Sketch of a Basheg grammar

A speedlang for Mareck

Tsṇa Hătsui

Contents

— 1 —

Introduction

Basheg (endonym: *bəšəg*) is spoken by some 30 000 people in a group of villages in the western part of the Ayasi mountains. This part of the mountain range is sometimes known as the Bashungur mountains, after the local name. Basheg has no obvious relation to any other language (while there are several interesting parallels with neighboring languages such as the well-known Tsuy, these are best explained as resulting from a long history of mutual contact).

An unusually sapient cat has suggested that I investigate this language, and I here present some sketched results of a very preliminary attempt at doing so. I have not yet had the chance to travel to the Bashungur mountains for fieldwork, but I have luckily been able to learn some of the language from the wonderful Aŋ Səlul, a native Basheg speaker who has resided here in Bonang these past few months.

Some preliminary findings from this endeavor that may be of particular interest to certain readers are listed as follows:

1. The system of 'lambdized' morphophonological forms (§2.7)
2. The phonological voicing distinction in plosives (§2.1)
3. The pattern of vowel (non-)epenthesis at certain word edges (§2.5)
4. The topic-prominent syntactic structure (§5.1)
5. The phasal polarity adverbs (§5.3)
6. The simple tense-aspect system with only one overt marker (§4.4), contrasted with the relative abundance of mood and evidentiality markers (§4.5)
7. Various modest insights into the Basheg lexicon (distributed throughout the sketch)
8. Several glossed example sentences

Finally, a note to the particularly observant reader: while it is a common recommendation to use the venerable Syntacticians's Manual of Useful Diagnostic Sentences (SMOUDS) when documenting an unknown language, I have eschewed this in favor of eliciting improvised sentences exploring whatever structures I felt like investigating at the time, because I have no respect for authority. Sorry.

Notes for Mareck: I expect that you may receive this document later than hoped for. This is, ahem, largely due to an unforeseen disruption in interdimensional communications, and may be further exacerbated if my intermediary in your realm does not deliver the sketch promptly. I do apologize for this.

Yours, Tsṇa Hătsui

— 2 —

Phonology

2.1 Phonemes

There are 19 consonant phonemes. Only the labiovelar plosives stand out as being unusual, though such consonants are not unheard of in the region.

	Labial	Alveolar	Palatal	Velar	Labiovelar	Glottal
Nasal	m	n		ŋ		
Plosive	p	t		k	k͡p	
	b	d		g	g͡b	
Fricative	f	s	ʃ	x		(h)
Tap		ɾ				
Approximant		l	j		w	

The stop system in unusual in that neither series seems to be articulatorily unmarked: the voiceless stops are clearly aspirated [pʰ tʰ kʰ k͡pʰ], and the voiced stops are modal [b d g g͡b]. There is no lexically specified /h/ phoneme, but words otherwise beginning with a vowel are pronounced with a [h-] onset if utterance-initial or prosodically emphasized. There is also no /ŋ͡m/ phoneme, but there are phonetic [-ŋ͡m.g͡b-] and [-ŋ͡m.k͡p-] clusters resulting from assimilation.

The vowel system is simple, albeit with an odd asymmetry. There is no /o/ or /ɔ/-like vowel, and the highly frequent schwa /ə/ cannot be considered any sort of '/ɔ/ in disguise'. Basheg has a form of vowel harmony, detailed later, in which the schwa acts clearly as a higher counterpart to /a/, not a lower counterpart to /u/.

	Front	Central	Back
High	i		u
Mid	ɛ	ə	
Low		a	

The orthographic transcription is as given in the tables above, except that /ɾ/ is written *r*, /ʃ j/ are written *š y*, /k͡p g͡b/ are written *ṗ ḅ*, and /ɛ/ is written *e*. The consonants /k͡p g͡b/ must not be confused with the clusters /k.p g.b/, as there are contrasts like *səṗu* [sə.k͡pʰu] 'eagle' vs. *səkpu* [səkʰ.pʰu] 'large basket'. The phonetically assimilated [ŋ͡m] is considered a form of /ŋ/, and written accordingly.

The only phonologically distinctive elements in Basheg are the consonants and vowels presented above. In particular, length, stress, and tone are not distinctive. There is regular phrase-final prosodic prominence that may be heard as stress, but this is not a word-level phenomenon.

2.2 Syllable and word structure

Basheg is fairly fond of codas and consonant clusters, at least more so than its major neighboring languages. A general schematic of the Basheg phonological word can be given as $[C^{?}VC^{?}]^{+}O^{?}$, i.e. one or more syllables of the shape (C)V(C), optionally followed by a word-final consonant which must be an obstruent (O). Ignoring the question of vowel hiatus for now, an exhaustive listing of bisyllabic word shapes without hiatus is given below.

Shape	Example	Shape	Example
V.CV	*iri* 'wisdom'	CV.CV	*ḅəli* 'bone'
VC.CV	*elte* 'meat'	CVC.CV	*yašta* 'soft'
V.CVC	*ayar* 'needle'	CV.CVC	*bəšəg* 'Basheg'
VC.CVC	*ilsəm* 'sun hat'	CVC.CVC	*teldaŋ* 'rope'
V.CVCC	*emaŋk* 'type of stone'	CV.CVCC	*deŋelt* '2744'
VC.CVCC	*usturs* 'pasture'	CVC.CVCC	*faŋḅekt* 'bridge'

Vowel hiatus is allowed, but it appears to be very rare. Only /ə ɛ a/ may begin a vowel sequence, as /i u/ productively epenthesize semivowels /j w/ between themselves and a following vowel. Hiatus sometimes results from morphology, and in very few cases exists prespecified in roots such as /-a.ɛk͡p/ *-aeḅ* 'to steal' or /nə.i/ *nəi* 'liver'.

There is no moraic minimum on words, and monosyllabic words may have the shape of any legal syllable, including a bare CV as in *fe* 'tent'.

2.3 Consonant processes

There are two phonological processes in which a consonant assimilates to another, both process both processes being quite frequent in morphology and optionally applicable as cross-word sandhi. There is also a dissimilatory process that is limited to a few affixes and never applies across words, as well as a geminate avoiding process.

- **Progressive devoicing** in obstruent-obstruent clusters: voiced obstruents become voiceless after voiceless obstruents.
- **Regressive place assimilation** in nasal-stop clusters: when a nasal consonant precedes a plosive, the former assumes the place of articulation of the latter.
- **Sibilant-stop dissimilation:** when created by suffixation, the clusters /s-s ʃ-s s-ʃ ʃ-ʃ/ resolve to [st^{h} ʃt^{h} ʃt^{h} ʃt^{h}].
- **Degemination:** clusters of two identical consonants reduce to a single consonant (except suffixal sibilants, which dissimilate as stated).

Implicit in the statement of the devoicing rule is that voiced-voiceless clusters are unaffected. This is true, and words like *ags* 'table' and *šədku* 'song' with such clusters are not uncommon.

The devoicing rule commonly affects inflectional suffixes like instrumental /-dE/, passive /-b/, and inferential /-gE/, which become devoiced after voiceless obstruents:

(1) /anak-dE/ [anak^{h}**t^{h}**ɛ] *anakte* (hammer-INS)
/ʃ-Ark-b-i/ [ʃaɾak^{h}**p^{h}**i] *šarakpi* (3SG.M-see-PASS-PFV)
/ʃ-ɛnɛs-gE/ [ʃɛnɛs**k^{h}**ɛ] *šeneske* (3SG.M-be_married-INFR)

The other assimilation rule is about nasal consonants. In Basheg, the contrast between different nasal consonants is neutralized before a following plosive consonant. This gives rise to alternations in roots and inflectional suffixes:

(2)

/g͡bəli-m-dE/	[g͡bəli**n**di]	*ḅəlindi*	(bone-PL-INS)
/sAŋ-dE/	[sa**n**dɛ]	*sande*	(rain-INS)
/fəxən = pi/	[fəxə**m**p^{h}i]	*fəxəmpi*	(table = also)
/š-in-b/	[ʃi**m**b]	*šimb*	(3SG.M-buy-PASS)
/š-in-gE/	[ʃi**ŋ**gi]	*šiŋgi*	(3SG.M-buy-INFR)

Nasals only assimilate before plosives. Heterorganic clusters such as *mš*, *ŋs*, *nx*, *ŋn*, *mŋ*, *ŋr*, and so on are common.

