The Majang Language

Joswig, A.

Part II: Phonological Inventory and Orthography

This section introduces the inventories of the Majang phonology – vowels, consonants and suprasegmentals. It also displays the rules which govern the use of the various elements of these inventories.

In this section, data presented without any kind of bracketing is understood to be in its surface-phonemic representation, transcribed according to the phoneme inventories shown in tables 1 and 3. Only when further phonetic detail needs to be shown for greater clarity, phonetic data is supplemented in square brackets […]. On the other hand, in some places information is given about the individual morphemes making up a word. These are always shown in their underlying representation, and the brackets {…} are used to surround such lexical units. This is different from the other parts of this language description, where the underlying representation is shown without brackets.

Individual sounds are shown by /**x**/ to be phonemic, by [**x**] to be phonetic and by <**x**> to be in orthographic representation.

II.1 Vowels

This section first assesses the vowel inventory, then shows proof of contrast for the various identified phonemes, and finally looks into distributional restrictions.

II.1.1 Phonemic inventory

The vowel inventory of Majang is similar to that of the Southeast-Surmic languages, which according to Moges (2008, p. 260) have seven contrastive basic vowels, with a height contrast affecting the mid vowels .

	Front	Central	Back
High	**i**		**u**
High-Mid	**e**		**o**
Low-Mid	**ɛ**		**ɔ**
Low		**a**	

Table 1: vowel phonemes

Bender differs from Unseth and Moges on the Majang vowel inventory. He (1983, p. 114) felt only confident to present six vowels; he noticed the seventh vowel /**ɔ**/, but did not see enough contrastive evidence to confirm its phonemic status. Unseth (1991, 2007) proposed nine vowels, and Moges (2008) even ten vowels for Majang. The chart presented above is the same as the one proposed in Joswig (2012) and Getachew (2014).

In a 2879-word text sample, /**a**/ is the most frequent vowel with 1336 occurrences, closely followed by /**ɛ**/ with 1237 occurrences. /**i**/ appears 867 times, /**o**/ 506 times, and /**ɔ**/ 401 times. The least-frequent vowels are /**e**/ with 347 occurrences and /**u**/ with only 262 occurrences. This means that the vowel /**a**/ occurs five times more often than the vowel /**u**/.

II.1.2 Contrasts and phonetic realizations

Bender (1983, p. 114) was not able to prove contrast between /**o**/ and /**ɔ**/. This was merely due to lack of data, as such contrast does exist in the language: **mòːréŋ** *'he boils'* vs. **mɔ̀ːrɛ́ŋ** *'it is shriveled'* is an incontestable minimal pair between the two vowels. Such a contrast would also be expected for reasons of symmetry.

Example II.1: vowel contrasts[6]

i vs. **a**	**tĭmŋ**	*he wounds*	**tìjɛ́ŋ**	*he hears*
	tămŋ	*it drips*	**tàjɛ́ŋ**	*he opens*
i vs. **ɛ**	**kúrí**	*tree, sp.*	**pàːríŋ**	*he tries*
	kúrɛ́	*hunting net*	**pàrɛ́ŋ**	*he chops*
i vs. **e**	**kóndì**	*fish trap*	**tíːmân**	*cloud*
	kóndé	*bottle*	**tèmâːn**	*firewood*

[6] In this and all other following contrast charts, the words are presented in their citation form. This means that verbs are accompanied by the *SFT*-marker **=ŋ**.

i vs. **ɔ**	**íré**^{L}	*footprint*	**cɔ́:bí**	*hoof*
	ɔ̀rɛ́:^{L}	*ancestor*	**ɔ̀bɔ́**	*knot*
i vs. **o**	**ílán**^{L}	*udder*	**tìjɛ́ŋ**	*he hears*
	ó:lán	*ability*	**tò:jɛ́ŋ**	*he pierces*
i vs. **u**	**kírí**^{L}	*thread*	**tì:ɟɛ́ŋ**	*he takes revenge*
	kúrí	*tree, sp.*	**tù:ɟɛ́ŋ**	*he roasts*
e vs. **a**	**émɗ**^{L}	*canoe*	**jímé**^{L}	*cemetery*
	ámɗ^{L}	*abdomen*	**jímá**^{L}	*back*
e vs. **ɛ**	**ɟé:mɛ́**^{L}	*before*	**ɗéjŋ**	*cooking stone*
	ɟɛ́:mɛ́^{L}	*tree, sp.*	**ɗɛ̀jíŋ**	*he desires*
e vs. **ɔ**	**rémɛ́**^{L}	*duty*	**ké:r**	*courtyard*
	rɔ́mɛ́^{L}	*proverb*	**kɔ́:r**	*middle*
e vs. **o**	**cè:d**	*here*	**ké:r**	*courtyard*
	cò:d	*there*	**kó:r**	*ditch*
e vs. **u**	**té:l**	*lake*	**ɗé:k**	*under*
	tù:l	*five*	**ɗúk**^{L}	*forest*
ɛ vs. **a**	**mɛ̀lɛ́ŋ**	*he arrives*	**ɓɛ̌rŋ**	*he crushes*
	màlɛ́ŋ	*he strikes*	**ɓǎrŋ**	*he forbids*
ɛ vs. **ɔ**	**wɛ́:r**	*storm*	**ɗépɛ́**^{L}	*lion*
	wɔ́:r	*feast*	**ɗɛ̀pɔ́**	*entrance hall*
ɛ vs. **o**	**ɛ́lt**^{L}	*grassland*	**kɛ̌ŋ**	*he pounds*
	ólt^{L}	*fish*	**kǒŋ**	*he gathers*
ɛ vs. **u**	**kòrɛ́ŋ**	*he peels*	**kúrɛ́**	*hunting net*
	kòrúŋ	*he closes*	**kùrù**	*foam*
a vs. **ɔ**	**bàdɛ́ŋ**	*it disappears*	**pà:kɛ́ŋ**	*it is hot*
	bɔ̀dɛ́ŋ	*he escapes*	**pɔ̀:kɛ́ŋ**	*he tears*
a vs. **o**	**tà:jɛ́ŋ**	*he harvests*	**kàwɛ́ŋ**	*he bites*
	tò:jɛ́ŋ	*he pierces*	**kòwɛ́ŋ**	*it is sour*
a vs. **u**	**máɲá**^{L}	*sister*	**tâ:r**	*meat*
	múɲá	*earthworm*	**tùr**	*garbage dump*
ɔ vs. **o**	**kɔ́:r**	*middle*	**pɔ̀:cɛ́ŋ**	*he praises*
	kó:r	*ditch*	**pò:cɛ́ŋ**	*he polishes*

ɔ vs. **u**	**bɔ́lɔ́ŋúrkúŋ**	*he is old*	**ɗɛ̀pɔ́**	*entrance hall*
	bùlúnkùr	*bubble*	**tàɗápú**^{L}	*ashes*
o vs. **u**	**gò:mòj**	*trap*	**tò:r**	*smoke*
	gó:múj^{L}	*tree, sp.*	**tùr**	*garbage dump*

There is significant phonetic variation for most of the vowels. The individual vowels show an enormous bandwidth with respect to their formants, which is caused by various phonological factors, such as syllable structure, position in the word, and consonants in the syllable. Therefore it is easily possible to hear all kinds of vowels in the raw phonetic data. This obviously prompted Unseth and Moges to claim the row of *[-ATR]* high vowels /**ɪ, ʊ**/ for Majang.

Current speakers of Majang, however, do not perceive any difference between [ɪ] and [**i**], or between [ʊ] and [**u**], respectively. These phones represent the same sound unit to them. Words like **dìldílŋ** *'he is fat'* or **gìrgìɗí:ŋ** *'he rolls'* certainly have instances of [ɪ] in them (the latter only in the first syllable). They can be explained as the positional variant of /**i**/, when it happens to be short and preceding a lateral or flap consonant. More conspicuous, however, are cases where the unsuspecting ear perceives an [**e**] – these could be analyzed as instances of /**ɪ**/. The two sounds are difficult to distinguish in many languages with nine- or ten-vowel ATR-contrast systems (Casali, 2008, p. 509). I have indeed heard words like **pé:t** *'girl'* or **ké:r** *'courtyard'* with different vowel qualities from different speakers. Some of these pronunciations prompted me to transcribe them as [**pɪ́:t**] and [**kɪ́:r**]. When pressed hard for these differences, the speakers did not hear them. They agreed that the difference, if there was any, is of no consequence at all to the meaning of the word, and all variations are perfectly acceptable. The same is true for words like **ɲón** *'lie'* or **ɗó:** *'land',* where there is a similar potential confusion between [**o**] and [**ʊ**]. Some phonetic variation can be heard, but it does not matter at all to the speakers.

None of these considerations present any difficulty for a nine-vowel system analysis. Two phonetically identical vowels could still be variants of two different phonemes, which are neutralized in their phonetic realization. Vowel pairs such as /**o**/ and /**ʊ**/ or /**e**/ and /**ɪ**/ easily lend themselves to such a scenario. But the underlying difference needs to show elsewhere in the language, through an ATR-harmony in the suffixes following the stem. There is no such harmony in Majang. Therefore, with no solid phonetic evidence, and no data pointing to an ATR-based vowel-harmony system, it needs to be concluded that Majang has no ATR opposition among high vowels.

The tenth vowel of Moges' description, which he transcribed as [**ʌ**], does not exist in Majang as an independent sound unit either. There is, however, an enormous bandwidth of /**a**/-like vowels phonetically. The first formant *(F1)* can be as low as 600 Hz and as high as 900 Hz with the same speaker, and the second formant *(F2)* can be between 1250 and 1650 Hz. Again, this considerable variation can be explained by environmental factors. A sonorant consonant following the vowel in the same syllable has a lowering effect on the first formant of the phoneme. A long /**aː**/ has a higher chance of being pronounced with a high *F1*. Palatal glides or nasals in the same syllable have a tendency of raising the second formant. Short /**ɛ**/ is often pronounced as [**ə**], that is as a central unrounded mid vowel.

To illustrate the seven vowels of Majang, Figure 3 below (taken from Joswig 2012, p. 273) gives an indication of the first two formants of each vowel. To make things comparable, the chart represents only the vowels of monosyllabic nouns and verbs (3rd person singular) of the language (a total of 96 vowel tokens). Each symbol in figure 3 stands for one token. The vertical dimension displays the frequency of *F1* in Hertz, and the horizontal dimension shows the frequency of *F2*.

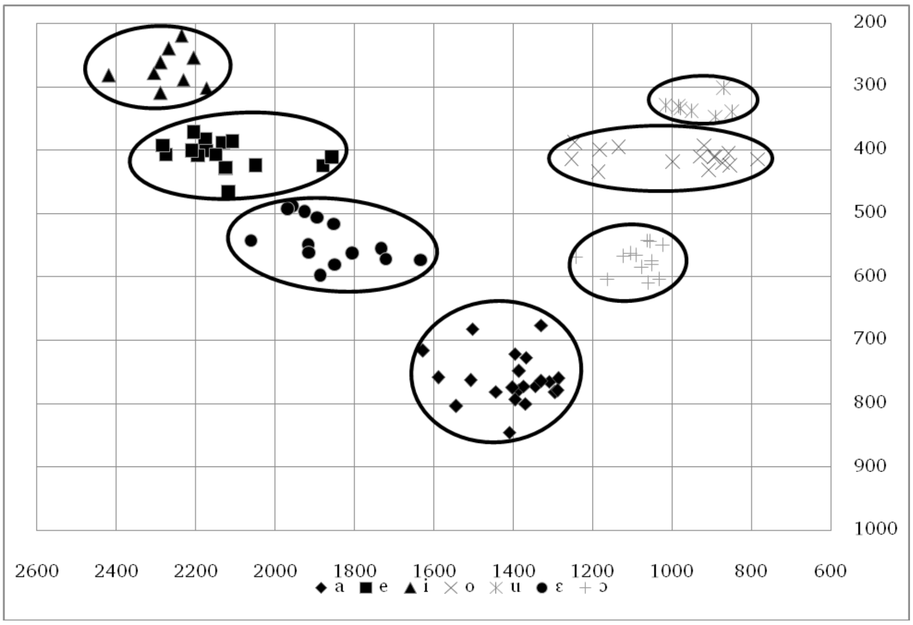

Figure 3: vowel formants of monosyllabic words

In such a controlled environment, the seven vowels show up as surprisingly discrete groupings. Except for the vowel /**a**/, there is no significant variance with respect to *F1*. Casali (2008, p. 507f) states that *"[…] all else equal, a [+ATR] vowel will have lower F1 than its [-ATR] counterpart. […] F1 has consistently been found to be a very robust cue for distinguishing [+ATR] vowels from their [-ATR] counterparts […]."* Therefore, it would be very difficult to interpret a nine- or ten-vowel system from the picture presented here. If there are only seven vowel phonemes in monosyllabic words, there is no reason to expect a larger number of vowel phonemes in polysyllabic words, even if the vowels in these polysyllabic words show a more significant phonetic variation regarding *F1* and *F2*.

II.1.3 Vowel length

Vowel length plays an important role in Majang. This is typologically quite different from the Southeast-Surmic languages, which seem to make use of vowel length only sparingly; in both Mursi and Suri, the few long vowels are attributed to the loss of intervocalic consonants (Bryant, 2013, p. 28f; Mütze, 2014, p. 39f). The distinction between long and short vowels cannot be analyzed in these terms in Majang. It is frequently used and lexically contrastive, as can be easily confirmed by numerous minimal pairs:

Example II.2: short and long vowels

a) **ùtéŋ**	**{ùt-ɛ́=ŋ}**	*he drinks*	vs.	**ù:téŋ**	**{ù:t-ɛ́=ŋ}**	*it rusts*
b) **òlán**	**{òlán}**	*husband*	vs.	**ó:lán^{L}**	**{ó:l-an^{L}}**	*be able (INF)*
c) **gɔ̀gɔ̀j**	**{gɔ̀gɔ̀j}**	*he crawls*	vs.	**gɔ̀:gɔ̀j**	**{gɔ̀:gɔ̀j}**	*ford*
d) **tàjá:ŋ**	**{tàj-á=ŋ}**	*I open*	vs.	**tà:já:ŋ**	**{tà:j-á=ŋ}**	*I harvest*
e) **kɛ̀tɛ́ŋ**	**{kɛ̀t-ɛ́=ŋ}**	*he cuts*	vs.	**kɛ̀:tɛ́ŋ**	**{kɛ̀:t-ɛ́=ŋ}**	*he scatters*
f) **mènéŋ**	**{mèn-ɛ́=ŋ}**	*he twists*	vs.	**mè:néŋ**	**{mè:n-ɛ́=ŋ}**	*he steers*
g) **ìn**	**{ìn}**	HORT	*vs.*	**ì:n**	**{ì:n}**	*you*

In phonetic terms, the length difference can be quite pronounced. The following two charts show the words of example II.2a) with the duration of their first vowel measured. The vowel in question is shaded.

