# Xup, a language of Xupux

**M.M.N.H.**

*A descriptive grammar*

2024

*Dedicated to gan Minhó; you were a good one*

```
Class: artlang
Version: 0.5
Date: 21 October, 2024
```

# Contents

# 0 | Introduction

In this book I shall explore and describe the Xup language of the Xup people.

## 0.1 | Conventions

In this book, I shall use blue text for Xup words, whether they be in orthographic transcription or non-bracketed phonemic transcription (common).

Forward slashes with blue text (/example/) are used for phonemic transcription, square brackets ([example]) are used for phonetic transcription, and blue-text angle brackets (⟨example⟩) are used for orthographic transcription.

Underlined text (which may sometimes be enclosed by 'single quotes') is used for translations, sans-serif text is used for important terms, *italicized* text is used for normal emphasis, and SMALL CAPS is used for glossed terms. "Scare quotes" are used for non-standard, ironic, or otherwise deviant usages of terms; and ‹chevrons› are used for certain notations.

Glosses are structured as follows:

(0.1) transcription
native script
morphemic transcription (object language)
morphemic transcription (metalanguage)
'translation'

Ungrammatical, infelicitous, or otherwise "bad" glosses are preceded by an asterisk ‹*›.

When used as examples to demonstrate a particular grammatical feature, the morphemic metalanguage transcription will usually only contain the relevant information.

## 0.2 | External history

The Xup language is a speedlang (a conlang created within a time restraint) created by me, mareck (M.M.N.H.). It was created within the timeframe of Saturday, October 4th, 2024, to Monday, October 21st, 2024. The challenge was proposed by me.

The following creative restraints have been made:

- have ‹ƛ›
  - bonus: have ‹ƛ̓›
- have a phonological domain edge effect
- have word order

- have phasal polarity
- have $>4$ xor $<2$ tense/aspect morphemes
  - bonus: have $>4$ xor $<2$ mood/evidentiality morphemes

With the following tasks:

- do a lexicon showcase
- translate and gloss five (5) acceptably-sourced sentences
- document and showcase the language
- submit
- bonus: do something spooky
- bonus: say 'hi' to Miacomet!
- bonus: submit a sentence to the 5MOYD Sentence Submission Form

The letter ‹ƛ› is used for the coronal lateral plosive /ƛ/, and ‹λ› for the parasitic lateral morpheme /λ/ (Ch. 1 and § 2.4). The domain edge effect is the deletion of lexical high tones at the right edge of a phrase (§ 1.3). Word order is determined by focus (§ 2.2). Phasal polarity is explained in § 2.6. There is one tense/aspect morpheme (§ 2.4), and there are six evidentiality/mood morphemes (§ 2.5).

This document in and of itself documents and showcases the language, satisfying the relevant task; acceptably-sourced example sentences are found throughout, and the lexicon showcase is found in App. C.

During this speedlang, I kind of got distracted by my other current conlang, Khiw, which is chugging along nicely. In fact, this whole speedlang was a bit of an excuse to play around with potential Khiw things (specifically, getting comfortable with phasal polarity stuff). Also, getting in 5MOYD translations *in situ* rather than making up rather boring example sentences.

Regardless, I'm afraid this document is rather sparse due to my aforementioned issues. I have a lot of incoherent notes in a `.txt` file. It was fun playing around with a looser phonology (I am usually very picky about clusters). I don't imagine I'll revisit this conlang, but anything could happen.

# 1 | Phonology

In this chapter, I explore the sounds and related phenomena of Xup. This includes abstract (phonemic[1]) and concrete (phonetic) forms, as well as suprasegmental units. Orthography is detailed in the next chapter. I shall use (a modified) $^{off}$*IPA* for phonemic transcription, and $^{can}$*IPA*[2] for phonetic transcription.

