

lang₂

grammar of a constructed language



Seth Thompson
a.k.a. kilenc

September 30, 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	 9
5 Nouns	10
5.1 Deixis	10
5.1.1 Generic	10
5.1.2 Proximal	10
5.1.3 Distal	11
5.2 Number	11
5.2.1 Additive plurals	12
5.2.2 Other plural strategies	12
5.2.3 Associative plurals	13
5.3 Gender	13
5.4 Apposition	13
5.4.1 Inalienable possession	14
5.5 Phrasal syntax	14
6 Pronouns	15

7	Verbs	16
7.1	Agreement	16
7.1.1	Person	16
7.1.2	Deixis	17
7.2	Converb	18
7.3	Transitivity	18
7.4	Negation	20
7.5	Aspect and mood	20
7.5.1	Perfective	20
7.5.2	Irrealis	20
7.5.3	Subjunctive	21
8	Adpositions	22
8.0.1	<i>im</i>	22
8.0.2	<i>ez</i>	22
8.0.3	<i>tə</i>	22
8.0.4	<i>ah</i>	22
8.0.5	<i>osc</i>	22
8.0.6	<i>u</i> and <i>su</i>	23
9	Adjectives	24
10	Particles	25
11	Clauses	26
11.1	Fronting	26
III	APPENDIX	27
A	Culture	28
A.1	Metaphors	28
B	Registers	29
B.1	Poetry	29
C	Lexicon	30
A	31
Ȧ	31
C	31
E	31
Ė	31
H	31
I	31
K	31
L	32

M	32
N	32
O	32
Q	32
P	32
R	32
S	32
T	32
U	33
V	33
Y	33
Z	33

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 18 consonant phonemes. There are four distinguished places of articulation: *labial*, *plain alveolar*, *sibilant alveolar*, and *dorsal*, and four distinguished methods of articulation: *stop*, *fricative*, *nasal*, and *approximant*. Stops can be further divided into tenuis and voiced, although the voiced counterparts are marginal.

	Labial		Alveolar		Sibilant		Dorsal	
Stop	p	<i>p</i>	t	<i>t</i>	ʈʂ	<i>c</i>	k	<i>k</i>
	b	<i>b</i>	d	<i>d</i>	ɖʐ	<i>j</i>	g	<i>g</i>
Fricative	f	<i>f</i>	ɸ	<i>l</i>	s	<i>s</i>	x	<i>h</i>
Nasal	m	<i>m</i>	n	<i>n</i>	ɳ	<i>z</i>		
Approximant	w	<i>v</i>	ɹ	<i>r</i>			j	<i>y</i>

1.1 Consonants . . . 2

1.2 Vowels 3

1.3 Allophony 3

Table 1.1: Consonants

Nasal Sibilant The nasal sibilant /*ɳ*/ is a typologically odd segment. The prototypical representation of this sound is [ʐ̃], 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 [ʐ̥̃], although some speakers may realize it as a weakly nasalized [ʐ̃] or a fully nasalized approximant [ɹ̃]. Nevertheless it patterns as a nasal sibilant in distribution and morphophonemic processes and is thus phonemically represented as /*ɳ*/.

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

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

TODO:

This next section about the sibilant affricate could use some wording cleanup. It should also cover all the instances of /*ʈʂ*/ that are [s].

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

Sibilant Affricate Although the phone [t͡s] is rather common in the corpus, its underlying form is not always the phoneme /t͡s/. The underlying phoneme is usually elucidated by inflection. For example, both *tsekla* “steering wheel” and *cekla* “great uncle” share the same surface form, [t͡sekɫə]. However, when inflected for the plural, the former becomes [tasək'ɫaʒəɫ] and the latter [t͡sək'ɫaʒəɫ]. As such *tsekla* is analyzed as /tasekɫa/, whereas *cekla* is analyzed as /t͡sekɫa/. This analysis parallels liaison 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>ẽ</i>	(ə)	<i>a</i>
Low					a	<i>a</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.

