

lang₂

sketch grammar of a constructed language



Seth Thompson
a.k.a. kilenc

July 9, 2021

Contents

Contents	ii
Introduction	v
1 Origins	v
2 Goals	v
3 Lore	v
 I PHONOLOGY	 1
1 Segments	2
1.1 Consonants	2
1.2 Vowels	3
1.3 Allophony	3
2 Supersegments	4
2.1 Stress	4
3 Morphophonology	5
3.1 Consonants	5
3.2 Vowels	6
4 Phonotactics	7
4.1 Roots	7
4.2 Frequencies	7
 II MORPHOSYNTAX	 8
5 Nouns	9
5.1 Stems	9
5.1.1 Vocalic stems	9
5.1.2 Neuter stems	10
5.2 Deixis	10
5.2.1 Generic	10
5.2.2 Proximal	10
5.2.3 Distal	11
5.3 Number	12
5.3.1 Additive plurals	12
5.3.2 Associative plurals	13
5.4 Gender	13
5.5 Apposition	14
5.5.1 Inalienable possession	14
5.5.2 Reducing syntactic complexity	14

5.6	Phrasal syntax	15
6	Pronouns	16
7	Verbs	17
7.1	Stems	17
7.1.1	<i>t</i> -stems	17
7.1.2	<i>r</i> -stems	17
7.1.3	<i>m</i> -stems	18
7.2	Agreement	18
7.2.1	Polypersonal	18
7.2.2	Deictic	19
7.3	Negation	19
7.4	Converb	19
7.5	Transitivity	20
7.6	Aspect and mood	20
7.6.1	<i>ot</i>	21
7.6.2	<i>sesam</i>	21
7.6.3	<i>nenat</i>	21
7.6.4	<i>het</i>	22
7.6.5	<i>pit</i>	23
8	Adpositions	24
8.0.1	<i>im</i>	24
8.0.2	<i>ez</i>	24
8.0.3	<i>tɛ</i>	24
8.0.4	<i>ah</i>	24
8.0.5	<i>osc</i>	24
8.0.6	<i>u</i> and <i>su</i>	25
9	Adjectives	26
10	Particles	27
10.1	Imperative	27
11	Clauses	28
11.1	Fronting	28
11.2	Extraposition	28
III	APPENDIX	29
A	Culture	30
A.1	Metaphors	30

B	Registers	31
B.1	Poetry	31
C	Lexicon	32
A.	33
Ȧ.	33
C	33
E.	33
Ė.	33
H	33
I	33
K.	33
L	34
M	34
N	34
O	34
Q	34
P.	34
R.	34
S.	34
T.	34
U	35
V.	35
Y.	35
Z.	35

Introduction

1 Origins

lang₂ is an *a priori* artlang originally conceived to fulfill a speedlang challenge, then modified to handle a relay, then finally molded into its own project. Although I had no clear motivations in mind when beginning it, I found that the confluence of random decisions—the Gleb-generated inventory, challenge stipulations, the early translation choices—made the project into something I really enjoyed.

1 Origins v

2 Goals v

3 Lore v

2 Goals

The overarching goal is to create something that is cool to me. The project is meant to be naturalistic, but when naturalism conflicts with aesthetic, aesthetic will be prioritized.

A primary way to grow this language and develop new ideas will be through translation of poetry and scientific journals. I hope that these will push the limits of my syntactical rules while also developing an interesting corpus. I'll also try to use 5MOYDS and hopefully at some point a journal to expand my ability to speak the language and get an intuitive sense for what constructions it prefers.

As I develop this conlang, my goals will be to explore analytic constructions and syntactic minutiae, write in-depth documentation of how information structure manifests in the language, and produce a robust dictionary and corpus.

3 Lore

lang₂ is set in a worldbuilding project I create as a hobby. The language is an isolate spoken on the eastern coast of a peninsula, once the language of city-states and now a common language throughout a number of coastal countries. It is heavily influenced by East Cape, the *de jure* language of the peninsula, and other languages of trade. The world the speakers know is analogous to our 1930s.

Part I

PHONOLOGY

Segments

1

1.1 Consonants

lang₂ has 14 consonant phonemes, although some newer analyses also analyze voiced stops as phonemic as wells. There are four distinguished places of articulation: labial, plain alveolar, sibilant alveolar, and dorsal. There are four distinguisheds manner of articulation: stops, fricatives, approximants, and nasals.

	Labial	Alveolar	Sibilant	Dorsal
Stop	p <i>p</i> (b) <i>b</i>	t <i>t</i> (d) <i>d</i>	ʈs <i>c</i> (ɖʒ) <i>j</i>	k <i>k</i> (g) <i>g</i>
Fricative	f <i>f</i>	ʈ <i>l</i>	s <i>s</i>	x <i>h</i>
Nasal	m <i>m</i>	n <i>n</i>	ɳ̥ <i>z</i>	
Approximant	w <i>v</i>	ɹ		j <i>y</i>

Nasal Sibilant The nasal sibilant /ɳ̥/ is a typologically odd segment. The prototypical representation of this sound is [ʒ̃], a voiced fricative with simultaneous nasal release; however, phonetic research suggests that true fricatives cannot be fully nasalized. Thus, this segment’s typical realization may be better represented as a slightly fricated approximant [ʳ̥̃], although some speakers may realize it as a weakly nasalized [ʳ̃] or a fully nasalized approximant [ʳ̃̃]. Nevertheless it patterns as a nasal sibilant in both distribution and allophony and is thus phonemically represented as /ɳ̥/.

Labial Fricative The fricative /f/ has a marginal distribution, mostly appearing in loan words and onomatopoeia. Most instances of historical /f/ underwent debuccalization and were subsequently lost. Although /x/ also appears frequently in onomatopoeia, its distribution is more widespread and it is not considered marginal.

Sibilant Affricate Although the phone [ʈs] is rather common in the corpus, its underlying form is not always the phoneme /ʈs/. The underlying phoneme is usually elucidated by inflection. For example, both *tsékla* “steering wheel” and *cekla* “great uncle” share the same surface form, [ʈsɛkʌ]. However, when inflected for the plural, the former becomes [ʈsɛkʌʔaʒəɹ] and the latter [ʈsɛkʌʔaʒəɹ]. As such

1.1 Consonants . . .	2
1.2 Vowels	3
1.3 Allophony	3

The *sibilant* place of articulation is a sprachbund feature, present in other Peninsular languages including East Cape.