## 1.1 | Consonants

There are seventeen consonant phonemes in Xup:

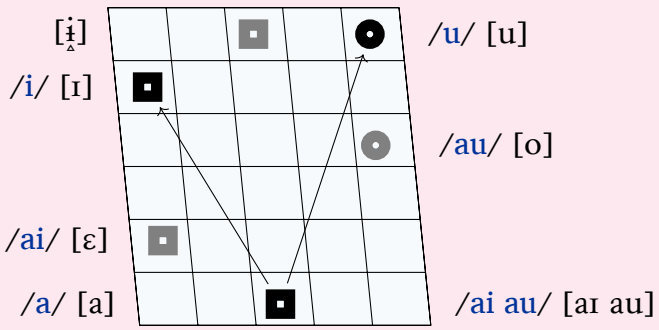
| | *labial* | | *coronal* | | | | | | *dorsal* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *lateral* | | *central* | | *sibilant* | | *velar* | | *glottal* | |
| *plosive* | p | [p] | ƛ | [tɬ] | t | [ʈ] | c | [tʂ] | k | [k̟] | ʔ | [ʔ̟] |
| *voiceless continuant* | f | [f] | ɬ | [ɬ̠] | ɾ | [rˤ] | s | [ʂ] | x | [x] | h | [ɦ̟] |
| *voiced continuant* | v | [v] | l | [l̠] | r | [r] | z | [ʐ] | ɣ | [ɣ̟] | | |

- /p/ is bilabial; /f v/ are labiodental
- /ƛ ɬ l/ are laminodental; /t ɾ r c s z/ are apicoälveolar
- /k x ɣ/ are velar; /ʔ h/ are glottal

Consonants experience no significant taxophony.

## 1.2 | Vowels

There are three phonemic vowels and two diphthongs in Xup:




Monophthings experience little significant taxophony.

[1] Wherein a phoneme is a strictly *contrastive unit* that is abstracted to succinctly represent various but related phonetic surface forms.

[2] See Natural Phonetics on canipa.net.

### 1.2.1 Vowel taxophony

Diphthongs coalesce.

- /ai au/ surface as [ɛ ɔ] before a coda consonant.
- else, /ai au/ surface as [aɪ au]
- an epenthetic voiceless vowel [i̥] is inserted between plosive-plosive clusters.

## 1.3 | Tone

There is one phonemic tone in Xup, the high tone, as well as an unmarked zero tone. The high tone /◌́/ (H) surfaces as high [¯], while the unmarked zero /◌/ (∅) tone surfaces as mid [-]. The tone-bearing unit is the syllable.

A lexical high tone cannot occur at the right edge of a phrase: it is deleted. Independent verb phrases take a high tone at the left edge.

## 1.4 | Phonotactics

Phonotactics describe the ways phonemes are organized in relation to each other, and how they are structured within domains. The profile of the phonological word is as follows[3]:

$$\#\left[_{\omega}\left[_{\sigma}\ \mathrm{T}^{?}\mathrm{C}_{1}^{?}\mathrm{VV}^{?}\mathrm{C}_{2}^{?}\mathrm{C}_{3}^{?}\right]\sigma^{*}\mathrm{C}_{4}^{?}\right]\#$$

- # a word boundary; [ ] a domain
- $\omega$ a phonological word; $\sigma$ a syllable
- ◌$^{?}$ zero or one; ◌$^{*}$ zero or more
- T tone (§ 1.3)
- $\mathrm{C}_{1-4}$ consonants; V a vowel; VV a diphthong

Singleton coda consonants may be any consonant. Consonant clusters are as follows:

[3] I shall use a modified (i.e., in conjunction with regex-like conventions) version of *Recursive Baerian Phonotactics Notation* (RBPN), a non-standard but infinitely more useful notation; see Blumire & Baer (2017).

