

Mum grammar

John Mansfield
3 April 2024

General matters

This document is a work-in-progress, documenting my understanding of the Mum language, Papua New Guinea (glottocode: kqa). I may make copies of this available, not because it represents high-quality language documentation, but since it may be better to share even rough notes since very little other documentation on this language is available.

This document is based on about 5 days of elicitation, transcription and translation of Mum language. The vast majority of this was with Samuel Ambos, who is an outstanding language consultant. This allowed us to cover quite a lot of ground in 5 days. Samuel is about 40 years old, is from Kimbugor language, and I believe that he has significant standing as a local representative. Samuel has a good knowledge of Mum, e.g. he can provide translations of words or sentences from English (or Tok Pisin) without much hesitation. He can also tell stories in Mum. At the same time, Samuel defers to other older people as the true experts. In Samuel's everyday language use with his peers and younger kin, there is a lot of code-switching into Tok Pisin. I was able to hear this for myself, and they also told me this. Samuel's younger kin in their teens, 20s and 30s can also speak Mum fairly well.

The name 'Mum' means 'what'. This is also the case for some nearby languages, such as Anamuxra (Ingram 2001).

The word *yasi* '(older) brother' has a special regional significance. People from the Mum area are known for addressing each other as *yasi*, and indeed soon after meeting Samuel we began addressing each other as *yasi*. It appears to be a kind of locality marker. In the same way, people from Madang town area are known for addressing each other as *mam*, which I was told means 'father' in the Bel language of Madang. Male staff at Binatang Research Centre, just outside of Madang, do indeed often address each other as *mam*, though neither speaker or addressee has Bel language heritage.

Relations to other languages and dialects

Mum has been identified as a member of the Sogeram language family by Daniels. His work on reconstructing Sogeram phonology includes some information on Mum, which has analyses as a sister of Sirva, a neighbouring language to the south.

The most systematic documentary material available on Mum is a list of almost 400 words recorded by Hans (John) Z'graggen, alongside several neighbouring languages in his South Adelbert wordlist collection. Note that Z'graggen identifies the Mum wordlist as 'Katiati', named after the village in which he recorded it. Z'graggen's typical practice seems to have been to use village names as language identifiers. The speakers I worked with from Kimbugor village were very clear about calling their language 'Mum', and they also identify this unambiguously as the language of Katiati village, and several other villages in the area.

There is evidently a certain amount of dialectal variation between Mum speaking villages. In wordlist elicitation, Samuel sometimes told me a few variants, mentioning villages where they might be used. For example:

‘road’	<i>knd</i>	Kimbugor
	<i>kumb</i>	Amjaivivu
‘knife’	<i>tuki</i>	Kimbugor
	<i>mai</i>	Amjaivivu
‘language’	<i>vana</i>	Kimbugor
	<i>kuyu</i>	Katiati
‘today’	<i>nmati</i>	Kimbugor
	<i>nmatira</i>	Amjaivivu
	<i>amuskira</i>	Usimbugor

Comparing words elicited from Samuel (Kimbugor village) against Z’graggen’s 1980 wordlist (Katiati village) also reveals substantial differences. I have compared about 100 words from these lists, most are either the same form, or related form with sound change. A smaller group consists of those where the forms are unrelated. In the related-form group, the phonological difference almost always involves vowel reduction or deletion in the Kimbugor forms, compared to the Katiati forms. Among the unrelated forms, some of these pairs may well not be dialectal variation at all, but merely different responses to an English prompt that allows more than one interpretation or translation. But this seems unlikely to account for all of them, especially when we consider very basic vocabulary such as ‘ear’. The three tables below illustrate examples of each group.

Same form in Kimbugor and Katiati (identical or almost-identical)

	Katiati	Kimbugor
woman	<i>nawundi</i>	<i>nawundi</i>
head	<i>miku</i>	<i>miku</i>
sugar cane	<i>aga</i>	<i>aga</i>
taro	<i>sis</i>	<i>sis</i>
yam	<i>nyambara</i>	<i>nyambara</i>
tobacco	<i>asara</i>	<i>asara</i>
sago	<i>maβa</i>	<i>maβa</i>
coconut.tree	<i>koima</i>	<i>kuima</i>

Related forms with sound changes

	Katiati	Kimbugor
2PL.NOM	<i>narɥ</i>	<i>nar</i>
3PL.NOM	<i>nurɥ</i>	<i>nur</i>
man	<i>kuru</i>	<i>kru</i>
mouth	<i>mɥka</i>	<i>mka</i> [<i>mɥka</i>]
tooth	<i>mɥka gurs</i>	<i>mka grs</i> [<i>mɥkagrɥs</i>]
child	<i>ningi</i>	<i>nyingi</i>
brain	<i>mekwi</i>	<i>nymik</i>
neck	<i>punggu</i>	<i>pungg</i>
hair	<i>mɥnɥ</i>	<i>mn</i>
left arm	<i>kakag</i>	<i>kakai</i>
navel	<i>sumbirp</i>	<i>smbirp</i>

liver	<i>mavunggrəsu</i>	<i>mavnggrs</i>
sweat	<i>pumbə</i>	<i>pmb</i>
walk	<i>kund-</i>	<i>knd-</i>
sit down	<i>mənsə-</i>	<i>mnj-</i>
sleep	<i>kar-</i>	<i>karg-</i>

Unrelated forms

	Katiati	Kimbugor
face, forehead	<i>kpumka</i>	<i>mirmba</i>
ear	<i>kəngsigi</i>	<i>kuv</i>
upper leg	<i>tanola</i>	<i>punggna</i>
belly	<i>kəpə</i>	<i>mav</i>
heart	<i>kusingə</i>	<i>kutm</i>

Phonology

Vowel inventory

i	(i, u)	u
	a	

The central vowels are highly frequent, but as I will argue below, should be considered ‘intrusive vowels’, not part of the phonological structure of the word.

