Chapter 2
Segmental phonology

2.1 Consonants

2.1.1 Consonant phoneme inventory

Kokota presents a set of consonant phonemes which are remarkably symmetrical. Three place classes exist distinguished by the features [±labial] and [±coronal]. While some fine grained place variation exists within these classes, each class may be characterised as follows:

(2.1)	[+labial, -coronal]	(bilabials and labiodentals)
	[-labial, +coronal]	(post-alveolars)
	[-labial, -coronal]	(velars and glottals)

Five manner classes exist: two obstruent classes (plosive and fricative), and three sonorant classes (nasal, lateral and rhotic). Of these, plosives, fricatives and nasals occur in each of the three place classes. Laterals and rhotics occur in only one place. Kokota is unusual in that a corresponding voice pair exists in every place/manner class. This means that Kokota has a full set of voiceless counterparts to every voiced consonant phoneme, including sonorants. There are thus 22 consonant phonemes in 11 place and manner pairs. The consonant phoneme inventory is as follows:

Table 2.1: Consonant phonemes.

	[+labial, -coronal]		[-labial, +coronal]	[-labial, -coronal]	
	Bilabial	Labiodental	Post-alveolar	Velar	Glottal
Plosive	p b		t d	k g	
Fricative		f v	s z	ɣ	h
Nasal	m̥ m		n̥ n	ŋ̊ ŋ	
Lateral			l̥ l		
Rhotic			ɾ̥ ɾ		

2.1.1.1 Consonant phoneme feature matrices

The crucial feature distinctions for these consonants give the following matrices:

Table 2.2: Consonant feature matrices.

	p, b	t, d	k, g	f, v	s, z	h, ɣ	m̥, m	n̥, n	ŋ̊, ŋ	l̥, l	ɾ̥, ɾ
labial	+	-	-	+	-	-	+	-	-	-	-
coronal	-	+	-	-	+	-	-	+	-	+	+
continuant	-	-	-	+	+	+	+	+	+	+	+
sonorant	-	-	-	-	-	-	+	+	+	+	+
nasal							+	+	+	-	-
lateral							-	-	-	+	-
rhotic							-	-	-	-	+

2.1.1.2 Evidence for consonant phoneme status

The following tables demonstrate consonant phoneme contrasts by voicing status, and manner and place of articulation.

Table 2.3: Contrastive sets demonstrating voicing and manner of articulation distinctions.

	labials	coronals	nonlabial noncoronals
voicing	/putu/ 'Barringtonia tree' /butu/ 'stomach'	/tafa/ 'encounter' /dafa/ 'clear by burning'	/kulu/ 'be first' /gulu/ 'thunder'
	/fila/ 'thunderclap' /vilai/ 'knife'[*]	/siku/ 'lawyer cane' /ziku/ 'cone shell'	/tahina/ 'his/her salt water' /taɣina/ 'himself, herself'
	/nom̥i/ 'hear' /nomi/ 'ourEXC'	/n̥iɣo/ 'completive aspect' /niɣo/ '2SG object'	/naŋ̊o/ 'class of in-law' /naŋo/ 'shadow'
	-	/bubul̥i/ 'clam sp.' /bubulo/ 'morning mist'[*]	-
	-	/ɾ̥uta/ 'untangle' /ɾuta/ 'inedible swamp taro'	-
plosive/ fricative	/fofoto/ 'tree sp.' /popoto/ 'fish sp.'	/papase/ 'ginger, turmeric' /papate/ 'crustacean sp.'	/gasi/ 'torch' /ɣase/ 'girl'[*]
	/vavau/ 'kapok tree' /babao/ 'be tired'[*]	/izo/ 'point' /ido/'mother'	/huhu/ 'question' /kuku/ 'defecate'
plosive/ nasal	/mahai/ 'eat' /bahai/ 'tree sp.'	/na/ '1EXC subject realis' /da/ '1INC subject'	/ŋa/ 'but' /ga/ '1EXC subject neutral'
plosive/ rhotic	-	/ade/ 'here' /aɾe/ 'those'	-
lateral/ rhotic	-	/ɾeɾeo/ 'shield' /leleo/ 'tawny shark'	-

Table 2.4: Contrastive sets demonstrating place of articulation distinctions.

	labials	coronals	nonlabial noncoronals
voiced plosives	/ba/ 'Possibilitative marker'	/da/ '1INC subject'	/ga/ '1EXC subject neutral'
voiceless plosives	/puku/ 'be short'	/tuku/ 'wait'	/kuku/ 'defecate'
voiced fricatives	/viviviɾi/ 'propeller' /vilai/ 'knife'	/ziziɾi/ 'tree sp.'[*] /izo/ 'point'	/iɣo/ '2SG object' /ɣilai/ 'until'
voiceless fricatives	/fodu/ 'be full'	/sodu/ 'be tall' /soda/ 'clam'	/hoda/ 'take'[*]
voiced nasals	/ma/ 'father'	/na/ '1EXC subject realis'	/ŋa/ 'but'
voiceless nasals	/m̥am̥aɣu/ 'be fearful' /m̥em̥e/ 'be difficult'	/n̥an̥aɣaɾai/ 'lobster sp.'[*] /n̥ae/ 'be clear'	/ŋ̊eŋ̊e/ 'be separate' /ŋ̊au/ 'eat'[*]

2.1.1.3 Consonant phoneme frequencies

On the basis of a representative sample of Kokota words the relative frequencies of consonant phonemes has been calculated. The word list contains 335 basic lexical items giving a total of 748 consonant phoneme tokens. The word list and the raw frequency figures for each phoneme is contained in Appendix 2. These figures give the following frequency percentages.

[*] No minimal pair has been identified. This near minimal pair demonstrates the contrast.

2.1.1.3.1 Proportion of consonant phonemes by place and manner of articulation

Table 2.5: Proportion of consonant phoneme tokens by manner of articulation.

Plosives	Fricatives	Nasals	Laterals	Rhotics	Total
37.7% (282)	27.3% (204)	17.2% (129)	8.2% (61)	9.6% (72)	100% (748)

Table 2.6: Proportion by place class and proportion of obstruents to sonorants.