| → | p | ƛ | t | c | k | ʔ |
|---|---|---|---|---|---|---|
| p | | pƛ | pt | pc | pk | pʔ |
| ƛ | ƛp | | ƛt | ƛc | ƛk | ƛʔ |
| t | tp | tƛ | | tc | tk | tʔ |
| c | cp | cƛ | ct | | ck | cʔ |
| k | kp | kƛ | kt | kc | | kʔ |
| ʔ | ʔp | ʔƛ | ʔt | ʔc | ʔk | |
| f | | fƛ | ft | fc | fk | fʔ |
| ł | łp | | łt | łc | łk | łʔ |
| ɾ | ɾp | ɾƛ | | ɾc | ɾk | ɾʔ |
| s | sp | sƛ | st | | sk | sʔ |
| x | xp | xƛ | xt | xc | | xʔ |
| h | hp | hƛ | ht | hc | hk | |

| → | v | l | r | z | ɣ |
|---|---|---|---|---|---|
| v | | l | r | z | ɣ |
| l | v | | r | z | ɣ |
| r | v | l | | z | ɣ |
| z | v | l | r | | ɣ |
| ɣ | v | l | r | z | |

Voiced clusters can only occur word-medially; voiceless clusters may occur both word-medially and -finally. Triple clusters are of the shape plosive/continuant-plosive-plosive, composed of legal continuant-plosive and/or plosive-plosive sequences.

Cluster resolutions:

| → | p | ƛ | t | c | k | ʔ | f | ł | ɾ | s | x | h | v | l | r | z | ɣ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | p | | | | | | f | pƛ | pt | pc | pk | pʔ | v | łp | ɾp | sp | xp |
| ƛ | | ƛ | | | | | ƛp | ł | ƛt | ƛc | ƛk | ƛʔ | fƛ | l | ɾƛ | sƛ | xƛ |
| t | | | t | | | | tp | tƛ | ɾ | tc | tk | tʔ | ft | łt | r | st | xt |
| c | | | | c | | | cp | cƛ | ct | s | ck | cʔ | fc | łc | ɾc | z | xc |
| k | | | | | k | | kp | kƛ | kt | kc | x | kʔ | fk | łk | ɾk | sk | x |
| ʔ | | | | | | ʔ | ʔp | ʔƛ | ʔt | ʔc | ʔk | h | fʔ | łʔ | ɾʔ | sʔ | xʔ |
| f | p | | | | | | f | fƛ | ft | fc | fk | fʔ | v | łp | ɾp | sp | xp |
| ł | | ƛ | | | | | łp | ł | łt | łc | łk | łʔ | fƛ | l | ɾƛ | sƛ | xƛ |
| ɾ | | | t | | | | ɾp | ɾƛ | ɾ | ɾc | ɾk | ɾʔ | ft | łt | r | st | xt |
| s | | | | c | | | sp | sƛ | st | s | sk | sʔ | fc | łc | ɾc | z | xc |
| x | | | | | k | | xp | xƛ | xt | xc | x | xʔ | fk | łk | ɾk | sk | x |
| h | | | | | | ʔ | hp | hƛ | ht | hc | hk | h | fʔ | łʔ | ɾʔ | sʔ | xʔ |
| v | p | fƛ | ft | fc | fk | fʔ | f | fƛ | ft | fc | fk | fʔ | v | | | | |
| l | łp | ƛ | łt | łc | łk | łʔ | łp | ł | łt | łc | łk | łʔ | | l | | | |
| r | ɾp | ɾƛ | t | ɾc | ɾk | ɾʔ | ɾp | ɾ ƛ | ɾ | ɾc | ɾk | ɾʔ | | | r | | |
| z | sp | sƛ | st | c | sk | sʔ | sp | sƛ | st | s | sk | sʔ | | | | z | |
| ɣ | xp | xƛ | xt | xc | k | xʔ | xp | xƛ | xt | xc | x | xʔ | | | | | ɣ |

## 1.5 | Lenition

Lenition is a process that occurs. It's seen mainly in the inverse suffix (§ 2.2).

| | | | | | | |
|---|---|---|---|---|---|---|
| ↓ | p | ƛ | t | c | k | ʔ |
| L | f | ɬ | ɾ | s | x | h |
| ↓ | v | l | r | z | ɣ | ∅ |