Table 1.1: Consonants

See [Ohala et al. \(1998\)](#) for further discussion of nasal fricatives.

Affrication of stop-fricative clusters and pre-[s] schwa deletion are common processes that yield [ʈs], discussed further in §3.1.

tsekla is analyzed as /tasekɫa/, whereas *cekla* is analyzed as /tsekɫa/. This analysis parallels liason in some languages.

1.2 Vowels

lang₂ has 8 vowel phonemes, 5 plain and 3 nasalized.

	Front		Mid				Back					
High	i	<i>i</i>					u	<i>u</i>				
Mid	e	<i>e</i>	ẽ	<i>ɛ</i>	(ə)	<i>a</i>	(ǣ)	<i>q</i>	o	<i>o</i>	õ	<i>q</i>
Low					a	<i>a</i>	ã	<i>q</i>				

Table 1.2: Vowels

TODO:

This section about vowel neutralization probably belongs in morphophonology? It also has to do with the supersegment stress.

Vowel Neutralization Mid and low vowels /e a o/ and their nasal counterparts are reduced to [ə ə̃] in unstressed syllables. Schwa is not phonemic, but neutralization is common, so it appears frequently throughout the corpus. In speech, the underlying vowel becomes evident when stress is shifted due to morphological processes.

The surface form is typically romanized *a* even for underlying /e o/. Dictionaries typically denote the *shadow vowels* as *nbm.*, an abbreviation of *nabam* “shadow.”

1.3 Allophony

Supersegments

2

2.1 Stress

2.1 Stress 4

Stress in **lang**₂ is lexically and morphologically productive. Stress typically falls on the penultimate syllable of a word. Atypical vowel stress, most common in loan words, is marked with an acute. Stress is also morphologically productive, distinguishing between unmarked and distal deixis in nouns. Affixation, compounding, and other stress-shifting processes often cause stress-based minimal pairs to become homophones.

For example, *tqka* “tree” has typical stress and is unmarked, but *tqká* “reindeer” has atypical ultimate stress and is thus marked.

Stressed vowels have three phonetic differences from unstressed vowels. First, they typically have a rising pitch. Second, they are typically longer than unstressed vowels. Third, onsets before long vowels have a longer VOT than other onsets, a manifestation of slight aspiration or breathiness.

Secondary stress falls on alternating syllables starting from primary stress and spreading left. For example, *kagəsa* “army” has regular stress on the penultimate syllable, but when inflected in plural form, it surfaces as *kegqəsaʒar* [ke.gə'sa.ʒəɹ]. Secondary stress prevents the reduction of /e a o/ to [ə], but, unlike primary stress, does not cause length or VOT increase.

TODO:

This all just got recently moved from allophony because it's better analyzed as a morphophonological phenomenon, probably. So some of it will have to be re-written with that framing in mind. Probably means a lot more double slashes and "it only occurs across morpheme boundaries" type stuff.

3.1 Consonants . . . 5

3.2 Vowels 6

3.1 Consonants

Voicing Nasal and stop clusters are realized as voiced stops, occasionally prenasalized. The resulting phones $[(^{(m)}b\ ^{(n)}d\ ^{(n)}\widehat{d}z\ ^{(n)}g)]$ are romanized *b d j g*. Clusters where the nasal is the onset of a syllable, not a coda, do not assimilate, so /pn/ is still realized [pn].

These clusters assimilate in place to the stop, so /np/ surfaces as [b], not as [d].

Although the assimilation process is most common across morpheme boundaries or in loan words, voiced plosives can be found in some native morphemes. However, since these phones only occur word-medially in limited and predictable distribution, they are not traditionally considered phonemic. Most native words with voiced plosives are transparent compounds, such as *ebar* "below" (← *ez* + *par*) or *kageša* "army" (← reduplication of *kəša*).

Some scholars argue that voiced plosives are phonemic or becoming phonemic because of their presence in loan words and verb conjugations.

Some words, especially some conjugations of *m*-stem verbs, are spelled with a word-final voiced stop, but the stop is still typically realized as a medial cluster. For example, *sed* "they tell me" is underlying /semt/ and realized $[se^{(n)}d(ə)]$. For some speakers, especially younger speakers or those in informal contexts, the final schwa is elided.

TODO:

Affrication and assibilation would make more sense if it was only across morpheme boundaries, but I have some words that I like where it's morpheme-internal. How should I handle that? It could be reworked as a diachronic process, made phonemic, a morphophonemic process—but it needs more thought.

Affrication Alveolar plosive and fricative clusters are realized as a sibilant affricate. Clusters of /ts/, /tʃ/, /tʃ/ and /tx/ all neutralize to [tʃ], romanized as *c*. The /tʃn/ cluster likewise affricates, but is realized as [tʃʃ], romanized as *cz*.

The voiceless nasal affricate is notoriously hard for non-native speakers to pronounce and is often used as a shibboleth.

Assibilation Sibilant and non-sibilant fricative clusters are realized as sibilants. Unvoiced sibilants /t͡s s/ clustering with /f x ʃ/ are realized as [s:], romanized as *cc* or *ss* depending on the underlying phoneme. The voiced sibilant /n͡z/ instead is realized as [ʒ:] in such clusters, romanized as *zz*.

3.2 Vowels

Schwa Deletion The reduced vowel [ə] is often deleted between consonants, especially non-alveolar stops, and the sibilants /t͡s s n͡z/. For example, *ksarat* /kosɑɹat/ is commonly realized as [ksɑɹət], and spelled accordingly. The elision process results in word-final or word-initial [s] or [ʒ] being the only syllable-internal clusters. However, these clusters are not consistently realized, and occasionally have an epenthetic schwa, especially word-finally.

Nasalized schwa [ɐ̃] rarely undergoes deletion as it is typically longer than [ə].

Note that /t͡s/ surfaces [s] in these environments.

TODO:

Something here about word-final clusters. I haven't been consistent about when it's spelled as a cluster or when it's spelled with the vowel, and I need to either commit to the inconsistency, or figure out how that works.

TODO:

There might end up being an isogloss map of schwa deletion: most dialects delete ahead of sibilants, some ahead of sibilants and approximants, and others in all environments?

4.1 Roots

4.1 Roots 7

4.2 Frequencies . . . 7

Nouns Most nominal roots are of the form CV(C)CV. Nominal roots infrequently end in a consonant, typically /s/ or /r/ for neuter nouns. Rarely, nominal roots end in /m n/ or /k/.