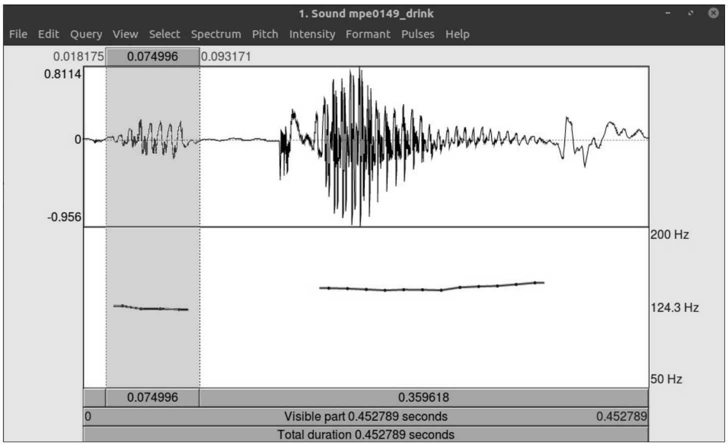

Figure 4: duration of short vowel in **ùtéŋ** *'he drinks'*

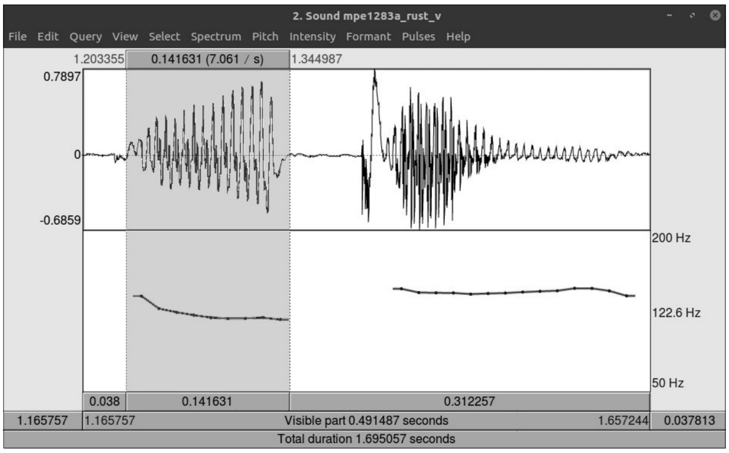

Figure 5: duration of long vowel in **ù:téŋ** *'it rusts'*

In these recordings, which were made on separate occasions, the duration of the long vowel (about 140 ms) is almost twice as long as the duration of the short vowel (about 75 ms). This evidence might tempt the observer to the conclusion that vowel length is a clear-cut phonological feature that speakers of Majang can always safely discern. But this is not the case. Already the data in example II.2 shows that lexical minimal pairs based on vowel length can only be found involving word-initial syllables. The situation with non-initial syllables gives rise to the suspicion that vowel length may be less distinctive beyond the first syllable, except possibly as a device for grammatical distinctions (see examples II.3-II.5). Although for example II.2 a) all speakers would readily agree that the first vowel in *'drink'* is short, and in *'rust'* it is long, there are some suffixes where the evidence is less clear. For example, regarding the *1s* suffixes **-á** and **-à**, my consultants sometimes pronounced them long, sometimes short. In the same way other suffixes, usually those ending in a vowel, had conflicting length evaluations. A good example of that is the impersonal suffix **-ɛ́ː**, for which the consultants insisted on a long vowel for most words. In some cases, however, they would always tell me to transcribe it as a short vowel, as in **rìjɛ́**^{L} *'he is called'*. In other cases where the consultants insist on a short vowel, it can clearly be seen that the vowel goes back to a long vowel underlyingly. This is the case with the short version of the *TF.3s* suffix **-gêːɗ**, which is **-gé**^{L} according to the perception of the speakers. Similar evidence can be found for the infinitive markers **-ɛ́ː**^{L} and **-íː**^{L}, which again are truncated versions of a sequence of long vowel plus /ɗ/ (see section IV.2.2.1 for more details). Indeed the phonetic evidence is inconclusive, as the duration of these vowels is somewhere in the span between 80 and 110 ms, as in the following illustration of the word **tònúrgé**^{L} *'he shouted at them'*, taken from a recorded text.

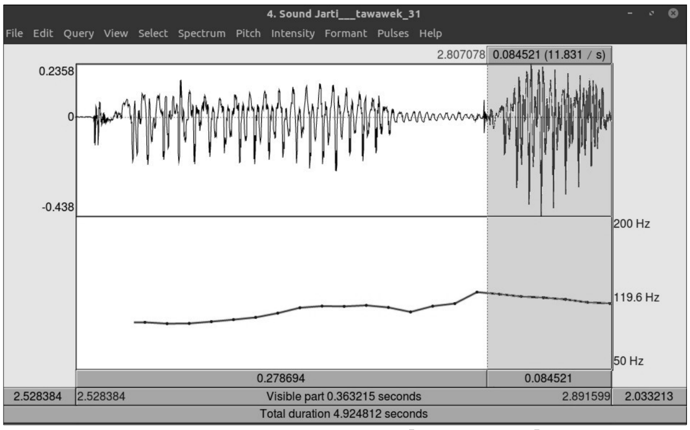

Figure 6: duration of **-gé**^{L} in **tònúrgé**^{L}

Accordingly, there is some lack of clarity regarding the length of vowels in some morphemes, and some of these could probably have been represented differently in this study.

But in spite of all this, length is crucially important in Majang; not only is it used to distinguish lexical items, but also to modify the lexical meaning of a root, e.g. to indicate different aspectual shades:

Example II.3: aspectual categories expressed by vowel length

tòróꜜkáːŋ	*I step*	vs.	**tòróːꜜkáːŋ**	*I trample*

Some verbs display stems with changing vowel lengths across the same paradigm:

Example II.4: stems with different vowel lengths in verb paradigms

gǎjŋ	*he imitates*	vs.	**gàːjáːŋ**	*I imitate*

The same is true for nouns, where length is used to express number and case differences:

Example II.5: stems with different vowel lengths in noun paradigms

a)	**kàːrí**	coffee leaf *(SG.ABS)*	vs.	**káríkᴸ**	coffee leaves *(PL.ABS)*
b)	**tɔ̂ːn**	child *(SG.ABS)*	vs.	**tɔ̂n**	child *(SG.ERG)*

Short vowels are considerably more frequent than their long counterparts. In the 2879-word text sample, 3817 short vowels are balanced by 1139 long vowels, a ratio of more than 3:1. But this ratio is not evenly spread out for all vowels. Whereas 243 short vowels /**e**/ occur compared to 104 long counterparts, a ratio of 2.33:1, against 245 short vowels /**u**/ there are only 17 long vowels /**uː**/, a ratio of 14.4:1. The other five vowels are much closer to the established average. One possible explanation for the surprisingly high ratio between the /**u**/ vowels is that many of the short ones go back to a vowel /**i**/ undergoing labial harmony, which does not affect long vowels /**iː**/.

There is a difference between a long vowel and two identical short vowels meeting at morpheme boundaries. All previous examples show long vowels belonging to one morpheme. They are then pronounced as just that – long vowels. But there are other possibilities:

Example II.6: apparent long vowel resulting from two adjacent morphemes

ɛ́rcɛ́ɛ́ᴸ	*milk*	<	{**ɛ́rcɛ́**	**-ɛᴸ**}
			milk	*-PL.ABS*

In this example the plural suffix **-ɛᴸ** is added to a stem ending in the vowel **ɛ**. This results in a sound that on the surface may look like a long vowel. Phonologically, however, there is a difference between this /**ɛɛ**/ and the first vowel in **kɛ̀ːtɛ́ŋ** *'he scatters'* (example II.2d). When two identical vowels belonging to two different morphemes come together, speakers often insert a phonetic separator, which could either be an approximant ([**j**] or [**w**], depending on the vowels involved) or the glottal stop [**ʔ**]. In fast speech, however, the two-vowel sequence is often indistinguishable from a long vowel. So example II.6 may be pronounced as either [**érsɛ́ː**], [**érsɛ́jɛ́**] or [**érsɛ́ʔɛ́**], with no difference in meaning. Throughout this work, I show both vowels spelled out when they belong to two different morphemes, and use the phonetic length symbol /**ː**/ to indicate a long vowel belonging to only one morpheme.

The Majang language allows vowel sequences involving long vowels.

Example II.7: vowel sequences

a) **{dàràː-ɛ́ː^{L}}**
despise-IMPS
it is despised

b) **{áːbɛ́-ɛːk-ɛ̀}**
fig_tree-PL-LOC
at the fig trees

Example a) has a long vowel (the impersonal marker) following a long stem vowel. In example b) the long vowel of the plural morpheme follows a short stem vowel of the same quality. If no glottal stop is used to separate the two vowels, the resulting vowel length of /**ɛ́ɛ́ː**/ appears to not be significantly longer than a regular long vowel /**ɛː**/, or the union of two short vowels /**ɛ**/ as in **{áːbɛ́-ɛ^{L}}** '*fig trees' (PL.ABS)*.

II.1.4 Distribution

Most of the vowel phonemes can appear anywhere in a word. Vowels can be found word-initially, word-finally and inside a word.

Monosyllabic words with only a short vowel and no coda are very rare. Only function words like **à** *'with'*, **ɓá^{L}** *'REMPST'* or the conjunction **kɛ́** belong to that category. Nearly as infrequent are monosyllabic words ending in a long vowel: **ɗɔ́ː** *'world'*, **pɛ́ː** *'soup'* and **wàː** *'house (DAT)'* are the only nouns found in that category. The only verbs encountered with a stem resulting in such words are **ŋàː** *'he stinks'*, **mèː** *'it hurts'*, **dɛ́ː** *'it is red'*, **kɛ̀ː** *'he goes'* and **wɛ̀ː** *'he breathes'*.

As already stated earlier, there is no ATR-based vowel-harmony system in Majang. If there were, one would expect some variation within the suffixes, with basically two allomorphs for each suffix morpheme, based on the ATR value of the stem vowel. There are also languages where the affix morpheme has an effect on the stem vowel (Casali, 2008, p. 500). Neither case can be observed in Majang. On the contrary, the verbal system allows combinations of vowels which would be deemed impossible in a language with ATR-based vowel harmony. Here are some examples:

Example II.8: vowel combinations with different ATR values

a)	**ɛ̀ːkîːŋ**	b)	**mɔ́ːꜜrún**
	{ɛ̀ːk-ǐː^{L}=ŋ}		**{mɔ́ːr-ín}**
	draw-1P.DJ=SFT		*wrinkled-2S.DJ*
	we draw water		*you are wrinkled*

Both the stems **ɛ̀ːk** and **mɔ́ːr** are completely stable, regardless of the following suffixes. The same is true about the *1P.DJ*-suffix **-ǐː**^{L}, which undergoes some tonal variation, but remains stable with respect to the vowel. The suffix for the 2nd person singular disjoint is **-ín,** as in **làːrín** *'you lose'*, but when following a labial consonant or a syllable with a round (or labial) vowel, it is rounded to **-ún**, as in the previous example. No vowel changes can be attributed to an ATR vowel harmony.

Even within the roots of nouns or verbs, it is possible to find vowel combinations which defy any attempt to describe them in terms of an ATR-based vowel-co-occurrence restriction. Roots like **ɛ̀ːbɔ̀d** *'heal'* and **éːmɛ́j** *'honor'* should not be encountered in a language with ATR-based vowel harmony.

V1/V2	**i**	**e**	**ɛ**	**a**	**ɔ**	**o**	**u**	total
i	64	24	36	48	9	14	1	196
e	28	40	15	32	1	4	2	122
ɛ	43	4	72	33	4	6	2	164
a	90	19	81	119	3	15	26	353
ɔ	21		17	26	16		6	86
o	24	12	25	36		26	27	150
u	10	14	13	8	1	7	24	77
total	280	113	259	302	34	72	88	1148

Table 2: vowel co-occurrences in polysyllabic words

Table 2 is a chart of all vowel co-occurrences in different polysyllabic words in the analyzed Majang text sample. The table counts whenever a vowel of column 1 is followed by a vowel of row 1, regardless of whether this vowel is part of the stem or of a suffix.

Some conclusions can be drawn from this table: almost any vowel can be followed by any other vowel, except for the two back-mid vowels /**o**/ and /**ɔ**/, which were never found following each other, and the sequence /**ɔ-e**/. These look like structural constraints which may indeed be going back to an old ATR-based vowel-harmony system at an earlier stage of the language's history. For many vowels, the preferred combination appears to be with the same vowel. The vowels /**a, ɛ, e, i, u**/ show a high occurrence as follow-up vowels in a vowel sequence, which is best explained by their prominent use in suffixes. The vowel /**ɔ**/, although not at all infrequent in the language, has a disproportionately low occurrence rate as a second vowel, because it is rarely used in suffixes, except in demonstratives – and each different demonstrative is only counted once in the above chart.

Some combinations are very infrequent. The one token filling the slot /**e-ɔ**/ is the short relative pronoun **écɔ̀**. The slot /**i-u**/ is taken by the complex verb **pàːrírkúndɔ́**[L] *'while you*$_{PL}$ *are trying'*. This same word is also the only place where the sequence /**u-ɔ**/ was encountered, as this contains the rare *2P* marker **-ɔ**[L].

The picture is not very different when only monomorphemic stems are taken into account. The vowel /**ɛ**/ can still follow /**e**/ inside a root, but not the other way around. The only other restrictions in addition to those in Table 2 concern the vowel /**u**/, which never seems to follow any of the vowels /**i**, **ɛ**, **ɔ**/ inside of roots.