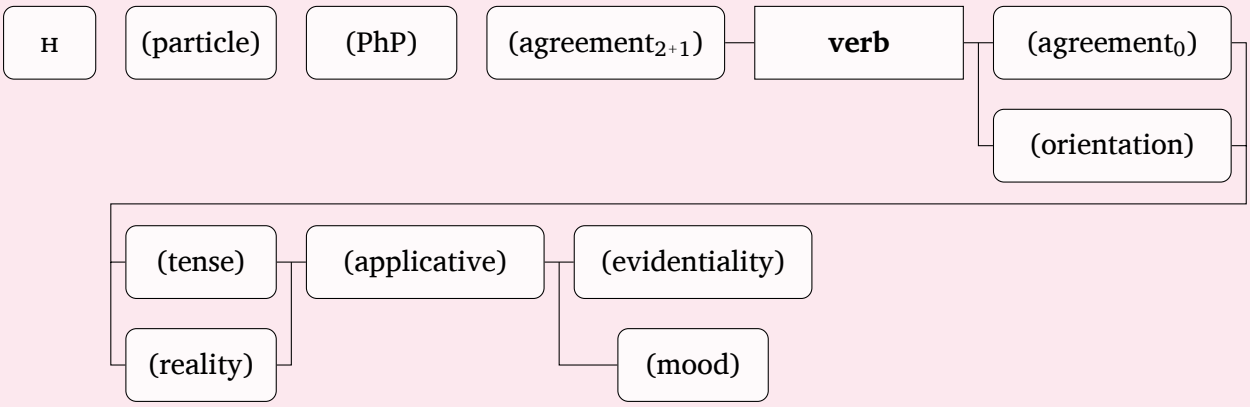
It is also seen in compounds, where the final consonant of the first root undergoes lenition.

(1.1) xup 'person' + ƛáixc 'father' = xufƛaixc 'dog-man'

Note how this is not simply a result of cluster resolution, as /p-ƛ/ would return /pƛ/; also note that the second root loses its tone.

# 2 | Verbs

Verbs are content words that describe eventualities. They inflect for many things.




The independent verb complex obligatorily takes a high tone ◌́ (H) at the left edge, which attaches to the first syllable at the left edge (and is obliterated if there is already a high tone). This shall be glossed as PRE (predicate), when/if it is relevant/I remember to.

## 2.1 | Transitivity

All verbs are by default intransitive and unaccusative, and may be made unergative or transitive via agreement slots and/or applicatives.

## 2.2 | Agreement & orientation

Agreement occurs in three places on the verb: immediately before the verb, either an agreement$_1$ (person/number) or a fused agreement$_2$+agreement$_1$ (number + person/number) prefix occurs; immediately after the verb, an agreement$_0$ (person/number) suffix occurs, which is exclusive with orientation.

| *agreement*$_2$ \ *agreement*$_1$ | 1SG | 1PL | 2SG | 2PL | 3SG | 3PL |
|---|---|---|---|---|---|---|
| ∅ | ɣ(a)- | x(u)- | p(u)- | uv(u)- | ∅ | t(i)- |
| SG | ɣas- | xus- | pus- | uf- | ∅ | ur(u)- |
| PL | ɣat- | xit- | pit- | ut- | it(u)- | ir(i)- |

| *agreement*$_0$ | SG | PL |
|---|---|---|
| 1 | -x, -k | -xp, -ap |
| 2 | -f, -p | -p, up |
| 3 | ∅ | -s, -c |

Unaccusative intransitive verbs take only agreement$_0$, and unergative intransitive verbs take only agreement$_1$.

Transitive verbs take fused agreement$_{2+1}$ and an orientation suffix.