There is a kind of dissimilation: clusters of two sibilants that are formed through suffixation are avoided by turning the latter into a stop, and in the case of /s-ʃ/, spreading palatalization to the preceding sibilant, giving [ʃt^{h}]:

(3)

/kaɾas-sA/	[k^{h}aɾa**st^{h}**a]	*karasta*	(pouch-INE)
/buʃ-sA/	[bu**ʃt^{h}**ə]	*buštə*	(sky-INE)
/nis-ʃ/	[ni**ʃt^{h}**]	*ništ*	(son-DEF.M)
/ŋɛɾaʃ-ʃ/	[ŋɛɾa**ʃt^{h}**]	*ŋerašt*	(hunter-DEF.M)

Only suffixes ever participate in dissimilation: prefixes do not undergo or trigger it.

Finally, Basheg has no geminate consonants, and when two identical consonants would otherwise be adjacent, one of the two is deleted. In productive formations, this rule only applies after other rules have taken place, including epenthesis (§2.5).

(4)

/ɾɛbad-dE/	[ɾɛba**d**ɛ]	*rebade*	(cucumber-INS)
/ʃ-Ek-gE/	[ʃɛ**k^{h}**ɛ]	*šeke*	(3SG.M-be_away-INFR; fed by devoicing)
/nɛs = s-Aŋ/	[nɛ**s**aŋ]	*nesaŋ*	(so_that = 3SG.N-fall; no dissimilation)
/d-AfAn-n-i/	[dafa**n**i]	*dafani*	(NEG-fly-3PL-PFV)
/d-AfAn-n/	[dafa**n**a**n**]	*dafanan*	(NEG-fly-3PL; blocked by epenthesis)

The example above illustrates that devoicing in a cluster like /-kg-/ feeds degemination (through an intermediate /-kk-/ stage). However, Basheg in general allows unassimilated voiced-voiceless obstruent clusters. A cluster like /-gk-/, thus, can in fact surface as [-gk^{h}-], though at least in the speech of my consultant, forms like /d-uɾəg-k-i/ *durəgki* 'I didn't sleep' only clearly surface as [du.ɾəg.k^{h}i] in very careful speech, with the more natural form being [du.ɾə.k^{h}i], and similarly for /-bp-/ and /-dt-/ clusters.

2.4 Vowel harmony

Basheg has a form of vowel harmony, though it is not absolute: disharmonic words are common and regular results of certain morphological configurations.

The harmony can be considered height harmony or tongue root harmony (the phonetic details of its realization demand further study, and it will be treated as height harmony here for convenience). The vowels /i ə u/ belong to the 'high' group, and /ɛ a/ to the 'low' group. There are two alternating vowel pairs: /i ɛ/, notated as an underspecified phoneme /E/, and /ə a/, notated as /A/. The high vowel /u/ does not have a low counterpart, and is never present in any alternations. Vowel harmony is manifested in four rules:

- **Root specification:** root morphemes have only one harmonic value: all vowels in a single morpheme are either specified high /i ə u/, underspecified /E A/, or specified low /ɛ a/. This is a static constraint that does not in itself lead to any (productive) alternations.
- **Progressive spreading:** harmonic values spread to the right onto underspecified vowels.
- **Regressive spreading:** harmonic values may spread to the left from a stem onto a prefix that remains underspecified after progressive spreading.
- **Default low:** underspecified vowels surface as low if not otherwise specified as high.

The default-low rule means that only spreading of [+high] is directly detectable, but there is reason to believe that [−high] can also spread onto underspecified vowels, as it in doing so may have the observable effect of blocking regressive spreading.

Most suffixes have underspecified vowels, such as locative /-sA/ and instrumental /-dE/, meaning that they will surface with a high or a low vowel depending on the root:

	High			Low	
	'branch'	'shoulder'	'surface'	'tent'	'pig'
Locative	*siksə*	*məŋsə*	*udsə*	*fesa*	*šapsa*
Instrumental	*sikti*	*məndi*	*udi*	*fede*	*šapte*

Some suffixes are uniformly high, such as past tense /-i/. When attached to a low stem, the high value does not spread, and the resulting word is disharmonic, such as /ʃ-Alt-i/ [ʃalt**i**] 'he ate'. This is the primary way in which disharmonic words surface.

Monosyllabic roots always have high or underspecified vowels. The latter type is susceptible to prefix-controlled harmony, caused for example by the third person plural neuter agreement marker *id-*.

(5) /ʃ-Aŋ/ [ʃ**aŋ**] *šaŋ* 'he falls'
/id-Aŋ/ [id**əŋ**] *idəŋ* 'they (N) fall'
/ʃ-Eɾp/ [ʃ**ɛɾp**^{h}] *šerp* 'he dies'
/id-Eɾp/ [id**iɾp**^{h}] *idirp* 'they (N) die'

This prefix-controlled harmony can propagate through the root and onto suffixes:

(6) /ʃ-Aŋ-ɾA/ [ʃaŋ**ɾa**] *šaŋra* 'he might fall'
/id-Aŋ-rA/ [idəŋ**ɾə**] *idəŋrə* 'they (N) might fall'

The verb roots in the above examples are underspecified /-Aŋ/ and /-Eɾp/. There are monosyllabic roots that are invariably high, such as /-us/ 'to gather', and /-əg/ 'to be', but there are no invariably low monosyllabic roots such as */-aŋ/.

Polysyllabic roots are more diverse. Some bisyllabic roots such as /-Ag͡bAn/ 'to slide' (most of which have two /A/ vowels) are susceptible to harmony:

(7) /ʃ-Ag͡bAn/ [ʃ**ag͡ban**] *šaḃan* 'he slides'
/id-Ag͡bAn/ [id**əg͡bən**] *idəḃən* 'they (N) slide'

However, most polysyllabic roots are either specified high or low and are not affected by harmony, e.g. /-atsɛɾ/ 'to stink'. There are also some specified-low suffixes that are unaffected by root-controlled harmony, such as the hearsay evidential /-ɛla/. This is another variation on how disharmonic words can surface.

(8) /id-atsɛɾ/ [id**atsɛɾ**] *idatser* 'they (N) stink' (not *[id**ətsiɾ**])
/ʃ-iɾip-ɛla/ [ʃiɾip^{h}**ɛla**] *širipela* 'he is said to be wise' (not *[ʃiɾip^{h}**ilə**])

Finally, some polysyllabic roots are affected by harmony only in idiomatic phrases and lexicalized derivations, and not in productive formations.

So far, only progressive spreading has been demonstrated. Regressive spreading is less common, but also an essential part of Basheg vowel harmony. Certain prefixes, such as first person plural /Ag-/ are affected by regressive harmony, raising before a high root:

(9) /Ag-Aŋ/ [**ag**aŋ] *agaŋ* 'we fall'
/Ag-Eɾp/ [**ag**ɛɾp^{h}] *agerp* 'we die'
/Ag-us/ [**əg**us] *əgus* 'we gather'
/Ag-iɾip/ [**əg**iɾip^{h}] *əgirip* 'we are wise'

Regressive spreading is restricted to prefix-root configurations as the above, which is why it does not apply in root-suffix configurations like the disharmonic [ʃalti] discussed above.

Other prefixes, such as /ak͡p-/ 'while', are unaffected by regressive harmony, so it is analyzed with a specified-low vowel: /ak͡p-id-us/ [**ak͡p**idus] 'while they (N) gather' (not *[**ək͡p**idus]). This is the final way that disharmonic words can surface. What makes this configuration particularly interesting is that when the low prefix /ak͡p-/ is used before an underspecified prefix like /Ag-/, the latter always surfaces as low [ag-], even before a high root. This is explained as progressive spreading of the specified [−high] value from /ak͡p-/ onto /Ag-/, before regressive spreading from the root has a chance to apply. Compare the following with the earlier example 9:

(10) /ak͡p-id-us/ [ak͡p**id**us] *ạ̀pidus* 'while they (N) gather'
/ak͡p-Ag-us/ [ak͡p**ag**us] *ạ̀pagus* 'while we gather' (not *[ak͡p**əg**us])

2.5 Epenthesis

Basheg has two broad types of epenthesis: regular epenthesis repairing phonotactic violations, and exceptional forms of epenthesis in reduplication. This section is concerned with the former.