1.3 Allophony

Supersegments

2

2.1 Stress

2.1 Stress 4

Stress in **lang₂** 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, *tąka* “tree” has typical stress and is unmarked, but *tąká* “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, *kagęsa* “army” has regular stress on the penultimate syllable, but when inflected in plural form, it surfaces as *kegęsazar* [ke.gǝ'sa.žǝ]. Secondary stress prevents the reduction of /e a o/ to [ǝ], 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 /tn̩z/ cluster likewise affricates, but is realized as [tʃ̩], 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 / \widehat{ts} s/ clustering with / \widehat{fx} ɬ/ are realized as [s:], romanized as *cc* or *ss* depending on the underlying phoneme. The voiced sibilant / \widehat{nz} / 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 / \widehat{ts} s \widehat{nz} /. For example, *ksarat* /kosəɹət/ is commonly realized as [ksəɹət], 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 [ə̃] rarely undergoes deletion as it is typically longer than [ə].

Note that / \widehat{ts} / 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

Modern native lemma derive from CV or CV(C)CV roots. Some roots have null initials, but they are comparatively rare. Roots allow CC clusters that aren't allowed across morpheme boundaries.

Roots are nominal by default, and in the corpus many roots appear in bare form as nouns. Stems reflect reduced derivational morphemes that transformed roots into verbs or other nouns.

The most prominent verb stem is *-t*, forming both transitive and intransitive verbs with a variety of meanings. Another common stem is *-m*, which also derives stems of either valency class and has a vaguely frequentative connotation. The other rarer verb stems are the intransitive *-k* and the transitive *-l* and *-y*, which encoded some kind of telicity distinction.

Outside of bare roots, *-s* stem is the most common for nouns. There are also *-c* stems and *-z* stems, which likely derived collective and agentive nouns, respectively. All nouns belonging to the sibilant stems are neuter gender. The non-sibilant *-r* stem can form nouns of either gender.

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].

4.1 Roots 7

4.2 Frequencies . . . 7

Stem endings can often be reconstructed with vague meanings.