Orientation describes the ranking of arguments by animacy.

| | |
|---|---|
| DIR | ∅ |
| INV | -zí, -ˡí |
| LCL | -ʔ, -u |

Within a clause, constituents are ordered by focus, with more-focal constituents coming first, and less-focal constituents (e.g., topics) coming last. Orientation describes how the arguments are ranked, as ordered by focus, by animacy. The animacy hierarchy is as follows:

| 2 | > | 1 | > | 3 | > | 3′ |
|---|---|---|---|---|---|---|
| ← *hi* ← | | | | → *lo* → | | |

With the direct (DIR), the focus (‹[+F]›) is less animate than the non-focus (‹[−F]›); with the inverse (INV), the focus is more animate than the non-focus.

Pronouns that are [−F] may be dropped.

(2.1) ƛáixc ɣásípku (kai)

ƛáixc ɣás- íp -ku kái
dog SG›1SG- see.DIR -EGO 1SG

‘I$_{[-F]}$ saw the dog$_{[+F]}$’
‘the dog$_{[+F]}$ saw me$_{[-F]}$’

(2.2) ƛáixc kái ɣásípku

ƛáixc kái ɣás- íp -ku
dog 1SG SG›1SG- see.DIR -EGO

‘I saw$_{[-F]}$ the dog$_{[+F]}$’
‘the dog$_{[+F]}$ saw$_{[-F]}$ me’

(2.3) kái ɣásíʔíku ƛaixc

kái ɣás- íp -í -ku ƛáixc
1SG SG‹1SG- see -INV -EGO dog

‘I$_{[+F]}$ saw the dog$_{[-F]}$’
‘the dog$_{[-F]}$ saw me$_{[+F]}$’

(2.4) kái ƛáixc ɣásíʔíku

kái ƛáixc ɣás- íp -í -ku
1SG dog SG‹1SG- see -INV -EGO

‘I$_{[+F]}$ saw$_{[-F]}$ the dog’
‘the dog saw$_{[-F]}$ me$_{[+F]}$’

The local (LCL) orientation is used with speech act participant (SAP) on speech act participant agreement, and with reflexives. With agreement$_{2+1}$ with a SAP in agreement$_1$ position, the local is used to indicate that agreement$_2$ tracks the other SAP (i.e., with 1 or 2 in agreement$_1$, 2 or 1 in agreement$_2$, respectively).

(2.5) tíc ɣásíp?ku kai

tíc ɣás- íp -ʔ -ku kái
2SG SG›1SG- see -LCL -EGO 1SG

‘I saw you’

(2.6) kái púsíp?í tic

kái pús- íp -ʔ -hí tíc
1SG SG›1SG- see -LCL -EXP 2SG

‘you saw me’

With a lone agreement$_1$, the local is used to form reflexives.

(2.7) ɣíp?ku kai

ɣ- íp -ʔ -ku kai
1SG- see -LCL -EGO 1SG

‘I saw myself’

(2.8) típ?kái cíca?

t- íp -ʔ -xái cícá?
3PL- see -LCL -REP 3PL

‘they saw themselves’

## 2.3 | Applicatives

Applicatives. Not much to say, they just need to be here.

| | |
|---|---|
| CAU | -xic |
| BEN | -sau |
| COM | -xi? |
| LOC | -tí |
| LAT | -sús |
| ABL | -các |

## 2.4 | Tense & reality

Tense situates eventualities in time, and reality situates eventualities in worlds.

| | | |
|---|---|---|
| *tense* | DIS | -ƛ, -ƛ̓ |
| *reality* | IRR | -(a)ri |
| | NEG | -ulu, -úlú, -ł |

The discontinuous /-ƛ/ is a parasitic lateral morpheme, which scans the verb root right to left, lateralizing coronal consonants (/t c, ʀ s, r z/ → /ƛ, ł, l/). Laterals block this scanning, and if there is no such consonant (or if it is blocked), it surfaces as a suffix /-ƛ̓/.

The discontinuous tense (and, the only tense marker) is used for past eventualities that (or whose result states) no longer hold true. It cannot take mood.