	[+labial]	[+coronal]	[-labial, -coronal]	Total
Obstruents	15.5% (116)	20.5% (153)	29.0% (217)	65.0% (486)
Sonorants	6.0% (45)	27.5% (206)	1.5% (11)	35.0% (262)
Total	21.5% (161)	48.0% (359)	30.5% (228)	100% (748)

These figures are interesting in that while coronals are well represented as both obstruents and sonorants, the nonlabial noncoronals are the smallest class of sonorants, but the largest class of obstruents. As discussed in 2.1.2.3.3, nonlabial noncoronal fricatives frequently lenite to zero in casual speech, so the actually proportion of obstruents in this place class is smaller in normal discourse. Further, it seems possible that this class of fricatives is in the process of being lost. If this process was complete, the figures for nonlabial noncoronal obstruents would be almost halved to a figure lower than that for coronal obstruents. Nonetheless, due to the very large numbers of plosive tokens in this class the figures would remain high.

Broadly speaking the language displays a preference for the production of obstruents to tend towards the back of the mouth but for the production of sonorants to tend towards the front.

2.1.1.3.2 Proportion of consonant phonemes by voicing and sonority

The proportion of voiced to voiceless tokens overall is roughly equal (51.7% (387) voiced, 48.3% (361) voiceless). However this does not accurately reflect the voicing situation, as a further breakdown of the figures into obstruent and sonorant categories shows:

Table 2.7: Proportion by voicing and sonority.

	Obstruents	Sonorants
Voiced	31.7% (154)	88.9% (233)
Voiceless	68.3% (332)	11.1% (29)
Total	100% (486)	100% (262)

2.1.1.3.3 Summary of consonant phoneme frequency proportions

These figures reflect a marked preference for voiceless obstruents and for voiced sonorants. Overall the language displays a strong preference for voiceless obstruents weighted towards the back of the mouth, and for voiced sonorants weighted towards the front.

2.1.2 Consonant phonemes

In this section each phoneme will be briefly described.

2.1.2.1 Sonorants

As indicated above, a voiced and voiceless pair occurs for each place and manner category of sonorants. Nasals are represented at each of the three place classes. There are thus labial (bilabial), coronal (post-alveolar), and nonlabial noncoronal (velar) nasals. Laterals and rhotics are produced only in post-alveolar place.

The rhotics are underlyingly taps, and are almost always realised on the surface as such. However when they immediately precede a lateral they are realised as a rhotic approximant. This adjacency occurs in rapid speech as the result of vowel syncope:

(2.2) /aɾe+lau/ → /aɾelau/ → /aɾlau/ → [aɹlau]

2.1.2.2 Plosives

Within the class [-sonorant, -continuant] a voiced and a voiceless plosive exist in each place class, produced as bilabial, post-alveolar and velar. There is no prenasalisation of voiced plosives, and voiceless plosives are unaspirated in all positions.

2.1.2.3 Fricatives

As with plosives, within the class [-sonorant, +continuant] voiced and voiceless counterparts exist in each place class. However, place of articulation features for fricatives are not all identical to those of the corresponding plosives, and there is less symmetry than with other categories.

2.1.2.3.1 Labial fricatives

The labial fricatives are labiodental and differ from each other only in voicing status.

2.1.2.3.2 Coronal fricatives

The coronal fricatives are post-alveolar, being produced further back than the English equivalents. Some variation occurs in production of the voiced phoneme. When preceding a high front vowel /i/ the feature [+high] spreads from the vowel to /z/, increasing the length of obstruction:

```
(2.3)        [+high]
               / |
              /  |
            [z]  V
```

This effect is palatalisation of the /z/, however the resulting variant is not apico-palatal like the English /ʒ/. Instead it remains apico-post-alveolar, but the tongue body height is raised to its position for the vowel.

This process only occurs under the influence of a following, not a preceding vowel, the feature apparently only spreading within a syllable. However, it also does not occur word initially. Thus in (2.4)a. palatalisation occurs, while in (2.4)b. and c. it does not.

(2.4)	a.	/tazi/ 'lamp'	→	[taz^{j}i]
	b.	/ɣizuna/ 'lamp'	→	[ɣizuna]
	c.	/ziku/ 'lamp'	→	[ziku]

In addition, a lesser degree of tongue body raising seems to occur before other non-low vowels, possibly commensurate with the height of the vowel. Thus a small amount of raising occurs when /z/ precedes /e/, /o/ and /u/, with possibly slightly more raising before /u/ than the two mid vowels.

The extent to which variation in realising /z/ occurs differs between speakers from the same village. It is possible that this phoneme is unstable in Kokota and may be in the process of shifting place of articulation. The evidence from speaker age-based variation is at this stage inconclusive, though it may weakly suggest a move from [ʒ] to [z]. Alternatively this variation may be stable in the language, as it appears to be in a number of languages, Papuan and Oceanic, across northwest Melanesia. (Ross pers. comm.)

Variation with this phoneme is present in neighbouring languages. Fitzsimons (1989:16) reports that in Zabana the phoneme "has two allophones: the voiced dental fricative [**z**] and the alveo-palatal fricative [**ž**]. The two are in free variation, as in the following place name: [**taimaihaza**] ~ [**taimaihaža**]..." It is possible that his phonetic description of the variants may on closer inspection require some modification, however it is clear that variation occurs. While his comments suggest that there is no hard and fast phonologically motivated variation, the possibilities of greater or lesser tendencies in certain environments is not canvassed, nor is the possibility of age or regional variation.

In Cheke Holo more marked variation occurs, regarded by White et al as diachronic change in progress: "The voiced fricative /z/ appears to be disappearing from the language, increasingly replaced by /j/ [i.e. /dʒ/] in the idiolects of younger speakers." (1988:x) The direction of this change seems unlikely, given that in all three languages /z/ reflects the Proto Oceanic, and according to Ross (1988:221) also Proto Isabel, phoneme **j*. The actual basis and distribution of this variation in Cheke Holo therefore remains unclear.