Therefore, although there are some apparent restrictions regarding the co-occurrence of vowels, there is enough evidence discounting any ongoing ATR-based vowel harmony in Majang. The picture is, of course, complicated by the fact that four Proto-Surmic vowels /**ʊ**, **o**, **ɪ**, **e**/ (Moges, 2002, p. 209) have collapsed into the two vowels /**e**/ and /**o**/, so that what now may look like a violation of ATR-based vowel harmony could in fact go back to a difference that was present in an earlier stage of the language. For example, a root such as **èːbɔ̀d** *'heal'* is likely to go back to a proto-language root ***ɪːbɔd**, which would not violate an ATR-based vowel-harmony system[7]. Still, it is surprisingly difficult to find traces of the old ATR-based vowel harmony in Ma-

[7] This word is probably an old causative of **bɔ̀d** *'be well'*. See section V.5.3 for more on the old prefix **-i**.

jang[8]. A good example for this are the two vowels /**e**/ and /**ɛ**/. They look very much like *[+ATR]* and *[-ATR]* counterparts of the same vowel, so one should expect them to show some kind of distribution which follows old ATR-based patterns. Indeed, the two vowels alternate in some suffixes which contain them, but based on a principle unrelated to ATR: /**e**/ only occurs in words with a high vowel in the previous syllable, whereas /**ɛ**/ shows up in all other environments. Therefore the variation can be described in terms of a height assimilation. This becomes clear from the following examples:

Example II.9: height assimilation

a) **cù:wěrŋ**
{cù:w-er=ŋ}
sting-3P.DJ=SFT
they sting

b) **cù:wéŋ**
{cù:w-é=ŋ}
sting-3S.DJ=SFT
it stings

c) **ɔ̀:jɛ̌rŋ**
{ɔ̀:j-ɛr=ŋ}
obstruct-3P.DJ=SFT
they obstruct

d) **ɔ̀:jɛ́ŋ**
{ɔ̀:j-ɛ́=ŋ}
obstruct-3S.DJ=SFT
he obstructs

e) **rù:měrŋ**
{rù:m-er=ŋ}
decide-3P.DJ=SFT
they decide

f) **rù:méŋ**
{rù:m-é=ŋ}
decide-3S.DJ=SFT
he decides

g) **ò:dɛ̌rŋ**
{ò:d-ɛr=ŋ}
difficult-3P.DJ=SFT
they are difficult

h) **ò:dɛ́ŋ**
{ò:d-ɛ́=ŋ}
difficult-3S.DJ=SFT
it is difficult

i) **tì:ɟěrŋ**
{tì:ɟ-er=ŋ}
avenge-3P.DJ=SFT
they avenge

j) **tì:ɟéŋ**
{tì:ɟ-é=ŋ}
avenge-3S.DJ=SFT
he avenges

k) **ɓɛ́:cɛ̂rŋ**
{ɓɛ́:c-ɛr=ŋ}
touch-3P.DJ=SFT
they touch

l) **ɓɛ̀:cɛ́ŋ**
{ɓɛ̀:c-ɛ́=ŋ}
touch-3S.DJ=SFT
he touches

[8] Examples IV.107 and IV.119 of two infinitive allomorphs may show traces of the old ATR-harmony system.

m) **à:mɛ̌rŋ**
{**à:m-ɛr=ŋ**}
yawn-3P.DJ=SFT
they yawn

n) **à:mɛ́ŋ**
{**à:m-ɛ́=ŋ**}
yawn-3S.DJ=SFT
he yawns

As can be seen, the two allomorphs of the morphemes for 3rd person singular and 3rd person plural in this class of verbs obviously depend on the preceding stem vowel. The /**ɛ**/-variant can co-occur with any vowel, as long as it is not high. High vowels require the presence of the /**e**/-variant of the respective suffix (see section II.5.4). Again, words like **ɔ̀:dɛ́ŋ** or **ɓé:cêrŋ** would not be expected in an ATR-based vowel-harmony system.

The behavior of the *2s* suffix in example II.8b) lends itself to a description in terms of a rounding harmony:

Example II.10: rounding harmony

a) **ìbá:lîrŋ**
{**ìbá:l-ir=ŋ**}
play-3P.DJ=SFT
they play

b) **ɓòkòtǔrŋ**
{**ɓòkòt-ur=ŋ**}
kill-3P.DJ=SFT
they kill

For a detailed description of the rounding harmony process, see section II.5.3.

II.2 Consonants

Just like the previous section on vowels, this section on consonants first assesses the consonant inventory, then shows proof of contrast for the various identified phonemes, and finally looks into distributional restrictions.

II.2.1 Phonemic inventory

The consonant inventory of Majang is relatively small, containing only 18 phonemes. This is surprising when comparing Majang with other Surmic languages: Didinga (de Jong, 2004, p. 145) has a total of 36 consonants, and other Southwest Surmic languages have similar inventories. Southeast Surmic languages, such as Suri and Mursi, tend to have smaller inventories.

Mütze (2014, p. 26) counts 21 consonants for Mursi, and Bryant (1999, p. 16) 22 for Suri, which is only slightly more than found in Majang. Majang uses four places of articulation, and a maximum of six different manners of articulation can be found for the alveolar sounds. Only the obstruents and the nasals cover all four places of articulation.

	Labial	Alveolar	Palatal	Velar
Voiceless Obstruents	p	t	c	k
Voiced Obstruents	b	d	ɟ	g
Implosives	ɓ	ɗ		
Nasals	m	n	ɲ	ŋ
Oral Sonorants	w	r	j	
Lateral Sonorant		l		

Table 3: Majang consonant phonemes

The consonant inventory of Majang appears to be very tidy at first glance. For each of the four places of articulation, there is a voiced and a voiceless obstruent, as well as a nasal. At the palatal place of articulation, though, this tidiness is achieved only through a notable abstraction from the phonetic facts, which will be discussed and defended in detail in section II.2.2.3 on palatal consonants below.

Bender (1983, p. 116) shows the exact same consonant inventory, but with the addition of the glottal stop [ʔ], which he claims to be a phoneme of the language. Getachew (2014, p. 49) appears to agree with Bender, calling the glottal stop a phoneme; then again, he also seems to be aware of its complete predictability. Unseth (1988a), who otherwise fully shares Bender's inventory, rejects the phonemic status of the glottal stop on the grounds that it only occurs in word-initial position or between a few prefixes[9] ending in a vowel and the following stems. In the analysis proposed here, the glottal stop is only a phonetic device for separating syllables with a vowel onset, and by no means the only device (see example II.6 for a case where the glottal stop may be inserted). The syllabification rules of Majang only allow the insertion of a glottal stop when the preceding syllable ends in a vowel, or if the syllable is at the beginning of the word. Therefore the sound is excluded from the phoneme inventory. Apart from this small difference and a few different choices for the phoneme characters, the consonant inventory presented

[9] According to his analysis. I don't analyze them as prefixes, but as unbound particles.

above is the same as that proposed not only by Bender (1983), but also by Unseth (2007, p. 628). In addition, Getachew (2014, p. 49) includes /**h**/ in the phonemic inventory of Majang, qualifying this with the observation that it only seems to appear in loan words.

The distinction between fricatives, affricates and plosives is not relevant in the Majang phonological system, as there is no opposition between these at any place of articulation. Implosives are firmly established in the system, although only at the labial and alveolar places of articulation. The consonant inventory is completed by the two liquids /**l**/ and /**r**/ and by the two glides /**j**/ and /**w**/.

The following chart provides a frequency count of all consonant phonemes of Majang based on a 2879-word text sample:

Example II.11: frequency of consonant phonemes

rank	phoneme	count	rank	phoneme	count
1	**k**	1190	10	**l**	208
2	**n**	987	11	**ɓ**	187
3	**r**	487	12	**j**	169
4	**c**	448	13	**ɟ**	151
5	**t**	445	14	**d**	140
6	**g**	351	15	**w**	119
7	**ɗ**	308	16	**b**	109
8	**m**	306	17	**p**	87
9	**ŋ**	254	18	**ɲ**	35

It can be seen that the implosive phonemes are much more frequent than voiced obstruents. Only the voiced obstruent /**g**/ has a higher count than the two implosives. Another fact worth noting is that the labial voiceless obstruent /**p**/ has the second-lowest count of all phonemes, much less frequent than the very similar phoneme /**k**/, which tops the list with a good margin. Because of the nasal assimilation rule (section II.5.1) and the decision to transcribe all morpheme-internal combinations of /**ŋk**/ as <**nk**>, the numbers for the phonemes /**n**/ and /**ŋ**/ in this ranking could be slightly different. Any adjustment to this would not change the overall second place in the ranking for /**n**/, and probably also not the standing of /**ŋ**/.

II.2.2 Contrasts and phonetic realizations

This section provides proof of contrast between all possible pairs of phonemes that can be reasonably suspected of being variants of the same phoneme. Furthermore, information is given on how the phonemes are pronounced; if this is different depending on context or environment, information about allophones and free variation is also provided.

II.2.2.1 Labials /p, b, ɓ, m, w/

Most phonemes of the labial place of articulation are pronounced according to the IPA value of their symbols chosen here. The voiceless plosive /**p**/ is pronounced with a considerable amount of aspiration. The obstruent /**p**/ may also be realized by some speakers as a labiodental fricative [**f**] (Bender, 1983, p. 116). The implosive /**ɓ**/ is just that phonetically: [**ɓ**][10]. In the Godare variety of Majang, the voiceless plosive /**p**/ often is not fully released at the end of a word, to the point that it is barely audible at all.

Example II.12: contrasts between labials

p vs. **b**	**pédtàn**	*end*	**ápɛ́ː**[L]	*grandfather*
	bɛ́ːdtàn	*inhabitant*	**áːbɛ́**	*fig tree*
p vs. **ɓ**	**pàkàtíŋ**	*he decreases*	**tèpér**	*shoulder blade*
	ɓàkàtíŋ	*he unwraps*	**tɛ́ɓér**[L]	*thunder*
p vs. **m**	**pàcɛ́ŋ**	*he carves*	**páɲá**[L]	*cousin*
	màcɛ́ŋ	*he borrows*	**máɲá**[L]	*sister*
p vs. **w**	**pàrɛ́ŋ**	*he chops*	**ápɛ́ː**[L]	*grandfather*
	wàrɛ́ŋ	*he looks for*	**áwɛ́**	*iron*
b vs. **ɓ**	**bèɗíŋ**	*he sits*	**bàːlɗìːɗíŋ**	*he throws*
	ɓèɗíŋ	*he is awake*	**ɓálɗìɗíŋ**	*he sells*
b vs. **m**	**àbìː**	*cloth*	**bòːrò**	*gecko*
	àːmì	*hair*	**mòːr**	*anvil*
b vs. **w**	**áːbɛ́**	*fig tree*	**bɔ́dɛ́**[L]	*oil palm*
	áwɛ́	*iron*	**wɔ́ːr**	*feast*

[10] See Moges (2006, p. 823ff) for a very careful description of the implosive sounds of Majang, based on acoustic measurements. He states that in word-final position the implosives have a devoiced allophone, and are sometimes just represented by a glottal stop.

ɓ vs. **m**	**gàɓɛ́ŋ**	*he accuses*	**ɓù:kɛ́ŋ**	*he uncovers*
	gàmɛ́ŋ	*it is a fetus*	**mùkɛ́ŋ**	*he stabs*
ɓ vs. **w**	**cúɓój**	*clay*	**ɓèɗíŋ**	*he awakes*
	cúwǒj	*eel*	**wènɛ̀**	*antenna*
m vs. **w**	**màcɛ́ŋ**	*he borrows*	**kèmúŋ**	*straighten*
	wà:cɛ́ŋ	*he speaks*	**kè:wúŋ**	*sharpen*

II.2.2.2 Alveolars /t, d, ɗ, n, l, r/

The voiceless plosive /**t**/ is realized with noticeable aspiration [**tʰ**]. The phoneme /**r**/ is pronounced as a flap [**ɾ**]. The implosive /**ɗ**/ is pronounced according to the IPA value of its phonemic character [**ɗ**][11], and likewise the other phonemes /**d**, **n**, **l**/ as [**d**, **n**, **l**]. In the Godare variety of Majang, the voiceless plosive /**t**/ often is not fully released at the end of a word, to the point that it is barely audible at all.

Example II.13: contrasts between alveolars

t vs. **d**	**tá:mɛ́**^{L}	*face*	**tǎmŋ**	*it drips*
	dá:mɛ́	*yellow*	**dǎ:mŋ**	*he chooses*
t vs. **ɗ**	**tǎmŋ**	*it drips*	**tòɲɛ́ŋ**	*he stops up*
	ɗǎmŋ	*he eats*	**ɗòɲɛ́ŋ**	*he smears*
t vs. **n**	**mèrmét**^{L}	*red pepper*	**àpátí**^{L}	*breast*
	mèrmén	*notice*	**ápání**	*current*
t vs. **l**	**tɔ́kɔ́j**	*lazy*	**kàtàmɛ́**	*town*
	lɔ́kɔ̀j	*plate*	**pàlámɛ́**^{L}	*argument*
t vs. **r**	**ɗò:túŋ**	*he harvests*	**mót**	*blind person*
	ɗò:rúŋ	*it crows*	**mò:r**	*anvil*
d vs. **ɗ**	**dǎ:mŋ**	*he chooses*	**pèdíŋ**	*it is consumed*
	ɗǎmŋ	*he eats*	**ɓèɗíŋ**	*he is awake*
d vs. **n**	**dò:mɛ́ŋ**	*he allows*	**ŋàdíŋ**	*he is sad*
	nòmɛ́ŋ	*he follows*	**ŋàníŋ**	*he turns around*
d vs. **l**	**dì:léŋ**	*he carries*	**dɔ̀:kɛ́ŋ**	*he wanders*
	lì:léŋ	*he sinks*	**lɔ̀kɛ́ŋ**	*he overturns*

[11] See again Moges (2006) for details on the phonetic variation.

d vs. **r**	**à:díŋ**	*he washes*	**dò:méŋ**	*he allows*
	àríŋ	*he weaves*	**rómí**^{L}	*morning*
ɗ vs. **n**	**ɓèɗíŋ**	*he is awake*	**wô:ɗ**	*who?*
	ɓè:níŋ	*he sews*	**wón**	*which?*
ɗ vs. **l**	**ɗúk**^{L}	*forest*	**ò:ɗéŋ**	*it is difficult*
	lúk^{L}	*bastard*	**ò:léŋ**	*he is able*
ɗ vs. **r**	**ɗà:wúŋ**	*it is bright*	**pì:ɗéŋ**	*he shivers*
	rà:wúŋ	*he singes*	**pìréŋ**	*he flies*
n vs. **l**	**nòméŋ**	*he follows*	**tɔ̂:n**	*new*
	lòtéŋ	*he adds*	**bɔ̂:l**	*weak*
n vs. **r**	**tépén**^{L}	*forehead*	**wóní**^{L}	*tree, sp.*
	tèpér	*shoulder blade*	**wòrí**	*money*
l vs. **r**	**wè:léŋ**	*he robs*	**tùl**	*rain*
	wè:réŋ	*he alters*	**tùr**	*garbage dump*

II.2.2.3 Palatals /c, ɟ, ɲ, j/

The two palatal obstruents /**c**, **ɟ**/ show the most phonetic variation of all consonants in Majang: the voiceless obstruent /**c**/ is most often realized as a voiceless alveolar fricative [**s**], whereas its voiced counterpart /**ɟ**/ (/**j**/ in Bender's inventory) is usually pronounced as a voiced post-alveolar affricate [**d͡ʒ**]. The voiceless palatal /**c**/ may be pronounced as a voiceless post-alveolar fricative [**ʃ**] by some speakers, which supports its analysis as a palatal sound in the system. A regular allophonic rule applies for /**c**/ following the palatal nasal /**ɲ**/, which turns /**c**/ into [**t͡ʃ**]: **ɓéɲcè** [**ɓéɲt͡ʃè**] *'today'*. The voiced palatal obstruent also has a less frequent variant [**ʒ**] (voiced post-alveolar fricative).