Regular epenthesis is triggered as a repair mechanism in three cases, all of which represent phonotactically impermissible structures:

- **Onset clusters:** Basheg has no complex onsets, so if more than one consonant is present before the first vowel, an epenthetic vowel is inserted to break the cluster up.
- **Medial complex codas:** Basheg allows complex codas of two consonants, but only word-finally, so if more than two consonants are present between two vowels, an epenthetic vowel is inserted.
- **Final overcomplex codas:** Basheg allows complex codas of two consonants word-finally, so if more than two consonants are present after the last vowel, an epenthetic vowel is inserted.

The above rules intentionally do not say what quality the epenthetic vowel has, or exactly where it is placed, as this is a matter of some complexity.

In general, the epenthetic vowel may be characterized as a vowel underspecified for both height and frontness. In the absence of preceding vowels, i.e. when epenthesis occurs before the first vowel, the epenthetic vowel defaults to not being front, behaving as if it were /A/, i.e. surfacing as /a/ before a low stem and /ə/ before a high stem. When there is a preceding vowel, the [+front] feature of /i/ and /ɛ/ may spread onto the epenthetic vowel, in which case the epenthetic vowel receives the same vowel quality as the vowel before it. However, for whatever reason, the [+back] feature of /u/ cannot spread. Only the [+high] feature of /u/ spreads onto an epenthetic vowel, resulting in epenthesis of /ə/. These possibilities will be illustrated below.

The placement of the epenthetic vowel is also subject to a consideration of several cases. In the most general case, epenthesis occurs after the first consonant of the offending cluster. Exceptions to this have to do with certain consonant clusters that epenthesis does not generally break up.

Examples of the general behavior in word-medial and word-final epenthesis are shown in the table below, triggered by the addition of the ablative /-ŋA/ and the genitive /-t/ to nouns with -RC and -NS stems.

	'raven'	'bag'	'paint'	'caterpillar'	'scar'
Nominative	*niŋs*	*šərg*	*ulb*	*sert*	*malš*
Ablative	*niŋ**i**sŋə*	*šərəgŋə*	*uləbŋə*	*sere**t**ŋa*	*mal**a**šŋa*
Genitive	*niŋ**i**st*	*šərəgt*	*uləbt*	*sere**t***	*mal**a**št*

The underlying form of the stems above is assumed to be that found in the nominative, with the epenthetic vowel being productively added stem-internally in the process of inflection.

(11)	/mAlʃ/	[malʃ]	*malš*	‘scar (NOM)’
	/mAlʃ-ŋA/	[mal**a**ʃŋa]	*malašŋa*	‘scar (ABL)’
	/mAlʃ-t/	[mal**a**ʃt^{h}]	*malašt*	‘scar (GEN)’

Note also that /sEɾt-t/ triggers regular stem-internal epenthesis (to an intermediate form *sɛɾɛtt), before undergoing degemination to produce the output form [sɛɾɛt^{h}], with ‘opaque’ epenthesis.

The pattern in the above examples applies to stems ending in a liquid followed by an obstruent, or a nasal followed by a fricative, as demonstrated above, as well as to stems ending in an consonant-sonorant cluster (for which see below). Stems ending in other clusters (viz. obstruent-obstruent or nasal-stop) are not themselves broken up by epenthesis, and an epenthetic vowel is instead forced to apply between the stem and the suffix:

(12)	/tuŋəʃt-ŋA/	[tuŋəʃt^{h}**ə**ŋə]	*tuŋəštəŋə*	‘cooking pot (ABL)’
	/tuŋəʃt-t/	[tuŋəʃt^{h}**ə**t^{h}]	*tuŋəštət*	‘cooking pot (GEN)’
	/ɾAnt-ŋA/	[rant^{h}**a**ŋa]	*rantaŋa*	‘edge (ABL)’
	/ɾAnt-t/	[rant^{h}**a**t^{h}]	*rantat*	‘edge (GEN)’

Word-final epenthesis is subject to the principles laid out above, as well as the consideration of whether a two-member cluster is legal word-finally. Word-final clusters of two consonants are legal (after any applicable assimilation and dissimilation) precisely if the latter consonant is an obstruent. If an underlying form contains a word-final cluster instead ending in a sonorant, that triggers epenthesis. This is commonly seen as a result of adding suffixes like plural /-m/, and lative /-ɾ/:

(13)	/sul-m/	[sulə**m**]	*suləm*	‘pigs’
	/dAp-ɾ/	[dap^{h}**a**r]	*dapar*	‘river (LAT)’

Epenthesis of a consonant-sonorant cluster can also occur within a stem, without the presence of any affixes. In this case, the unmarked form of the word contains epenthesis, and this epenthetic vowel may not be necessary in other forms of the word. This possibility is illustrated for the stems /ifɾ/ ‘wind’ and /ʃAtm/ ‘mortar’:

(14)	/ifɾ/	[if**i**ɾ]	*ifir*	‘wind’
	/ifɾ-A/	[ifɾə]	*ifrə*	‘wind (ACC)’
	/ʃAtm/	[ʃat^{h}**a**m]	*šatam*	‘mortar’
	/ʃAtm-A/	[ʃat^{h}ma]	*šatma*	‘mortar (ACC)’

Epenthesis in word-initial clusters operates by similar general principles. There is no preceding vowel to copy frontness from, and the epenthetic vowel always acts as an underspecified /A/, i.e. surfacing as [ə] or [a] depending on vowel harmony.

(15)	/d-x-aŋ-ʃ/	[d**a**xaŋʃ]	*daxaŋš*	‘may he not fall’
	/d-x-ib-ʃ/	[dəxibʃ]	*dəxibš*	‘may he not lie’

2.6 Reduplication

Basheg has several reduplicative processes, mostly having to do with lexical derivation. Full reduplication of nouns is common with a variety of more or less lexicalized meanings, and various forms of partial reduplication are used for different purposes.

Several forms of reduplication employ a phonologically anomalous form of epenthesis. The examples of epenthesis in the previous section were characterized as necessary repair of phonotactic violations, but in reduplication, epenthesis occurs sometimes where it does not appear strictly necessary in order to ensure phonotactic compliance. In fact, any stem that both begins and ends

in a consonant (or consonant cluster) triggers epenthesis between the stem and the reduplicand in full reduplication, even where a consonant cluster would be permissible:

	Base	Reduplicated
Epenthesis	*bin* ‘nest’	*bini**b**in* ‘hatchling’
	bəšəg ‘Basheg’	*bəšəg**ə**bəšəg* ‘Basheg talk, usage’
	šur ‘fire (mass n.)’	*šur**ə**šur* ‘flame’
	kes ‘path’	*kes**e**kes* ‘trail, trace’
	teldaŋ ‘rope’	*teldaŋ**a**teldaŋ* ‘act of dragging with rope’
No epenthesis	*firə* ‘saliva’	*firəfirə* ‘act of spitting’
	anak ‘hammer’	*anakanak* ‘act of hammering’

A word like *keskes* would be phonotactically legal, but the reduplication process enforces an additional epenthesis. Something similar is evident in a form of intensivizing partial reduplication of verb stems:

(16) *-us* ‘to gather’ > *-usus* ‘to reunite’
-eḅar ‘to use’ > *-eḅarar* ‘to exhaust’
-urš ‘to slumber’ > *-urəšurš* ‘to sleep deeply’
-emelb ‘to tremble’ > *-emelebelb* ‘to really shake’

In the above examples, the rime of the final syllable is reduplicated, but in the cases that this includes a consonant cluster, the medial copy of the cluster is broken up by epenthesis: [-urəʃurʃ] is not *[-urʃurʃ], and [-ɛmɛlɛbɛlb] is not *[-ɛmɛlbɛlb], even though these word forms are all legal from a strictly phonotactic perspective.

In addition to the full reduplication and right-edge (suffixing) partial reduplication seen above, Basheg also has left-edge (prefixing) partial reduplication in some derivational patterns (§3.1).

