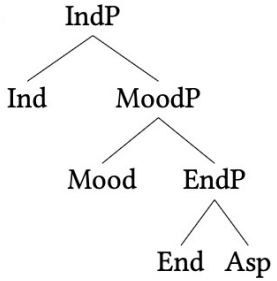
(9) V ⇔ 

/-s/ IS LOST

(10) o ⇔ 

/d-/ IS LOST

(11) n ⇔ 

(12) b ⇔ 

Anderson, Stephen R. (1992). A-Morphous Morphology. Cambridge University Press.

Blevins, James P. (2016). Word and Paradigm Morphology. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199593545.001.0001

Cortiula, Maria. (2023). The Nanosyntax of Friulian verbs: An analysis of the present and past in Tualis Friulian. PhD thesis at Masaryk University, Czech Republic. https://ling.auf.net/lingbuzz/007617

De Clercq, Karen. (2020). The morphosyntax of negative markers: A Nanosyntactic account. De Gruyter.

De Clercq, Karen; Caha, Pavel; Starke, Michal, & Vanden Wyngaerd, Guido. (2025). Nanosyntax: State of the art and recent developments. In P. Caha; K. De Clercq, & G. V. Wyngaerd (Eds.), Nanosyntax and the lexicalisation algorithm (pp. 1–44). Oxford University Press. https://doi.org/10.1093/9780198947158.003.0001

Demonie, Anne-Li. (2025). Complications of Complex Left Branches: The case of Dutch verbal prefixes. In P. Caha; K. De Clercq, & G. Vanden Wyngaerd (Eds.), Nanosyntax and the lexicalisation algorithm (pp. 212–247). Oxford University Press. https://doi.org/10.1093/9780198947158.003.0007

Demonie, Anne-Li & De Clercq, Karen. (2023). The Addressee in the present imperative in Latin. Rivista di Grammatica Generativa. https://ling.auf.net/lingbuzz/007533

Haller, Felix. (2004). Dialekt und Erzählungen von Themchen [Dialect and folktales from Themchen]. VGH Wissenschaftsverlag GmbH.

Hill, Nathan W. (2010). A lexicon of Tibetan verb stems as reported by the grammatical tradition. Bayerische Akademie der Wissenschaften. https://github.com/tibetan-nlp/lexicon-of-tibetan-verb-stems

Liu, Kitty Wenying. (2024). Verb paradigms In Tibetic: Morphophonology In diachronic perspective. MPhil thesis at the University of Cambridge, UK. https://doi.org/10.17605/osf.io/z9sr2

Matthews, Peter H. (1991). Morphology: An introduction to the theory of word-structure. Cambridge University Press.

Starke, Michal. (2018). Complex left branches, spellout, and prefixes. In L. Baunaz; L. Haegeman; K. De

Clercq, & E. Lander (Eds.), Exploring nanosyntax (pp. 239–249). Oxford University Press. https://doi.org/0.1093/oso/9780190876746.003.0009

Starke, Michal & Cortiula, Maria. (2021). Bearing the Brazilian cross. ReVEL, 19(18). http://www.revel.inf.br/files/0ff31931c754c02b7c337272c39ede87.pdf

Tournadre, Nicolas & Suzuki, Hiroyuki. (2023). The Tibetic languages: An introduction to the family of

languages derived from Old Tibetan. LACITO Publications. https://lacito.cnrs.fr/the-tibetic-languages/

Wiland, Bartosz. (2025). Polish prefix stacking redux. In P. Caha; K. De Clercq, & G. V. Wyngaerd (Eds.), Nanosyntax and the lexicalisation algorithm (pp. 165–211). Oxford University Press. https://doi.org/10.1093/9780198947158.003.0006

Zeisler, Bettina. (2004). Relative tense and aspectual values in Tibetan languages. De Gruyter. https://doi.org/10.1515/9783110908183