The variation found with this phoneme in Kokota, and its relationship with variation found in Zabana and Cheke Holo, warrants further investigation.

2.1.2.3.3 Nonlabial noncoronal fricatives

The most considerable place differentiation within a place class pair applies to the [-labial, -coronal] class of fricatives. While the voiced member of the pair is the velar /ɣ/, the voiceless counterpart is the glottal /h/. This is primarily a distinction of tongue height and backness, with the tongue raised and backed ([+high, +back]) for the voiced member of the pair, but lowered ([+low]) for the voiceless member.

The most significant factor relating to this class is its widespread lenition to zero. This occurs both as regular morphophonemic deletion of /ɣ/, and as widespread phonetic deletion of both /ɣ/ and /h/, particularly in rapid speech. This appears to reflect a process of diachronic phoneme loss, at least of the voiced velar fricative.[1]

2.1.2.3.3.1 Phonetic deletion

Both /ɣ/ and /h/ optionally lenite to zero in casual speech, particularly rapid speech. This may occur morpheme medially, so for example /zoɣu/ 'drop' may be realised as [zou] and /glehe/ 'very' as [glee]. The greater the frequency of a word the more likely it is to display this lenition. So for example the ritualised greeting /gɾuɣu keli/ 'good night' is almost always realised as /gɾuu keli/. This is also found with the reflexive base in one person/number category. Reflexivisation is expressed by means of a reflexive base lexeme, to which inalienable possessor enclitics are attached (see 4.1.2.3). This base has the underlying form /taɣi-/, so for example /taɣi-mu/ 'yourself', /taɣi-di/ 'themselves' etc. However, for first person singular reflexive, the normal surface form is [tai-gu], to the extent that speakers usually give this as the citation form.

Deletion of this class of fricatives also occurs, though less frequently, with word initial segments, so /ɣilai/ 'until' may be realised as [ilai], and /huhuga/ 'lie' as [uuga]. Equally, it occurs occasionally word medially but morpheme initially, for example /n-e-ɣe/ 'RL-3-PRES' may occur as [nee].

However, it appears that the most widespread phonetic deletion of these phonemes is intervocalically between identical vowels, especially the high back vowel /u/. Articulation of these phonemes between identical vowels in casual speech ranges freely from full articulation, through momentary devoicing in the case of /h/, to complete deletion.

[1] In addition to this apparent loss of /ɣ/, a small number of attested examples exist of an apparent diachronic shift from /ɣ/ to /ŋ/, with the archaic forms /maɣava/ 'be hot' and /ɣoɣozi/ 'whistle' now replaced by /maŋava/ and /ŋoŋozi/. It is not clear whether these are idiosyncratic alternations, cognate borrowings, or evidence of a shift, with /ŋ/ in some other items also reflecting proto /ɣ/.

2.1.2.3.3.2 Morphophonemic deletion

Deletion of /ɣ/ occurs regularly as phonologically conditioned allomorphic variation in at least three morphemes, all with the underlying form /ɣu/. These are in an enclitic marking verbs with progressive aspect, a suffix attaching to numeral roots to form cardinals, and extremely high use verb 'be thus'.[2]

Phonological conditioning with the enclitic and the suffix is identical. The /ɣ/ lenites to zero in every environment except between identical vowels (effectively, except when the preceding vowel is /u/). So for example /zaho+ɣu/ 'go-PROG' is always realised as [zahou] and /kota+ɣu/ 'go ashore-PROG' as [kotau], but /au+ɣu/ 'be at-PROG' is normally realised as [auɣu] and /n-e-u+ɣu/ 'was being thus' as [neuɣu]. Equally, with numeral roots ending in vowels other than /u/ cardinals are formed with the suffix realised as /u/. So /tilo+ɣu/ 'three-CRD' is realised as [tilou] and /gaha+ɣu/ 'five-CRD' as [gahau], but /palu+ɣu/ 'two-CRD' is normally realised as [paluɣu].

Interestingly, when tested with verb surface forms such as /zahoɣu/ and cardinal forms such as /tiloɣu/, with the fricative present on the surface, informants identified the forms as acceptable, commenting that they were "very good Kokota", but observed that no-one pronounced the words in that way.

When referring to the lack of /ɣ/ deletion following /u/ I have said that the unlenited forms are "normally" realised. This is because while the morphophonemic process does not delete the segment in this environment, phonetic deletion may do so. Thus /au+ɣu/ may be realised in rapid speech as /auu/. The distinction between morphophonemic deletion and phonetic deletion is clear however. In individual elicitation of forms bearing the suffix or enclitic speakers will always give the full form when preceded by /u/, and never when not.

Phonological conditioning of /ɣ/ deletion with the verb /ɣu/ 'be thus' is similar to that of the enclitic and suffix. The form often functions as the head of a kind of tag clause consisting of a single word made up of the verb and a cliticised preverbal auxiliary coding modality and person. This means that the form normally does not occur word initially. Since none of the preverbal auxiliaries end with the vowel /u/ the fricative never occurs on the surface in these tags. Thus /n+e+ɣu/ 'it was thus' is always realised as [neu]. Occasionally the form does occur without an auxiliary as a monomorphemic word. Only in this environment is the verb realised as /ɣu/, with the fricative present on the surface.

2.1.2.3.3.3 Diachronic factors

The frequent synchronic deletion of this phoneme may reflect diachronic loss in progress. A similar phenomenon may be occurring in the neighbouring language Zabana. Fitzsimons (1989:17) notes that in Zabana: "The string **ghua** [i.e. /ɣua/] has undergone a historical sound change to become **wa** everywhere…" (His bold.) In one sense this is an odd claim, since he does not list /w/ as a phoneme in the language. He goes on to say: "The process is complete amongst most of the population, except a few of the older generation." Here the specific environment appears to have facilitated the loss. One possible explanation for this is that the fricative has lenited to zero, leaving the sequence [ua]. In Kokota the high back vowel /u/ is often realised as the glide [w] prevocalically. This is also true to some extent of Zabana, Fitzsimons reporting that it occurs when /u/ is "followed by a morpheme boundary and **i**." (p17) It seems plausible that this variation also occurs in other environments in Zabana. If so the shift described by Fitzsimons may be:

(2.5) /ɣua/ → /ua/ → [wa]

[2] On the basis of their respective functions it could be argued that the suffix and the enclitic actually reflect separate occurrences of a single morpheme.