2.7 Lambdization

Lambdization is the name for an idiosyncratic process in Basheg whereby many (but not all) content words (noun, verbs, adjectives) have an alternate ‘lambdized’ form usually characterized by the presence of an /l/ phoneme in the stem. Lambdization is not an inflectional form, being more similar to derivation, but it is not a typical derivational process either. It has no particularly consistent semantics of its own, and is often completely optional or used for phonological euphony. However, it sometimes conveys certain pragmatic or stylistic nuances that are highly conventionalized in particular (syntactic or discourse) contexts, and it often features in fixed expressions or lexicalized phrases where it has become obligatory or acquired a specific meaning for a particular lexeme. As lambdization has such diffuse usage patterns, lambdized forms will not be glossed with any meaningful abbreviation, being instead indicated with a lambda symbol ‘ƛ’.

The precise usage of the lambdized forms is not a phonological topic, and is left for other sections (§3.1, §4.2, §5.1). This section focuses on the phonological patterns of lambdization, which are entirely parallel between nouns, verbs, and adjectives. These patterns have a regular core area of application, but are characterized by a high degree of irregularity and optionality outside of that.

Stems ending in *-r* or *-n* always have a lambdized form in which the coda consonant is replaced with *-l*. The immediately preceding vowel may also undergo fronting (*a ə* becoming *e i*), though this is irregular, and is most common with monosyllables.

		Plain	Lambdized
Nouns	‘sun’	*xar*	*xel*
	‘pebble’	*nəŋur*	*nəŋul*
	‘arrow’	*isən*	*isəl*
Verbs	‘find’	*-ubur*	*-ubul*
	‘jump’	*-ayan*	*-ayal*
	‘speak’	*-ən*	*-il*
Adjectives	‘big’	*tur*	*tul*
	‘calm’	*afan*	*afal*

Stems ending in other consonants or vowels may or may not have a separate lambdized form. Many stems ending in a single stop have a lambdized form with an *-l-* infixed before the stop. Other stems may have lambdized forms where an *-l-* replaces a medial *-r-* or *-n-*, or is found infixed next to another stem-internal consonant, with little regularity. The remark about vowel fronting applies to these cases as well.

		Plain	Lambdized
Nouns	‘branch’	*sik*	*silk*
	‘father’	*kare*	*kele*
Verbs	‘say’	*-ag*	*-elg*
	‘put’	*-ip*	*-ilp*
	‘cook’	*-əfən*	*-əflən*
Adjectives	‘hard’	*uŋim*	*ulŋim*

In some words (monosyllabic words ending in a vowel and certain forms regularly carrying an *l*), only a change from central to front vowels can be observed.

— 3 —

Nouns and noun phrases

Basheg nouns are morphologically complex, inflecting for number, definiteness, and case. In addition to this, Basheg allows (and sometimes requires) Suffixaufnahme. The Basheg noun phrase hosts a number of elements, most of which agree with the head noun in gender and number.

3.1 Noun stems, derivation, and gender

Every Basheg noun has a gender (masculine, feminine, or neuter), but unlike some other languages, Basheg gender is entirely predictable on the basis of nominal semantics: the masculine and feminine genders contain male and female humans, and the neuter gender contains everything else, including animals and gods, and sometimes humans whose gender is unknown (my consultant variously uses masculine or neuter in such situations, and accepts both when I ask).

Basheg noun stems often consist of a single root, but are also commonly derived from verbs or other nouns through prefixation, suffixation, or some form of reduplication. Prefixation is a particularly common way of signalling gender: many masculine-feminine noun pairs are related by prefixing *f-* in the masculine form, and *r-* in the feminine form, e.g. *fale* 'boy', *rale* 'girl', and in derivations (see below). Many masculine and feminine nouns are unprefixed, however.

Agent nouns can be derived from verb roots by prefixing *f-* or *r-* followed by reduplicating the first -VC- sequence of the root: *-inub* 'to sing' becomes *fininub* 'singer (M)' or *rininub* 'singer (F)', and *-əlt* 'to shoot (an arrow)' becomes *fələlt* 'archer (M)' or *rələlt* 'archer (F)'.

Complete reduplication of Basheg nouns (often with an epenthetic vowel as seen in §2.5) can be used for several lexicalized and semi-productive purposes. Commonly the reduplicated form is a form of action nominal corresponding to what one prototypically does with the reduplicated noun, such as *bəšəgəbəšəg* 'Basheg talk, usage' from *bəšəg* 'Basheg'. Another common function is as part of a part-whole relationship, in which the reduplicated form refers to a part of the (unreduplicated) whole, such as *šurəšur* 'flame' from *šur* 'fire'.

Todo: Derivational prefixes between gender and root

Todo: Derivational suffixes

Many noun stems have lambdized forms, which are usually entirely optional. For example, my consultant accepts *kəlti isənə* 'I shot an arrow' and *kəlti isələ* 'I shot an arrow.λ' as entirely synonymous and completely interchangeable. However, some lambdized forms have dedicated idiomatic uses: for example, while *xar* 'sun' and *xel* 'sun.λ' are usually synonymous, the latter forms part of an idiomatic construction: *xelmer damas* 'to four suns.λ' idiomatically means 'in four days' (and similarly with other numbers), but *xarmar damas* 'to four suns' only has the (unlikely) literal reading. Additionally, nouns and adjectives seem to appear in the lambdized form more often when focalized (especially when this results in a discontinuous noun phrase), but in such situations lambdization remains optional, if preferred. My consultant tells me that some phrases sound formal, almost condescendingly so, when read with lambdized forms.

3.2 Basic inflection of nouns and adjectives

Basheg inflects nouns and adjectives for the categories of gender, number, definiteness, and case. Adjectives are generally nounlike, but can be distinguished on the basis of morphology. Both nouns and adjectives exhibit Suffixaufnahme,[1] a type of far-reaching agreement in which an adjective or adnominal noun is marked for the number and case of its head noun as well as that of any nouns the head noun itself agrees with (recursively). Suffixaufnahme is detailed in §3.4, with this section focusing on simple, unembedded forms. The basic morphological structure of nouns and adjectives is as follows:

- **Nouns** are schematically: Stem-(PL)-(DEF)-(Case)-(Agreement by Suffixaufnahme...)
- **Adjectives** are schematically: Stem-Agreement-(Agreement by Suffixaufnahme...)
- **Agreement** is marked by a fusional suffix sensitive to gender, number, definiteness and case, with a very high degree of syncretism.

In addition to phonologically regular morphophonology, there is some amount of morphologically specified allomorphy, especially in pluralization of nouns and the expression of case.

Noun inflection

This section runs through the morphemes used in noun inflection. To start, an example paradigm of the basic forms of a typical masculine noun is shown in the following table. These forms follow the schema sketched above, with the morphemes listed below.

fana 'friend (M)'

	Singular		Plural	
	Indefinite	Definite	Indefinite	Definite
Nominative	*fana*	*fanaš*	*fanam*	*fanamiš*
Accusative	*fana*	*fanaša*	*fanama*	*fanamišə*
Genitive	*fanat*	*fanašt*	*fanant*	*fanamišt*
Lative	*fanar*	*fanašar*	*fanamar*	*fanamišir*
Inessive	*fanasa*	*fanašta*	*fanamsa*	*fanamištə*
Adessive	*fanawa*	*fanašu*	*fanamu*	*fanamišu*
Ablative	*fanaŋa*	*fanašŋa*	*fanamŋa*	*fanamišŋə*
Instrumental	*fanade*	*fanašte*	*fanande*	*fanamišti*

Pluralization is the most irregular aspect of noun inflection. Plural forms are regularly marked by the suffix *-m*, which gains an epenthetic vowel after consonants. As this epenthesis can also suppress stem-internal epenthesis that would otherwise occur, there are essentially three regular types of plural forms, as shown here:

(17) a. Plurals in *-m*:

ḅəli	>	*ḅəli-m*	'bones'
səxə	>	*səxə-m*	'hands'
tame	>	*tame-m*	'fields'

[1] Despite the unusual character of this term, I prefer it to the equally established alternative term 'case stacking'. The latter is inaccurate, as not only agreement in case but also in number (et cetera) may be stacked (or *aufgenommen*, as it were). One might call it 'agreement stacking', but this would be an unnecessary invention when there is already a perfectly accurate and well-known word for the phenomenon.

b. Plurals in *-Vm*:

sul	>	*sul-əm*	‘pigs’
anak	>	*anak-am*	‘hammers’
nərt	>	*nərt-əm*	‘words, things said’
anseŋ	>	*anseŋ-em*	‘coins’

c. Plurals in *-Vm* on epenthesis stems:

ifir	>	*ifr-im*	‘winds, gusts of wind’
katal	>	*katl-am*	‘drops’

The underlying stem for ‘wind’ is /ifɾ/, and the second [i] in the singular [ifiɾ] is epenthetic, which is why it does not surface in the plural, and similarly for ‘drop’.