Consonant inventory

p	t		k	k ^w
β	s		ɣ	ɣ ^w
^m b	ⁿ d	^ɲ ɟ	^ŋ g	^ŋ g ^w
m	n	ɲ	ŋ	ŋ ^w
	r			
w		j		

I generally have the impression of a single rhotic with flap/trill realisation. But worth checking the word *aru* ‘old (person)’, since the recordings I have of this seem to be glides.

Minimal word

There are many monosyllabic words in Mum (including the name of the language). These mostly have a coda e.g. *taβ* ‘house’. There are fewer monosyllabic words with a single open syllable CV, but there are some, e.g. *su* ‘shit’, and personal pronouns e.g. *ji* ‘1SG’, *na* ‘2SG’. These are all pronounced with phonetically long vowels. Open syllables are more common in multisyllabic words, e.g. *tapa* ‘fence’.

There are also monosyllabic words with no onset, e.g. *uβ* ‘smoke’. The description below of intrusive vowels will further enrich the analysis of word minimality

Word edges

Word-initial: No /β, ɣ, ŋ, w/, and /r/ is rare or absent (only *rm* ‘?PURP’)

I’m not sure whether /w/ occurs in syllable onsets at all; there are various possible cases, but hard to distinguish here between /β, w/

Word-final: No /p/, and /ɣ/ is rare or absent (only *tuʷgiɣ* ‘waterhole’, which needs a better recording to confirm)

Syllable structure

Maximal CCVCC, but complex codas and onsets are rare;

Phonetic offglides in [au], [ai] are more parsimoniously treated as consonantly glide codas, e.g. *tawsi* ‘he shot it’, *taw* ‘snake’, *kajyura* ‘wallaby’, *kakaj* ‘left’. The alternative would be to treat them as complex nuclei (e.g. *taʷsi*), but notably no examples have been identified of such a purported complex nucleus with an additional consonant in the coda. This absence is explained if we propose that offglides fill the coda position.

Complex codas are rare, and the non-final consonant is always /r/:

si^mbirp ‘belly button’, *-ma-rŋ* ‘DIST.PST-1PL’, *na-rŋ* ‘eat-1PL’.

Complex onsets are also rare, and the second consonant is /r/:

kru ‘man’, *ʷgris* ‘seed’, *sram* [sra:m] ‘many (children, dogs)’,

(*psa* ‘skin’ is almost cluster-like, but has a shortish intrusive vowel)

Resonant nuclei

Resonants quite frequently serve as syllable nuclei, though sometimes with slight accompanying vocalic excrescence.

The liquid /r/ is often the nucleus, e.g. *a^mbrka* ‘flying fox’, *mukr* ‘possum’, *kr-ma-n* ‘stay-FUT-1SG’.

Also nasals, e.g. *sigm* ‘pig’, *mka* [m̥ka] ‘mouth’, *agutm* ‘small stick’, *nmati* [nm̥.a.ti] ‘today’ (Katiati dialect), *nmik* [n̥m̥.i:k] ‘brain’ *kna* [k̥n̥.a] ‘poison’, *k-n* [k̥n̥] ‘stay-1SG.CURR’, *k-na* [k̥n̥.a] ‘stay-2sg.curr’,

Also some prenasalised stop break across syllables, e.g. *kⁿdi* [k̥n̥.di] ‘road’, *t^mba* [t̥m̥.ba] ‘stone’, *s^mbirp* [s̥m̥.birp] ‘belly button’

And one with a fricative, *ms-ta* [m̥s̥.ta] ‘stay-ss’.

There is often brief (around 30ms) vocalic articulation in the transition between preceding obstruent and these resonant nuclei. But this is not always present.

Also sequences like [s̥], e.g. in *sp* [s̥p^{hə}] ‘unripe’ and *sgm* [s̥ɣm̥] ‘pig’. Kind of sesquisyllabic articulation, not clear if there is a distinct initial syllable. See example of two different articulations of *puknya spe* (2024-02-11_Samuel-Ambos_05).

One could instead consider this to be the syllable nucleus, though it would be a rather flimsy one.

Note that *ʷgris* ‘seed’ can be analysed as having a complex onset, as presented here, or as a resonant nucleus, i.e. *ʷgrs*. The vowel-like portion is very brief. Since there are other words that have complex onsets (*kru* ‘man’, *psa* ‘skin’) and again others with resonant nuclei (*mukr* ‘possum’), either analysis seems reasonable for *ʷgr(i)s* ‘seed’.

A word can lack any vowels, having only a nuclear /r/, e.g. *mr* ‘tree sap’.

There is also an unusual word, *prmr* ‘small brown and black bird sp.’, which has two syllables (intensity and pitch peaks), [pr.mɾ].

Syllable boundary clusters

These are rather rare

yvragi ‘(s)he arrived’, *mir^mba* ‘face’, *tar-maⁿd* ‘(s)he will shoot’, and many others involving /ɾ/ in either c1 or c2 position;

also *jimjik* ‘brain’, *tam^uga* [tamga] ‘eye’, *kupsun* ‘morning’, *kptiti* ‘afternoon’, *sukta^mbu* ‘short’, *pukna* ‘mango’, *jibni* ‘cane’

There are also several arising from non-final coda /j/, e.g. *kujma* ‘coconut tree’, *miⁿajkru-* ‘push’. These may have slight intrusive vowels, perhaps only in careful speech; but I analyse the lexical phonology of these words as including sequences of consonant segments /jm, jk/, with variable phonetic articulation.

Also in some reduplicative words *piⁿdviⁿd* ‘bird sp.’, *mirmir* ‘lungs’, *manman* ‘?lack’

atⁿda ‘a long time ago’ has a phonological sequence of two consonant segments, but this is realised phonetically with vowel epenthesis, i.e. [atinda].

Word-final consonant articulation

Word-final voiceless stops have an aspirated release, e.g. *sip* [sip^h] ‘mucus’, *kut* [kut^h] ‘back’, *sijat* ‘bird sp’, *muk* ‘egg’, *jimjik* ‘brain’. In phrase medial positions this is also followed by an intrusive vowel, e.g. *kut kambu* [kut^hi kambu] ‘back down’.