Verbs Most verbal roots are of the form (C)VC or CV(C)CVC. /t/ is the most common root-final phoneme, although there are some /m/ and /r/ stems as well.

Affixes A smaller set of phonemic segments are allowed in affixes. Neither the labials /p f m w/ nor the high vowel /u/ appear in true affixes. In the fossilized partial reduplication process, the reflexes of labial are either /k/ (← /p f/) or /n/ (← /m w/). The high vowel /u/ lowers to /o/, often simply realized as [ə].

Because /u/ and /w/ pattern similarly, some phonemic analyses conflate them.

4.2 Frequencies

WIP Lexifer file

with: std-ipa-features coronal-metathesis

letters: a ɔ b c d e ɛ f g h i k l m n o ɒ p r s t u v y z

C = t s k n y h r c m z l p v f

N = s r m n k

P = t m r

V = a e i ɔ o ɛ ɒ u

random-rate: 40

words: CVC?CVN? C?VC?CVP C?VP C?VN?

filter: mt > d; nt > d; zt > d; mc > j; nc > j; zc > j; mk > g; nk > g; zk > g; mp > b; np > b; zp >

Part II

MORPHOSYNTAX

Nouns

5

The **lang₂** noun phrase is largely analytic, but nouns do inflect for deixis and number. Nouns have three broad inflection patterns, largely related to the way they inflect for plurality.

5.1 Stems

Nouns are broadly divided into three stems based on their inflectional patterns: *vocalic*, *s-stems*, and *r-stems*. Rarely noun stems will end in other consonants, but these have no discernible shared patterns and often comprise newer loans.

Most noun stems are vocalic stems; typically these end in mid vowels, although some do end in a high vowel. Vocalic stems are usually common gender, excluding loanwords, which are prescriptively assigned neuter gender even if they end in a vowel.

Neuter nouns, on the other hand, are often either *r-stems* or *s-stems*, the latter being more common. Although no longer morphologically productive, the endings on *s-stems* and *r-stems* derive from historical derivation processes. Many roots are reflected in both endings, but the shared meaning between them is not always transparent.

- 5.1 Stems 9
- 5.2 Deixis 10
- 5.3 Number 12
- 5.4 Gender 13
- 5.5 Apposition . . 14
- 5.6 Phrasal syntax 15

S-stems can end in any sibilant, typically *-s* but also *-c* or *-z*.

5.1.1 Vocalic stems

Vocalic stems have fairly regular, agglunative inflection patterns. However, the proximal plural is shortened to *-rran* for most speakers.

	Generic	Proximal	Distal
SG	<i>metka</i>	<i>metkan</i>	<i>matkó</i>
PL	<i>matkozar</i>	<i>matkorran</i>	<i>matkazár</i>

Table 5.1: Inflection of vowel-stem *metka* “bowl”

All vocalic forms share the same endings, but many will have an unpredictable final vowel due to stress-based mid vowel neutralization. There is no way to predict the final vowel from the uninflected lemma, so learners often resort to memorization.

5.1.2 Neuter stems

Neuter stems share a common inflection pattern. For both neuter stems, the proximal surfaces as /on/ instead of /n/. The proximal plural also shortens for neuter stems, but takes the form *-zza-*, influenced by the assibilation morphophonological process. The shared plural form for both stems can create homophones.

	Generic	Proximal	Distal
SG	<i>retus</i>	<i>ratusan</i>	<i>ratús</i>
PL	<i>ratuzzar</i>	<i>ratuzzan</i>	<i>ratuzzár</i>

	Generic	Proximal	Distal
SG	<i>pebar</i>	<i>pabaran</i>	<i>pabár</i>
PL	<i>pabazzar</i>	<i>pabazzan</i>	<i>pabazzár</i>

Some strategies to resolve homophones are covered in §5.3.1.

Table 5.2: Inflection of *s*-stem *retus* “blade”

Table 5.3: Inflection of *r*-stem *pebar* “garden”

5.2 Deixis

Nominal deixis has a variety of uses, including evidentiality, distance, familiarity, and topicality. Verbs and adjectives exhibit agreement for deictic reference. There are three deictic categories: *generic*, *proximal*, and *distal*.

5.2.1 Generic

The unmarked or dictionary form of a noun is used when the noun is widely understood or well-known, for immaterial referents that cannot be deictically located, or if evidence of the referent is not known. If direct or reported evidence exists, it’s infelicitous or questionably grammatical to use unmarked form.

5.2.2 Proximal

The proximal form of a noun is used when the speaker is certain, nearby, or familiar with the noun. This form most commonly denotes direct evidence, meaning the speaker has personal experience with the marked noun. It is marked with the suffix *-n*.

I had this idea, then found out that, as usual, a natural language had it first. Read [Huijsmans, Reisinger, and Matthewson \(2020\)](#) for more about the Salishan languages.

-n is morphophonemically //on//, where //o// doesn’t surface for vocalic stems.

Direct evidence The canonical meaning of the proximal form is direct evidence, often translated as “I saw.”

TODO:

The definiteness constructions probably need to be reworked to square better with (a) the stuff I’ve learned about definiteness and (b) the use of the deictic forms for topic/focus.

Definiteness Proximal forms can be used to describe the definiteness of a referent. This construction is only used for weak, uniqueness-based definiteness (eg. “the Moon”), never for strong, anaphoric definiteness (eg. “the book”). For strong definiteness, the noun *sin* “???” is used alongside distal form, as in (1b).

- (1) a. **Nassoi-n kqstecik su kageša su kagéstapa.**
nassoi-n kqstecik su kageša su kagéstapa
 king-PROX command and army and navy
 “The king (that we know) commands both army and navy.”
- b. **sah ez-Rosaṃ pít ató sín.**
sah tē=rosaṃ pít ató sín
 soon PREP=COOK hold\DIST grain\DIST ???\DIST
 “The rice (that you mentioned) is about to be cooked.”

(5MOYD #1381)

5.2.3 Distal

The distal form of a noun is used when the speaker is uncertain, far, or unfamiliar with the noun. This form typically denotes indirect evidence, including inference, meaning the speaker has heard of or can make an educated guess about the existence of the marked noun. Reported deixis is marked by shifting stress to the ultimate syllable of the word.

Indirect Evidence The prototypical meaning of the distal form is indirect evidence, often translated as “heard about” or “they said.” As in (2), this evidence is encoded into the clause via the subject and the predicate that agrees with it.