It is necessary to justify why the phoneme **/c/** is treated here as a palatal obstruent, when in fact its most frequent realization is the alveolar fricative [**s**], without an apparent allophonic rule governing the change to [**s**]. This decision follows Bender (1983, p. 117), who attributes the variation in the palatal sounds to the removal of the bottom incisors, which apparently distorts the consonants beyond their position in the system. Like Bender, I therefore assume an underlying palatal nature of the two phonemes /**c**/ and /**ɟ**/, although this is not necessarily justified by their phonetic realization. Accordingly it would be possible to treat /**c**/ as an alveolar fricative /**s**/, as was done by Getachew (2014, p. 48), but this phonetically driven classification of sounds would add another dimension of complexity into the consonant inventory by claiming an opposition between plosives and fricatives, for which there is no further evidence in the language. At the same time, it leaves a glaring gap for the voiceless counterpart of the phoneme /**ɟ**/, which in all its phonetic realizations is clearly post-alveolar. In effect, the decision to follow Bender's treatment of /**c**/ as a palatal phoneme implies that the palatal space in Majang is wider than in other languages, including the edge of the alveoles at which the front-most sibilants can be formed. This is quite in line with Trubetzkoy (1939), who claims (p. 35) that the phoneme is best defined as the total of the phonologically relevant features of a sound entity, and that its content entirely depends on its place in the system, defined by the distinctive oppositions found in the language (p. 39). The same idea is built upon more recently by Clements (2003) in his concept of *Feature Economy*.

There is, however, no phonological rule in Majang that changes /**ɟ**/ to /**c**/ or vice versa, as such a rule would clearly show that both phonemes belong to the same place of articulation. In the same way, there is also no rule that shows /**c**/ to be associated with the alveolar sounds of the language.

The palatal nasal /**ɲ**/ and the palatal glide /**j**/, however, are pronounced, like their IPA symbols suggest, as [**ɲ**] and [**j**][12]. In the other places of articulation, the Godare dialect has an unreleased variety of the voiceless obstruents. This is not applicable for the palatal obstruent /**c**/, as this is usually realized as a phonetic fricative.

Example II.14: contrasts between palatals

c vs. **ɟ**	**mácɛ́**[L]	*debt*	**gàcój**	*hoe*
	máːɟɛ́[L]	*dwarf*	**gáꜜɟój**	*courageous man*
c vs. **ɲ**	**còːn**	*towards*	**kòcɛ̀**	*bag*
	ɲòn	*place*	**kòːɲɛ́ŋ**	*he curses*
c vs. **j**	**càwɛ́ŋ**	*he sprinkles*	**pàcɛ́ŋ**	*he carves*
	jàwɛ́ŋ	*he shortens*	**pàjɛ́ŋ**	*he vomits*

[12] Note that I do not follow the Ethiopianist/Africanist tradition of transcribing the palatal glide /**j**/ as /**y**/. In the same manner, the transcription /**j**/ never serves as the representation of the phonetic sound [**d͡ʒ**], as this is phonemically represented as /**ɟ**/ throughout this study.

ɟ vs. ɲ	**tòɟέŋ**	*he disturbs*	**ɟɔ́nkɔ́lέ**	*porcupine*
	tòɲέŋ	*he stops up*	**ɲɔ́:ná**[L]	*insult*
ɟ vs. j	**tòɟέŋ**	*he disturbs*	**ɟámὲ**	*machete*
	tò:jέŋ	*he pierces*	**jáwέ**[L]	*circumcision*
ɲ vs. j	**tòɲέŋ**	*he stops up*	**kà:ɲ**	*brideprice*
	tò:jέŋ	*he pierces*	**ká:j**	*night*

II.2.2.4 Velars /k, g, ŋ/

The three velars /**k**, **g**, **ŋ**/ show minimal phonetic variation. Their pronunciation corresponds to the IPA value of their respective symbols, except that the voiceless obstruent /**k**/ is pronounced with considerable aspiration: [**kʰ**]. In the Godare variety of Majang, the voiceless plosive /**k**/ is often not fully released at the end of a word, to the point that it is barely audible at all.

Example II.15: contrasts between velars

k vs. **g**	**kòrúŋ**	*he shuts*	**é:kê:r**	*truth*
	gòrúŋ	*he is sick*	**ègèr**	*how many?*
k vs. **ŋ**	**kǎ:rŋ**	*he fights*	**tàk**	*inside*
	ŋǎ:rŋ	*he goes*	**táŋ**[L]	*cow*
g vs. **ŋ**	**dùgéŋ**	*he hides*	**tágá**[L]	*camel*
	dù:ŋéŋ	*it evaporates*	**tà:ŋá:ŋín**[L]	*heron*

II.2.2.5 Nasals /m, n, ɲ, ŋ/

Because of their phonetic similarity, it is appropriate to show examples for contrasts between the four nasal phonemes of Majang.

Example II.16: contrasts between nasals

m vs. **n**	**cèm**	*straight*	**dá:mέ**	*yellow*
	cè:n	*pronoun 3s*	**dà:né**	*beehive*
m vs. **ɲ**	**màlέŋ**	*he strikes*	**kɔ́mí**[L]	*zebra*
	ɲà:léŋ	*it is light*	**kɔ̀ɲέŋ**	*he persuades*
m vs. **ŋ**	**mùkéŋ**	*he stabs*	**tém**	*small*
	ŋù:kéŋ	*he pulls*	**ὲtéŋ**	*he stands*

n vs. **ɲ**	**órpán**^{L}	*naming ceremony*	**nòmɛ́ŋ**	*he follows*
	pá:ɲ^{L}	*mortar*	**ɲòn**	*place*
n vs. **ŋ**	**dó:kún**	*bride-price*	**nácí**^{L}	*bread*
	dòkúŋ	*it lands*	**ŋádí**^{L}	*pity*
ɲ vs. **ŋ**	**ɲá:j**	*liver*	**kó:ɲ**	*curse*
	ŋâ:j	*old woman*	**kǒŋ**	*he gathers*

Majang has an active nasal assimilation rule, which turns each alveolar nasal placed in front of a velar consonant into a velar nasal, as in **àŋàn** *'four'* versus [**àŋàŋ-k**] *'four (POSS)'*.

II.2.3 Distribution

Before discussing the distribution of consonants, it is necessary to give some information on syllable, root and word structure in anticipation of sections II.4 and II.7. Syllables can be open and closed, and there can be two-consonant clusters both word-medially and word-finally. Two-consonant clusters can also appear inside a morpheme, as in the verb stem **dèrt** *'slide'*. Words can be monomorphemic, but frequently include suffixes and even enclitics. Prefixes are very rarely encountered; these always end in a vowel. Noun and verb roots are often monosyllabic, but can consist of several syllables.

Most consonants can appear in all environments. There are, however, some limited distributions which are difficult to explain. Only a very small number of words end in a labial consonant, and except for the adverb **òtó:p** *'often'*, and the noun **ɟó:p**^{L} *'people'*, no words end in a labial obstruent. The labial nasal /**m**/ does appear word-finally in about 20 words, but this is balanced by the more than 200 words ending in the alveolar nasal /**n**/. Between vowels and in word-initial position /**m**/ is by far the most frequent nasal.

More transparent is the fact that at the end of a word, non-palatal voiced obstruents are not permitted phonetically. If an underlying voiced obstruent appears in that position, it is devoiced, such as **ìɟá:g**^{L} *'work'*, which is pronounced [**iɟa:k**^{h}], in this way creating a neutralization not unlike obstruent hardening in other languages, such as German. The speakers do not appear to be aware of that alternation; so this process seems to be of a postlexical nature – see chapter 7 of Mohanan (1986) for evidence that speaker judgments are linked to the lexical level of phonological representations.

All consonants appear in good numbers word-initially and between vowels, except for the alveolar nasal /**n**/, which surprisingly appears word-initially in only about a dozen lexemes, as opposed to nearly ninety for /**m**/, 21 for /**ɲ**/ and 34 for /**ŋ**/. Except for the assimilation of alveolar nasals to the velar place of articulation (see section II.5.1), nasals keep their place of articulation, both inside the root and at morpheme boundaries: for example, **ámɗ**^{L} *'stomach'* and **ɗàmɗú** *'after he ate'* show /**m**/ maintaining its labial place of articulation preceding alveolar phonemes. The two reduplicated verb stems **tíntím** *'fold'* and **tɔ́ntɔ́m** *'incubate'*, however, suggest that some assimilation happens preceding alveolar consonants inside a stem. No other instances were encountered where any nasal assimilates to another place of articulation. **tɛ́mk** *'it is small (SUB)'* shows /**m**/ stable in front of a velar consonant. The alveolar nasal can be encountered preceding palatal consonants, as in **pàránɟí**^{L} *'white person'*. The palatal nasal was encountered preceding /**k**/ in **kó:ɲkàn** *'cursed person'*. The velar nasal was sometimes found preceding alveolar phonemes, as in the stem **cɔ̀:líláŋt**^{L} *'vulture (NOM)'*. It appears that Majang can be characterized as a split system regarding nasal assimilation, with a clear assimilation of alveolar nasals towards velar nasals, and a general lack of assimilation between other places of articulation.

In consonant clusters across syllable boundaries, there is a high tendency for having sonorants as first consonant *(C1)* and obstruents as second consonant *(C2)*, as in words like **tàmà:rɛ̀rkí** *'they learn'* or, twice, in **ɓànkàwkà** *'power'*. Obstruents can only be found as *C1* when *C2* is also an obstruent, as in **tá:pꜜtá:nák**^{L} *'letters'* or **ìɟá:gkɛ**^{L} *'when they started to work'*. In the same way, sonorants also only appear as *C2* when *C1* is another sonorant, as in **mérménáːꜜrá** *'I notice'* or **ɓòkó:rjántá**^{L} *'tortoise (DAT)'*.

Implosives usually appear as *C2* in a consonant cluster. An alveolar implosive /**ɗ**/ cannot be found preceding another consonant within the same word. It is changed to the flap /**r**/, as in **è:ŋàɗá:ŋ** *'I smell'* vs. **è:ŋárŋ** *'he smells'*. The labial implosive /**ɓ**/ is rarely found preceding consonants. The only example encountered in the text corpus is **gàgáɓkɛ́jgíɗá**^{L} *'whenever I give her'*. Moges (2006, p. 829) points out that in spite of these restrictions applying to Majang implosives, they are much more flexible in appearance than predicted by Greenberg (1970, p. 131), who stated that implosives are universally not found in word-final position, and not in consonant clusters, particularly involving nasals. Both situations do happen in Majang.

In the few cases of monomorphemic and polymorphemic word-final consonant clusters, *C1* needs to be a *[+SONORANT]*[13] consonant. If *C1* is /**l**/ or a nasal, then *C2* can be a *[-SONORANT]* consonant (including implosives), or the nasal /**ŋ**/ (such as in **ɛ́lt**ᴸ *'grassland'*, **é<u>mɗ</u>** *'canoe'* or **tǎ<u>mŋ</u>** *'it drips'*). If *C1* is a *[-NASAL]*, *[-LATERAL]* sonorant /**r**, **j**, **w**/, then only /**ŋ**/ and /**n**/ can be found in the *C2* position, as in **kǒwŋ** *'he digs'*, **káːrn**ᴸ *'war'*, **ɓɛ́jn**ᴸ *'thorn'*, or **ɗɛ́jŋ**ᴸ *'cooking stone'*. The restrictions on word-final consonant clusters are entirely in line with the *sonority hierarchy* proposed by Lowenstamm (1981), which predicts the following order for syllable codas: *approximant – sonorant – nasal – obstruent*. Two phonemes from the same level can combine to a consonant cluster. This is exactly what can be seen in Majang, giving the implosives a special status as being neither a sonorant nor an obstruent, placing them between nasals and obstruents.

The language strongly disfavors a syllable rhyme containing both a long vowel and a consonant cluster[14]. If these two things come together in any configuration, the vowel is shortened. This can be observed for example in the locative form of **ŋédán** *'bee'*, which makes use of the non-central case stem /**ŋédáːn**/, to which the locative stem extension **-t** is added, resulting in the form ***ŋédáːnt**, which is shortened to **ŋédánt**.

II.3 Distinctive Features

When describing phonological rules, I occasionally make reference to the distinctive features of the phonological system. It is therefore necessary to determine what the distinctive features are for each phoneme. The discussion of the features presented here is not intended to be a contribution to any phonological model interested in features. No attempt is made to validate or disprove any existing feature inventories. The inventory proposed here is entirely language-specific and was chosen because it reflects the phonological realities of Majang. The vowels of Majang can be classified as seen in Table 4, with the understanding that they are all further marked as *[+SYLLABIC]*:

[13] See section II.3 for a definition of each distinctive feature.

[14] There are exceptions to this rule, such as the three possessive pronouns **gáːnk** *(1S.PL)*, **góːnk** *(2S.PL)* and **géːnk** *(3S.PL)*, or the nominative singular form **wéːnt** of **wéːná**ᴸ *'ear'*.

		[-LABIAL]		*[+LABIAL]*
[+HIGH]		**i**		**u**
[-HIGH]	*[+ATR]*	**e**		**o**
	[-ATR]	**ɛ**		**ɔ**
			a	
		[-BACK]	*[+BACK]*	

Table 4: distinctive features of Majang vowel phonemes

It may appear surprising that in such a feature representation the *[-HIGH]* vowels are not further distinguished for vowel height, including the low vowel /**a**/. The introduction of the feature *[LOW]* would introduce an unnecessary redundancy into the feature system. But it needs to be pointed out that the representation in table 4 masks another lower-level redundancy: the feature *[-ATR]* is neither required nor desired to characterize the vowel /**a**/.

I prefer the use of the feature *[LABIAL]* over the feature *[ROUND]*, because the labial consonants and the round vowels trigger the same process, and therefore it may be assumed that the same phonological feature is responsible. See section II.5.3 for a detailed discussion on how this feature *[LABIAL]* governs labial harmony.