(2.9) ƛuɣíh sú?áfkái @laina

ƛuɣíh sú?áv -xái @laina
window open -REP Lena

'Lena opened the window...' (the window is still open)

(2.10) ƛuɣíh łú?áfkái @laina, kasixi?kái łíłí cíkt (5MOYD #873)

ƛuɣíh sú?áv -ƛ -xái @laina kasi -xi? -xái łíłí cíkt
window open -DIS -REP Lena close -COM -REP quick 3SG

'Lena opened the window, but closed it immediately.' (the window is no longer open)

The irrealis reality is used for modal things, futures and the like. It cannot take evidentiality.

(2.11) tíłúpt itciƛ̓úri cíkt (5MOYD #2059)

tí- łúpt it- siƛ̓ú -ri cíkt
2SG- tongue PL›3SG- remove -IRR 3SG

'he will cut out your tongues'

The negated reality is used for nonexistent eventualities (negation).

(2.12) cakuru?aru ɣasípúlúku kai (5MOYD #1326)

cakuru?aru ɣas- íp -úlú -ku kái
jakuruaru lizard SG›1SG- see -NEG -EGO 1SG

'I did not see a jakuruaru lizard'

## 2.5 | Evidentiality & mood

Evidentiality and mood.

| *evidentiality* | |
|---|---|
| EGO | -ku |
| EXP | -hí |
| FAC | -káu |
| REP | -xái |

| *mood* | |
|---|---|
| IMP | -tai |
| INT | -ƛul |

The egophoric evidentiality (EGO) is used when the speaker is highly-involved in the event, while the experiential (EXP) is used for any direct sensory evidence (regardless of and usually denoting a lesser degree of involvement). Compare:

(2.13) lú ika ɣákítúlúhí kai (5MOYD #1243)

lú ika ɣá- kít -úlú -hí kái
NEG already 1SG- study -NEG -EXP 1SG

'I didn't really study yet' (it wasn't my fault)

(2.14) lú ika ɣákítúlúku kai (5MOYD #1243)

lú ika ɣá- kít -úlú -ku kái
NEG already 1SG- study -NEG -EGO 1SG

'I didn't really study yet' (intentionally)

The factual (FAC) is used for general knowledge and things the speaker is reasonably certain of.

(2.15) kítúskáu xup

kitú -s -káu xup
eat -3PL -FAC person

'people eat' (in general)

(2.16) lú cili ɣatpúrácáckáu kułáu píƛ kułupíƛ savait (5MOYD #2000)

lu cili ɣat- fúrá -các -káu kułau píƛ kułupiƛ savait
NEG still PL›1SG- hang -ABL -FAC 2000 smoyd

'i plan on stop smoyds after 2000'

The reported (REP) is used for reported information, hearsay, and so on.

(2.17) úrusáiríxái @tupai cífúf pafckái taiki (5MOYD #2068)

uru- sáiɾ -í -xái @tupai cí- fúf pav -c -xái taikí
SG‹3PL- dry -INV -REP Dupe 3- parent be flat -3PL -REP clothes

'Dupe's mother sun-dried the clothes'

The imperative mood (IMP) is used for commands and such. It varies depending on person.

(2.18) ɣákitúritai

ɣa- kitú -ri -tai
1SG- eat -IRR -IMP

'I shall eat!'

(2.19) xúkitúritai

xu- kitú -ri -tai
1PL- eat -IRR -IMP

'let's eat!'

(2.20) púkitúritai

pu- kitú -ri -tai
2SG- eat -IRR -IMP

'(you$_{SG}$) eat!'

(2.21) úvukitúritai

uvu- kitú -ri -tai
2PL- eat -IRR -IMP

'(you$_{PL}$) eat!'

(2.22) kítúritai

kitú -ri -tai
eat -IRR -IMP

'may they$_{SG}$ eat!'

(2.23) tíkitúritai

ti- kitú -ri -tai
3PL- eat -IRR -IMP

'may they$_{PL}$ eat!'

The interrogative mood (INT) is used for questions.