The fact that Fitzsimons does not identify [w] as a phoneme supports this hypothesis. If it is correct, then the Zabana /ɣ/ is in the early stages of being lost. It is therefore possible that as a result of mutual influence or parallel development /ɣ/ is in the process of being lost in these languages.

2.1.3 Phonological processes involving consonants

2.1.3.1 Labialisation and velarisation

Consonant variation occurs in Kokota when a consonant anticipates rounding and/or backness features of a following high back vowel, these features spreading from the vowel to the consonant.

```
(2.6)              [+round]
                   [+back]
                   /   |
                  /    |
                C      u
```

It appears that all classes of consonants are affected in this environment. With consonants that already have the feature [+labial] the effect is velarisation, as in (2.7) a. and b. With consonants that already have the feature [+back] the effect is rounding, as in (2.7) c. and d.[3]

(2.7)	a.	/fufuɣo/ 'tomorrow'	→	[fˠufˠuɣo]
	b.	/m̥um̥ui/ 'be wet'	→	[m̥ˠum̥ˠui]
	c.	/kukuti/ 'eel'	→	[kʷukʷuti]
	d.	/ɣuɾa/ 'be boiling'	→	[ɣʷuɾa]

The effect on coronals, which are both [-labial] and [-back], is primarily rounding, accompanied by an increase in dorsal height.

(2.8)	a.	/tulufulu/ 'thirty'	→	[tʷulʷufˠulʷu]
	b.	/ɾ̥uɾ̥uta/ 'untangle'	→	[ɾ̥ʷuɾ̥ʷuta]

The degree of raising of the tongue body with both labials and coronals varies depending on whether the environment also involves a preceding /u/. Consonants which occur intervocalically where both vowels are /u/ involve a greater degree of tongue body raising then those which only precede /u/. So for example the second /f/ in (2.7) a. is produced with more dorsal height than the first.

Labialisation appears to apply to medial consonant clusters occurring between two instances of the high back vowel. So for example /bubluse/ 'be easy' appears to be realised as [bʷubʷlʷuse]. It does not, however, seem to be true of initial clusters. Thus /gɾui/ appears to be realised as [gɾʷui], not *[gʷɾʷui]. It is not clear what happens with medial clusters that precede but do not also follow /u/.

2.1.3.2 Palatalisation

Most consonant phonemes of the place class [-labial - coronal] undergo palatalisation under the influence of a following front vowel. It appears that the process does not occur with /h/. The process thus appears to apply to velar consonants only, not to the entire nonlabial noncoronal place class.

In this process the tongue position assimilates to some extent to the position of the following front vowel by moving forward slightly. However, while tongue position on upper articulator moves towards palatal position, the resulting allophone is very distinct from a palatal fricative since it remains dorsal. The resulting allophone is dorso-palatal, not lamino- or apico-palatal.

[3] The preponderance of the high back vowel throughout these examples is not significant, they were chosen to maximally display the phenomenon.

The degree of palatalisation appears to vary on two parameters: the height of the front vowel, and whether the preceding vowel is also a front vowel.

The effect of the first of these parameters is that the higher the front vowel the greater the degree of palatalisation. Thus a velar consonant before /i/ will tend to palatalise to a greater degree than before /e/. So for example the initial consonants in /ŋinoi/ 'stir', /ɣilai/ 'until' and /kilo/ 'finger, toe' are produced with the tongue fronted to a greater degree than the initial consonant in /ŋehe/ 'umbrella', /ɣeɾi/ 'beside' and /keli/ 'be good'.

The second parameter means that a velar consonant that follows as well as precedes a front vowel will palatalise to a greater degree than one that follows a non-front vowel. So a velar consonant occurring between two front vowels will be palatalised to a greater degree than one which only precedes a front vowel. So for example the second /ŋ/ in /ŋeŋe/ 'be separate' is more strongly palatalised than the first, as is the second /k/ in /kekeli/ 'be pleased'.

2.1.3.3 Glottal epenthesis

There is no phonemic glottal stop in Kokota. However, glottal epenthesis occurs in careful speech in a number of environments.

When carefully or emphatically producing words which commence with a vowel, a voiced plosive, or a rhotic tap, some speakers initiate the production with glottal closure. Thus when reading a word list, for example, these speakers will produce ohai 'be tame' as [ʔohai], biɾ̥o 'sleep' as [ʔbiɾ̥o], and ɾeha 'shout' as [ʔɾeha].

In casual speech glottal epenthesis may occur between vowels at a word or morpheme boundary, or in certain circumstances morpheme internally.

So for example /eu a lao/ 'thus we go' may be realised as [eu ʔa lao], and /na ohadi/ 'I feed them' as [na ʔohadi]. This frequently occurs between a preposed particle and a vowel initial root, for example between the preposed causative particle and a vowel initial verb. Thus /fa aui/ 'cause it to be present' may be realised as [fa ʔaui], and /fa ikoai/ 'make it small' as [fa ʔikoai]. Glottal epenthesis often occurs between a prefix, proclitic or reduplicated syllable and vowel initial roots. Thus /i+ipi/ 'be wearing clothes' (from /ipi/ 'wear (clothes)') is often realised as [iʔipi].

While intervocalic glottal epenthesis typically involves at least a morpheme boundary, it may occur between morpheme internal vowels in a syllable sequence where the second vowel is stressed. This is shown dramatically when a cliticised tag clause affects the stress so that an underlyingly unstressed second vowel in a sequence becomes stressed. For example in isolation the personal name /fáknoe/ has the stress pattern shown. However, once the tag clause /nekeu/ is cliticised, giving the sequence /faknoenekeu/ '…Faknoe, it was like that', primary stress shifts to the final vowel of the name. Glottal epenthesis may then intervene in the final vowel sequence of the name: [fàknoʔénekeu].