The feature *[ATR]* may also be a surprising choice, as it was established that there is no ATR harmony in the language. Still the feature *[ATR]* is the only distinctive feature provided by the standard literature of generative phonology with the power to distinguish between the *[-HIGH]* vowels, even if a feature *[LOW]* were invoked, which would only apply to /**a**/ in Majang. Hall (2007, p. 329) observes that *"the feature [ATR] is used to capture the contrast between /i e o/ ([+ATR]) and /ɪ ɛ ɔ/ ([-ATR]) – both in West African languages with [ATR] harmony, as well as Germanic languages like English."* This is in agreement with Casali's (2008, p. 507) explanation that *[+ATR]* vowels are characterized by a lower F1 value than *[-ATR]* vowels (and are therefore higher in traditional phonetic terms). The only other feature previously used for this purpose, *[TENSE]*, was abandoned in the early days of generative phonology. Ladefoged's (2005, p. 8 f) suggestion to use a feature *[MID]* for languages without ATR-based vowel harmony was not taken up by phonologists, possibly because it would still fail to distinguish between the two rows of mid vowels in an intuitive understanding of the feature. It was also seen in Table 2 that reference to an ATR-based vowel harmony in an older stage of the language explains some co-occurrence restrictions between some vowels; further traces can be seen in some morpheme alternations such as the nominalization markers **-ɔn** and **-an** (see section IV.2.2.1). If it was the feature [ATR] that caused this harmony in the past, it appears logical that the same feature still distinguishes the same vowels, even if a harmony is no longer based on this distinction.

[ATR] is the only vowel feature which is not also used for consonants.

The following table helps to identify the consonant features, which are all needed in addition to the feature *[-SYLLABIC]*.

		[+LABIAL]	*[-LABIAL]*			
			[-HIGH]	*[+HIGH]*		
		p	**t**	**c**	**k**	*[-VOICED]*
[+OBSTRUENT]		**b**	**d**	**ɟ**	**g**	*[+VOICED]*
[-OBSTRUENT]	*[-SONORANT]*	**ɓ**	**ɗ**			
	[+SONORANT]	**w**	**r**	**j**		*[-LATERAL]*
	[-NASAL]		**l**			*[+LATERAL]*
	[+NASAL]	**m**	**n**	**ɲ**	**ŋ**	
			[-BACK]		*[+BACK]*	

Table 5: distinctive features of Majang consonant phonemes

The classification with respect to the manner of articulation follows Clements and Osu (2002, p. 308), who state that *"[...] the common property distinguishing implosives from explosives is the absence of air pressure buildup in the oral cavity. [...] this property is exactly the correlate of the feature [-OBSTRUENT]."* Indeed the two implosives have a different distribution from both obstruents and sonorants in Majang, and therefore the *[OBSTRUENT]* feature provides the main dividing line between the Majang consonants. The eight *[+OBSTRUENT]* phonemes only need to be further classified as *[-VOICED]* and *[+VOICED]*. For the ten *[-OBSTRUENT]* sounds, a further division is made by the feature *[SONORANT]*, which cannot just be seen as the inversion of the feature *[OBSTRUENT]* (Clements & Osu, 2002, p. 337f), but which in turn singles out the implosives from the other *[-OBSTRUENT]* phonemes. The *[+SONORANT]* phonemes are either nasals *[+NASAL]* or approximants *[-NASAL]*. The approximants are either lateral (/**l**/) or non-lateral (/**w**, **r**, **j**/). This distinction is motivated by the fact that the three non-lateral approximants on the one hand and the lateral approximant on the other display different phonotactic behavior (see section II.2.3).

The Majang consonant system provides no need for using the feature *[CORONAL]* to specify any place of articulation, as the feature *[LABIAL]* is needed anyway in the language and therefore is sufficient to distinguish between labial and alveolar sounds. The labial harmony (section II.5.3) proves that the feature *[LABIAL]* is applied to both vowels and consonants; the labial consonants and round vowels clearly form a natural class. This renders the feature *[CORONAL]* redundant in Majang. The three remaining places of articulation – alveolar, palatal, and velar – can easily be distinguished by the features *[HIGH]* and *[BACK]*, which are also used for distinguishing vowels.

In total, all 25 consonants and vowels of Majang can be distinguished by the following ten non-prosodic distinctive features *[SYLLABIC]*, *[OBSTRUENT]*, *[SONORANT]*, *[NASAL]*, *[VOICED]*, *[LATERAL]*, *[LABIAL]*, *[HIGH]*, *[BACK]* and *[ATR]*.

To be able to fully accommodate all phonological phenomena of Majang in a feature notation, three prosodic features would have to be introduced: *[LONG],* a tone and a register feature. But apart from the feature *[LONG]*, prosodic features are not used in this language description.

II.4 Syllable and Root Structure

II.4.1 Syllable structure

The Majang maximal-syllable template describing the possible phonetic syllables of the language is CVC, which allows for consonant clusters across syllable boundaries word-medially. Only at the end of a word this template can be exceeded by monosyllabic consonant clusters, but usually only following short vowels[15], and if they do not violate restrictions caused by the sonority hierarchy. For more details on consonant clusters, see section II.2.3.

Unlike Bender (1983, p. 115), the analysis of this study does not assume the presence of diphthongs in Majang. The examples given by Bender are in-

[15] See footnote 14 for some exceptions to this.

stead interpreted as sequences of a vowel and an approximant: **kuroi** (Bender) vs. **kúꜜrój** *'donkey'*; **waikun** (Bender) vs. **wájkún** *'seed'*.

In Majang there are many places where the transition from a syllable ending in a vowel to a syllable beginning with a vowel invites hearing a glide [**j**] or [**w**], depending on the nature of the vowels involved. This study follows the convention of only writing a glide between vowels if its underlying existence can be established from additional evidence, like if in a different morphological context the glide is found at the word boundary or in front of a consonant. Thus I write **cáːkòjὲ** *'valleys'* because of **cáːkòj** *'valley'*, but **ɓòὲ** *'antelope'*, because in this word, the final /**ɛ**/ is part of the noun root, and there is no independent evidence for a glide between the two vowels, besides the possibility to transcribe this word phonetically as [**ɓòjὲ**]. The final two syllables in the two words **cáːkòjὲ** and **ɓòὲ** are phonetically not easily distinguishable[16].

In order to maintain the Majang syllable structure, the language uses a number of devices to break up unacceptable consonant clusters. One of these, involving the presence of alveolar implosives preceding other consonants, is introduced as rule 3 in section II.5.2. Less specific is a vowel-epenthesis rule which inserts the default vowel /**i**/ between stems ending in a *[-SONORANT]* consonant and a consonantal suffix or enclitic. A very rough representation of this rule might look like this:

Rule 1: vowel epenthesis

$$\mathbf{Ø} \rightarrow \mathbf{i} \;/\; \underset{[\textit{-SONORANT}]}{C} \;\_\_ + C$$

This rule is applied for example with the subordinate-clause clitic **=k** and with the sentence-final topicality *(SFT)* clitic **=ŋ**, which attach themselves without epenthesis to preceding vowels and most sonorants, but require this epenthetic vowel in a position following obstruents:

16 With **ɓòὲ** a speaker may also choose to separate the two vowels by a glottal stop [**ɓòʔὲ**]. This glottal stop may in principle be inserted between any two vowels in a word, but most often it is not (see section II.1.3).

Example II.17: contrasts between nasals

a) **ɗúŋêŋ** <	**{ɗúŋé^{L}=ŋ}** *hyena*\SG.NOM=SFT	b) **kúꜜrôjŋ** <	**{kúꜜrój^{L}=ŋ}** *donkey*\SG.NOM=SFT
c) **cɔ̀:líláŋtîŋ** <	**{cɔ̀:líláŋt^{L}=ŋ}** *vulture*\SG.NOM=SFT	d) **kùrɓútûŋ** <	**{kùrɓút=ŋ}** *maggot*\SG.NOM=SFT

In examples a) and b), there is no epenthetic vowel, because the *SFT*-clitic finds a vowel or approximant (/**j, w**/) to attach itself to. Examples c) and d) have an epenthetic vowel following a stem ending in an obstruent. In example d), the epenthetic vowel changes from /**i**/ to /**u**/ as a result of the labial-harmony rule 4, introduced in section II.5.3 below. This shows that the epenthetic vowel is subject to this rule which otherwise only applies to verbal suffixes.

The feature *[-SONORANT]* is a simplification of the actual environments triggering this rule. Quite often the epenthesis also happens following nasals and even some other oral sonorants. For the *SFT*-clitic the following generalizations hold:

- A stem ending in a vowel or an approximant will never have the epenthetic vowel (as it would serve no purpose): **ɗúmâ:ŋ** *'owner (NOM)'*, **ɗúŋêŋ** *'hyena (NOM)'*, **kúꜜrôjŋ** *'donkey (NOM)'*.
- A stem ending in a *[-SONORANT]* sound will almost always have the epenthetic vowel: **wɛ́:ntîŋ** *'ear (NOM)'*, **émɗíŋ** *'canoe (NOM)'*, **ɗó:kûŋ** *'land (NOM)'*. Stems ending in an alveolar implosive which is not part of a CC-sequence change this implosive to the phoneme /**r**/ (see section II.5.2). This explains the alternation between **è:ŋàɗ̠á:ŋ**^{L} *'I smell'* and **è:ŋár̠ŋ** *'he smells'*.
- Stems ending in a nasal use the epenthetic vowel only following a long stem vowel: **dɔ̀mɔ́:nûŋ** *'leopard (NOM)'*, **wà:já:nîŋ** *'plant (NOM)'*, **ɛ́dɛ́:nîŋ** *'mountain (NOM)'*. Following a short vowel, the *SFT*-marker is dropped entirely: **tɔ̂n** *'child (NOM)'*, **tèkán** *'aunt (NOM)'*, **cà:kóm** *'friend (NOM)'*. This criterion looks more straightforward than it actually is. One of these long vowels is apparently *caused* by the *SFT*-marker; without it, the plain nominative form of **ɛ́dɛ́n** *'mountain (ABS)'* is **ɛ́dɛ́n**^{L} with a short vowel. The *SFT*-marker therefore seems to create its own condition where it can appear. But it does not do so on other words.

- For other sonorant environments, the placement of the epenthetic element appears to be even more random, as illustrated by word pairs such as **kùtúr** *'hog (ABS)'* vs. **kùtúrùŋ** *'hog (NOM)'* as opposed to **wár**^L *'dog (ABS)'* vs. **wârŋ** *'dog (NOM)'*. For lack of a better generalization, it needs to be concluded that stems ending in /**l**/ and /**r**/ can sometimes be found taking a vowel, and sometimes not.

II.4.2 Root structure

In Majang, the synchronic distinction between roots and stems is often difficult to make. Some derivation affixes clearly identify many stems as derived from particular roots, as in the case of **ɓó:ján**^L '*hate* (*INF*)', which is derived from the verb root **ɓó:j**. There are, however, many lexemes where the different stem forms cannot be easily traced back to an identifiable underlying root. This is the case when the differences between the stems are either of a purely tonal nature, or when there are differences in vowel length or quality, or when the stems differ in the absence or presence of a particular (usually stem-final) consonant, which cannot be identified as a productive derivation marker. All these factors can be illustrated by the stems **ɛ́dɛ́n** (*SG.ABS*), **ɛ́dɛ́n**^L (*SG.NOM.MOD*), **ɛ̀dɛ̂:n** (*SG.ERG*), **ɛ́dɛ́n**^L (*SG.NOM*), **ɛ̀dɛ̀:n** (*SG.LOC*), **ɛ̀dɛ̀:n** (*SG.DAT*), **ɛ́dɛ́nk** (*PL.ABS*) and **ɛ́dɛ̀n** (*PL.ERG*). These forms all refer to the lexeme for '*mountain*', and some of them have additional grammatical uses to those listed here. Each stem is treated as a simple stem in this study, which implies that they are not treated differently from a root. This is particularly the case when looking at the phonological structure of roots, as this section attempts.

II.4.2.1 Noun roots

Simple noun roots in Majang consist of either one, two or three syllables. The root syllables conform to the maximal-syllable template, with the additional possibility of having consonant clusters at the end of the root. Beyond this, there does not seem to be any restriction as to what can constitute a noun root. Monosyllabic roots with a CV pattern are non-existent, and monosyllabic CV: roots are extremely rare. The example **wà:** below is the dative variant of the CVC root **wɛ́j**^L *'house'*.

Example II.18: examples of noun roots

1-syllabic	gloss	2-syllabic	gloss	3-syllabic	gloss
ká:rn	*war*	**kántὲ**	*basket*	**kácíkír**	*stump*
mò:r	*anvil*	**mó:rέ**	*fat*	**mὲkέlέm**^{L}	*hawk*
tὲk	*in-law*	**tálój**	*swarm*	**tὲŋɔ̀ní:**	*locust*
wà:	*house (DAT)*	**wòrí**	*money*	**wà:kójót**	*god*
ɓà:j	*heart*	**ɓɔ́dέ**^{L}	*palm tree*	**bùcùlέ**	*puppy*

II.4.2.2 Verb roots

Verb roots in Majang have a very strong tendency to end in a consonant. This is compatible with the fact that practically all person suffixes begin with a vowel. Roots ending in a consonant therefore enable easy syllabification. This is certainly true for all roots of the i- and ε-classes of verbs (see section IV.2.1.1 on verb classes). The various a-classes, however, contain many roots ending in a vowel. This may be the reason for the development of the **-k** extension morpheme used in these classes (see p. 248), which does not seem to carry any meaning. Its purpose appears to be to create well-formed stems which otherwise would end in a vowel.

Monomorphemic perfective verb stems usually consist of one syllable, but a smaller number of roots has two syllables.

Example II.19: examples of monomorphemic verb roots

1-syllabic	gloss	2-syllabic	gloss
à:m	*yawn*	**àgàl**	*steal*
càn	*lose*	**ɓɔ̀lɔ̀:r**	*grow*
dὲrt	*slide*	**dìgòj**	*greet*
gòt	*blow*	**gùpàt**	*spill*
ó:j	*shout*	**ògàr**	*cut hair*

Imperfective verb stems are formed by the reduplication of the first CV sequence of the first stem syllable, and therefore have at least two syllables. This reduplication does not appear to be a straightforward productive phonological process, as the following examples illustrate – tones may change and vowel length is usually lost. Rather, the imperfective verb stems need to be treated as lexicalized forms.