There are several other pluralization patterns of varying degrees of regularity, all of which are less frequent. All plural forms have in common that they end in a consonant. Nouns ending in nasals and voiced stops often do not change in the plural form, some nouns ending in *-r* or *-l* take a denasalized consonantal allomorph *-b* without triggering epenthesis, and some nouns are completely irregular or suppletive.

(18) a. Unmarked plurals:

nəm	>	*nəm*	‘candles’
seŋ	>	*seŋ*	‘flies’
ṗirəd	>	*ṗirəd*	‘ears’

b. Plurals in *-b*:

ayar	>	*ayar-b*	‘needles’
bil	>	*bil-b*	‘bears’

c. Irregular plurals:

iġ	>	*uyəg*	‘men’
nelge	>	*axad*	‘cows’

Zero-pluralization after nasals and voiced stops as seen above is quite frequent in the lexicon, but must be considered irregular since there is no way to derive e.g. [sɛŋ] as a regular result of /sEŋ-m/, and indeed, a few nasal-final words like *anseŋ* ‘coin’ in 17b do show the expected pattern of pluralizing with an overt suffix.

Definite forms are marked with a suffix (attaching after the plural suffix when both are present). The basic form of the definite suffix is determined by the gender and number of the inflecting noun, as in the table below.

	M	F	N
SG	*-š*	*-n*	*-s*
PL	*-iš*	*-iš*	*-id*

The morphology of the definite suffixes is entirely regular and exceptionless, but the particular forms of the suffixes as well as the phonological rules of the language conspire to create a large degree of (predictable) morphophonological variation in definite forms. The definite singular suffixes trigger epenthesis after consonants according to the regular patterns outlined in §2.5 (with the feminine definites patterning differently by virtue of ending in a sonorant), the masculine and neuter suffixes are susceptible to sibilant dissimilation, and the vowel in the definite plural suffixes can suppress epenthesis in plural forms that otherwise require it. A sample of the morphophonological possibilities is shown below:

(19) a. Masculine singular definites:

iġ	>	*iġ-š*	‘the man’
felp	>	*felep-š*	‘the stranger (M)’
xars	>	*xaraš-t*	‘the brother’

b. Feminine singular definites:

unə	>	*unə-n*	‘the woman’
ral	>	*ral-an*	‘the aunt’
relp	>	*relp-en*	‘the stranger (F)’

c. Neuter singular definites:

geda	>	*geda-s*	‘the mango’
šərg	>	*šərəg-s*	‘the bag’
lis	>	*lis-t*	‘the thought, the idea’

d. Masculine and feminine plural definites:

uyəg	>	*uyəg-iš*	‘the men’
unə-m	>	*unə-m-iš*	‘the women’
ral-am	>	*ral-m-iš*	‘the aunts’

e. Neuter plural definites:

nəri-m	>	*nəri-m-id*	‘the oars’
anak-am	>	*anak-m-id*	‘the hammers’
šərg-əm	>	*šərg-əm-id*	‘the bags’

The definite suffix is similar to the corresponding agreement marker used on verbs (§4.3), but differs in the feminine singular as well as the masculine and feminine plural. There is also a resemblance to the agreement suffix used on indefinite adjectives, though it is not as close.

Each noun is finally also marked for case. There are eight cases: nominative, accusative, genitive, lative, inessive, adessive, ablative, and instrumental. The nominative is always unmarked. Case is expressed by a suffix attached after the plural and definite suffixes, commonly including an underspecified vowel. Two suffixes (accusative and adessive) have morphologically specified allomorphy distinguishing between vowel-final and consonant-final roots (i.e. phonologically predictable allomorphy that is nevertheless not attributable to otherwise regular phonological processes). Other case suffixes also have different surface forms in different conditions, but that is in those cases due to regular phonology. The table below shows the basic forms of the case suffixes:

	NOM	ACC	GEN	LAT	INE	ADE	ABL	INS
After a vowel	(zero)	(zero)	*-t*	*-r*	*-sA*	*-wA*	*-ŋA*	*-dE*
After a consonant		*-A*				*-u*		

As accusative is zero-marked after a vowel, nominative and accusative are only distinct after a vowel. However, since plural stems and definite suffixes always end in a consonant, a nominative-accusative merger only happens for some nouns in the singular indefinite form as was shown in the earlie example paradigm of *fana*.

The genitive and lative suffixes exhibit the expected difference in epenthesis patterns stemming from the latter being a sonorant (and thus unable to end a word-final consonant cluster), and the inessive suffix is susceptible to sibilant dissimilation, and the instrumental suffix is susceptible to devoicing.

Adjective inflection

Adjective inflection is in a sense both more complex and less complex than noun inflection. Adjectives agree with their head noun in all its grammatical categories (gender, number, definiteness, and case), and since nouns have an inherent gender but do not inflect for gender, adjectives theoretically have more forms than nouns. However, adjective agreement is very highly syncretic, and the number of distinct surface forms of any adjective is far lower than that of a noun. Adjectives always express case in a similar way to nouns, but the sensitivity of adjectives to the other

categories of gender, number, and definiteness decreases dramatically in cases other than the nominative and accusative. A sample paradigm for the adjective *yəl* 'good' is shown in the table below. Cases after genitive have been left out to save space (they share their pattern of syncretism with the genitive).

yəl 'good'

		Indefinite			Definite
		M	F	N	
Nominative	SG	*yəli*	*yələr*	*yəls*	*yəl*
	PL	*yəlu*	*yəlu*	*yəld*	
Accusative	SG	*yəli*	*yələ*	*yələ*	*yələ*
	PL	*yəlu*	*yəlu*	*yələ*	
Genitive	SG	*yəlt*	*yəlt*	*yəlt*	*yəlt*
	PL	*yəlit*	*yəlit*	*yəlit*	

As is evident from the above table, the inflection of adjectives trades the agglutinative character of noun inflection for a more reduced paradigm where each form may be characterized as consisting of the root followed by only a single fusional suffix. Definite forms of adjectives are the most reduced, systematically only being sensitive to case.

The underlying forms of each morpheme are not evident from a single paradigm. The table below shows the analysis of the nominative and accusative adjective suffixes (other cases are treated further down). Normal phonological processes apply to these suffixes (notably epenthesis and vowel harmony), in addition to the *-A* suffixes having the same allomorphy as the standard accusative suffix, becoming zero after a vowel.

		Indefinite			Definite
		M	F	N	
Nominative	SG	*-i*	*-r*	*-s*	(zero)
	PL	*-u*	*-u*	*-d*	
Accusative	SG	*-i*	*-(A)*	*-(A)*	*-(A)*
	PL	*-u*	*-u*	*-(A)*	

Indefinite forms are maximally distinct in the nominative (M.PL and F.PL are merged, but those are never distinct in any Basheg agreement paradigm). In the accusative, the forms which are represented by consonant suffixes in the nominative (F.SG, N.SG, and N.PL) all merge. Building on the intuition that adjectives can only host a single overt suffix, one could in fact characterize the accusative forms as applications of the regular accusative suffix *-(A)* on top of the nominative forms, with the *-r*, *-s*, and *-d* suffixes being deleted.

In cases other than nominative and accusative, only two agreement forms are used: a basic form for singular indefinites and all definites (e.g. genitive *-t*), and a form for plural indefinites (e.g. genitive *-it*). The default suffixes are completely identical to the corresponding case suffixes for nouns, and the plural indefinites can be characterized as an *-i* followed by the same case suffix, except that the final vowel in suffixes that have it becomes volatile in this form, represented by parentheses in the table below.