- (2) **egi Matkó aczé sém lar tē-het.**
egi matkó aczé sém lar tē-het
 just basket\DIST two\DIST will:be\DIST there PREP=be:at
 “(She said) there will be just two baskets.”

(5MOYD #1314)

5.3 Number

Nouns inflect morphologically for an additive plural, but there is also a syntactic construction used to form associative plurals. The unmarked form of a noun encodes expected number, e.g. *parsa* “eyes” which defaults to a pair and must be modified by a numeral or appositive to specify a singulative.



Figure 5.1: Additive vs. associative plurals

The primary difference in the two plurals is the composition of the set: additive plurals refer to largely homogenous referents, whereas associative plurals refer to largely heterogenous referents.

5.3.1 Additive plurals

Additive plurals are used for a set of homogenous referents and never heterogenous referents; e.g. *matkozar* is “a set of the same (or similar) bowl” and never “a set of diverse bowls.”

Additive plurality is indicated through the suffix *-zar*. Morphological marking is optional and a noun can be inferred additively plural from context. As such marking is less common for small, discrete or easily countable sets or when a referent has been established plural in prior conversation. However, speakers are not always consistent with marking.

Morphonologically, plural marking can be thought to precede deictic marking; plural distal nouns have accent placed on the plural suffix. As seen in Table 5.1, *matkozar* becomes *matkazár*, not **matkózar*.

The second meaning would use the associative plural.

Mandatory plural marking is stylistically preferred in formal contexts.

Combating homophony Because the suffixes of *s*-stems and *r*-stems merge in the plural, some minimal pairs are rendered homophones when inflected. To combat this, speakers sometimes employ *tevi* “many” as an appositive modifier for the singular form.

Explicit number When number is specified with a numeral, the noun is not marked for plurality, as in (3). This can also be used as another strategy to combat homophony.

(3) **vęci ocza**

vęci *ocza*

mercenary two

“two mercenaries”

(cf. *vęcizar* “mercenaries”)

5.3.2 Associative plurals

Unlike the additive plural, the associative plural is not marked morphologically; it is nonetheless rather prevalent. The preposition *ezzu* conveys the associative meaning.

The associative is used for a set of heterogeneous referents. For animate (especially human) referents, the meaning is canonically “a person and their associates,” as in (4). The focal referent (*i.e.* most important) is the marked noun.

(4) **Sayanarnat otik ezzu Kanyi ez-laran.**

sayenar-nat *ot-ik* *ezzu* *Kanyi* *ez-lar-n*

be.ignorant-CVB be-PROX ASSOC NAME PREP=EXPL-PROX

“For this, Kanyi and his friends won’t be much *help*.”

Like *su*, *ezzu* can appear stranded. See §8.0.6.

Terminology in this section adapted from [Daniel and Moravcsik \(2007\)](#).

However, the associative can also have a number of idiomatic, context-specific meanings, usually referring to diverse groups.

5.4 Gender

Although nouns traditionally distinguish *common* and *neuter* gender, this has largely become a prescriptive convention. Loanwords, especially technical loanwords, are typically assigned neuter gender. Some words only distinguish gender for certain uses or contexts, thus dictionaries typically denote if a given usage is expected to require neuter gender.

The gender distinction is more common in literature or academia.

5.5 Apposition

Apposition is fairly common due to the small adjective class present in **lang₂**. Apposition is used in a number of collocative constructions but is distinct from compounding largely because of stress patterns and morphophonological processes. Furthermore, appositive nouns shown agreement with the noun they modify, like adjectives.

Compounds can shift stress and also show cross-morpheme sound changes.

5.5.1 Inalienable possession

Apposition is a fairly common strategy for possession of all types, but inalienably possessed nouns require apposition, as in (5b).

Alienable possession via apposition is dispreferred in formal registers.

- (5) a. ?**hora im-kęsa**
 hora im=kęsa
 wrist of=soldier
 Intended: “soldier’s wrist”
- b. **hora kęsa**
 hora kęsa
 wrist soldier
 “soldier’s wrist”

The sentence in (5a) more readily lends itself to a reading indicating inalienable possession, such as “soldier’s craftsmanship” or “soldier’s weaponsmithing.”

5.5.2 Reducing syntactic complexity

Appositives are also used as a way of combating heaviness in complex noun phrases. This often occurs in noun phrases with multiple prepositional modifiers, especially when one of the modifiers is **im** “of.” Other prepositions can be reduced this way as well.

When multiple prepositional phrases modify a head noun the phrase is syntactically heavy. For example (6) has three modifiers, an adjective and two prepositional phrases, one with its own adjective.

- (6) **rat Hecik pakné taá im-lalasan ez-helazzaran acoan.**
 rat hecik pakné taá im=lalasa-n ez=helazzaran acoan
 I be.at house\DIST blue\DIST of=auntie-PROX PREP=mountains tall
 “I’ll be at auntie’s blue house near the tall mountains.”

To reduce the syntactic load of the sentence, a speaker might instead render it as (7). The phrase is clearly appositional; **lalasa** switches

from proximal marking (indicating the speaker has met her) to distal marking (to agree with its head noun).

(7) **rat Hecik pakné taá lalasá ez-helazzaran acoan.**

rat hecik pakné taá lalasá ez=helazzaran acoan
 I be.at house\DIST blue\DIST auntie\DIST PREP=mountains tall
 “I’ll be at auntie’s blue house near the tall mountains.”

The new sentence has fewer function words and less disparate inflections, reducing some of its complexity.

5.6 Phrasal syntax

Noun phrases are predominantly head-initial. Generally speaking, syntactically simpler constituents occur before more syntactically complex ones, as shown in (8).

- (8) head → adjective → number → appositive → prepositional phrase

Pronouns are morphologically and syntactically similar to other nouns—they share inflectional patterns and can be modified by adjectives. The notable differences are that pronouns rarely inflect for deixis, and some appositive constructions are ungrammatical.

TODO:

Settle on pronominal forms—right now it's *rat* 1, *a(f)* 2, *sec* 3c, and *moc* 3n. There's some isogloss map about whether 2 is *a* or *af*.

Possession Like nominal possession, pronominal possession can be expressed either through apposition or prepositional phrases. Thus both *lar im-sec* and *lar sec* can mean “his thing.” However, apposition is more common in isolation for pronouns than it is for nouns.

Formal registers prefer the prepositional construction.