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 ą b c d e ę f g h i k l m n o ɔ 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 ə o ɛ ɔ 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
```

Part II

MORPHOSYNTAX

The **lang₂** noun phrase is largely analytic, but nouns do inflect for deixis and number.

5.1 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*.

Although diachronically related to demonstratives, spacial deixis is only a secondary use of the proximal and distal forms. Instead, they are primarily indicators of non-propositional evidentiality, *i.e.* the speaker's evidence of a nominal referent. Evidence in this sense is visual or non-visual, the latter largely encoding hearsay or inference. The non-propositional evidentiality system in **lang₂** is semantically based on the speaker's knowledge at or prior to speech time.

5.1.1 Generic

The generic form of the noun is morphologically least marked and used as the citation form.

The generic form is most often used for gnomic statements.

(1) **Kəsa esyi oc.**

kəsa esyi oc

hero is good

“Heroes are good.”

If direct or indirect evidence exists, it's infelicitous to use the generic form.

5.1.2 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**.

5.1	Deixis	10
5.2	Number	11
5.3	Gender	13
5.4	Apposition	13
5.5	Phrasal syntax	14

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.

[Bhadra \(2020\)](#) terms this a Type I system.

-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.

5.1.3 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 samés lar t-het.**

egi matkó aczé sem-és lar tɛ=het
just basket\DIST two\DIST will-3.DIST there to=be at

“She said there’ll be just two baskets.”

(5MOYD #1314)

5.2 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.

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.



Figure 5.1: Additive vs. associative plurals

5.2.1 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 ??, *matkozar* becomes *matkazár*, not **matkózar*.

The second meaning would use the associative plural.

Mandatory plural marking is stylistically preferred in formal contexts.

5.2.2 Other plural strategies

Because the suffixes of *s*-stems and *r*-stems merge in the plural, some minimal pairs are rendered homophones when inflected. There are a few strategies to combat homophony.

Adjectives Speakers sometimes employ *tevi* “many” as an appositive modifier for the singular form, as in (3a)

Explicit number When number is specified with a numeral, the noun is not marked for plurality, as in (3b).

- (3) a. *vęci tevi* “some mercenaries” b. *vęci ocza* “two mercenaries”

5.2.3 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.

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

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

(4) **Sayanalal oti ezzu Kanyi z-laran.**

sayenal-al ot-i ezzu Kanyi ez=laran
be ignorant-CVB be-PROX ASSOC NAME in=this

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

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

(5) **ezzu Mázzirran segi s-yiat saska hakra.**

ezzu mázzirran segi su=yiat saska hakra
ASSOC ballers want:REFL to=steal comeback

“Each and every player wants to pull of the upset.”

5.3 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.

5.4 Apposition

Apposition is fairly common due to the small adjective class present in **lang**₂. Apposition is used in a number of collocative constructions but is distinct from compounding largely because of stress patterns and morphophonological processes. Furthermore, appositionive nouns shown agreement with the noun they modify, like adjectives.

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

5.4.1 Inalienable possession

Apposition is a fairly common strategy for possession of all types, but inalienably possessed nouns require apposition, as in (6a).

Alienable possession via apposition is dispreferred in formal registers.

- (6) a. **hora kəsa** b. ?**hora im-kəsa**
 “soldier’s wrist” “soldier’s wrist”

The prepositional construction more readily lends itself to a reading indicating inalienable possession, such as “soldier’s craftsmanship” or “soldier’s weaponsmithing.”

5.5 Phrasal syntax

Noun phrases are predominantly head-initial. Generally speaking, syntactically simpler constituents occur before more syntactically complex ones.

- (7) 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.

Like the noun phrase, the **lang₂** verb phrase is mostly analytic, and inflection is largely reserved for agreement. Other parts of the verb complex are handled by periphrastic constructions, notably including valency operations and TAM.

7.1 Agreement

Verbs display agreement along two axes: *deixis* and *person*. Deictic agreement is only with the subject of the clause, but person agreement is with both subject and the object, if the verb is transitive. Only the sole finite verb of a clause bears agreement; verbs demoted to adjuncts or arguments are always uninflected.

7.1.1 Person

In old **lang₂**, verb agreement was transparently derived from cliticized pronouns. However, sound change reduced agreement endings, leading to synchronic forms that are often actually patient agreement.

TODO:

How to cover morphophonological stuff? The two main ones are affrication of /t/ + /s z/ and cluster-final /s ts/ merging to [s], spelled *s*.