	GEN	LAT	INE	ADE	ABL	INS
Default	*-t*	*-r*	*-sA*	*-u, -wA*	*-ŋA*	*-dE*
Plural indefinite	*-it*	*-ir*	*-is(ə)*	*-iw(ə)*	*-iŋ(ə)*	*-id(i)*

The suffixes *-is(ə)*, *-iw(ə)*, *-iŋ(ə)*, and *-id(i)* may optionally lose their final vowels before a prosodic break, but more importantly always lose their final vowels when immediately followed by another

vowel-initial agreement suffix (i.e. via Suffixaufnahme).

3.3 Overview of the noun phrase

Apart from the head noun itself (or pronoun or head adjective), no element in the Basheg noun phrase is mandatory. Demonstratives, possessives, adjectives, numerals, and various other adnominal modifiers are often present for added specificity. The overall word order within noun phrases is somewhat free, and noun phrases sometimes even become discontinous when parts of it are fronted for focus or other reasons (§5.1), but in neutral situations the most common order (at least in my consultant's speech) is (Demonstrative)–Noun–(Adjective)–(Numeral), leaving possession and relative clauses aside for the moment. Examples of typical noun phrases with several of these elements present are shown below, all in the unmarked nominative case.

(20) **salem ŋarad šəŋər**

sale-m *ŋar-ad* *šəŋər*
door-PL big-N.PL two.N

'two big doors'

(21) **ni ṗəlus aband**

ni *ṗəlus* *ab-and*
PROX bowl wood-like

'this wooden bowl'

(22) **nam uyəg yənəp damas**

na-m *uyəg* *yənəp* *damas*
DIST-PL man.PL long four

'those four tall men'

(23) **šiks metax**

šik-s *metax*
knife-DEF sharp

'the sharp knife'

Apart from some possessives, the most common prenominal elements are the demonstratives *ni* 'this' and *na* 'that' shown above, which (uniquely for adnominal modifiers) agree in number with the head noun using the usual nominal plural suffix *-m*. Other modifiers also agree in number, but using other suffixes. When modified by a demonstrative, the noun is not explicitly marked with a definite suffix (though adjectives still take the unmarked definite form). Other elements can also appear in the prenominal position, especially when they have a determining function, or in general fully optionally when modifying an indefinite noun.

(24) **tum kubmid**

tum *kub-m-id*
three.N brush-PL-DEF

'the three brushes (distinguishing against other brushes in smaller or larger groups)'

(25) **ŋara sales**

ŋar-a *sale-s*
big-DEF door-DEF.N

'the big door (distinguishing against a small door)'

(26) **ŋars sale** (= **sale ŋars**)

ŋar-s *sale*
big-N door

'a big door'

There is no obligatory article in Basheg noun phrases, as shown by the indefinite examples above. However, the numeral *fa* 'one' may be on its way to grammaticalizing into an article, as it is sometimes used with an apparently bleached semantic contribution.

(27) **fa šik**

fa *šik*
one.N knife

'one knife, ?a knife'

The Basheg noun phrase is marked for one of eight cases depending on its role or function. Nominative and accusative marks subject and objects (§5.2), among other uses. The genitive marks possessors and some experiencers, and is required after some prepositions, among limited other uses. The other five cases (lative, inessive, adessive, ablative, and instrumental) have a core adverbial function (§5.3), but all also have various extended uses (§5.2, 5.4). This section briefly discusses the marking of case.

Todo: More detail

Case may be thought of as a nominal property marked on nouns, that adjectives (and sometimes other elements) secondarily agree with, but it may be more helpful to think of case as applying to the noun phrase as a whole, with its marking on nouns and adjectives being coequal expressions of this.

(28) a. **šiksə metaxa**

*šik-s-**ə*** *metax-**a***
knife-DEF-ACC sharp-ACC

'the sharp knife (ACC)'

b. **šikist metaxt**

*šik-is-**t*** *metax-**t***
knife-DEF-GEN sharp-GEN

'the sharp knife (GEN)'

Todo: Possessors... It is a mess

3.4 Suffixaufnahme

Basheg nouns in certain cases can directly modify another noun, and in this case, they may agree with their head in the same way as an adjective would. If the modifying noun itself has an adjective, the adjective may receive stacked agreement. Suffixaufnahme is mandatory for prenominal modifiers, and optional for postnominal modifiers.

Prototypical examples of Suffixaufnahme involve possessors in genitive case. An example is shown in 29, where an accusative noun phrase contains a genitive attributive modifier marked both with genitive *-t* and accusative *-a*. The same phrase could equally well be *xarsa Ẹpant* (without stacking) or *Ẹpanta xarsa* (opposite order). With the modifier before the noun, however, stacking (i.e. full agreement) becomes mandatory.

(29) **xarsa Ėpanta**

*xars-**a*** *Ėpan-**t-a***
brother-ACC Ekpan-GEN-ACC

'Ekpan's brother (ACC)'

Other cases can also undergo stacking, as shown in example 30, where an ablative modifier designating origin agrees with its head noun in plural number and instrumental case.

(30) **Gatasŋaidi firibəndi**

*Gatas-**ŋa-idi*** *firibə-n-**di***
Gathas-ABL-PL.INS sword-PL-INS

'[with] swords from Gathas (INS)'

Case stacking may be iterated, as shown in example 31, where *karetsar* modifies *šiŋsər*, which itself modifies *fəxənər*, leading to three case suffixes surfacing on the first noun.

(31) **karetsar šiŋsər fəxənər**

*kare-**t-sa-r*** *šiŋ-**sə-r*** *fəxən-**ər***
father-GEN-INE-LAT house-INE-LAT table-DEF-LAT

'onto a table in father's house'

3.5 Pronouns and demonstratives

Basheg personal pronouns are as shown in the table below. The first-person pronouns *ka* and *ŋa* bear a striking resemblance to the corresponding Tsuy pronouns, which are identical. Since a direct borrowing is unlikely, it is not clear whether this should be interpreted as a coincidence, or some sort of ancient cognate. More work is certainly necessary on the historical development of Basheg. The personal pronouns should be compared with verb-agreement exponents (§4.3), which have a clear, if incomplete, resemblance to the personal pronouns.

	1	1+2	2	3M	3F	3N
SG	*ka*	–	*ye*	*šini*	*rini*	*sini*
PL	*ŋa*	*ŋi*	*yam*	*nənəm*	*nənəm*	*idəm*

Todo: Personal pronoun inflection. Almost regular, but not quite

There are also various interrogative and indefinite pronouns. Interrogative pronouns include *inə* 'what' and *ebe* 'who'. Catch-all indefinite pronouns are formed (as in Tsuy) by a suffix on the corresponding interrogative pronoun: *inək* 'something, anything, nothing' and *ebek* 'someone, anyone, noone'. The negative sense emerges in negative contexts.

The demonstrative determiners *ni* 'this' and *na* 'that' (plurals *nim*, *nam*) were mentioned earlier. Uncommonly for languages of this region, they may also stand on their own as demonstrative pronouns. They both also have derived adjectives: *nit* and *nat*, which are chiefly used with indefinite nouns: *šik nats* 'a knife there, one of those knives'.

3.6 Numerals

Basheg has an idiosyncratic counting system, using numerals expressed in base 14. The single-digit numerals show evidence of earlier counting based on the numbers 5, 7, and 10, surfacing with inconsistent and lexicalized patterns in contemporary Basheg (as produced by my consultant, who insists that all Basheg speakers count in just this way).

There are simplex words for the digits 1: *fa*, 2: *šəŋər*, 3: *tum*, 4: *damas*, 5: *nus*, 7: *sed*, the base 14: *rups*, and two powers of the base: 196: *kamsa* (equal to 14×14), and 2744: *deŋelt* (equal to $14 \times 14 \times 14$). The digit 6 is formed as a lexicalized compound *nu-fə*, corresponding to $5 + 1$. Similarly, 8 is *nu-tum* ($5 + 3$). Three numbers are formed by adding to 7: 10 is *si-tum* ($7 + 3$), 11 is *se-damas* ($7 + 4$), and 12 is *si-nus* ($7 + 5$). The last digit 13 is formed from 14 through the addition of a 'predecessor' prefix: *aŋ-rups* ($14 - 1$). This leaves only 9 unaddressed, and it is the most complex digit: it is formed with the same predecessor prefix, as *aŋ-situm* ($10 - 1$), but since 10 is already complex, the expression for 9 is essentially equivalent to $(7 + 3) - 1$.