The **lang₂** verb phrase is more inflectional than the noun phrase, but periphrastic constructions handle most temporal marking, with inflection largely reserved for agreement. Because verbs have fixed valency, prepositional adjuncts, periphrastic constructions, and dummy nouns are common.

7.1 Stems

7.1.1 *t*-stems

T-stems are fairly predictable, although there is some affrication of /t/ in certain forms.

	GEN	PROX	DIST
→ 0	<i>het</i>	<i>hetik</i>	<i>hét</i>
→ 1	<i>hetta</i>	<i>hettik</i>	<i>hattá</i>
→ 2	<i>hetca</i>	<i>hetcik</i>	<i>hatcá</i>
→ 3C	<i>hec</i>	<i>hecik</i>	<i>hacá</i>
→ 3N	<i>hetna</i>	<i>hetnik</i>	<i>hatná</i>
3N →	<i>hecz</i>	<i>heczik</i>	<i>haczá</i>
REFL	<i>hetka</i>	<i>hetkik</i>	<i>hatká</i>

7.1	Stems	17
7.2	Agreement . . .	18
7.3	Negation	19
7.4	Converb	19
7.5	Transitivity . .	20
7.6	Aspect and mood	20

Table 7.1: Conjugation of *t*-stem *het*

7.1.2 *r*-stems

R-stem verbs are generally predictable in most forms. However, they share morphology for both neuter forms agent and patient.

	GEN	PROX	DIST
→ 0	<i>par</i>	<i>parik</i>	<i>pár</i>
→ 1	<i>parta</i>	<i>partik</i>	<i>partá</i>
→ 2	<i>parca</i>	<i>parcik</i>	<i>parcá</i>
→ 3C	<i>pars</i>	<i>parsik</i>	<i>parsá</i>
3N	<i>paz</i>	<i>pazzik</i>	<i>páz</i>
REFL	<i>parka</i>	<i>parkik</i>	<i>parká</i>

Table 7.2: Conjugation of *r*-stem *par*

7.1.3 m-stems

M-stems are notable for widespread voicing caused by /m/ clusters.

	GEN	PROX	DIST
→ 0	<i>sem</i>	<i>semik</i>	<i>sém</i>
→ 1	<i>sed</i>	<i>sedik</i>	<i>sadá</i>
→ 2	<i>sej</i>	<i>sejik</i>	<i>sajā</i>
→ 3C	<i>sems</i>	<i>semsik</i>	<i>samsá</i>
→ 3N	<i>semna</i>	<i>semnik</i>	<i>samná</i>
3N →	<i>semz</i>	<i>semzik</i>	<i>samzá</i>
REFL	<i>seg</i>	<i>segik</i>	<i>sagá</i>

Table 7.3: Conjugation of *m*-stem *sem*

7.2 Agreement

Verbs agree with both the deictic position of the subject noun and the person of the two least oblique arguments of transitive verbs. Only the sole finite verb of a clause bears agreement.

7.2.1 Polypersonal

Transitive verbs exhibit polypersonal agreement via a suffix to the verb root. Intransitive verbs don't require person marking, but it can be used for emphasis or clarification; in such cases, either the reflexive or 3c patient morpheme is used.

		Patient			
		1	2	3C	3N
Agent	1	-	c	s	n
	2	t	-	s	n
	3C	t	s	s	n
	3N	n	z	z	z
	REFL	k			

Table 7.4: Person agreement

For the most part verb agreement is actually patient agreement, but some isolated cases are actually polypersonal. Because so many agreement suffixes share the same form, **lang₂** is only occasionally dropped.

7.2.2 Deictic

Generic noun forms do not have agreement morphemes, but proximal nouns demand the verbal suffix *-ik* and distal nouns shift stress to the final syllable. These suffixes are attached after polypersonal agreement suffixes.

TODO:

Rework agreement with distal nouns: it should be more interesting, perhaps shifting stress in some places but not others. For example 3C.PAT forms might not shift stress because [s] is pretty cluster-happy.

The stress shifting on both nouns and verbs derives from the same, now elided, morpheme.

7.3 Negation

Negation is handled with a verbal suffix *-res* which occurs before agreement affixes. Verbal negation is the primary form of negation in **lang**₂, so the forms are rather common. It has a number of allomorphs shown in Table 7.5.

	GEN	PROX	DIST
→ 0/2/3c	<i>-ras</i>	<i>-réks</i>	<i>-rés</i>
→ 1	<i>-rac</i>		<i>-réc</i>
→ 3N	<i>-rans</i>	<i>-régs</i>	<i>-réns</i>
3N →	<i>-raz</i>	<i>-rékz</i>	<i>-réz</i>
REFL	<i>-raks</i>	<i>-rék</i>	<i>-réks</i>

The affix's position is from its origin as an auxiliary which bore agreement.

Table 7.5: Negative suffixes

The affix is morphophonemically //*ɛs*//.

Because of the merger of the intransitive, second-person and third-person patient forms, **lang**₂ is considered to have asymmetrical negation. As a consequence, explicit person marking via pronouns is more common with negated verbs, especially in the proximal, where all common persons shared the same phonetic form.

7.4 Converb

The converb form of a verb is used for simultaneous action. The converb is commonly used to describe the manner of the main clause, and is also commonly used in periphrastic constructions. The converbial suffix is *-nat*, although some roots select *-os*.

TODO:

Work more on the converb and flesh this out with examples. Might change the morphophonemic form.

7.5 Transitivity

Transitivity is lexically fixed, but transitive verbs can still be functionally intransitive with the use of dummy objects.

Prescriptive convention holds that verbs exhibit polypersonal agreement with their dummy objects, but speakers commonly omit polypersonal marking in these constructions, as in (9b).

(9) a. **sec Sasamsik tasa.**

sec sesam-s-ik tasa

3C say-3C.P-PROX letter

“He’s saying something.”

(Formal)

b. **sec Semik lar.**

sec sem-ik lar

3C say-PROX EXP

“He’s talking.”

(Informal)

Informal constructions often use the more grammaticalized *lar* instead of the dummy collocative.

Not every verb has a collocated intransitive form. For these verbs, periphrastic constructions can also serve as valency-changing operations.

7.6 Aspect and mood

Periphrastic constructions handle most temporal marking in **lang**₂, covering aspects and moods. The core verb of the periphrastic construction is the only finite verb of a clause, bearing all agreement, and the semantic verb is demoted to an adjunct in converbial form or as a bare infinitive with an adpositional clitic.