In careful speech glottal epenthesis also occasionally occurs following a vowel when the following syllable is stressed and has an initial non-continuant (i.e. either a plosive or a nasal). Again this typically occurs across a word or morpheme boundary, with /fa+tóli/ 'cause to be open' occasionally realised as [faʔtóli] and /gá kózei/ 'I/we sing it' as [gáʔ kózei]. Again it also occasionally occurs morpheme internally, but only in very careful or emphatic speech, /nakóni/ 'person', for example occasionally realised as [naʔkóni].

2.1.3.4 Glides

There are no glide phonemes in Kokota. The glides [w] and [j] occur phonetically as allophones of non-low vowels, as discussed in 2.2.4.3.

There are a small number of Pijin loans which have [w] in initial position. These are [wasi] 'wash', [wiki] 'week', [wokobaoti] 'stroll about' and possibly [wida] 'window'. Of these [wasi] has almost certainly been borrowed as an underlying /uasi/, with regular glide formation occurring on the initial segment. The extent to which this is true of the other examples is not clear at this stage. If they appear to be underlying glide initial, they may reflect code switching to Pijin, though this seems unlikely since [wiki] in particular occurs with high frequency. Alternatively these forms may reflect the development through borrowing of /w/ as a phoneme in Kokota, although if that is so this nascent phoneme is still highly marginal, occurring only in the examples cited.

2.2 Vowels

2.2.1 Vowel phoneme inventory

The Kokota vowel inventory is wholly unremarkable, reflecting the widespread Oceanic five vowel system:

Table 2.8: Vowel phonemes.

	front	central	back
high	i		u
mid	e		o
low		a	

These are primary vowels, with non-back vowels unrounded, slight rounding on the mid back vowel, and maximal rounding on the high back vowel. The system presents a triangular maximal differentiation between the high and low vowels, with the mid vowels equidistant between these three points.

The high vowels /i/ and /u/ are produced slightly lower than the cardinal vowel positions 1 and 8 (and slightly lower than their English counterparts). Within that there is a certain amount of height variation, with a further lowering of /i/ and /u/ occurring regularly in word final position and occasionally elsewhere.

No phonemic length distinctions exist in Kokota. While it is possible to find pairs distinguished by vowel length, this represents a distinction between the presence of a single vowel and the presence of a VV sequence in which the two vowels happen to be identical, not between a short and a long vowel. Thus /ipi/ 'wear' reduplicates to form /iipi/ 'be wearing clothes'. The distinction is between a single occurrence of /i/ and a VV sequence. This is demonstrated by the fact that an optional epenthetic glottal stop may occur between the segments, giving the surface forms [iʔipi] (as discussed in 2.1.3.3).

Phonetic length variation does however occur with all vowels, with unstressed non-word-final vowels typically reducing in length. Stressed and word final vowels may also reduce in length to a lesser extent and lesser degree.

2.2.2 Evidence for phoneme status - vowels

The following set demonstrates the contrastive status of all vowels except /o/:

(2.9) /toɾu/ 'bêche-de-mer' /toɾi/ 'open (tr)' /toɾa/ 'open (itr) /toɾe/ 'ask'

The following set contrasts /o/ with /a/ and /e/

(2.10) /aɾa/ '1SG pronoun' /aɾe/ 'those (nearby)' /aɾo/ 'these (touching)'

The following set contrasts /o/ with /u/ and /i/

(2.11) a. /hoɾe/ 'dugout canoe' /huɾe/ 'carry on shoulder'
b. /nodo/ 'stop' /nodi/ 'their'

2.2.3 Vowel phoneme frequencies

As discussed in 2.1.1.3 in relation to consonant phonemes, a word list of 335 basic lexical items (see Appendix 2) was used to calculate the relative frequencies of vowel phonemes. The list contains a total of 797 vowel phoneme tokens. The relative frequencies of each are:

Table 2.9: Vowel phoneme frequencies.

i	e	a	o	u	Total
15.3% (122)	13.3% (106)	35.4% (282)	18.6% (148)	17.4% (139)	100% (797)

Once the phonemes are grouped by features, either by height or on the front-back parameter, the figures are remarkably even.

Table 2.10: Relative frequencies by height features.

[+high] (/i/ and /u/)	[-high, -low] (/e/ and /o/)	[+low] (/a/)	Total
32.7% (261)	31.9% (254)	35.4% (282)	100% (797)

Table 2.11: Relative frequencies on the front-back parameter.

[+front] (/i/ and /e/)	[-front, -back] (/a/)	[+back] (/u/ and /o/)	Total
28.6% (228)	35.4% (282)	36.0% (287)	100% (797)

2.2.4 Phonological processes involving vowels.

Widespread vowel syncope occurs in Kokota. This is discussed in 3.2. Phonetic variation in vowel length and height is discussed in 2.2.1.

Further processes involving vowels include glide formation, and vowel devoicing. In addition, certain VV sequences are eligible for a process of diphthong formation, while in casual speech diphthongs are reduced to a single segment. Finally, a process occurs involving a transition matrix which superficially resembles epenthesis of the vowel /a/ between /h/ and /o/.

2.2.4.1 Diphthong formation

Kokota has no phonemic diphthongs. This is demonstrated by speaker syllabifications, in which every vowel in a sequence is syllabified separately. However, in normal speech certain non-identical VV sequences regularly undergo a process of diphthong formation. In this process the two vowels combine to form the nucleus of a single syllable, creating a heavy syllable with the structure shown in (2.28)b., but reducing overall syllable number.

2.2.4.1.1 Eligible sequences

A crucial criterion in the eligibility of a VV sequence to undergo diphthong formation is relative height: the second vowel in the sequence must higher than the first. Sequences of vowels involving no increase in height are ineligible for diphthong formation. Thus /hohoa/ 'yawn', /sarie/ 'nut sp.' /beata/ 'be calm (of sea)', and /tegeo/ 'thank' are trisyllabic.