Multiples of the base are formed through simple concatenation (*širups*, *tumrups*, ...), though when a digit follows, *rups* alone becomes *rupi*, and higher multiples ending in *-rups* become *-rupə*. The number before a multiple of the base is always formed with the *aŋ-* prefix. Larger numbers follow this basic logic. It should be noted that *kamsa* '196' and *deŋelt* '2744' are susceptible to vowel harmony, raising to *-kəmsə* and *-diŋilt* after high prefixes. The numeral forms are summarized in the following table.

1	*fa*	15	*rupi fa*	29	*širupə fa*	43	*tumrupə fa*
2	*šəŋər*	16	*rupi šəŋər*	30	*širupə šəŋər*	44	*tumrupə šəŋər*
3	*tum*	17	*rupi tum*	31	*širupə tum*		...
4	*damas*	18	*rupi damas*	32	*širupə damas*	56	*damasrups*
5	*nus*	19	*rupi nus*	33	*širupə nus*	70	*nusrups*
6	*nufə*	20	*rupi nufə*	34	*širupə nufə*	84	*nufərups*
7	*sed*	21	*rupi sed*	35	*širupə sed*		...
8	*nutum*	22	*rupi nutum*	36	*širupə nutum*	168	*sinusrups*
9	*aŋsitum*	23	*rupi aŋsitum*	37	*širupə aŋsitum*	182	*rukəmsə*
10	*situm*	24	*rupi situm*	38	*širupə situm*	196	*kamsa*
11	*sedamas*	25	*rupi sedamas*	39	*širupə sedamas*		...
12	*sinus*	26	*rupi sinus*	40	*širupə sinus*	392	*šikəmsə*
13	*aŋrups*	27	*aŋširups*	41	*antumrups*		...
14	*rups*	28	*širups*	42	*tumrups*	2744	*deŋelt*

The numerals above are shown in the counting form, which is identical to the neuter form. A few numerals ('one', 'two', 'three', and 'five') show a separate form when directly modifying a masculine or feminine noun:

	Masculine	Feminine	Neuter
'one'	*xa*	*ra*	*fa*
'two'	*šəŋi*	*šəŋər*	*šəŋər*
'three'	*tu*	*tu*	*tum*
'five'	*nun*	*nun*	*nus*

Todo: Do numerals inflect for case? Probably not?

— 4 —

Verbs

The Basheg verb is agglutinative, being able to host a relatively large number of prefixes and suffixes, though most forms found in practice have no more than three or four affixes.

4.1 Inflection overview

The structure of the finite verb can be well described in terms of a fixed template of morpheme slots. A schematic presentation of this template is shown in the table below. Slots are numbered negative before the root, and positive after the root. One special slot labeled X contains agreement morphemes, and can move around in the template, appearing either as a prefix or as a suffix depending on the presence of certain other affixes. The slot numbered 0' contains derivational morphemes that are often lexicalized and only semi-productive, which is why it has only been partially integrated into the template. Only the root (slot 0) and agreement (slot X) are mandatory in every verb.

Slot	Content	Morphemes
−4	Subordinator	*aṗ*= 'while', *nes*= 'so that', ...
−3	Negation	*d-* (NEG), *b-* (PROH)
−2	Mood	*x-* (OPT), *lEs-* (NEC), *l-* (IMP), *g-* (IMP)
X	*Agreement*	*k-* (1SG), *iš-* (2SG), *s-* (3SG.N), ...
−1	Voice	*t-* (REFL), *ḅ-* (RECP)
0	**Root**	*-əgum-* 'choose', *-Ark-* 'see', *-emen-* 'run' ...
0'	Derivations	(various affixes, lambdization, reduplication)
1	Motion	*-əŋ* (AND), *-əl* (VEN)
2	Causative	*-Ad*, *-Axt*, *-Ans* (CAUS)
3	Passive	*-b* (PASS)
X	*Agreement*	*-Ak* (1SG), *-Eš* (2SG), *-s* (3SG.N), ...
4	Perfective	*-i* (PFV)
5	Mood	*-d* (COND), *-rA* (POT), *-ugə* (PROB), *-ame* (DUB)
6	Evidential	*-ŋA* (VIS), *-gE* (INFR), *-ela* (HSY)

The morphology and semantics of the grammatical categories in the above presentation is described in more detail in later sections.

The root and derivations together with slots 1 and 2 can be considered to form a verb stem, on the basis of which nonfinite forms are created, and certain derivational processes apply. Nonfinite forms are typically formed with a single prefix attached to the verb stem.

4.2 Verb stems

Todo: Derivations to and from

A few adjectives derive into stative verbs through the addition of a *-p* suffix: *iri* 'wise' becomes *-irip* 'to be wise'.

Todo: Lambdization

One interesting example of lambdization used for euphony comes from the expression for 'we (excl.) see'. This would regularly be *agag*, from to root *-ag* 'to see', but my consultant explained that this repetition 'sounds bad', and it is much preferred to instead say *agelg*, using the lambdized root *-elg* 'to see.λ'.

Todo: Andative, venitive

Todo: Causatives

The three different causative suffixes *-Ad*, *-Axt*, and *-Ans* have slightly different selection requirements and semantic force.

4.3 Agreement

The verb agrees with its subject, the nominative argument. This agreement is manifested as either a prefix or a suffix. In general, agreement is a prefix if the verb carries no other prefixes, though some prefixes (chiefly imperative *l-*, *g-* and reflexive *t-*) are compatible with prefixing agreement, and agreement unconditionally becomes a suffix when the verb is preceded by a focused constituent (§5.1). Agreement exponents are shown in the following table (compare with pronouns in §3.5). Prefixes and suffixes differ in that high /i/ in prefixes corresponds to underspecified /E/ in suffixes, the 1SG suffix gains an /A/ vowel, and /g/ in 1PL prefixes becomes /ŋ/.

		1	1+2	2	3M	3F	3N
Prefix	SG	*k-*		*iš-*	*š-*	*r-*	*s-*
	PL	*Ag-*	*ig-*	*in-*	*n-*	*n-*	*id-*
Suffix	SG	*-Ak*		*-Eš*	*-š*	*-r*	*-s*
	PL	*-Aŋ*	*-Eŋ*	*-En*	*-n*	*-n*	*-Ed*

A paradigm with the verb *-AŋArp* 'to leave' is shown below, in affirmative (prefixed) and negative (suffixed) forms:

		1	1+2	2	3M	3F	3N
Plain	SG	*kaŋarp*		*išəŋərp*	*šaŋarp*	*raŋarp*	*saŋarp*
	PL	*agaŋarp*	*igəŋərp*	*inəŋərp*	*naŋarp*	*naŋarp*	*idəŋərp*
NEG	SG	*daŋarpak*		*daŋarpeš*	*daŋarapš*	*daŋarpar*	*daŋaraps*
	PL	*daŋarpaŋ*	*daŋarpeŋ*	*daŋarpen*	*daŋarpan*	*daŋarpan*	*daŋarped*

4.4 Perfective

Despite its complex verb structure, Basheg only has a single tense-aspect suffix, namely the perfective marker *-i*. This form is clearly aspectual in nature, being used for fully realized events situated either in the past or in the future.

Todo: Some Aktionsart interaction probably

4.5 Mood

Basheg has a wide variety of mood markers. Leaving aside imperatives (which are described in the next section), there are two mood prefixes and four mood suffixes employed in various configurations to mark different ways in which events are connected to possible and actual realities.

In general, prefixing moods are the more speaker-attitude-oriented moods. Optative *x-* is used by the speaker to convey a positive or hopeful attitude towards an unrealized event, and necessitive *lEs-* is used for situations that the speaker deems necessary.

The suffixing moods generally relate events to their conditions, potential, or estimated likelihood of occurence. Conditional *d-* marks events whose realization depends on another event. Potential *-rA* marks that an event is possible without taking likelihood into account, and can be used to describe somebody's ability to perform a certain action. Probabilitive *-ugə* describes probable events, and finally dubitative *-ame* describes unlikely events.

Todo: Cease with the listposting and get some examples in here...

Todo: Another day.

4.6 Imperative and prohibitive

There are two prefixes (*l-* and *g-*) labeled as imperatives. The *l-* prefix may be used with first person plural to form hortative expressions ('let's go'...) and with second person forms to form a direct imperative that may be found rude in many circumstances. The other prefix *g-* is used in the second person to form a somewhat more polite imperative, and with third persons to form a sort of shifted imperative ('have him go, make sure that he goes'...).