There is also a word *uk^w* ‘sore’ with an interesting citation form. Phrase-medially this has just a small intrusive vowel *uk^w nanga* [uk^unanga], but in citation form it has a much stronger intrusive vowel, like [uk^hu], with progressive rounding, longer and louder than a typical intrusive vowel. But I guess this is an effect of hyper-articulation in the citation form.

The prenasalised stops are like the voiceless ones in that they demand some kind of release. This results in a following short intrusive vowel when they are in word-final position, though these are shorter than other types of intrusive vowels (30-40ms; more on this below). For example *pi^mb* [pimb^o] ‘sweat’, *kaⁿd* [kand^o] ‘hand’, *tiβi^ug* [tiβiŋg^o] ‘shoulder’, *na^ug^w* [na^uŋg^u] ‘blood’, *na^ug^w* [na^uŋg^u] ‘back of neck’; or followed by another consonant, e.g. *piⁿdviⁿd* [pind^ovind^o] ‘bird sp.’

Plain nasals also end words with a flourish. They may either end with a slight stop articulation, e.g. *sram* [sram^b] ‘many’, or occasionally a slight intrusive vowel, e.g. *ki-n* [kinⁱ] ‘stay-1SG’, though this is less frequent than the prenasalised stop release.

Prenasalised stops

The nasal feature of prenasalised stops is generally articulated in any non-initial position; but in absolute initial position there tends to be little or no prenasal realisation, i.e. *na^mbuyu* ‘they’ being realised as [ndambuyu] phrase-medially versus [dambuyu] phrase-initially. The palatal prenasalised stop is somewhat rare, so far only identified in the words *pu^wy* ‘bone’, *ka^wy* ‘??’, and in most forms of the verb ‘sit’, e.g. *miⁿye* ‘sit.3SG.C’

Rounded dorsals

As shown in examples above, the rounded dorsal obstruents /k^w, ŋg^w, ŋ^w, ɣ^w/ are distinguished by rounding they add to a following intrusive vowel, and an offglide on a preceding low

vowel. Rounded dorsals seem fairly rare. A stop example is *tak^w* ‘moon’. A prenasalised example is *na^{ŋg^w}* ‘back of neck’; a nasal example is *kiŋ^w* ‘pandanus’. The only fricative example thus far only attested is *jaɣ^{w-i}* [jaɣwi] ‘go.up-3SG’.

Fricatives

The peripheral fricatives /β, ɣ/ tend to be articulated very lightly, but otherwise show no notable allophony. The coronal /s/ is more strongly and more variably articulated, showing degrees of palatalisation, and also occasionally some affrication, e.g. *suru* [tsuru] ‘leak’, *na^{ŋg^w} siki* [na^uŋɟiki] ‘blood rope = vein’.

The fricatives are the only obstruent type that does not demand final release, e.g. *maβ* ‘belly’, *kuβ* ‘ear’, *ŋgris* ‘seed’.

Contrasts between non-low vowels

The two peripheral non-low vowels, /i, u/, cannot be consistently explained as conditioned allophones of /i/. There are some instances where these occur without the relevant palatal or labial consonants, for /u/ including *nu* ‘3SG.NOM’, *nur*, ‘3PL.ACC’, *u-ta* ‘go-SS’, *kuyu* ‘language’, and for /i/ including *kitr* ‘black’, *mir* ‘tongue’, *ŋipu* ‘ash’, *ipu* ‘small’.

Nonetheless, it is worth noting that the majority of [u] vowels appear adjacent to a labial or dorsal (potentially rounded-dorsal) consonant, while many [i] vowels are adjacent to a palatal (especially if we consider /s/ to be phonologically a palatal, see Foley 2018). Perhaps there was vertical vowel system at some point in the history of the language.

But there are also words where something phonetically like [i, u] occurs, but it appears to have a more centralised articulation, or variable articulation, plus a relevant contextual consonant, suggesting that it should be treated as an allophone of /i/. For example:

mi^{ŋga}- [muŋga ~ minga] ‘go down’, *jiβ*- [jiv ~ jiβ ~ juβ] ‘hit’

Sometimes morphological alternations suggest that a round vowel is dependent on a rounded velar consonant. For example

Intrusive central vowels

Mum has two quite different types of vowels. ‘Full’ vowels, /a, i, u/ are part of the lexical representation of a word. ‘Intrusive’ vowels (which we could also call epenthetic (Blevins & Pawley 2010)) are not contrastive or informative parts of a word’s phonology, but rather vocalic gestures that arise from prosodic patterns of articulation. Full vowels are more consistently produced (modulo morpho-phonological alternations), are phonetically longer, and usually more peripheral. Intrusive vowels are inconsistently produced, have brief duration, and are central unrounded unless a neighbouring consonant induces a distinctive colouring. I will generally annotate intrusive vowels as [i] (following other linguistic documentation from the region), though for strongly rounded examples I annotate [u], and I indicate particularly brief examples in a superscript position as [i^u]. However it should be kept in mind that these are phonetic details rather than phonological contrasts.

Mum epenthetic vowels are overall fairly similar to those found in nearby Anamuxra (Ingram 2001: 43–45). This is interesting because Mum is not closely related to Anamuxra; meanwhile the Sogeram languages most closely related to Mum, such as Sirva, have a fully phonemic central vowel, rather than epenthesis (Daniels 2015: 664).

[Summarise main principles]

Below I will describe the patterns of vowel epenthesis found in different word shapes, in with respect to different consonant types. I begin with the most phonologically minimal words and work up to those with two full vowels. The following should not be read as strict rules. Intrusive vowels by their nature are variable, for example noticeably distinct forms have been recorded for p^mb^nd [pɪmbɪnd ~ pumbɪndi] ‘hot’. Nonetheless, some fairly clear patterns are discernable.

CV: There is no vowel intrusion in words that consist of a single open syllable with a lexical vowel. The lexical vowel is phonetically long, fulfilling the bimoraic minimum, e.g. *su* [su:] ‘shit’, *na* [na:] ‘you’. The same applies with a complex onset, *kru* [kru:] ‘man’.