A further criterion involves movement on the front-back parameter: VV sequences involving tongue movement from front to back or back to front are ineligible. Consequently the sequences /iu/, /ui/, /eo/, /oe/, /eu/ and /oi/ are ineligible. Only sequences of front vowels, sequences of back vowels, or sequences with /a/ as the first vowel are eligible.

Diphthong formation occurs with a sequence of the low vowel /a/ plus any other vowel. Thus /hae/ 'where', /mai/ 'come', /lao/ 'go', /pau/ 'head' are all monosyllabic. It also occurs with a height increasing sequence of vowels with the same status on the front/back parameter, with /hei/ 'who' and /dou/ 'be big' also monosyllabic.

The vowel sequences which occur as diphthongs are thus /ae/, /ai/, /ao/, /au/, /ei/ and /ou/.

2.2.4.1.2 Diphthong frequencies

The possible diphthongs described above do not all occur with equal frequency. In the word list in Appendix 2, a total of 49 diphthongs occur. The distribution of possible diphthongs within that is:

Table 2.12: Diphthong frequencies.

ai	au	ae	ao	ei	ou	Total
16 (32.7%)	14 (28.6%)	9 (18.4%)	6 (12.2%)	3 (6.1%)	1 (2.0%)	49 (100%)

It is interesting that there is a very strong preference for maximal differentiation between the two vowels. The two which involve the least shared features (/ai/ and /au/) – where the height difference is the greatest and the vowels differ on the front/back parameter, account for over half (61.3%) of all diphthongs present in the list. Those where the height difference is less but there is a difference on the front/back parameter (/ae/, and /ao/) account for a smaller group (30.6%). Those where the height difference is less and there is no difference in front/back status (/ei/ and /ou/) (8.1%) account for the smallest group. It is also interesting to note that there is a slight preference for front V2s (57.2%).

However, while these frequency tendencies are significant, they do not necessarily correspond to occurrence in normal discourse. For example while /ou/ only occurs in one item in the word list, that item, /dou/ 'be big', is a very high use item, generating high use of the diphthong. As a further example, an allomorph of the 3SG object enclitic is /-i/. Consequently the 3SG marked form of any verb with the final vowel /e/ has the /ei/ diphthong finally, giving a high use to this relatively low frequency diphthong.

2.2.4.1.3 Restrictions on diphthong formation

Diphthong formation is not restricted by morpheme boundaries. Any eligible VV sequence may undergo diphthong formation, regardless of whether the vowels are adjacent in a single morpheme, or brought together by concatenation. Thus in (2.12)a. the presence of the 3SG object enclitic generates a VV sequence which is ineligible for diphthong formation and the resulting word is trisyllabic, but in (2.12)b., where diphthong formation may occur, the resulting word is disyllabic:

(2.12) a. /huzu-i/ 'push it' → [huzui]
b. /friŋe-i/ 'work it' → [friŋei]

Diphthong formation occurs in rapid speech across word boundaries. In (2.13) the final vowel of the first word in the sequence combines with the initial onsetless syllable vowel of the second word to generate a diphthong.

(2.13) /nona uke ana/ 'that mother his' → [nonaukeana]

2.2.4.2 Diphthong reduction

Two processes of diphthong reduction optionally occur in casual speech. In one the second vowel is deleted, in the other the two vowels in the sequence coalesce to form a single monophthong produced in a position intermediate between the positions of the two vowels in the original sequence. These processes are discussed in 3.2.8 and 3.2.9.

2.2.4.3 Glide formation

With certain constraints, non-low vowels are realised as glides in casual speech when they occur as the first vowel in a VV sequence which is not eligible for diphthong formation. The front vowels /e/ and /i/ are realised as the palatal glide [j], while the back vowels /o/ and /u/ are realised as the labial glide [w]. This may occur where the vowel is preceded by a single consonant. The result is a surface cluster (in which any consonant may be C3).

(2.14)	a.	/tegeo/ 'thank'	→	[tegjo]
	b.	/n̥iaklo/ 'tree sp.'	→	[n̥jaklo]
	c.	/ikoa/ 'be small'	→	[ikwa]
	d.	/kuiti/ 'lie, deceive, trick'	→	[kwiti]

However, an absolute constraint exists in the language prohibiting onset clusters of more than two consonants. Consequently, where the first of the two relevant syllables already has an onset cluster, the process is blocked, as the glide would add a third consonant to the onset. Consequently in the old form /baknoa/ 'be slow' glide formation from the /o/ is prevented, but in the new form of this lexeme, /bnakoa/, it is not:

(2.15)	a.	/baknoa/ 'be slow'	→	*[baknwa]
	b.	/bnakoa/ 'be slow'	→	[bnakwa]

Just as glide formation may generate an onset cluster C2, it may also generate a single C onset where there is no preceding consonant. This may be because the VV sequence occurs word initially, as in (2.16) a., or because a preceding vowel will be realised in a separate syllable.

(2.16)	a.	/iaɾo/ 'those (distant)'	→	[jaɾo]
	b.	/ka-ia/ 'at the'	→	[kaja]
	c.	/n-e-ke-u-o/ 'it was thus'	→	[nekewo]

The implication of the above environments for glide formation, is that a non-low vowel will undergo glide formation if it can form an onset. The motivation for this is to reduce the prosodic complexity of the surface form, in terms of number of morae or syllables.

Glide formation is prevented where it would create a word consisting of a single light syllable, violating a word minimality constraint. So in isolation /kuo/ 'break' always occurs on the surface as a disyllable. Once the form is suffixed the risk to word minimality is removed and the process occurs.