The imperative prefixes have an idiosyncratic way of combining with agreement. In the second person singular, no overt agreement is expressed at all, uniquely for any verb form. In the first and second person plural, agreement is expressed between the prefix and the verb stem, and in the third person, regular suffixing agreement is used.

The two imperatives are also united by being incapable of combining with regular negation, instead corresponding to a prohibitive form in *b-*, which neutralizes the distinction between *l-* and *g-*, and also necessarily combines with a suffix *-i* (it is unclear whether this should be identified with the perfective suffix). The prohibitive is only possible for second person, never first or third person. The table below illustrates the imperative and prohibitive paradigm of *-ast* 'to take'.

		1	1+2	2	3M	3F	3N
IMP (*l-*)	SG	–		*l-ast*	–	–	–
	PL	(*l-ag-ast*)	*l-ig-ast*	*l-in-ast*	–	–	–
IMP (*g-*)	SG	–		*g-ast*	*g-ast-aš*	*g-ast-ar*	*g-ast-as*
	PL	–	–	*g-in-ast*	*g-ast-an*	*g-ast-an*	*g-ast-ed*
PROH	SG	–		*b-ast-i*	–	–	–
	PL	–	–	*b-in-ast-i*	–	–	–

First person plural hortatives are almost always inclusive of the addressee, but my consultant nevertheless finds *lagast* and *bagast* to be well-formed expressions, which is why they have been included in the table in parentheses.

4.7 Evidentiality

Todo: Talk about this, I guess. Evidentiality is somewhat optional.

4.8 Nonfinite forms

Todo: Yeah, I don't know. Going to improvise on this for now.

— 5 —

Clauses

Basheg clauses are characterized by a word order determined primarily on pragmatic factors, and heavy pro-drop tendencies. Schematically, finite clauses are structured as (Topic)-(Adv)-(Focus)-Verb-(Remainder). Only some adverbs go in the preverbal position. Everything except the verb is optional.

Todo: Actually, there are probably nonverbal clauses too. A tale for another day.

5.1 Topic and focus

Topics, when present, are expressed at the beginning of a clause, and always in the unmarked nominative case. Focused constituents are expressed immediately before the verb, retaining their expected case marking, and triggering suffixing agreement on verbs when present. Other constituents are expressed after the verb. The following example is a pragmatically neutral clause part of a longer narrative about the witch (the topic).

(32) **Dəxrən rəgumiela ni laŋaška.**

dəxr-ən	*r-əgum-i-ela*	*ni*	*laŋaš-k-a*
witch-DEF	3SG.F-choose-PFV-HSY	PROX	peak-DIM-ACC

'[They say that] the witch chose this hilltop.'

The effect of focalization can be seen clearly in a pair of examples like the following. The first example is neutral, with a topical subject. In the second, the object *fədələ* has been focused, forcing verb agreement to be expressed in a suffixing manner.

(33) a. **Igš sinŋə fədələ.**

ig-š	*s-in-ŋə*	*fədəl-ə*
man-DEF	3SG.M-eat-VIS	apple.λ-ACC

'[I saw that] the man was eating an apple.'

b. **Igš fədələ inisŋə.**

ig-š	*fədəl-ə*	*in-is-ŋə*
man-DEF	apple.λ-ACC	eat-3SG.M-VIS

'[I saw that] it was an *apple* that the man was eating.'

Even though topics are always nominative, they may fill the role of an object, as in the following example, where there are two nominative arguments. The verb still agrees with the subject, and not the (nominative, topicalized) object.

(34) **Meteg narpigi xayan falemiš.**

Meteg	*n-arp-i-gi*	*xayan*	*fale-m-iš*
Meteg	3PL-see-PFV-INFR	yesterday	boy-PL-DEF

'As for Meteg, the boys apparently saw him yesterday.'

Any constituent, including adverbial phrases, may be topicalized.

When there is neither an overt topic nor a focused element in a clause, leaving the verb clause-initial, a verb focus reading is possible.

Focalization can apply to subparts of a noun phrase, leading to discontinuous noun phrase. This is particularly common with adjectives, as in the following example. The adjective coheres with its head noun in case marking, despite being linearly separated from it. My consultant appears to find topicalized constituents, and especially discontinous ones, especially pleasing to the ear in their lambdized form, though he notes that the plain version remains possible.

(35) a. **Ŋelta imisk kubə.**

ŋelt-a *imis-k* *kub-ə*
better.λ-ACC show-1SG brush-ACC

'I will show [him] a *better* brush.'

5.2 Valency and arguments

Despite theoretically allowing any argument in a clause to be dropped, Basheg imposes fairly strict interpretation rules on valency, and unexpressed arguments must receive a definite interpretation from context. The passive voice is commonly used to demote unknown or unspecific arguments.

(36) a. **Šəgumi kə.**

š-əgum-i *kə*
3SG.M-choose-PFV 1SG.ACC

'{He/*someone} chose me.'

b. **Kəgumbi.**

k-əgum-b-i
1SG-choose-PASS-PFV

'I was chosen {by him/by her/by someone}.'

Arguments are typically expressed in nominative case (subjects) and accusative case (objects), though ditransitive verbs add a lative indirect object, and some verbs have a lative experiencer with no accusative argument.

(37) **Gedamid kiḅi Eḍanar.**

geda-m-id *k-iḅ-i* *Eḍan-ar*
mango-PL-DEF 1SG-give-PFV Ekpan-LAT

'I gave the mangoes to Ekpan.'

(38) **Kar idilp ŋərum.**

ka-r *id-ilp* *ŋəru-m*
1SG-LAT 3PL.N-be_good cat-PL

'I like cats.'

5.3 Adverbials

Basheg adverbs are found in all structural positions of the clause: they can be topicalized and focalized, and they can appear postverbally. There is also a fourth adverb position located linearly between the topic and focus positions that is capable of hosting a particular subset of adverbs.

Basheg has several underived adverbs, such as time expressions (*xayan* ‘yesterday’, *gi* ‘today’, *aṗe* ‘tomorrow’, *sələ* ‘now’), and spatial expressions (*idəxə* ‘close by’, *əmp* ‘far away’). Nouns or noun phrases inflected in spatial cases (lative, inessive, adessive, ablative) or instrumental are also commonly used in an adverbial function.

(39) **Kəli Sələtsə eŋetersa.**

k-əl-i	*Sələ-t-sə*	*eŋeter-sa*
1SG-come-PFV	Sala-GEN-INE	ship-INE

‘I came on Sala’s ship.’

The ‘phase polarity’ adverbs *xili* ‘already’ and *bašt* ‘still’ deserve special mention for their relation to negation. They are most naturally positioned in the preverbal adverb position, but it can be shown that clause-level negation nonetheless scopes over these adverbs. Consider the following four examples. If preverbal adverbs had scope over the verb, the semantics of the final two examples would be expected to correspond in the opposite way, but this is not what is observed.

(40) a. **Xili nurəg.**

xili	*n-urəg*
already	3PL-sleep

‘They are already sleeping.’

b. **Bašt nurəg.**

bašt	*n-urəg*
still	3PL-sleep

‘They are still sleeping.’

c. **Xili durəgən.**

xili	*d-urəg-ən*
already	NEG-sleep-3PL

‘They are not yet sleeping.’

d. **Bašt durəgən.**

bašt	*d-urəg-ən*
still	NEG-sleep-3PL

‘They are no longer sleeping.’

5.4 Nonverbal predication

Todo: Predicate adjectives agreement as usual

Todo: Copula verb? Several?

Todo: Comparatives

5.5 Complex clauses

(41) **Šagəŋame məgusir.**

š-ag-əŋ-ame	*m = əg-usir*
3SG.M-say-AND-DUB	that = 1PL-need

‘He might say that we need [them].’

(42) **Ilxišidəŋ sələ, farunuyəld igumu dəxrən.**

ilx = iš-id-əŋ	*sələ*	*fa = r-unuy-əl-d*	*igum-u*	*dəxr-ən*
if = 2SG-go-AND	now	then = 3SG.F-change-VEN-COND	flatbread-ADE	witch-DEF

‘If you go there now, the witch will turn you into flatbread.’