TODO:

Flesh this section out with more examples and more specific uses.

7.6.1 ot

The verb *ot* “be” has two periphrastic constructions, a *perfective* and a *support* construction used for focus-fronting verbs.

Perfective In the perfective construction, the semantic verb is demoted to adjunct with the preposition *tə*. The perfective can be used in any time frame, although by default it does have a past-time connotation.

The perfective construction generally used for events that occurred over a fixed time frame, especially when a duration is given. In contrast with an unmarked verb, the perfective emphasizes sequences of events or actions done a finite number of times.

Support In the support construction, the semantic verb is demoted to adjunct as a converb. This construction is used for fronting verbs for emphasis, and does not convey any aspectual information.

7.6.2 sesam

The verb *sesam* “say” has one periphrastic construction, an irrealis.

Sesam is often shorted to *sem* informally.

Irrealis In the irrealis construction, the semantic verb is demoted to adjunct with the preposition *tə*. The irrealis can be used in any time frame, although by default it does have future-time connotation.

The irrealis is used for all events that the speaker supposes might or might’ve occurred. The construction is very general, and has broad semantic meaning—including conditional, jussive, and optative senses. In contrast with the unmarked verb, the irrealis emphasizes that the situation is not factual, but is hoped or posited to occur or have occurred.

7.6.3 nenat

The verb *nenat* “crouch” has one periphrastic construction, a subjunctive.

Subjunctive In the subjunctive construction, the semantic verb is demoted to adjunct with the preposition *ah*. The subjunctive construction has a more limited scope than the *sesam* irrealis construction, typically expressing counterfactuals or doubt.

7.6.4 *het*

The verb *het* “be at” has two periphrastic constructions, a *passive* and an *imperfective*, the latter typically restricted to narrative contexts.

Passive In the passive construction, the semantic verb is demoted to adjunct as a converb. The *het* passive is not a true passive because the verb does not change valency (*i.e.* the A-like argument cannot be omitted). Instead of valency operations, the role of this construction is typically to change verbal agreement, as in (10).

- (10) a. ***Azzár hossusarsik nassoïn.**
as-zâr hassusar-s-ik nassoï-n
 foreigner-PL.DIST exalt-3C»3C-PROX king-PROX
Intended: “(I see) the foreigners (I’ve heard about) praising the king.”
- b. **Nassoïn hecïk azzâr hossusarnat.**
nassoï-n het-s-ik as-zâr hassusar-nat
 king-PROX be.at-3C»3C-PROX foreigner-PL.DIST exalt-CVB
 “(I see) the king being praised by the foreigners (I’ve heard about).”

In (10a), the hypothetical speaker intends to mark the verb as proximal to convey direct evidence, but the utterance is ungrammatical because the verb doesn’t agree with its subject, *azzâr*. To correct this, the construction in (10b) is used, which takes advantage of the passive to mark the verb phrase as proximal.

In addition to its role shuffling agreement, the *het* passive can also be used to clarify sentences that become ambiguous due to focus fronting, as in (11).

- (11) a. **Kanyin akakvatcïk Arpatan.**
Kanyï-n akekvat-s-ik Arpat-n
 NAME-PROX tag-3C»3C-PROX NAME-PROX
 “Kanyï tagged Arpat.”
 or “Who Arpat tagged was Kanyï.”
- b. **Kanyin hecïk Arpatan akakvatnat.**
Kanyï-n het-s-ik Arpat-n akekvat-nat
 NAME-PROX be.at-3C»3C-PROX NAME-PROX tag-CVB
 “Kanyï was tagged by Arpat.”

In (11a), it's not clear if Kanyi is the semantic agent or a semantic patient that's been fronted for focus. Without context, both interpretations are grammatical. The use of the passive in (11b) is less ambiguously interpreted, almost always meaning that Arpat was the semantic agent.

7.6.5 *pit*

TODO:

Rework the *pit* passive, which currently doesn't make a lot of sense—how is it demoting stuff, when semantically you'd expect “hold” to be transitive? Maybe *pit* means something else, maybe it will demote in a different way, maybe the valency is weirder...

The verb *pit* “hold” has one periphrastic construction, the mediopassive.

Mediopassive In the mediopassive construction, the semantic verb is demoted to adjunct with the preposition *ez*. Unlike the *het* passive, the *pit* mediopassive is a true passive; the A-like argument does not appear.

Unlike most other periphrastic verbs, *pit* is rarely used outside periphrasis, having largely been replaced by *akrar*.

Adpositions

8

Adpositions are syntactically bound morphemes that express some relationship (often spacial) between constituents. However, they are considered words, not affixes, because the stress pattern of the noun they bind to does not shift. Their phonological independence differentiates them from affixes.

Compare *kq̣sazar* “soldiers,” marked via affix, to *retus im-ḳ̰sa*, “soldier’s blade,” marked via adposition.

Adpositions are a closed class, composed of only 6 members; finer distinctions can be made with *compound adpositions*, such as *ṭ̰-kamc im* “after, to the back of.” Although many such constructions are common enough to be lexically set, they are not nearly as ubiquitous as lone prepositions.

8.0.1 *im*

The adposition *im* indicates possession, often alienable. It can also indicate origin.

8.0.2 *ez*

The adposition *ez* conveys location inside an object or large body. It can also be used for composition of manmade objects.

8.0.3 *ṭ̰*

The adposition *ṭ̰* conveys motion relative to a location, either towards or away from.

8.0.4 *ah*

The adposition *ah* conveys location on the surface of another object. It can also be used for general location.

8.0.5 *osc*

The adposition *osc* conveys location surrounding another object. It is commonly used in a temporal sense to indicate a time frame, often translated as “around the time of.”

8.0.6 **u** and **su**

U and **su** are more limited in semantic scope than other adpositions and are rarely used in compound prepositions to gain further nuance. However, they have similar syntactic distribution and are thus considered members of the preposition class. They typically convey association alongside.

While **u** is used largely to join two noun phrases as arguments of one head, such as *laran u lár* “this and that,” **su** is used for emphasis or in some fixed constructions, such as *su kagəsa su kagəstapa* “both army and navy.”

Stranded su Unlike other prepositions, **su** can appear *stranded*, i.e. without modifying. For example, while (12a) violates the valency of the verb, (12b) does not because the prepositional phrase functions as the argument of the verb.