Except in a few outlier cases, agreement is fairly predictable from Table 7.1. Neuter agreement takes precedent over other morphemes when applicable. Furthermore, *-a-* only appears on intransitive verbs; for transitive verbs, the appropriate agreement with the other argument is used instead. Finally, *-r-* is used for intransitives, not the agent form.

When attached to a finite verb without deictic agreement, person agreement morphemes are realized as short: unlike many other affixes, they do not shift stress. This realization preserves old stress patterns before the morphemes were reduced. Because of this quirk, some scholars argue that person agreement morphemes are actually clitics to which deictic affixes attach.

7.1	Agreement . .	16
7.2	Converb	18
7.3	Transitivity . .	18
7.4	Negation	20
7.5	Aspect and mood	20

Despite person agreement, **lang₂** is rarely pro-drop.

Morpheme	Meaning
<i>-rc-</i>	1.AG
<i>-r-</i>	1.PT
<i>-a-</i>	2
<i>-s-</i>	3C
<i>-z-</i>	3N
<i>-ns-</i>	1↔3N
<i>-k-</i>	REFL

Table 7.1: Summary of person agreement

1st-person agreement is ergative as an accident of sound change.

TODO:

Do I want to keep this or say they got analogized? Maybe there's some weird interactions with the stress-fixing rules I've been cooking up?

Because **lang**₂ allows frequent fronting, a clause's agent and patient can become ambiguous when the agreement morphemes are not sufficient. This especially occurs with common and neuter agreement. A mediopassive construction formed with *het* can be used to clarify such instances. In the mediopassive, the semantic verb is demoted to adjunct as a converb.

The *het* construction is not a true passive: the verb does not change valency (*i.e.* the A-like argument can't be omitted).

(8) **Kanyi akvací Arpat.**

Kanyi akvací Arpat
NAME tagged:3C NAME

"Kanyi tagged Arpat"

Or: "Who Arpat tagged was Kanyi"

(9) **Kanyi hací Arpat akvetnal.**

Kanyi hací Arpat akvetnal
NAME be:3C NAME tagging

"Kanyi was tagged by Arpat"

In (8), it's not clear if Kanyi is the tagger or the focus-fronted taggee; both interpretations are grammatical. The use of the passive in (9) is less ambiguously interpreted, always meaning that Arpat is the semantic agent.

7.1.2 Deixis

Generic noun forms do not have agreement morphemes, but proximal nouns demand the verbal suffix *-i* and distal nouns shift stress to the final syllable. These suffixes are attached after person agreement suffixes. The proximal suffix *-i* is often stressed itself.

The stress shift is a remnant of an elided morpheme.

TODO:

Same question as earlier: do I like the stress rules for this affix? Or the way this next one preserves things?

Due to historical stress rules, the distal agreement suffix is more fusional for some select verb forms. Although the rime was lost from the common and neuter agreement suffixes in other forms, it remains in the distal form.

Generic	Distal
-s7 Verbs- <i>séc</i> 18	
-z	-zóc

Table 7.2: Deictic forms of agreement

Deictic agreement requires verbs to encode the same evidentiality as the subject noun phrase, but this may be infelicitous in some contexts. The *het* construction can also be used to change verbal agreement.

(10) ***Azzár hossusassi nassoin.**

as-zár *hassusal-s-i* *nassoi-n*
 foreigner-PL.DIST exalt-3C-PROX king-PROX

Intended: “(I see) the foreigners (I’ve heard about) praising the king.”

(11) **Nassoin heci azzár hossusalal.**

nassoi-n *het-s-i* *as-zár* *hossusalal*
 king-PROX be-3C-PROX foreigner-PL.DIST exalting

“(I see) the king being praised by the foreigners (I’ve heard about).”

In (10), the 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, *azzár*. To correct this, the construction in (11) is used, which takes advantage of the passive to mark the verb phrase as proximal.

7.2 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.

7.3 Transitivity

Transitivity is lexically set. There are three valency classes a verb can fall into: *transitive*, *intransitive*, and *pseudo-transitive*. Transitive verbs always have two arguments and intransitive verbs always have one. Pseudo-transitive verbs also take more than one argument, but are morphologically intransitive, and as such their additional argument is a prepositional phrase that cannot be omitted.

All valency classes allow a number of optional but often collocated adjuncts, introduced as prepositional phrases or converbs.

Non-finite verbs of each type do not have fixed valence and can omit all arguments. In some periphrastic constructions, those arguments are still required by the new finite verb. However, other constructions, and general adjectival or adverbial use, often appear without overt arguments.

There are few methods to decrease a verb's valency, which is usually done when an argument is sufficiently clear from context. In contrast, there are no methods to increase a verb's valency.

The most common method of valency reduction is dummy objects. Most transitive verbs have a collocated dummy object. Prescriptive convention holds that verbs exhibit person agreement with these objects, but in speech these verbs are often treated as morphologically intransitive, as in (12b). Although colloquial, this phenomenon points to further grammaticalization of dummy objects.

Some verbs have multiple collocations for different senses.

- (12) a. **a Semsî sora.** b. **a Sémai lar.**
 a semsi sora *a sémai lar*
 you say:3C DMY you say:2 DMY
 “You're talking.” “You're talking.”