Example II.20: imperfective stem formation through reduplication

a)	**dɛ̀nɛ̀**	**dɛ̀dɛ̀n**
	{dɛ̀n-ɛ̀}	**{dɛ̀dɛ̀n}**
	see\PFV-3S.CJ	*see\IPFV.3S.CJ*
	he sees	*he is seeing*
b)	**ɗàm**	**ɗáɗámí:**^{L}
	{ɗàm}	**{ɗáɗám-i:**^{L}**}**
	eat\PFV.3S.CJ	*eat\IPFV-AP.3S*
	he eats	*he is eating*
c)	**ré:ꜜrá**^{L}	**rèrèrá**^{L}
	{ré:r-á$^{L\}}$	**{rèrèr-á**^{L}**}**
	run\PFV-1S.DJ	*run\IPFV-1S.DJ*
	I run	*I am running*

II.5 Phonological Processes

There are few phonological rules on segments which apply across the whole language. The two most regular rules are *nasal assimilation* and *alveolar implosive weakening*. Two other important rules are morphologically restricted to only some sets of suffixes: *height assimilation* and *labial harmony*.

II.5.1 Nasal assimilation

As seen in section II.2.2, alveolar nasals adapt their place of articulation when preceding velar consonants, as in **àŋàn** *'four'* versus [**àŋàŋk**] *'four' (POSS)*. This was not observed for alveolar nasals preceding palatal or labial consonants, for lack of words providing such an environment. Only alveolar nasals were found to assimilate their place of articulation. Section II.2.3 contains examples of possible clusters of nasals followed by non-homorganic consonants.

Rule 2: nasal assimilation

$$\mathbf{n} \rightarrow [+\text{BACK}] \;/\; \_\_ \underset{[+\text{BACK}]}{C}$$

This rule states that /**n**/ changes to /**ŋ**/ preceding a velar consonant. It is difficult to establish whether this rule is lexical, that is whether speakers are aware of this change (see Snider (2018, p. 95) for this component of the distinction between lexical and postlexical rules). Throughout this study it is treated as postlexical, which means that it is not reflected by the orthographic conventions adopted for this study. The possessive form of **àŋàn** is therefore written as **àŋànk**.

For many morphemes it is impossible to establish whether the phonetic combination [**ŋk**] goes back to an underlying cluster /**nk**/ or /**ŋk**/. This is true for many pronouns and determiners, such as **gánk** *'POSS\1P.PL'*, **cénk** *'2S.CONTR'*, **kɔ́nk** *'REF\RECPST'*, **mánk** *'or'*, or even the possessive case marker **-onk**. In all these cases, I decided to transcribe them consistently as <**nk**>.

II.5.2 Alveolar implosive weakening

The alveolar implosive /**ɗ**/ is not allowed to occur preceding a morpheme beginning with a consonant. In that environment it is weakened to /**r**/, as in **èːŋà<u>ɗ</u>áːŋ**^{L} *'I smell'* vs. **èːŋá<u>r</u>ŋ** *'he smells'*. A formal representation of this rule might look as follows:

Rule 3: alveolar implosive weakening

ɗ → **r** / __+*C*

There is no corresponding weakening rule for the labial implosive consonant.

II.5.3 Labial harmony

All verbal suffixes involving the short vowel /**i**/ in their first syllable are subject to a vowel-harmony rule, which, depending on the rhyme (*R*)[17] of the preceding syllable, changes the suffix vowel. If the rhyme contains a *[+LABIAL]* consonant or vowel, the vowel is /**u**/, otherwise it remains /**i**/.

[17] For the nomenclature regarding syllable structure, this study follows Blevins (1995).

Example II.21: labial harmony

a) **dí:ꜜlín**^{L}	b) **làŋìr**	c) **pá:ŋík**
{dí:l-ín^{L}**}**	**{làŋ-ir**^{L}**}**	**{pá:ŋ-ík}**
take-2S.DJ	*find-3P.DJ*	*slap-IMP.SG*
you take	*they find*	*slap!*
d) **ŋèɗè:mún**^{L}	e) **ɓòkòtùr**	f) **láptúk**
{ŋèɗè:m-ín^{L}**}**	**{ɓòkòt-ir**^{L}**}**	**{lápt-ík}**
smile-2S.DJ	*kill-3P.DJ*	*dive-imp.sg*
you smile	*they kill*	*dive!*

If the last stem syllable contains a *[+LABIAL]* vowel or consonant anywhere in the rhyme, the suffix vowel is /**u**/. This is the case in example e), where the labial vowel /**o**/ in the last stem syllable triggers the process. But, as stated, the process also applies to stems with a labial consonant in the coda, as seen in example II.21d). It is also important to note that the labial consonant triggering the harmony does not have to be the final consonant of the stem, but may be another consonant found in the coda, as the /**p**/ in example f). The nature of the onset of the preceding syllable has no impact on the choice of the suffix vowel, as can be seen from example c), where the onset consonant is *[+LABIAL]*, but all rhyme sounds are *[-LABIAL]*.

It is this labial harmony that prompted the choice of *[LABIAL]* as a feature operating for both consonants and vowels in Majang (section II.3). As can be seen from examples d) and e), both a round vowel and a labial consonant determine that the suffix vowel changes from /**i**/ to /**u**/. It is best to assume that the rounding of the vowel is the result of the same phonological rule (not two different rules), and describing this one rule is easier when it makes reference to only one phonological feature. Another crucial assumption is that this feature *[LABIAL]* has a suprasegmental nature, determining that a whole syllable rhyme takes over the feature *[+LABIAL]* when it is infected with it anywhere. In the labial-harmony rule it is therefore necessary to make reference to the syllable rhyme as a phonological unit:

Rule 4: labial harmony on verbal suffixes and epenthetic vowels

V	→	**V**	/	*R*	+	__
[+HIGH]		*[+ LABIAL]*		*[+ LABIAL]*		
[-LABIAL]						
[-LONG]						

This rule states that a suffix morpheme containing the short vowel /**i**/ changes it to /**u**/ when following a morpheme ending in a syllable rhyme containing a labial sound. This does not apply to the long vowel /**iː**/. The *1P* suffix **-ǐː**^{L} in the verb **ŋèɗɛ̀ːmíː**^{L} *'we smile'* is not affected by the rule. As stated in section II.1.3 above, it is not always easy to determine whether a particular suffix contains a long vowel or a short vowel. One of these suffixes is the class marker of i-class verbs in some complex verb forms, as in **gáɓíːgíɗɛ̀ {gáɓ-íː-gíɗ-ɛ̀}** *'it was given to'*. If this were a short vowel, however, the labial-harmony rule would render this word as ***gáɓúːgúɗɛ̀**. The fact that this form does not materialize is taken as proof that this vowel of doubtful length must be a long one.

Still, it needs to be pointed out that this labial-harmony rule is restricted to verbal suffixes and epenthetic vowels. No nominal suffixes were encountered that are subject to this phonological rule. The only exception to this is the possessive marker **-k**, which is only used on some nouns, and which is often introduced by the epenthetic vowel /**i**/. This epenthetic vowel is always subject to labial harmony, also when it precedes the sentence-final topicality clitic **=ŋ** and the subordination clitic **=k**.

II.5.4 Vowel-height harmony for ɛ-class verb suffixes

Another phonological rule is restricted to verbs of the ɛ-class (see section ɛ-class verbs, p. 248 below); the class-specific verbal-suffix allomorphs that begin with a mid-front *[-ATR]* vowel /**ɛ**/ will change this vowel to its *[+ATR]* counterpart /**e**/ if the suffix follows a syllable containing a high vowel.

Example II.22: vowel height harmony on ɛ-class suffixes

a)	*3S* suffix	**{kɛ̀ːt-ɛ́}**	*he scatters*	**{mìːɲ-é}**	*he covers*
b)	*2P* suffix	**{kɛ̀ːt-ɛ̌ːr}**	*you*$_{PL}$ *scatter*	**{mìːɲ-ěːr}**	*you*$_{PL}$ *cover*
c)	*3P* suffix	**{kɛ̀ːt-ɛ̀r}**	*they scatter*	**{mìːɲ-èr}**	*they cover*
d)	*INF, NEG*	**{kɛ̀ːt-ɛ̀ːt}**	*to scatter, scattering*	**{mìːɲ-èːt}**	*to cover, covering*

This vowel-height harmony process is morphologically restricted to some suffixes applying on ɛ-class verbs[18] and does not apply elsewhere in the language.

[18] The impersonal marker **-ɛ́ꜜ**, for example, is not affected by this rule, including when it appears on ɛ-class verbs.

Rule 5: vowel-height harmony for ε-class suffixes

$$\textbf{ε} \rightarrow \textbf{e} \;/\; \underset{[+\textit{HIGH}]}{V} \;\; (C^2) \;\; + \;\; \_\_$$

This height-harmony rule states that a suffix morpheme containing the vowel /**ε**/ changes this to /**e**/ when following another morpheme having a high vowel /**i**/ or /**u**/ in its final syllable. This rule is not sensitive to vowel length, as opposed to the labial-harmony rule. Example II.22 shows that it applies to both long and short suffix vowels.

II.6 Tone

Majang has two distinctive tone levels, which are called high and low (*H* and *L*). High tones can be both automatically and non-automatically down-stepped in various situations (see section II.6.2). Falling and rising tones are created by combining high and low tones on the same syllable, which happens only at the end of a phonological word.

Roots of nouns and verbs are accompanied by different tonal melodies, as seen in the following examples. The number following the tone pattern indicates the frequency of occurrence in a 485-item sample of noun stems.

Example II.23: noun-tone melodies, as seen on bisyllabic and monosyllabic roots

a)	*H*	(186)	**múɲá**	*earthworm*	[¯ ¯]	**póɲ**	*group*	[¯]
b)	*L*	(83)	**dàrì**	*sky*	[_ ˎ][19]	**tàŋ**	*abscess*	[ˎ]
c)	*LH*	(84)	**dàlí**	*hump*	[_ ¯]			
d)	*HL*	(44)	**dúbì**	*moth*	[¯ ˎ]	**tɔ̂ːn**[20]	*child*	[╲]
e)	H^L	(82)	**bógó**^{L}	*stutter (N)*	[¯ ¯]	**ɲón**^{L}	*lie (N)*	[¯]

[19] Pre-pausal low tones have a slight falling contour in Majang, which is of a purely phonetic nature. See also example d).

[20] A *HL* melody was only found on monosyllabic noun roots with a long vowel. Some ergative noun stems, such as **tɔ̂n** *'child (ERG)'* have short vowels with a *HL* sequence. *LH* melodies do not exist on monosyllabic noun roots at all.

Example II.24: verb-tone melodies, as seen on bisyllabic roots[21]

a) *H* **ɗégér** *know* [¯ ¯]
b) *L* **ɓàdὲj** *break* [_ ﹨]
c) *LH* **ràgád** *prepare* [_ ¯]

As can be seen here, the noun roots appear to be tonally more varied than the verb roots, which only come in the three different patterns *L*, *H* and *LH*. The pattern *L* occurs 235 times in a 355- item sample of verb stems, *H* 98 times, and *LH* 22 times.

Although all lexical roots are accompanied by lexical tones, some affixes are underlyingly toneless, and then copy the tone of the preceding syllable (see section II.6.3).

Based on the available evidence, it is very difficult to decide which of the two tones is marked or unmarked. The frequency of the tones certainly does not help in this matter. In the 2879-word text sample, the high tone appears only slightly more frequently on short vowels (1894 times) than the low tone (1847 times). On long vowels the frequency differs, though, with 668 high tones, compared to only 366 low tones. Contour tones are much less frequent, with 55 falling tones (*HL*) on short vowels balanced by 76 on long vowels, and 21 rising tones (*LH*) on short vowels compared to 29 on long vowels.

In order to read the examples given in this section, it may be necessary to first understand the tone-orthography conventions used in this study. These are explained in section II.9.

II.6.1 Tone association rules

Most morphemes have inherent tone melodies. The association of tones to the available syllables happens from left to right. The language counts syllables, not morae, so that a non-final long vowel regularly only receives one tone during the association process.

Example II.25: tone association

a) **{dùgíɗíːk-ǐː-ŋ}** *we hide ourselves* [_ ¯ ¯ ⁄]
| |⋰ ⋀
L H *L H*

[21] The pattern *LH* is not found on monosyllabic roots.

b) **á:tùj** *razor* [¯ ˎ]
| |
H L

Example a) displays the association of one tone to two syllables, and the association of two tones to one syllable.

Example b) is shown as evidence that the Majang language counts syllables, not morae. If morae were counted, the *L* would already associate with the second component of the long vowel of the first syllable, resulting in ***â:tùj**. But this word would violate the aforementioned constraint regarding contour tones on non-final syllables. This constraint is evidence against a mora-counting nature of Majang, as these two principles are not compatible with each other – a language counting morae would require contour tones being able to occur on non-final syllables. In any case, all tone associations of Majang work very well under the assumption of a syllable-counting system.

II.6.2 Downstep

Majang uses downstep, both automatic and non-automatic. Whenever a *H* follows a surface *L*, its register is lowered, so that the *H* is pronounced on a lower pitch than any preceding *H* in the same clause. At clause boundaries, the register is reset. This is called automatic downstep.

There are also many instances, however, when a high tone following a high tone has its register lowered. There appear to be two different causes for this:

1. Some morphemes are accompanied by a floating low tone, which is not audible on the word itself, but which lowers the register of the following word. This non-automatic downstep usually[22] operates across word boundaries.

Example II.26: non-automatic downstep caused by floating low tones

a) **béá**^{L} **ná:k** *my spear* [¯ ¯ –]
|/ |
H L H

b) **bólból ná:k** *my calf of leg* [¯ ¯ ¯]
\/ |
H H

[22] Example IV.24 presents an analysis where the floating *L* is assumed to manifest itself on an attached suffix.

In both examples, a noun with a final *H* is followed by the *1S.SG.ABS* possessive pronoun, which also has an inherent *H*. In example a), this high tone on the pronoun is downstepped, whereas in example b), it is on the same register as the preceding high tone. In order to account for this difference, it is necessary to assume that the noun in example a) is accompanied by some device that lowers the register for the rest of the clause. The noun in example b), however, is not accompanied by such a register-lowering device. For the purposes of this section, it is not important to decide whether this device lowering the register is part of the preceding noun stem or an affix that is only attached to the nouns of a particular inflection class (see the section on singular marking on p. 150 for more information). In either case, the proposed analysis assumes that the noun in example a) is accompanied by a floating *L* that causes the lowering of the register on the following possessive pronoun. This floating *L* is absent in example b), and no phonological reason beyond posing such a floating *L* can be identified that determines which noun causes a downstep and which doesn't. In this study, I call this kind of non-automatic downstep *word downstep*, as it crosses word boundaries. In the 2879-word text sample, this kind of downstep occurs 882 times. As this floating low tone is a part of the lexical material of the utterance, it is in this study always written (with the superscript ^{L}), even if it does not have a phonetic effect preceding words beginning with a low tone.