VC: There are a couple of words that have a coda but no onset, and these do have brief final intrusive vowels in citation form: *uβ* [uβi] ‘smoke’, *uk* [uk^{hu}] ‘sore’. This may be related to the general pattern of onsetless syllables lacking prosodic prominence (see §sect).

CC: There are quite a few words with two consonants and no lexical vowels, and these generally have one or two intrusive vowels. If the second consonant is nasal then a single intervening intrusion occurs: *mn* [mɪn] ‘hair’, *wn* [wun] ‘wind’, *km* [kim] ‘bow’. This suggests that the intrusive vowel and the coda nasal each contribute a mora. However the dorsal nasal again shows an exception here, with an additional word-final intrusion: *ɲɲ* [ɲɪɲi] ‘dirty’.

If the second consonant is a rhotic, we find a long trilled rhotic filling the syllable rhyme, *mr* [mr:] ‘swallow.3SG.CU’, *mr* [mr:] ‘sap’. This fits the pattern noted above of rhotics as syllable nuclei.

If the second consonant is an obstruent, we get a distinct pattern with two intrusive vowels: *mk* [miki] ‘leaf’, *ɣt* [ɣuti] ‘top’, *sp* [sipu] ‘unripe’, *p^{mb}* [pimbi] ‘sweat’, *knd* [kindi] ‘last born male child’, *m^y* [mɪɲɪ] ‘sit.IMP’, *s^g* [sɪŋgi] ‘cooking pot’, *kβ* [kuβi] ‘night’, *tɣ* [tɰɣu] ‘bottom part’. This could be interpreted to suggest that intrusive vowels contribute at maximum a single mora; they do not permit the bimoraic lengthening of full vowels. Thus in the absence of a moraic coda, two intrusive vowels are required.

There is also one closed monosyllabic word with a complex onset, and it has a single intrusive vowel: *krs* [kris:] ‘seed’. The lack of a final intrusive vowel is likely because of the final /s/, which tends to be long and is never followed by an intrusive vowel.

CVC: Like CV words, these have no intrusive vowels. A second consideration is whether these words have phonetic vowel lengthening, which could indicate whether the coda consonant contributes a mora or not. My inspection of spectrograms is inconclusive on this point, and if anything there may be a gradient of vowel lengths depending on coda type. With the most sonorant codas nasal and rhotic codas the vowel seems to be of standard length: *man* [man] ‘banana’, *kiŋ^w* [kiŋ^w] ‘pandanus’, *mir* [mir] ‘tongue’. This is also the case with a complex onset, *sram* [sram] ‘many (children, dogs)’. The vowel is also of standard length with a sibilant coda, which itself is notably long: *kas* [kas:] ‘sand’, *βis* [βis:] ‘unripe (banana)’, *sis* [sis:] ‘taro’.

With a dorsal fricative or stop in the coda, the single vowel appears to be of intermediate length and I cannot easily assign it to the type of lengthened vowel described above for CV words, and the standard length of other word types: *taβ* [taβ] ‘house’, *maβ* [maβ] ‘belly’, *kuβ* [kuβ] ‘ear’, *kut* [kuth] ‘back’. But a couple of words with sibilant onsets appear to have standard length vowels: *sip* [sip^h] ‘mucus’, *sik* [sik^h] ‘fire’. Where the coda is a voiceless stop, citation form has a strongly aspirated final release, though this does not show any voicing. Where the coda is a prenasalised stop, the pattern is similar to other

obstruent codas, except that there is also a slight release, which here is voiced and therefore resembles a very weak, brief vowel. However these are shorter and weaker than standard intrusive vowels: *kaⁿd* [kaⁿdi] ‘right (hand)’, *pu^wj* [pu^wɲi] ‘bone’, *pu^g* [pu^gɲi] ‘neck’, *ja^g* [ja^gɲi] ‘second born female child’, *na^gg^w* [na^gɲg^wu] ‘back of neck’. These form an interesting contrast with the equivalent CC words that lack a lexical vowel (see above), which instead form two clear syllables with a standard intrusive vowel in each, e.g. *p^mb* [pimbi] ‘sweat’, versus *pu^wj* [pu^wɲi] ‘bone’. Like the trade-off between coda consonant types and nucleus vowel length, these suggest a balancing of phonological weight: if the first vowel is intrusive then a second one is favoured to

Finally, one outlier has also been observed in the CVC category: *paŋ* [paŋi] ‘net bag’, was pronounced unexpectedly with a final intrusive vowel. This may be an unusual property of the dorsal nasal (see also below).

C.CV: While we saw above that a complex onset is possible in *kru* ‘man’, other consonant sequences induce an intrusive vowel to produce a disyllabic word: *kβa* [kiβa] ‘garden’, *kpa* [kipa] ‘chest’, *psa* [pi^sa] ‘skin’, *mka* [m^kka] ‘mouth’. The lexical vowel here is not lengthened, suggesting that the intrusive vowel contributes a mora, and indeed forms a syllable. Some words have a brief rounded vowel following a dorsal onset in the first syllable, which I interpret as a rounded dorsal consonant with coarticulation on the intrusive vowel: *k^wra* [k^ura] ‘forest’, *k^wran* [k^uran] ‘body’, *k^wsi* [k^usi] ‘spear’. These contrast with the greater duration of a full lexical vowel, as in *kuku* ‘water’.

The initial syllable of this word type can also be filled by a resonant nucleus, as in *t^mba* [tɪmba] ‘stone’.

VCC: There is just two identified word of this type, and both have a resonant to fill the second syllable: *akⁿd* [a.kɪdⁱ] ‘salty food’, *u-rŋ* [u.rɲ] ‘go.1PL.CURR’.

CCC: Triple-consonant words, as expected, are interpolated by two intrusive vowels: *sym* [siɣim] ‘pig’, *kpt* [k^hɪpɪt^h] ‘full (belly)’; or can have a resonant nucleus in the first syllable: *kŋr* [kɲri] ‘clean’, *p^mbⁿd* [pɪmbɪnd ~ p^umbɪnd^o] ‘hot’.