- (12) a. ***Nassoin kəstecik tɛ-kagəsa.**

nassoin kəstecik tɛ=kagəsa
king commands PREP=army

Intended: “The king commands the army.”

- b. **Nassoin kəstecik su kagəsa su kagəstapa.**

nassoin kəstecik su kagəsa su kagəstapa
king commands and army and navy

“The king commands both army and navy.”

TODO:

This “stranded” analysis is probably really sketchy, since it’s clearly not fulfilling the role of other prepositions, but I don’t feel like opening that can of worms just yet so take it at face value for now.

Adjectives are a small, closed class of noun-like morphemes that cannot be arguments of verbs or prepositions without some other constituent. There are approximately 20 adjectives. Adjectives agree with their head noun for deixis (but not number) and syntactically appear before appositives or prepositional modifiers in the noun phrase.

Particles are a small but open class of discourse markers that can appear at the beginning of a clause. The most common particles are *???*, a polar question marker, *???*, a content question marker, and *qm*, an imperative marker. They are syntactically privileged, able to occur before other constituents of a clause, including even fronted arguments or adjuncts.

10.1 Imperative . . . 27

10.1 Imperative

The imperative particle *qm* ...

For prohibitives, the most common construction is to negate the verb, as in (13).

(13) *qm af ltnatréks!*

qm af itnet-réks

IMP 2 fight-NEG.PROX

“Don’t *you* fight me!”

(5MOYD #1440)

However, constructions with *kers* are also common, especially in the southwest. *Kers* has a more admonitive meaning in standard **lang**₂, often used in warnings or chidings, but wouldn’t be used for commands or requests.

East Cape uses similar constructions for its prohibitives, likely influencing the increased modal force of *kers*

Although the base-generated word order in **lang**₂ is SVO, this order rarely surfaces due to aggressive focus fronting. The most proximal, most newsworthy information is placed in the front of an utterance in first position. As a consequence, **lang**₂ is V2 order, mandating that a finite verb always be in second position. Adjuncts, including demoted verbal constructions, typically come after the core arguments of the verb. Often, however, the order of elements in a clause is determined by focality and evidentiality.

11.1 Fronting 28

11.2 Extraposition . 28

In practice, the most common word order in declarative sentences is SXOV or VXOS.

11.1 Fronting

The most likely phrases to be fronted are proximal or directly evident noun phrases, followed by distal or indirectly evident noun phrases. Generic noun phrases are rarely fronted except in fixed constructions.

Often, the fronted element will be the conversational focus.

lang₂ doesn't have explicit role-marking, so it's not always clear if a fronted noun phrase is subject or object. Various structures, including verb agreement and passives, help provided redundancy when context is not enough.

11.2 Extraposition

When a content-heavy phrase needs to be fronted, a dummy noun is often used to allow right-branching extraposition. The dummy noun is lexically dependent, but the generic nouns *lar* or *manc* can also be used, although they may sound stilted for some verbs.

- (14) **osc Armê kirayamik isyusan ocoan im nassoî kęstat ezzar vęci.**

osc armê kirayamik isyusan ocoan im nassoî kęstat ezzar vęci
 PREP DMY investigate special:tribune PREP king lead ASSOC mercenary

“What the special tribune is investigating is the king’s use of mercenaries.”

Part III

APPENDIX

A.1 Metaphors

A.1 Metaphors . . . 30

Some common conceptual metaphors in **lang₂** are given below; most things are also noted in the dictionary.

WISDOM IS A TREE Trees are an important part of **lang₂** culture, representing the emotional and intellectual lifespan of a person. Old, tall deciduous trees are a symbol of wisdom and maturity; as a plant may *ossat* “grow,” so too may a person *ossat* “become wiser.” Young children attend a *pebar* “garden” for primary education. A student might lament that a former mentor *tqvāt* “becomes senile (*lit.* loses leaves),” or that an elder has “thinning branches.”

POLITICS IS A SHORE The coastline is a common conceptual metaphor for things involving governance. A common idiom for kingdom or state is *taspa u tesa* “sea and shore,” its citizens are *adahēs* “sands,” its influence *almanizar* “waves.” A tribune might *kirqyam* “investigate (*lit.* delve into)” an issue, or a *cunvarq* “columnist (*lit.* seagull)” may discuss it.

BODY PARTS ARE SKILLS Traditional **lang₂** cultural and spiritual beliefs posit that the person is divided into two parts, a perfect body and an imperfect mind. As such many body parts are sacred to certain skills or traits—*hora* “wrist” represents craftsmanship; *ahka* “foot” represents wisdom or learnedness; *kamc* “back” represents labor; *makra* “chest” represents responsibility. Someone with poor skill may be *orra* “feeble” or *???* “uncoordinated,” while someone irresponsible may be *oca* “thin.”

B.1 Poetry

B.1 Poetry 31

The common structure of a classical **lang**₂ poem has nine lines, with alternating pairs of short lines and long lines. Short lines are two trochees between two amphibrachs, and long lines are four amphibrachs. The middle three lines of the verse share a central, interrelated metaphor, and the first and last lines feature repetition. Common poetic features also include alliteration across lines and rhyming of stress syllables. Syncope is often used to fit the meter.

A **lang₂**-to-English dictionary is provided below.

How to use

Entries for lexical items are listed by their spelling in generic form, ignoring morphophonological alterations. Derived words are listed as separate entries, but their source word is given. On the other hand, idiomatic or fixed expressions are given under the lexical item.

Each sense of a word has three basic parts: a quick, single-word translation for ease-of-use; a more detailed explanation of the concept; and an example sentence. The sentences are usually designed to help the reader figure the word's meaning out from context, particularly for **lang₂** speakers and learners.

Pronunciations are given dictionary-style in a phonetic alphabet more intuitive to native **lang₂** speakers.