2. When two separate underlying high tones come together within a word across morpheme boundaries, the first syllable of the second morpheme is affected by a lowering of the register. This happens in order to avoid an obvious violation of the Obligatory Contour Principle. Cahill (2004, p. 5) lists a number of other African languages (Kishambaa, Supyire and Namwanga) which share this behavior. The situation in Namwanga particularly mirrors the evidence in Majang (Bickmore, 2000, p. 302f). This non-automatic downstep only operates between morphemes within a word. Two high tones meeting across word boundaries do not trigger downstep if no floating *L* is involved (see example II.26b). The analysis proposed here does not assume the presence of a floating *L* between the two underlying high tones triggering the morpheme downstep. Such a floating *L* is present neither underlyingly, nor as a product of a phonological rule creating the downstep. The downstep is a lowering of the register, and is not caused by the insertion of an unrealized low tone. This agrees in principle with Odden's (1986, p. 366) interpretation of the

facts in Kishambaa. This form of non-automatic downstep is therefore fundamentally different from the word downstep introduced above, and is accordingly indicated by a different orthographic symbol (ꜜ).

Example II.27: non-automatic downstep within a word

cóːꜜmój *tree, sp.* [¯ –]
| |
H H

This noun consists of the root morpheme plus the non-productive nominalizer **-ój** (see example IV.125). I call this kind of non-automatic downstep *morpheme downstep*, as it crosses morpheme boundaries only, but never word boundaries. This kind of downstep is much less frequent than the word downstep, with only 87 occurrences in the 2879-word text sample.

Both types of non-automatic downstep may involve the same word, as the following noun phrase illustrates:

Example II.28: combination of downsteps

wákáꜜcákɛ́ᴸ **gáːnɛ́** [¯ ¯ – – – –]
{wákác-ák-ɛᴸ **gáːn-ɛᴸ}**
crossroads-PL-LOC *POSS\1S.PL-LOC*
at my crossroads

This example has two downsteps operating on the noun **wákáꜜcákɛ́ᴸ**. The first downstep following the stem **wákác** is caused by the addition of the plural suffix **-ák**, which has its own underlying high tone. The high tone on **wákac** spreads out over both syllables of the stem and therefore stays on the same register. The presence of the new *H* on **-ák** is announced by the morpheme downstep between the two morphemes. The *H* on **-ák** spreads out to the next morpheme, the toneless locative marker **-ɛᴸ.** As no new tone is involved, the register stays the same on the two syllables **-ákɛ́ᴸ**. Although the locative marker **-ɛᴸ** is toneless as such, and therefore available for the copying of the *H* on **-ák**, it is accompanied by a floating low tone. This floating *L* then causes the *H* on **gáːnɛ́ᴸ** to be downstepped, according to the principle of word downstep.

Although minimal pairs entirely based on non-automatic downstep are rare, they do exist. Quite frequently the contrast is used for the paradigmatic distinction between absolutive nouns *(ABS)* and nominative nouns *(NOM)*.

Example II.29: minimal pair based on downstep

a) **ɗɛ̀nà wár^{L} kɛ́kàr.** [_ _ ¯ ¯↘]
{ɗɛ̀n-à wár^{L} kɛ́kàr}
see-1S.CJ dog\SG.ABS again
I see a dog again.

a) **káwɗíː^{L} wár kɛ́kàr.** [¯ ¯ ¯ ¯↘]
{káw-ɗiː^{L} wár kɛ́kàr}
bite-*AP.3S dog\SG.NOM again*
The dog bites again.

In both examples, the noun **wár** *'dog'* is followed by the same adverb **kɛ́kàr**, which begins with a *H* and therefore serves well as an indicator for downstep caused by preceding words. Downstep of **kɛ́kàr** happens in example a), but not in example b), which illustrates that the forms for absolutive and nominative can be subtly different for some nouns (other nouns do not display that difference through downstep, if at all – see section IV.1.3.1 for details on how the various noun forms are distinguished morphologically).

The two different kinds of non-automatic downstep in Majang provide an excellent tool for determining word boundaries in Majang. This is illustrated by the following examples, where the absence of a downstep between the words of example d) proves to be crucial:

Example II.30: downstep and word boundaries

a) **ɗéːgárìː** [¯ ¯↘] *we sleep*
{ɗéːgár-ǐː^{L}}
sleep-1*P.DJ*

b) **ɗéːgárìː kɔ́** [¯ ¯ _ ¯] *we just slept*
{ɗéːgár-ǐː^{L} kɔ́}
sleep-1P.DJ RECPST

c) **ɗéːgáràrí** [¯ ¯ _ ¯] *you$_{PL}$ sleep*
{ɗéːgár-àrí}
sleep-2P.DJ

d) **ɗéːgáràrí kɔ́** [¯ ¯ _ ¯ ¯] *you$_{PL}$ just slept*
{ɗéːgár-àrí kɔ́}
sleep-2P.DJ RECPST

These examples show that the recent past marker **kɔ́** always carries a high tone, regardless of the preceding tones, which implies that the high tone is part of the underlying representation of this marker. It is also shown that the *H* on **kɔ́** is not downstepped following a high-toned morpheme (example d). This proves that the morpheme downstep does not apply to the marker **kɔ́**, which in turn proves that **kɔ́** is neither a suffix nor a clitic, as it was analyzed by Bender (1983, p. 134) or Unseth (1989b, p. 108), but an independent word (the same can be assumed, by extension, for all other tense markers of the language).

Unfortunately, the neat picture presented so far is disturbed by a number of verbal paradigms where the application of the morpheme-downstep rule does not suffice to explain the tonal variation between stems lexically ending on a high tone and those ending on a low tone. The following examples are taken from the relative-past paradigms presented in example IV.165.

Example II.31: unexplained tonal behavior in some complex verbal paradigms

a) **{pèd}** *'finish'* *L*-melody
1s **pèdíːɗá**^{L} **{pèd-íːɗ-a**^{L}**}** *after I had finished*

b) **{ìbáːl}** *'play'* *LH*-melody
1s **ìbáːlíːɗá**^{L} **{ìbáːl-íːɗ-a**^{L}**}** *after I had played*

Both examples take i-class stems ending on a consonant and add the relative-past marker **-íːɗ**, followed by the toneless *1s* person marker **-a**^{L}. Example a) gives the impression that the relative-past marker has an underlying high tone, which is realized without modification following the low-toned stem **pèd** *'finish'*. But following a stem ending on a high tone, as in example b) following **ìbáːl** *'play'*, the suffix high tone is not downstepped, as one would expect after the application of the morpheme-downstep rule. These examples suggest the presence of a phonological word boundary between the stem and the relative-past marker, but such a boundary has no place in these examples, because the syllabification treats the relative-past verbs as single words – it is not reasonable to assume a phonological word boundary inside the syllables /**díː**/ (a) and /**líː**/ (b). One could also assume a tone merger of the two high tones as a strategy to avoid a violation of the Obligatory Contour Principle (OCP), but the language does not make use of this strategy elsewhere. Another unsatisfactory way to describe what is happening is to assume two tonal allomorphs of the relative-past marker: one allomorph has an underlying high tone in the environment of a stem-final low tone. The other allomorph is toneless in the environment of a stem-final high tone, and it therefore copies the stem tone without triggering the morpheme-downstep. This allomorphy rule is suffering from a lack of phonological motivation.

A better way to look at this assumes that morphemes with this behavior follow a tone-replacement rule that replaces all stem tones with the high tone of the morpheme until it is blocked by a stem *L*. The morpheme **-ɪ̋:ɗ** (from now on marked with the super-high tone marker to indicate its tone-replacement behavior) replaces the stem tones by its own high tone. In example II.31a), this is blocked by the low tone on the stem, and in example b), it succeeds for the second stem syllable, which has an underlying *H*, but not for the first syllable of the stem, which has a blocking *L*. The absence of word-internal downstep in this word makes any analysis that maintains the association of the original stem H untenable.

This behavior is restricted to the more complex verb paradigms involving subordinate-tense forms or direction markers (section IV.2.3). These will be marked by the super-high tone marker, just as some other morphemes (particularly infinitive and noun-plural markers) whose tone-replacement behavior is not blocked by low tones on the stem.

II.6.3 Toneless morphemes and polar tones

Not all morphemes of the Majang language have their own underlying tone. Although all stem morphemes of Majang and all particles have their own inherent tone or tone melody, some suffixes are not specified for tone in the lexicon.

Example II.32: toneless *3P.DJ* morpheme

a)	**ìbá:lír**^{L}	b)	**làŋìr**
	{ìbá:l-ir^{L}**}**		**{làŋ-ir**^{L}**}**
	play-3P.DJ		*find-3P.DJ*
	they play		*they find*

The *3P.DJ* morpheme **-ir**^{L} for the i-class verbs always behaves like in these examples, and in the same way the *3P.DJ* morpheme **-ɛr**^{L} for the ɛ-class verbs. Following a stem ending in a high tone (example a), it carries a non-downstepped high tone. Following a stem ending in a low tone (example b), it carries a low tone. This variation is best explained by a simple tone-spreading rule which assumes that the suffix is inherently toneless. The morpheme-downstep rule shown above does not allow the high suffix tone in example

a) to be analyzed as an underlying tone, as it would then be downstepped. It must be analyzed as the same tone as the high tone on the last stem syllable, spreading over to the suffix.

Example II.33: tone spreading to toneless morphemes

a)	**ìbá:lír**^{L}	b)	**làŋìr**
	{ìbá:l-ir^{L}**}**		**{làŋ-ir**^{L}**}**
	L H L		*L L*

As can be seen in these examples, the situation is somewhat complicated by the presence of a floating low tone on many of these toneless morphemes. This means that when the *3P.DJ* suffix **-ir**^{L} copies the preceding stem tone and this results in a high tone on the suffix, the next word will then be downstepped, which implies a word-downstep situation and therefore the involvement of a floating low tone. This is also the case in the following examples:

Example II.34: toneless *PL*-morpheme **-ɛ**^{L} with floating low tone

a)	**kúꜜrójɛ́**^{L}	**gá:nk**	b)	**kùtùrɛ̀**	**gá:nk**
	{kúꜜrój-ɛ^{L}**}**			**{kùtùr-ɛ**^{L}**}**	
	donkey-PL.ABS	*POSS\1S.PL.ABS*		*hog-PL.ABS*	*POSS\1S.PL.ABS*
	my donkeys			*my hogs*	

The plural suffix **-ɛ**^{L} copies the stem tone in both examples a) and b). But in example a), the high tone on the following possessive pronoun is downstepped. Since the tone on the suffix in example b) is already low, no non-automatic downstep is applied to the following possessive pronoun – the high tone on **gá:ŋk** is automatically downstepped compared to previous high tones in the utterance.

This behavior is not what one would expect according to established phonological universals. One of the constraints relating to tone according to *Autosegmental Phonology* is a strong dislike of languages for floating low tones and for toneless syllables. Such constraints[23] can of course be violated, as apparently the Majang language does here, but it is striking to see that in this situation, there should actually be no need for such a violation. The morpheme **-ɛ**^{L} is a syllable and has an available low tone attached to it; one

[23] Cahill (2004, p. 13) translates Goldsmith's (1976) *Well-Formedness Conditions* into *Basic Tone Mapping (Well-Formedness) Constraints*, of which the two constraints ***TONELESS** and ***(L)** are relevant in evaluating the behavior of Majang.

would therefore expect that this low tone just becomes associated with the syllable, in this way eliminating the offending presence of both the toneless syllable and the unassociated floating low tone. The Majang language instead opts for leaving the floating low tone unassociated, while dealing with the toneless syllable through tone copying from the stem. Other morphemes, such as the *3P.CJ* marker **-ɛ̀r** on **ɗɛ́jɛ̀r** *'they want'*, have a fixed low tone, which is always realized on the syllable itself. There is no apparent alternative to an analysis which assumes that there is an underlying difference between an associated low tone on some morphemes (such as on **-ɛ̀r** *'3P.CP'*) and a non-associated, floating low tone on other morphemes, such as the noun-plural marker **-ɛ**^{L}. The underlying representation of these two morphemes must be different, as they behave differently in word formation.

At least some phonologists claim that each floating tone *must* connect to an available toneless TBU. The *Tone Mapping Rule* in Halle and Vergnaud (1982, p. 67), which they adopted from Williams (1976), would clearly rule out Majang's behavior. One way to uphold the Tone Mapping Rule is an assumption that the floating *L* is created in the process of the tone-spreading. Such a conclusion might be drawn as a result of the fact that almost all toneless morphemes behave like the *3P.DJ* marker **-ir**^{L} or the plural marker **-ɛ**^{L} – they have an inherent floating *L*. This could invite an analysis involving a rule that somehow adds this floating low tone to a toneless morpheme whenever the resulting surface tone on the suffix is high. Unfortunately, this proposed rule is rendered untenable by the presence of a few toneless morphemes without a floating *L* when a *H* spreads on the toneless syllable. Some non-modified plural nominative suffixes, for example, clearly copy the stem tone without using a floating low tone, as in **càkómá** < **càkóm-a** *'friends (NOM)'* or **wártún** < **wár-tun** *'dogs (NOM)'*. There is therefore no rule in Majang that places a floating *L* following high-tone copying. All other attempts to maintain the constraint would require assumptions about tone replacement on what is here called a toneless morpheme, and such an analysis would make very complicated what can easily be described in terms of tone spreading. Other languages display similar violations, such as the Bantu languages Kikuria and Chiyao (Odden, 1995, p. 459), although the floating tone involved in those languages is a high tone. In the following, all morphemes with a similar behavior as the *3P.DJ* markers **-ɛr**^{L} and **-ir**^{L}, or the plural marker **-ɛ**^{L}, are represented in the examples as being toneless with an unassociated floating low tone. Other analyses of their behavior may be possible, but I leave these to be worked out by those who have a higher interest in maintaining the validity of the above universal that such morphemes should not exist in any language.

One particular morpheme of the Majang language, the *1P.DJ* suffix on the verb, also shows variation, but in a different manner than toneless suffixes.

Example II.35: polar tone on the *1P.DJ* suffix

a)	**tòníː**^{L}	b)	**gúːgúɲìː**
	{tòn-i̽ː^{L}**}**		**{gúːgúɲ-i̽ː**^{L}**}**
	tell-1P.DJ		*enter-1P.DJ*
	we tell		*we enter*

The *1P.DJ* suffix consistently has a high tone following a low-toned stem (example a), and a low tone following a high-toned stem (example b). If it receives a high tone, it again causes downstep on a following high-toned word. It is therefore assumed that the suffix in its underlying representation is equipped with both a polar tone and a floating low tone, and is therefore represented in the lexicon as **-i̽ː**^{L}, where the superscript cross symbol represents the polar tone.