CVCC: An intrusive vowel interpolates the last two consonants: *muyβ* [muɣiβ] ‘the next day’, *kutm* [kutim] ‘liver’, *kink* [kinik] ‘star’, *mukr* [mukri] ‘white’

CCVC : An intrusive vowel interpolates the first two consonants: *mtup* [mɪtup] ‘wet’, *jka* [jⁱk:] ‘loincloth’; or there is a resonant nucleus *s^mbirp* [sɪmbi:rp] ‘belly button’.

Longer words have no need for intrusive vowels to contribute mora, though in some instances they may still have intrusive vowels to enable syllabification of certain sequences. There follow some examples without any intrusive vowels:

VCCC: *u-s-rŋ* [us.rɲ] ‘go.1PL.YEST’

CVCV: *tapa* [tapa] ‘fence’, *maŋa* [maŋa] ‘dirt’, *ama* [ama] ‘woman’s breast’, *kuku* [kuku] ‘water’, *siki* [siki] ‘rope’, *kiβa* [kiβa] ‘claw’, *kuma* [kuma] ‘hand’, *tuki* ‘knife’, *ina* [ina] ‘sun’.

VCVCC: *u-ma-rŋ* [umarɲ] ‘go.1PL.CURR’

CVCCVCV: *kajyura* [kajyura] ‘wallaby’

And here are examples of longer word with intrusive vowels to syllabify the consonants. As shown in the third example, there is no restriction against having several consecutive intrusive vowels in a word

CCCV: *pkɲa* [p^u.kⁱ.ɲa] ‘mango’
 CVCCV: *pu^ugna* [puŋ.gi.na] ‘thigh’
 CCCC: *mk^mbsa* [mikⁱmbⁱsa] ‘lip’
 CVCCVC: *nu^mbwum* [numb^uwum] ‘left’

Comparison to Sirva

As mentioned above Mum’s sister language Sirva is analysed as having a fully phonemic central vowel, rather than epenthesis (Daniels 2015: 664). Daniels provides two types of evidence for this, but I do not believe that either applies in the case of Mum.

Firstly, Daniels analyses Sirva as having diphthongs. This includes some beginning with the central vowel, /ii, iu/, which he points out can not be motivated as epenthetic in this environment. But as mentioned above, for Mum I analyse diphthong-like sounds as simple vowels combining with glide consonants. In the case of [ii], for Mum I instead represent this as [ij], which is already assumed structures of Mum phonology, as in: *kj* [kijⁱ] ‘stay.CUR.3SG’.

Daniels also mentions some Sirva minimal pairs showing central vowels. Some of these would in Mum instead be analysed as a minimal contrast between a phonemic vowel and the lack of any phonemic vowel. For example:

Sirva

kadi ‘tree sp.’

kidi ‘platform’ (Daniels 2015: 664)

Mum

kiβa [kiβa] ‘claw’

kβa [kiβa] ‘claw’

Finally, Daniels also provides a Sirva minimal pair that contrast a central vowel with the lack of any vowel. This certainly does appear to constitute good evidence for the phonemic status of /i/ in Sirva. But I have not observed any contrasts of this type in Mum. Indeed, in this post-voiced-stop environment, Mum words appear to have free variation between presence or absence of a short intrusive vowel.

Sirva

kid ‘rheum’

kidi ‘knife’ (Daniels 2015: 664)

Mum

kaⁿd [kandⁱ ~ kand] ‘right (hand)’,

It would be interesting to know what the Mum reflexes might be for the Sirva minimal pair *kid*, *kidi*. However I have not recorded any Mum word for ‘rheum’, and for ‘knife’, Samuel gave two words, neither related to the Sirva: *tuki* (Kimbugor dial.), *maj* (Amjaivivu dial.).

One of the historical developments that has been identified for Mum is the weakening of some full vowels to the central vowel, e.g. Proto North-Central Sogeram (PNCS) **manka* ‘egg’ > [miŋga] (Katiati dial.) (Daniels 2015: 95). This may have overall increased the prevalence of phonetic central vowels in Mum, which may perhaps have contributed to them becoming lexically uninformative and thus epenthetic. The cited form is also interesting because the Kimbugor dialect has undergone a different pattern of vowel weakening compared to the Katiati dialect, with the final vowel deleted altogether, as well as rounding of the first vowel and obstruent devoicing: *muk* [muk^h] ‘egg’. Another Mum innovation is the word-final loss of the central vowel from PNCS (Daniels 2010: 179), after an earlier loss of final resonants. Daniels describes this as ‘sporadic’, for example the loss in *maniŋ* > *mani* > *man* ‘banana’, versus the retention in *kintir* > *kindi* ‘root’ (Katiati dial.). But note that this variation follows one of the patterns I presented above for central vowel intrusion, where

final nasals do not induce a following intrusive vowel, e.g. *man* [man] ‘banana’, *mn* [min] ‘hair’; but final voiced obstruents do induce vowel intrusion, e.g. *kⁿd* [kindi] ‘last born male child’.

Stress

Here

Morpho-phonological alternations

Initial voiceless stops often become prenasalised (voiced) stops in the second position of a noun compound. E.g. *kuku* ‘water’, *kujma* ^u*guku* ‘coconut water’, *anak* ^u*guku* ‘salt water’, *iyu* ^u*guku* ‘rain water’;

kurja ‘dry’, *kujma* ^u*gurja* ‘dry coconut’.

It can also occur with /s/, where there is also palatalisation (for reasons unknown), e.g. in *siki* ‘rope’, *jiβni* ^u*jiki* ‘cane rope’.

Sometimes the alternation is to a fricative: *psa* ‘skin’, *ayu* *βsa* ‘tree skin = bark’.

The noun *krs* ‘seed’ occurs in several compounds, showing alternation to both prenasalised and fricative manners:

maβ ^u*grs* ‘stomach seed = heart’, *tanda* ^u*grs* ‘leg seed = ankle’

puj *γrs* ‘breadfruit seed’, *majⁿda* *γrs* ‘betelnut seed’ *ukama* *γrs* ‘daka vine seed’.