A	33
Ą	33
C	33
E	33
Ę	33
H	33
I	33
K	33
L	34
M	34
N	34
O	34
Q	34
P	34
R	34
S	34
T	34
U	35
V	35
Y	35
Z	35

A

asyisaikli • *noun* • /ěsyisöikli/ • **1** organization, association, official body

avara • *noun* • /avâřě/ ← OL₂ *agğré*, CC *gamrî* • see *secya* • **1 ntr.** star or constellation used for navigation

Arméds • *noun* ← Farlands *rmédas* • **1 ntr.** the Gospels; a collection of poems sacred to the *Véajan* religion

A

akassar • *noun* • /økássör/ ← OL₂ *økássör* “payment craft” • **1 ntr.** mathematics

C

E

egi • *particle* • **1** just, only: *egi matkó aczé sém lar tē-het* “she says there’ll be just two baskets”

emas • *noun* • **1 ntr.** plot of land

emassoi • *noun* ← *emas* + *-soi* • **1** boss; mid-level employee that oversees other employees **2** coach, manager, trainer; person with authority over a sports teams or its players **3** **emassoi kastezi** coach; sports executive that makes substitutions and decides strategy **4** **emassoi qkas** general manager; sports executive that signs and trades players **5** *archaic* feudal lord

esyi • *adj.* • **1** good **2** correct, appropriate

E

H

hassusar • *verb tr.* • /hössúsar/ ← OL₂ *höś-húsār*, redup. of *húsār* • **1** exalt, praise (smn.) **2 refl.** boast about *tē* smth. **3** be a fan of, root for, cheer for (a team)

hakra • *noun* • **1** battle, skirmish **2 pl.** military campaign **3 pl.** semester

hakraif • *noun* ← *hakra* + *-if* • **1** field hockey; a ball game played on a pitch between two teams of seven, each attempting to win by scoring more points via goals than the opponent

husar • *verb tr.* • **1** compliment **2** *archaic* shout at

I

ista • *noun* • /ístö/ ← Farlands *isto*, *hísto* • **1** (*Véajan*) *ntr.* grace; a quite prayer ceremony for blessing food, travel, and other small events **2** **tahat osc-ista** give grace **3** **fala ista** chamomile tea; a kind of soothing herbal drink imported from the Farlands

K

kageša • *noun* • /kě(n)gěsă/ ← OL₂ *kěkěsă*, redup. of *kěsă* • **1** battalion, unit **2** army; a branch of the military comprising ground troops **3** *sports* team, club

kęstat • *verb tr.* • /kęstět/ • **1** lead (smn.) towards *tē* a goal **2** command: *nassoin kęstací su kageša su kagęstaspa* “the king commands both army and navy.” **3** train (an apprentice) in *tē* a skill: *tę-coryam kęstatrí rat lala* “auntie’s teaching me to sew.” **4** formally teach (a student) in *ez* a discipline: *rappahan picí kęstetnal ez-latya* “the minister was brought up in the faith.”

kəsa • *noun* • 1 soldier 2 literary protagonist, hero

L

M

mazzi • *noun* • 1 ballplayer

N

nahhar • *noun* • /náhhör/ ← OL₂ *náhhör*
“gem craft” • 1 *ntr.* ring, hoop 2 *ntr.* bracelet, bracer 3 *ntr. obs.* jewelry

O

Q

qkas • *noun* • 1 *ntr.* payment

P

palác • *noun* • /pălát(ă)s/ • 1 *ntr.* dinner

paknalác • *noun* • /paknălát(ă)s/ ← por-manteau of *pakna* *u-palác* • 1 inn, motel; short-term lodgings for a traveller between urban areas

R

S

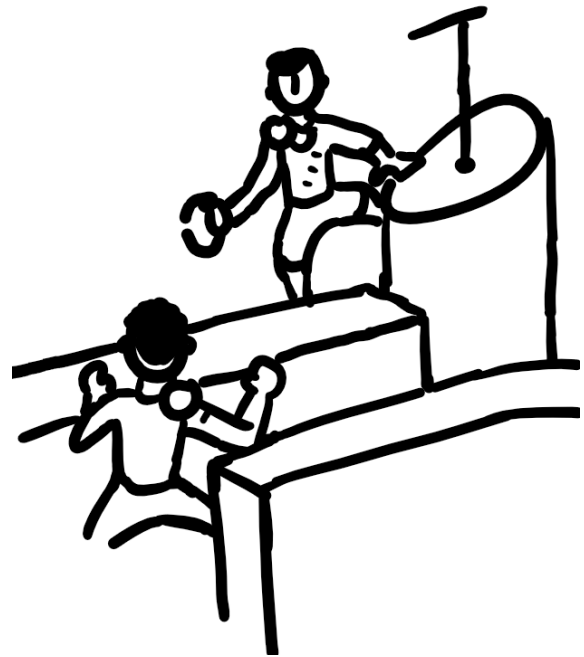
Samvéajan • *noun* ← compounding of *sem véaja* • always PROX • 1 *ntr.* the Gospels; a collection of poems sacred to the *Véajan* religion

sisik • *noun* ← clipping of *tahənar asyisaikli* “associated tag” • see *tahənar* • 1 *Inland* tag

T

tahət • *verb in.* • /táhět/ • 1 stretch (one’s limbs) 2 reach for *tə* something (to grab it) 3 aspire to *tə* a difficult goal: *tə-mazzi im-ot tahqí sec, manacan acoan raksí otnal kers* “he aspired to be a ballplayer, but he wasn’t tall enough” 4 pray about *osc* smth.

tahənar • *noun* • /tăhənhör/ ← OL₂ *tahənáhhör* “hoop reach” • 1 *ntr.* tag; a parkour game played on an obstacle course between two teams over a number of rounds, each attempting to win by grabbing rings before being tagged



taspa u-tesa • 1 *idiom* kingdom

tesa • *noun* • 1 shoreline, bank; where a moving body of water meets the land 2 *of people* hair line

teukkanas • *noun* • /teukkánös/ ← Far-lands *tiēk'kános* • 1 (*Véajan*) *ntr.* missionary, pastor; travelling evangelist who starts local chapters of faith centers 2 *humorous* tourist, vacationer

U

V

véaja • *noun* • /véă(n)jǝ/ ← Farlands *vészā*
 • **1** *ntr.* karma; recompense for virtuous behavior **2** *ntr. pl. prox.* a minority religion characterized by proselytization and belief in an afterlife **3** **sesam véaja** preach, orate: *rappahan sesamzí véaja ez-ǝkas esyi tǝ-laczin* “the minister preached of an eternal reward for the faithful”

Y

yerragín • *noun* • /yerrǝ(n)gín/ ← OL₂ *yer-vǝngín*, Farlands *iǝlunzşín* • **1** (*Véajan*) *ntr.* a prayer ceremony for greeting a visitor as a show of hospitality **2** **tahǝt osc-yerragín** lead the prayer ceremony **3** *idiom* **yiat yerragín** show off, grandstand, put on airs: *egi cuvarǝn yiací yerragín tǝ-ahka im-nassoin* “the king knew the jester was only showing off”

Z