(2.17)	a.	/kuo/ 'break'	→	*[kwo]
	b.	/kuo-di/ 'break them'	→	[kwodi]

This process applies to vowels that are adjacent morpheme internally, as in (2.14), (2.15) b., (2.16) a. and (2.17) b., and to vowels which are adjacent as the result of affixation or cliticisation, as in (2.16) b. and c. The process also occurs in rapid speech across word boundaries. In (2.18) the vowel of the object enclitic becomes a glide in the presence of the subsequent vowel initial subject pronoun:

(2.18) /ooeni ara/ 'I said it' → [ooenjara]

Where a non-low vowel occurs as the second vowel in a VV sequence which is eligible for diphthong formation it does not undergo glide formation, unless the opportunity exists for it to be realised as the onset of a subsequent vowel. Glide formation takes precedence over diphthong formation. So for example in /gauai/ 'be distant' the medial /u/ could potentially form a diphthong with the preceding /a/. However it also potentially forms an onset for a syllable containing the following vowels (which do form a diphthong). It is the formation of the glide as an onset for a second syllable that occurs:

(2.19) /gauai/ 'be distant' → [gawai]

The prosodic implications of glide formation are discussed in 3.2.10.

2.2.4.4 Interconsonantal vowel devoicing

Vowels assimilate to the voicing status of adjacent voiceless consonants when occurring between two identical voiceless consonants.

(2.20)

```
[-voice]     [-voice]
   |  \        /  |
   |    \    /    |
  C1      V      C1
```

This typically occurs with the vowel of echo syllables, and may involve plosives or fricatives as the environment.

(2.21) a. /kikibolo / 'football' → [ki̥kibolo]
b. /sasamala/ 'masturbate (of males)' → [sḁsamala]
c. /hahaglu/ 'broom' → [hḁhaglu]

This occurs regardless of whether the echo syllable reflects synchronic reduplication, as is the case with (2.21) c., or within a synchronically monomorphemic form, as in (2.21) a. and b. Indeed, (2.21) a. is a recent Pijin loan.

The process also occurs with vowels which are assigned secondary stress, as is the case with both (2.21) a. and b. It does not, however, appear to occur with vowels assigned primary stress.

It is not clear at this stage whether vowel devoicing occurs when the consonants which motivate the process are voiceless sonorants.

Vowel devoicing also appears to occur to a considerably lesser extent when the motivating consonants are not identical.

(2.22) /sisikama / 'palm sp.' → [si̥si̥kama]

2.2.4.5 A [+low, +voice] transition matrix in the sequence /h/ + /o/

The underlying sequence /h+o/ in casual speech usually generates the surface realisation [hao]. This is not in fact /a/ epenthesis, but is due to partial voicing of the /h/ in anticipation of the following vowel. To realise the underlying sequence the closure from [+low] /h/ to [-low] /o/ must be matched precisely with the onset of voicing in the shift from the [-voice] /h/ to the [+voice] /o/. However in normal speech these two changes are not matched precisely, the onset of voicing occurring fractionally before the change to [-low], while the [+low] feature of the /h/ remains. The result is a transitional matrix between the two segments which has the [+ voice, +low] features of the vowel /a/:

(2.23)

```
[+low]       [+voice]             [+low]        [+voice]
   |             |                   |   \        /   |
   |             |                   |     \    /     |

  /h/           /o/        ⇒        [h]     [a]     [o]
[+low]       [-low]               [+low]  [+low]  [-low]
[-voice]     [+voice]             [-voice][+voice][+voice]
```

The effect is that in normal speech words like /hoda/ 'take' and /zaho/ 'go' usually occur on the surface as [haoda] and [zahao]. This is not affected by stress - in /hoda/ the relevant syllable is stressed while in /zaho/ it is not.

2.3 Syllable structure

A number of phonological processes take place in casual speech which affect prosodic structure. These include vowel syncope, and the formation of glides and geminates. These processes often alter the surface syllable structure, particularly with vowel syncope generating non-underlying consonant clusters and syllabic consonants, as well as surface codas. However, these effects on syllable structure will be discussed in 3.2. In this section the underlying syllable structure of the language will be described.

Kokota syllable structure allows an onset and a nucleus. Codas do not occur (other than in a small number of recent Pijin loans), so all syllables are open. Onsets may consist of one consonant or a cluster of two consonants in certain configurations, and no onset syllables are allowed. Nuclei consist of one vowel or a diphthong of certain VV sequences.

2.3.1 Onsets

The overwhelming majority of Kokota syllables have a single consonant onset. However the language does permit syllables with no onset, and onset clusters of certain configurations of two consonants. In the phoneme frequency word list in Appendix 2, out of a total of 746 syllables present in 335 words, 657 (88.1%) have a single consonant onset, while 45 (6.0%) have no onset and 44 (5.9%) have a cluster onset.[4]

2.3.1.1 Syllables with no onset

As the figures above indicate, syllables with no onset occur in only a small proportion of words. However they occur in discourse to a disproportionate extent as a number of extremely high use words, particularly deictics, have an initial onsetless syllable. These include the pronouns /aɾa/ '1SG' and /aɣo/ '2SG', the locative /ade/ 'here', eight demonstratives, both articles, and three of the four irrealis subject indexing auxiliaries.[5]

2.3.1.2 Consonant clusters

Sequences of two onsonants in Kokota form complex onsets. Speakers invariably syllabify medial CC sequences as cluster onsets, and not as coda + onset sequences. Moreover, clusters freely occur in word initial position, eliminating the possibility that they represent a coda + onset. Indeed, there appears to be a preference for word initial clusters over medial clusters, as discussed in 2.3.4.

In casual speech a number of configurations of consonant clusters occur. Most are not underlying, being generated by vowel syncope or glide formation. These surface clusters do not occur in careful speech, and are revealed as non-underlying by speaker syllabifications. However, certain other cluster onsets are underlying. These conform to the following broad constraints: C1 must be an obstruent, and C2 must be a voiced coronal sonorant.

---

[4] For the cluster onset figures only underlying clusters have been counted (casual speech surface clusters generated by vowel syncope or glide formation have not). However, underlying clusters in initial echo syllables are excluded as these never occur on the surface (see 2.4.1.2). The no-onset figures do not include vowels which undergo diphthong formation. Kokota has no phonemic diphthongs, all diphthongs being generated from certain underlying VV sequences. The no-onset figure represents vowels which occur on the surface as the nucleus of a syllable nucleus with no onset, but not those which occur on the surface as the V2 of a diphthong nucleus.