Nouns

Here

Personal pronouns

	Nominative	Accusative	Genitive
1sg	yi	yang	yand
2sg	na	nang	nand
3sg	nu	nung	
1pl	ara	arang	
2pl	nar	narng	
3pl	nur, dambugu	nurng, dambugung	

Not clear whether *dambugu* etc should be treated as personal pronoun or a kind of demonstrative.

Complex nouns and noun phrases

Combinations of Adj and N have some flexibility of ordering; for example both *kuku pmb* ‘water hot’ and *pmb kuku* ‘hot water’ were judged acceptable. Note however that there is also a closely related verbalised structure, which then has a fixed order, *kuku pmb-nd* ‘water hot-VBLZ’.

N-Adj is attested more frequently, e.g. *mangg ipu* ‘valley little’, *puknya sepe* ‘mango unripe’, *agu kitr* ‘tree black’.

But the opposite is also found e.g. Adj-N in *kurnya meke* ‘dry leaf’.

Seemingly more lexicalised combinations appear to take the N-Adj order, e.g. *kuyma nggurnya* ‘coconut dry’ (note that this example also has a morpho-phonological voicing process at the boundary).

There are also many nominal compounds like *agu mav* ‘tree belly = tree trunk’, or *keva tapa* ‘garden fence’. These are semantically right-headed: the right element names the general type of object, and the left element expresses more specific properties or associations.

When combined with an adjective, the compounding elements remain contiguous, as in *agu mav kitr* ‘tree belly black = black tree-trunk’.

Complex nominal expressions are rare in the small amount of spontaneous speech collected so far. Both demonstratives and adjectives are attested spontaneously combining with nouns, but this occurs in only a minority of nominal expressions. Most spontaneous complex expressions have a noun and one modifier, either adjective or demonstrative. The only more complex expression attested so far is the double-adjectival *kuku meni ipu* ‘water cold few = a few cold drinks’ (2024-02-08_Samuel-Ambos_02). But most nominal expressions are either a simple noun, or simple demonstrative.

Elicitation of more complex nominal expressions produces the following orderings:

N-Adj

agu kitr ‘tree black’

N-Dem

agu nemdi ‘tree this’

N-Dem-Adj

agu nemdi kitr ‘tree this black’

N-Num

asa arkita ‘fish two’

N-Adj-Num

agu kitr arkita ‘tree black two’

Modification with both number and demonstrative caused some consternation, which may indicate that this construction is marginal or absent in the language. Eventually my consultant decided that either order is possible:

N-Num-Dem

agu arkita nembegu

~

N-Dem-Num

agu nembegu arkita

I also requested a translation for a four-word complex expression, which caused further consternation. He eventually came up with the following order, but it did not appear to come naturally to him, and I note that it is not fully consistent with the N-Adj-Num expression above:

N-Num-Dem-Adj

agu arkita nembegu kitr ‘tree two this black’

Genitive expressions

Peter Muni=ndu pugu
[name] =GEN village
'Peter Muni's village' (2024-02-08_Samuel-Ambos_02)

Kin terms

Siblings

	Kin: M, Older	Kin: M, Younger	Kin: F, Older	Kin: F, Younger
Ground: M	ya-si	ya-ra	ya-rma nungu-rma	ya-rma-ndak
Ground: F	ya-si	ya-ra-tak	ya-si-gat nugu-si-gat	ya-ra-gat

Verbs

Here

Tense

Current is used for events happening now, or that have already happened but today. Samuel calls it 'current'.

Yesterday is used just for yesterday.

Distant past is used for any time before yesterday.

There is also a 'Distant past 2', i.e. -s-m forms, which Samuel says has same meaning, and same temporal extension, as Distant past. Needs more checking.

Habitual is used for things that extend over a person's whole life, including people who are now deceased so the activity is past. But habitual is not used in some uses where it may have been expected, perhaps due to start/end points. E.g. describing someone going to school somewhere, or working somewhere, does not seem to provoke habitual, but instead uses the Current tense.

Verb suffix paradigms

Here are some of the more complete paradigms elicited
Underscore indicates notable irregularities

	CURRENT	YESTERDAY	DISTANT PAST
1SG	<i>tar-in</i>	<i>tara-s-n</i>	<i>tara-m-in</i>
2SG	<i>tara-na</i>	<i>tara-s-na</i>	<i>tara-ma-na</i>
3SG	<i>tar-i</i>	<i>tara-s-r</i>	<i>tara-m-i</i>
1PL	<i>tara-rng</i>	<i>tara-s-rng</i>	<i>tara-ma-rng</i>
2PL	<i>tara-ra</i>	<i>tara-s-ra</i>	<i>tara-ma-ra</i>
3PL	<i>tara-yu</i>	<i>tara-s-yu</i>	<i>tara-m-u</i>
	FUTURE	HABITUAL	? Distant past 2
1SG	<i>ta-rma-n</i>	<i>tara-nd-in</i>	<i>tara-s-m-in</i>
2SG	<i>ta-rma-na</i>	<i>tara-nda-na</i>	<i>tara-s-ma-na</i>
3SG	<i>ta-rmand</i>	<i>tara-nd-i</i>	
1PL	<i>ta-rm-dang</i>	<i>tara-nda-rng</i>	
2PL	<i>ta-rmand-ra</i>	<i>tara-nda-ra</i>	
3PL	<i>ta-rmand-yu</i>	<i>tara-nd-u</i>	

Table. Verb paradigm, *tara-* ‘shoot, spear’.