For the two Gur languages Moore and Lama the authors Kenstowicz, Nikiema and Ourso (1988) demonstrated that polar tone goes back to an underlying high tone that is dissimilated when following another high tone as an instance of *Meeussen's Rule* (*HH → HL*), which applies to avoid violations of the Obligatory Contour Principle (OCP). If this were the case in Majang, then the underlying representation of the *1P.DJ* suffix would be **-íː**^{L}. It would keep its shape when following a low-toned stem, and when following a high tone, OCP would force it to be realized as **-ìː**. Such an analysis would have the advantage of providing the suffix with a more regular position in the paradigm, where most other person markers also have a high tone, such as *1S.DJ* **-á** or *2S.DJ* **-ín**. The conjoint forms of these, including the 1st person plural, have a fixed low tone: **-à**, **-ìn** and **-ìː**.

In spite of these benefits, such an analysis is not tenable for Majang. There are countless morphemes in Majang with a lexical high tone that do not undergo Meeussen's Rule to avoid a violation of OCP. These include the other members of the disjoint person paradigm listed above. When these follow a high-toned stem, they are subject to morpheme downstep, which is the language's device of choice to deal with OCP violations. It could be argued that the long vowel of the *1P* suffix may trigger a different strategy, but there are other long-voweled morphemes such as the disjoint impersonal marker **-ɛ́ː**^{L}

(as in **ɓɔ̀lɔ́ːꜜrɛ́ː**ᴸ *'one grows'*) that refute this idea. It undergoes morpheme downstep like all other high-toned suffixes. Consequently, the only way to describe the *1P.DJ* morpheme is to assume an underlying representation that includes its polar-tone behavior[24]. Majang can therefore be listed among the languages that support Newman's (1995, p. 775) and Cahill's (2004) idea that at least for some languages one can assume genuine tonal polarity.

II.6.4 The functional load of tone in Majang

Lexical minimal pairs with regard to tone are not difficult to find in Majang:

Example II.36: tone minimal pairs

a)	**cóːꜜmój**	*HꜜH*	[¯ ¯]	*tree, sp.*	vs.	**còːmój**	*LH*	[_ ¯]	*quiver*
b)	**táŋ**ᴸ	*H*	[¯]	*cow*	vs.	**tàŋ**	*L*	[↘][25]	*abscess*
c)	**ŋédán**	*H*	[¯ ¯]	*bee*	vs.	**ŋèdàn**	*L*	[_ ↘]	*tooth*
d)	**kǎːrŋ**[26]	*LH*	[/]	*he fights*	vs.	**káːrŋ**	*H*	[¯]	*they go*

On top of the heavy lexical functional load of tone, Bender (1983, p. 117) states that *"the principal function of tone seems to be grammatical"* in Majang, without going into any details. It is indeed true that some morphological processes are characterized by tonal changes to the stem of a noun or verb (see section IV.1.3 for more details). More importantly, tone expresses some important syntactic functions – sometimes even as the only means:

Example II.37: ergative and absolutive distinguished by tone

a)	**èːŋáɗîr**	**kùtûr**	b)	**èːŋáɗîr**	**kùtúr**
	{èːŋáɗ-îr	**kùtûr}**		**{èːŋáɗ-îr**	**kùtúr}**
	smell-CF.3S	*hog\SG.ERG*		*smell-CF.3S*	*hog\SG.ABS*
	a hog smells (tr.)			*he smells a hog*	

[24] See Cahill (2004) for an analysis of polar tone in the Gur language Kɔnni which also is not OCP-driven. As he works in the Optimality Theory framework, he creates a morphologically restricted constraint **POLAR** to anchor the process outside of the realm of pure phonology.

[25] Pre-pausal low tone usually is phonetically realized as a tone falling from low.

[26] Because contour tones are much less frequent than level tones, minimal pairs involving contour tones are hard to find. Example d) therefore compares two different lexical roots in two different morphological contexts.

In examples a) and b), the only difference between *hog* in the absolutive and in the ergative case is the tone, which is *L-HL* for the ergative and *L-H* for the absolutive.

Another syntactic function expressed by tone is the conjoint-disjoint marking on the verb. The conjoint form is used for non-topical absolutive NPs:

Example II.38: conjoint and disjoint marking:

a) **ŋù:lè béá^{L}.**		b) **ŋù:lé tɔ̂n béá^{L}.**		
{ŋù:l-è	**béá^{L}}**	**{ŋù:l-ɛ́**	**tɔ́n**	**béá^{L}}**
break-3S.CJ	*spear\SG.ABS*	*break-3S.DJ*	*boy\SG.NOM*	*spear\SG.ABS*
He broke the spear.		*The boy broke the spear.*		

The conjoint form of the *3s* suffix in example a) has a low tone. The disjoint form instead has a fixed high tone (example b).

These examples show that tone plays a major role in the Majang language, and needs to be considered very carefully in order to understand not only the phonology, but also the morphology and the syntax of the language.

II.7 Word Structure

Beyond what was stated so far on syllable and stem/root structures of Majang, only a few observations need to be added for phonological words.

Majang phonological words are in principle not restricted in length or number of syllables, although, of course, there are practical limitations, based on the needs of the speakers to form meaningful words with the word-formation processes available to them. Words consisting of five syllables, such as **ɗégégéríkì:** ‘*we know each other*’ are fairly frequent. On the other hand, very short words are common in Majang, such as the vowel-only conjunction **à**.

As already pointed out, the Majang language provides some clues about phonological word boundaries. The most important of these is the word-internal downstep rule which lowers the second of two adjacent underlying high tones within a word (see section II.6.2). The demarcating function of this rule was illustrated by example II.30. Another tonal word-boundary marker is a contour tone, which can only appear on the final syllable of a phonological word.

II.8 Sentence Tone

Apart from the two tone levels and downstep, Majang also displays some tonal phenomena which operate on higher levels than the phonological word. I call these phenomena *sentence tone* in order to distinguish them from less predictable emotive pitch phenomena which are often called intonation in the literature and are unhelpfully lumped together with what is called sentence tone here. This study does not address intonation in this narrower sense of the word, but restricts the description to phenomena that can be linked to the syntactic structure of the utterance.

II.8.1 Connecting sentence tone

In complex sentences, a non-final clause and even a non-final noun phrase can be marked by a final sharply rising tone, which also resets the register to a higher position. This seems to happen at all clause boundaries; whether a particular phrase is involved in this appears to be less predictable. The function of this sentence tone is to warn the speaker that the sentence has not reached its end, but more is yet to come.

The following example from a natural text shows the effects of connecting sentence tone:

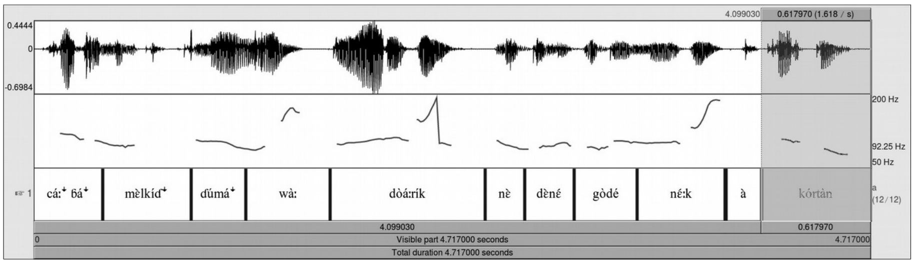

Figure 7: effects of connecting sentence tone

The sentence in the example above consists of several clauses and phrases:

Example II.39: connecting sentence tone

nɛ̀ cá:^{L} mɛ̀lkíɗ^{L} ɗúmá^{L} wà: dòá:rík, nɛ̀ dɛ̀nɛ́ gòdé nɛ́:k à kórtàn.

nɛ̀	**cá:^{L}**	**mɛ̀l-k-íɗ^{L}**	**ɗúmá^{L}**	**wà:**	**dòá:r=k**
CONJ	*then*	*arrive*-ANT-CP.3S.DJ	*owner*\NOM.SG.MOD	*house*\DAT.SG	*hunt*=SUB

nɛ̀ dɛ̀n-ɛ́ gòdé nɛ́ːk à kórtàn
conj see-3s.dj house poss\3s.sg.abs conj door\sg.abs

When the owner came home from hunting, he saw that his house was closed.

At the end of a non-final clause, the pitch rises to a frequency far exceeding that of a regular *H*. In the case of the first phrase, which ends in the low tone word **wàː** *'house (dat)'*, the resulting pitch contour extends from less than 90 Hz going all the way up to 170 Hz. The next two non-final units end with high tones on syllables with a consonantal coda. In these cases an extra syllable is included, consisting of the epenthetic vowel [**i**], with a pitch contour from 140-200 Hz and 130-190 Hz respectively. The beginning frequency of a high tone is usually between 120 and 130 Hz with this speaker. This tonal behavior of non-final clauses and phrases was observed with several different speakers and seems therefore to be well established in the grammar.

II.8.2 Interrogative sentence tone

Interrogative clauses receive a sentence-tone marking similar to the connecting sentence tone, with pitch levels rising well above the normal high tone frequency. This can be observed in the following example:

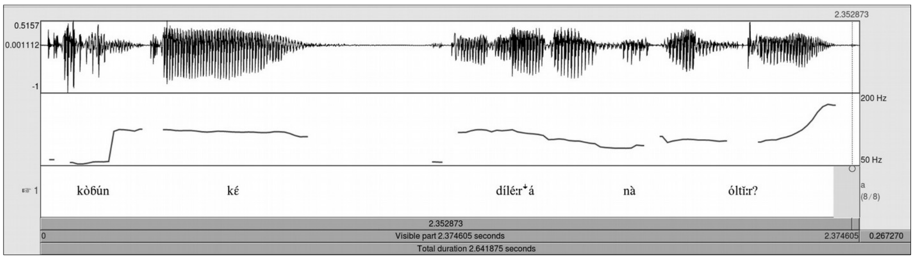

Figure 8: interrogative sentence tone

Example II.40: interrogative sentence tone

kòɓún kɛ́ díléːꜜrá nà óltíːr^{L}?
kòɓ-ín kɛ́ díl-eːr-á nà ólt-íːr^{L}
think-2s.dj quot bring-cf.1s.dj 2s.dat.cj fish-pl.abs
Do you think I bring you fish?

In this example of a yes-no question, the final word **óltíːr^{L}** *'fish (pl)'* has a *H* on both syllables. This *H* again rises to 175 Hz, a frequency normally out of range for regular high tones. This interrogative sentence tone is not only used for yes-no questions, but also for questions with interrogative pronouns, which in Majang usually appear at the end of the clause and therefore are the carrier of this sentence tone.

II.9 Transcription and Orthography

The orthography of Majang has not been officially settled, although literature production is already under way. For the purposes of this study, the following conventions are used to unambiguously transcribe all phonological distinctive features:

The consonants and vowels are represented as shown in the phoneme charts (Tables 1 and 3). Long vowels are represented by the vowel plus the phonetic length symbol, e.g. <**aː**>, <**ɔː**>; two identical consonants meeting at morpheme boundaries are represented by the doubling of the consonant character, e.g. <**kk**> – they are pronounced as geminated consonants. As seen in section II.1.3, the representation needs to distinguish between long vowels and sequences of two identical vowels, which are in this study indicated by doubling the vowel symbol, instead of using the IPA length marker.

Tones are represented according to the following conventions: a high tone is marked by an acute accent (e.g. <**á**>), and a low tone by a grave accent (e.g. <**à**>). The tone is marked where it is associated after the application of all lexical rules. Therefore it is not necessary to indicate automatic downstep. Automatic downstep happens when a low and a high tone are in sequence. Non-automatic morpheme-downstep is indicated by the symbol <ꜜ>, and floating low tones by <^{L}>. Rising tones (*LH*) are marked by the symbol <**ǎ**>, and falling tones (*HL*) by the symbol <**â**>.

Some morphemes of Majang are inherently toneless and copy the tone of the preceding syllable. In the surface representation used for most examples, these morphemes are shown with the resulting surface tones. But when reference is made to the morphemes as such, they are shown without any tone marks. Similarly, one morpheme, the *1P.DJ* suffix **-i̽ː**^{L} takes the opposite tone of the preceding syllable (see section II.6.3). In the surface representation used in this study, this suffix is shown with the resulting surface tone. But when reference is made to the suffix as such, it is shown as above with a superscript cross, to indicate that this suffix has a polar tone underlyingly.

A falling tone can be a tone falling from the same level as the *H* on the preceding syllable, or from a downstepped *H*. This is, for example, an important distinction between some conjoint and disjoint *3S.TF* verb forms, as in **ŋáːrgêːɗ** [¯ ＼] < **{ŋáːr-gêːɗ}** *'go-TF.3S.DJ'* and **ŋáːrꜜgêːɗ** [¯ ＼] < **{ŋáːr-gê᷆ːɗ}** *'go-TF.3S.CJ'*. This is caused by the different behaviors of the two morphemes; the first copies the *H* of the preceding syllable and builds the contour from that, and the second builds the contour from its own newly introduced *H*. To show the difference between the two forms of the *TF.3S* forms, the one starting with a copied *H* has the marker <ˆ>, and the one with its own *H* has the marker <᷆>. This difference is shown orthographically only for the underlying representation, which does not show morpheme downstep.

A final complication is presented by suffixes such as the infinitive marker /**-ɛ́ː**^{L}/, which places a single high tone on the complete word. To distinguish morphemes of this kind from those which copy an existing high tone, and from those which just have their own underlying high tone without affecting the word stem, any reference to such a suffix makes use of the super-high-tone marker <˝>, so that the underlying representation of the above infinitive marker is /**-ɛ̋ː**^{L}/. The same notation is used for morphemes with a tone-replacement behavior that is blocked by low tones on the stem, such as the relative-past markers **{-ɪ̋ːɗ}** and **{-ɛ̋ɗ}** seen in examples IV.165 d-i).

Majang publications that already exist use a practical orthography that is quite close to the one used here, with the following differences:

- Tone has so far not been marked in the orthography, but the Gambella Region authorities are considering ways in which the most important grammatical distinctions can be written.
- <**s**> is used for /**c**/, and <**j**> for /**ɟ**/, and <**y**> is used for /**j**/.
- The implosives are often written as digraphs <**bh**> and <**dh**>.
- Long vowels and double consonants are written by doubling the respective character.
- Publications vary in their treatment of the vowels /**ɛ**/ and /**ɔ**/, and the nasals /**ŋ**/ and /**ɲ**/. Some do not distinguish the open-mid vowels from /**e**/ and /**o**/, and they treat the nasals as digraphs <**ng**, **ny**> (see the texts in Getachew (2014) for this orthography). Other publications use special characters <**ɛ**, **ɔ**, **ŋ**, **ɲ**> for these sounds, which require advanced technical solutions to be available to the typists.