Informally, the more grammaticalized *lar* is more common than a collocated dummy noun.

For verbs that lack a distinct collocated intransitive form, or for certain pragmatic reasons, a periphrastic construction can also serve as a valency-changing operation. The intransitive copula *ot* is the most common, demoting the semantic verb to adjunct as a converb.

- (13) **Nassoin oci kqstetal.**
 nassoin oci kqstetal
 the king is leading
 “The king's in charge.”

Although rarer than object omission, subject omission is accomplished through the *pit* passive. The passive construction with *pit* promotes the object to subject and demotes the semantic verb to an adjunct with *ez*.

TODO:

Some more info about using the reflexive to reduce valency, too.

7.4 Negation

Negation can be handled in multiple ways. The typical method is a periphrastic construction with the verb *rek*. Other methods include the use of discourse particles ...

TODO:

Write the section about which particles can convey negation, when, and how etc.

7.5 Aspect and mood

7.5.1 Perfective

The perfective construction uses the intransitive copula *ot*, demoting the semantic verb 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.

7.5.2 Irrealis

The imperfective construction uses *sem* “say, want,” demoting the semantic verb 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.

7.5.3 Subjunctive

The subjunctive construction uses *nenat* “crouch,” demoting the semantic verb to adjunct with the preposition *ah*. The subjunctive construction has a more limited scope than the *sem* irrealis construction, typically expressing counterfactuals or doubt.

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 *kəşazar* “soldiers,” marked via affix, to *retus im-kəsa*, “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 *tə-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 *tə*

The adposition *tə* 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 emphasize or in some fixed constructions, such as *su kagəsa su kagəstapa* “both army and navy.”

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.

(14) *yiz*: Otr. vi: Mans pesayrci rat, vi: Mans kirayrci rat?

yiz *ot-r* *vi* *mans* *pesay-rci* *rat* *vi* *mans* *kiray-rci* *rat*
 then be-1 Q who marry-1.AG 1 Q who find-1.AG 1

“Now then—who do I marry, who do I find?”

(5MOYD #1518)

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 26

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.

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.

Often, the fronted element will be the conversational focus.

Part III

APPENDIX

A.1 Metaphors

A.1 Metaphors . . . 28

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 29

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	31
Ą	31
C	31
E	31
Ę	31
H	31
I	31
K	31
L	32
M	32
N	32
O	32
Q	32
P	32
R	32
S	32
T	32
U	33
V	33
Y	33
Z	33

A

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

avara • *noun* • /avără/ ← OL₂ *agăré*,
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 kəštezi** 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 *hös-
húsär*, redup. of *húsär* • **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 *ísto*, *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

kəstat • *verb tr.* • /kěstět/ • **1** lead (smn.)
towards *tē* a goal **2** command: *nassoin
kəstaci su kageša su kagėstaspa* “the king
commands both army and navy.” **3** train
(an apprentice) in *tē* a skill: *tē-coryam
kəstatrí rat lala* “auntie’s teaching me to
sew.” **4** formally teach (a student) in *ez* a
discipline: *rappahan picí kəstetnal 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/ ← pormanteau 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

sem • *verb tr.* • **1** say to, talk to **2** *refl.* tell that **3** *refl.* say **4** **sem ???** talk, speak **5** *coll., intr.* talk, speak **6** speak on, give an opinion on smth. concerning *ez* **7** **sem véaja** preach on *ez* **8** **sem ???** give a speech about *ez* **9** **sem ???** lecture on *ez* **10** want smn. to *im* do **11** *refl.* want to *su* do **12** **sem tẹ** would **13** **sem tẹ** ought **14** **sem tẹ** hopefully **15** (*of symbols, texts*) signify **16** (*of words, phrases*) mean, be defined as **17** (*of title, roles*) enable, allow to **18** **sagí tẹ** *coll., refl.* let's go to *tẹ: sagí tẹ-tesan* “let's go to the beach” **19** **sagí tẹ** *coll., refl.* let's *tẹ* do: *sagí tẹ-kęstat lala* “let's hire a nanny” **20** *coll.* looks like smn. **21** **samés** *coll., intj.* looks like!

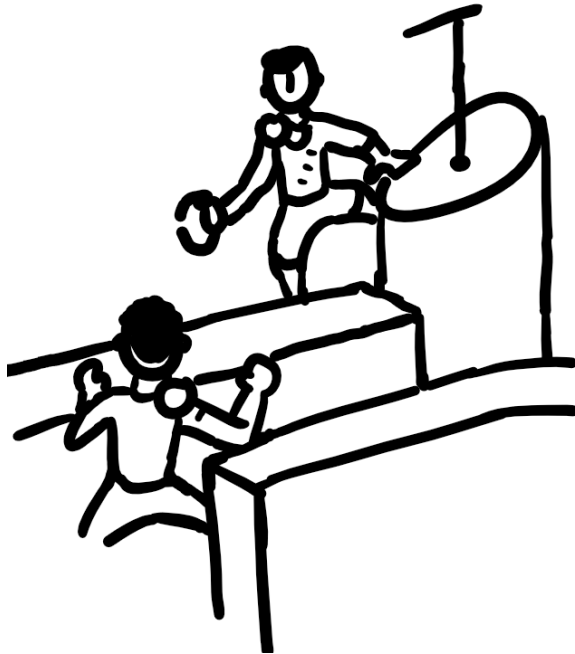
sisik • *noun* ← clipping of *tahęnar asyisaikli* “associated tag” • see *tahęnar* • **1** *Inland* tag

T

tahat • *verb in.* • /táhět/ • **1** stretch (one's limbs) **2** reach for *tẹ* something (to grab it) **3** aspire to *tẹ* a difficult goal: *tẹ-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.

tahēnhar • *noun* • /tăhénhǎ/ ← OL₂

tahēnâhhǎ “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* • /teukkânǎs/ ←

Farlands **tiēk'kânos** • **1** (**Véajan**) *ntr.* missionary, pastor; travelling evangelist who starts local chapters of faith centers **2** *humorous* tourist, vacationer

U

V

véaja • *noun* • /véă(n)jǎ/ ← Farlands **věžšā** • **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 sesamzî véaja ez-ǝkas esyi tē-laczin** “the minister preached of an eternal reward for the faithful”

Y

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

Z