	CURRENT	YESTERDAY	DISTANT PAST
1SG	<i>yvrág-in</i>	<i>yvrága-s-n</i>	<i>yvrága-m-in</i>
2SG	<i>yvrága-na</i>	<i>yvrága-s-na</i>	<i>yvrága-ma-na</i>
3SG	<i>yvrág-i</i>	<i>yvrága-s-r</i>	<i>yvrága-m-i</i>
1PL	<i>yvrága-rng</i>	<i>yvrága-s-rng</i>	<i>yvrága-ma-rng</i>
2PL	<i>yvrága-ra</i>	<i>yvrága-s-ra</i>	<i>yvrága-ma-ra</i>
3PL	<i>yvrága-yu</i>	<i>yvrága-s-yu</i>	<i>yvrága-m-u</i>
	FUTURE	HABITUAL	
1SG	<i>yvrág-rma-n</i>		
2SG	<i>yvrág-rma-na</i>		
3SG	<i>yvrág-rmand</i>		
1PL	<i>yvrág-rm-dang</i>		
2PL	<i>yvrág-rmand-ra</i>		
3PL	<i>yvrág-rmand-yu</i>		

Table. Verb paradigm, *yvrága-* ‘arrive, come’.

	CURRENT	YESTERDAY	DISTANT PAST
1SG	<u>u-n</u>	<i>u-s-n</i>	<i>u-m-in</i>
2SG	<i>u-na</i>	<i>u-s-na</i>	<i>u-ma-na</i>
3SG	<u>u-y</u>	<i>u-s-r</i>	<i>u-m-i</i>
1PL	<i>u-rng</i>	<i>u-s-rng</i>	<i>u-ma-rng</i>
2PL	<i>u-ra</i>	<i>u-s-ra</i>	<i>u-ma-ra</i>
3PL	<i>u-yu</i>	<i>u-s-yu</i>	<i>u-m-u</i>
	FUTURE	HABITUAL	IMPERATIVE
1SG	<i>u-rma-n</i>	<i>u-nd-in</i>	-
2SG	<i>u-rma-na</i>	<i>u-nda-na</i>	<u>u-gu</u>
3SG	<i>u-rmand</i>	<i>u-nd-i</i>	<i>u-m</i>
1PL	<u>u-m-dang</u>	<i>u-nda-rng</i>	<i>u-m</i>
2PL	<i>u-rmand-ra</i>	<i>u-nda-ra</i>	<i>u-ma-ra</i>
3PL	<i>u-rmand-yu</i>	<i>u-nd-u</i>	<i>u-m-u</i>
	MEDIAL DS		
1SG			
2SG	<i>u-ga-na</i>		
3SG			
1PL			
2PL			
3PL			
SS:			

Table. Verb paradigm, *u-* ‘go’.

	CURRENT	YESTERDAY	DISTANT PAST
1SG	<u>pa-n</u>		
2SG			
3SG	<i>pa-y</i>		
1PL			
2PL			
3PL			
	FUTURE	HABITUAL	IMPERATIVE
1SG	pa-rman		
2SG			<u>pay-ng</u>
3SG			
1PL			
2PL			
3PL			
ss: <i>pa-ta</i>			

Table. Verb paradigm, *pay-* ‘come’.

Sentences

Basic clause structure

Many utterances consist of just a verb, and a few consist of just a noun phrase. In each case some other part of the meaning is implied. Examples of both types can be seen in the following three consecutive utterances:

(X) a. *Megiyar Daka Market*
(at) Megiyar Daka Market

Stop-ta
stop-SS

(We) stopped.

Mainda ipu ukama ipu
betelnut little daka little
(We got) a little betelnut, a little daka.

Asara kanda indu mnga-ta
tobacco there ?DEM get-ss
(We) got tobacco there.

U-ma-rng
go-PST-1PL
We kept going.

Lines (b,d) have verbs but these are SS forms (§sect), which provide no information about the identity of the subject. The final line (e) has a finite verb that reveals the subject to be 1PL. Line (c) has only a coordinate noun phrase, and leave the type of event (getting) to be deduced, though the following line soon reveals this if it could not be guessed already. In fact the only one of these five lines that combines a noun phrase and verb is line (d), where ‘getting tobacco’ is fully expressed.

Preverbal and postverbal domains

The verb-final position in line (Xd) is a fairly consistent pattern, as in many Papuan languages that verb-final. Agents, themes, recipients, beneficiaries

kar pnggr-m-i
‘he pulled the car’
speid mngata
‘he got a spade’
kndi kvt-m-i
‘the road became dark’
draiva ambamin
‘I told the driver’
areng guramata
‘he helped us’

The main exception to verb finality, i.e. the main type of post-verbal element, is locative and temporal expressions. This is another trait shared with surrounding languages, at least in the Sogeram family (REF). There are many more examples beyond the ones shown here, and postverbal position appears to be highly consistent for locative (and temporal) expressions.

yvraga-ta pa-ta kt gaing
arrive-SS come-SS top DEM
‘We arrived at the top.’
mngga-ma-rng Rapak kuku gaing
descend-PST-1PL name river DEM
‘We went down to Rapak river there.’

ina ipu ntumrami taim gaing
A little sun shone at that time

kwnggra-g-u kupsung
cook-DS-3PL morning
'They cooked in the morning.'

The locative pattern can also result in a post-verbal personal pronoun, if the persons role is as a locative reference. Note that this construction uses the dative case:

kete gain yuvuragata nur-ng
we arrived at the top, to those people

Another post-verbal element is the word *garn*, which is as yet poorly understood but from Samuel's description ('when you saw it, you use this to make the story clear'), appears to have an evidential function.

kanda gaing k-y garen
there DEM stay-3SG.CUR ?EVID
'There it stayed.'

Ordering of preverbal NPs

None of the examples above have more than a single NP before the verb, and this is no accident as most clauses have at most one NP, or one core NP plus a locative NP. Thus far just a handful of spontaneous examples have been identified with two non-locative (and thus preverbal) NPs. Although these are not enough to make any strong proposals, it is notable that these examples have personal pronouns in initial position, despite those not being in agentive roles, and not being the verbal subjects.

yang sgm paymi

yang sgm tvata paigi
'The pig came towards me. The pig came running towards me.'

nur kandaing tmbagrs ipu guta
'we gave a little money to them'

There are a few further examples of multi-NP clauses from elicitation. However I am hesitant to read too much into these as elicitation prompts may effect the ordering. Verb-finality remains clear in these examples, and agents come before patients (A-P). However while verb-finality appears to be a robust pattern in the language, A-P ordering could well be primed by the English prompt, or alternatively could in fact reflect a tendency to put personal pronouns before other NPs. Further spontaneous speech is required on this point.

ndambgw samba tar-mand-yu
they will shoot a pig.'