(2.24) fi aƛáf púsƛípítcúsƛul (5MOYD #1983)

fi aƛáf pus- ƛípít -sús -ƛul
three day SG›2SG- stand -LAT -INT

'will you stay for three days?'

Mood interacts with irrealis vs. unmarked in some way.

## 2.6 | Phasal polarity

Phasal polarity markers.

| | |
|---|---|
| still | cili |
| already | ika |

Phasal polarity stuff.

(2.25) cíli caicaituluhi (5MOYD #1655)

cili caicait -ulu -hí
still snow -NEG -EXP

'there is still not snow'

(2.26) kákcku íka xitkípáizíku tútac (5MOYD #1215)

kákcku íka xit- xípáiz -í -ku tútac
1PL already PL‹1PL- take -INV -EGO seed

'we already took the seeds'

The marker ika already may also be used to mean now.

(2.27) kítúskáu xup, íka xukitúritai (5MOYD #1181)

kitú -s -káu xup ika xu- kitú - ri- tai
eat -3PL -FAC person already 1PL- eat -IRR- IMP

'people must eat, so let's eat now'

They are negated with the particle lu to form not yet and no longer:

(2.28) lú ika ɣakít?í kai (5MOYD #1243)

lú ika ɣa- kít -hí kái
NEG already 1SG- study -EXP 1SG

'I didn't really study yet'

(2.29) lú cili paicáxái @vihti (5MOYD #1454)

lú cili paicá -xái @vihti
NEG still be stupid -REP Windi

'Windi is no longer stupid'

# 3 | Nouns

Nouns are content words that describe entities. They take little in the way of morphology besides a prefix when possessed. There are pronouns, which may be dropped.

*possessive*

| | SG | PL |
|---|---|---|
| 1 | ki | ka- |
| 2 | tí- | tu- |
| 3 | cí- | |

*pronouns*

| | SG | PL |
|---|---|---|
| 1 | kái | kákcku |
| 2 | tíc | tuzái |
| 3 | cíkt | cícá? |

# | Appendices

In which Apps. A and B are lexicons of verbs and nouns, and App. C is a lexicon showcase.

Compounds, derivations, idioms, etc., are considered distinct lemmas. Definitions are separated by a double dagger ‹‡›.

Lemma entries are structured as follows:

root (CATEGORIES) : definition(s)

# A | Verbs

kitú : be eaten

íp : be seen

siƛú : be removed

sú?áv : be open

kasi : be closed

kít : be read ǂ be studied, learned

paicá : be stupid ǂ be angry

łíłí : be fast, quick, speedy

xípáiz : be held

ƛípít : be standing, stand

súvít : be sitting, sit

casái : be lying down, lie down

fúrá : be hanging, hang

xává : be leaning, lean

pav : be flat

sáiɾ : be dry

# B | Nouns

xup : person

ƛáixc : dog

ƛuɣíh : window ǂ portal ǂ glass

łúpt : tongue

caicait : snow

túmac : seed

xufƛaixc : dog-man

fúf : parent

taikí : clothing

aƛáf : sun; day

# C | Lexical highlights

This lexical highlights section serves to explore and describe certain words and word groups.

## C.1 | Posture verbs

Posture verbs are verbs that denote position and shape. There are five: ƛípít stand, súvít sit, casái lie, fúrá hang, and xává lean.

The verb ƛípít stand is used to refer to things taller than they are wide; súvít sit for things about as tall as they are wide; and casái lie for things wider than they are tall—standard posture verb stuff. Those are pretty straightforward.

The verb fúrá hang is used for things suspended from or in something (bats, posters, curtains, and the like; but also fish, things ensconced in glass, etc.); and xává lean is used for things dependent on at least two supports. I don't know whether things wedged in a material are fúrá or xává. Probably the latter, but an argument could be made for the former. I want to say that doors xává in their frames, and windows, too.

Like any other verb these can causativize meaning 'cause to be in X position'. Reddit posts are apparently fúrá'd.

You can imagine some nice diagrams demonstrating these various positions here.