Clause chains

Same subject chains use a non-finite verb with *-ta* for medial clauses.

Different-subject chains use a finite verb with *-g(a)* for medial clauses. This in itself is the main strategy, certainly covering sequential events, and also some events that appear to be simultaneous, e.g.:

ngga-gi-n smbra mnga-ta k-y

I'm watching him working (2024-02-08_Samuel-Ambos_02)

There are also instances of reduplication in the medial verb, and this appears to be used for some types of simultaneity. This needs more research.

yang mnga-ganda-ta ky-gi-gi maynda mengin
He's holding me up while I pick betelnuts

nung mnga-ganda-ta ky-gi-gin maynda mngi
I'm holding him up while he picks betelnuts

Appendices

Partial paradigms for selected verbs

I collate paradigms here for verbs where I have at least a few different forms. Parenthetic forms are for different-subject in medial verbs; DSB = different subject base Underscore indicates notable irregularities

	CURRENT	FUTURE	IMPERATIVE
1SG	<i>mnj-n</i>	<i>mnj-rma-n</i>	-
2SG	<i>mnj-na</i>	<i>mnj-rma-na</i>	<u><i>mnj</i></u>
3SG	<u><i>mnj</i></u>	<i>mnj-rmand</i>	<i>mnj-m</i>
1PL			<i>mnj-m</i>
2PL			<i>mnj-ma-ra</i>
3PL			<i>mnj-m-u</i>

ss: *ms-ta*

Table. Partial paradigm, *mnje*- 'sit'.

	CURRENT	MEDIAL DS
1SG	<i>yv-in</i>	<i>yvu-g-in</i>
2SG		<i>yvu-ga-na</i>
3SG	<i>yv-i</i>	<i>yvu-g-i</i>
1PL		<i>yvu-ga-rng</i>
2PL		<i>yvu-ga-ra</i>
3PL		<i>yvu-g-u</i>

ss: *yvu-ta*

Table. Partial paradigm, *yevu*- 'hit'.

	CURRENT	DISTANT PAST	MEDIAL DS
1SG	<i>mng-in</i>		
2SG			
3SG			<i>mnga-g-i</i>

ss: *mnga-ta*

Table. Partial paradigm, *menga*- 'hold, take'.

	CURRENT	DIST PAST	FUTURE	IMPERATIVE
1SG	<i>karg-in</i>	<i>karga-m-in</i>	<i>karg-rma-n</i>	-
2SG	<i>karga-na</i>		<i>karg-rma-na</i>	<i>karga</i>
3SG	<i>karg-i</i>		<i>karg-rmand</i>	<i>karga-m</i>
1PL				<i>karg-m</i>
2PL				<i>karga-ma-ra</i>
3PL				<i>karga-m-u</i>
MEDIAL DS				
1SG	<i>karga-g-in</i>			
2SG	<i>karga-ga-na</i>			
3SG				
1PL				
2PL				
3PL				

Table. Partial paradigm, *karga-* ‘sleep’.

	CURRENT	DISTANT PAST	FUTURE	MEDIAL DS
1SG	<i>"g-in</i>	<i>"ga-m-in</i>	<i>"ge-rma-n</i>	<i>"ga-g-in</i>
2SG	<i>"ga-na</i>		<i>"ge-rma-na</i>	
3SG	<i>"g-i</i>		<i>"ge-rmand</i>	<i>"ga-g-i</i>
ss: <i>"ga-ta</i>				

Table. Partial paradigm, *"ga-* ‘see’.

	CURRENT	DISTANT PAST
1SG		<i>nyndar-m-in</i>
2SG		
3SG	<i>nyndar</i>	
ss: <i>nyndar-ta</i>		

Table. Partial paradigm, *nyndar-* ‘hear’. see 2024-02-11_Samuel-Ambos_02

	CURRENT	FUTURE	DISTANT PAST
1SG	<i>k-n</i>	<i>k-rma-n</i>	
2SG	<i>k-na</i>	<i>k-rma-na</i>	
3SG	<i>k-y</i>	<i>k-rmand</i>	<i>ky-m-i</i>
IMPERATIVE		MEDIAL DS	
1SG			
2SG	<i>k-nyi</i>		
3SG		<i>ky-gi</i>	

Table. Partial paradigm, *ky-* ‘stay’.

	CURRENT	FUTURE	DISTANT PAST
1SG	<i>kwm-in</i>	<i>kwm-rma-n</i>	
2SG	<i>kwmu-na</i>	<i>kwm-rma-na</i>	
3SG	<i>kwm-i</i>	<i>kwm-rmand</i>	<i>kwmu-m-i</i>

Table. Partial paradigm, *kwmu-* ‘die’.

	CURRENT	FUTURE	DISTANT PAST
1SG			
2SG			
3SG	<i>yagw-i</i>		
3PL			<i>yaga-m-u</i>

ss: *yaga-ta*

Table. Partial paradigm, *yag-* ‘go up’.

	CURRENT	FUTURE	DISTANT PAST
1SG			
2SG			
3SG	<i>munggw-i</i>		
1PL			<i>munggu-ma-reng</i>

ss: *munggu-ta*

Table. Partial paradigm, *munggw-* ‘go down’.

	CURRENT	FUTURE	DISTANT PAST
1SG		<i>m-rma-n</i>	
2SG			
3SG	<i>?mr (?mr-ga)</i>	<i>m-rmand</i>	<i>mr-m-i</i>
	YESTERDAY	MEDIAL DS	
1SG			
2SG			
3SG	<i>mr-s-r</i>		

Table. Partial paradigm, *mr-* ‘swallow’. See 2024-02-11_Samuel-Ambos_02.

	CURRENT	FUTURE	DISTANT PAST
1SG	<i>irg-in</i>		
2SG			
3SG	<i>irg-i</i>		
	ss: <i>irga-ta</i>		

Table. Partial paradigm, *irga-* ‘cry (out)’.