[5] The articles are /ia/ 'theSG' and /iɾa/ 'thePL'. The demonstratives are /ao/ 'this (holding)', /aɾo/ 'these (holding)', /ine/ 'this (nearby)', /ide/ 'these (nearby)', /ana/ 'that', /aɾe/ 'those' and /iao/ 'that (distant)', /iaɾo/ 'those (distant)'. The auxiliaries are /a/ 'first person exclusive', /o/ 'second person' and /e/ 'third person'.

However, the C1 constraints are more complex, with place class constraints applying. These constraints differ for plosives and fricatives. With plosive C1s the constraint is on coronals, with only labials (/p/ and /b/) and noncoronal nonlabials (/k/ and /g/) occurring. With fricative C1s the constraint is on nonlabial noncoronals, with only labials (/f/ and /v/) and coronals (/s/ and /z/) occurring. Among the fricatives, however, all except /f/ are highly marginal, in some cases attested in only one lexical item.

Further, not all cluster combinations conforming to these constraints actually occur. Two of the plosives do not occur with a nasal C2, while the marginal nature of three of the fricative C1s means they do not occur with the full range of possible C2s. The attested clusters in non-loan lexical items are:

Table 2.13: Attested non-loan cluster combinations.

	ɾ	l	n
p	pɾ	pl	-
k	kɾ	kl	kn
b	bɾ	bl	bn
g	gɾ	gl	-
f	fɾ	fl	fn
s	?sɾ	-	sn
v	vɾ	-	-
z	zɾ	-	zn

The frequencies of the attested clusters vary. In the representative word list contained in Appendix 2 a total of 44 cluster onsets occur, reflecting the following cluster frequencies:

Table 2.14: Cluster frequencies.

	ɾ	l	n	total
p	3	2	-	5
k	2	3	2	7
b	2	2	1	5
g	8	5	-	13
f	9	2	2	13
s	-	-	0	0
v	1	-	-	1
z	0	-	0	0
total	25	14	5	44

As these figures indicate, clusters with fricative C1s other than /f/ are marginal. In the approximately 1800 lexical items currently recorded, /s/ is only attested with /n/ as C2. This cluster occurs in four attested lexical items: the relatively high use items /snakɾe/ 'allow' and /sasna/ 'be willing', and the relatively low use /snekɾi/ 'tear fingernail, earlobe etc' and /nasnuɾi/ 'sea urchin sp.'. Interestingly, /s/ is the only C1 not unequivocally attested with /ɾ/ as C2. However, some evidence suggests that while it is not clearly attested, it is possible. Older speakers have the verb /tasuɾu/ (meaning unclear), which due to regular vowel syncope normally occurs on the surface in casual speech as /tasɾu/. However for at least some younger speakers the underlying form of this item is the disyllable, with the second syllable onset cluster /sɾ/. It is possible that this is an established cluster which is sufficiently marginal for unequivocal examples to have not yet been recorded. Given the fact that /ɾ/ occurs with all other possible C1s this seems likely. Alternatively it may be that the age variation is evidence of a change underway in the language which allows this cluster. The voiced coronal /z/ is attested with both /ɾ/ and /n/ as C2, but in only one low use item each: /zniɾi/ 'be tangled' and /zozozɾo/ 'be pouring with water'. The voiced labial fricative /v/ is only attested with /ɾ/ as C2, and then only in two relatively low use items: /vɾaha/ 'tree sp.' and /viviviɾi/ 'propeller'. It is not clear whether the unattested clusters represented by gaps in Table 2.13 are possible but occur in so few lexical items they have yet to be attested, or whether they do not occur in the language.

The cluster possibilities represented in Table 2.13, and the frequencies represented in Table 2.14 both reflect a preference for clusters with a substantial difference in sonority between the C1 and C2. Broadly speaking, the strongest C1s (the plosives) can occur with any of the possible C2s, while the weakest (the voiced fricatives) can occur with the fewest possible C2s. Equally, the C2 with the highest sonority (the rhotic), can occur with the widest range of C1s. This is particularly noticeable if /s/ and /z/, which are normally regarded as extrametrical, are excluded. This sensitivity to a sonority hierarchy is strongly apparent in the frequencies shown in Table 2.14, represented in Table 2.15.

Table 2.15: Cluster frequencies by sonority.

		highest sonority ɾ	l	lowest sonority n	total
strongest	voiceless plosives (/p/ and /k/)	5 (42%)	5 (42%)	2 (16%)	12
	voiced plosives (/b/ and /g/)	10 (56%)	7 (39%)	1 (5%)	18
	voiceless fricatives (/f/)	9 (70%)	2 (15%)	2 (15%)	13
weakest	voiced fricatives (/v/)	1 (100%)	-	-	1

Broadly speaking, the stronger the C1 obstruent, the lower the preferred acceptable sonority of the C2.

2.3.1.2.1 Violations of cluster constraints in loan words

A number of marginal violations of the cluster constraints discussed above occur in loan words.

2.3.1.2.1.1 A coronal C1

One cluster, /tɾ/, violates the constraint on coronal plosive C1s. This cluster is attested in only one item, the Pijin loan /tɾiki/ 'lie, deceive, trick'. It is apparent that the cluster is underlying in this item for at least some speakers, as these speakers syllabify the item as a disyllable with the initial cluster.

2.3.1.2.1.2 Obstruent plus obstruent clusters

Apparently underlying obstruent plus obstruent consonant clusters are attested in four items, all Pijin or English loans. In two of these instances the cluster is initial. Both involve /s/ plus a voiceless plosive: /spika/ 'Speaker'[6] and /stiŋibili/ 'beetle sp.'[7]. That such exceptions should involve the C1 /s/ is not surprising, given its extrametrical status.

More problematic are the two attested medial obstruent plus obstruent clusters. One, /kastom/, again involves a C1 /s/. However there is considerable variability in whether the CC sequence is treated as an onset cluster, or the C1 is treated as a coda of the preceding syllable (in which case there is a CC sequence but no cluster). An elderly speaker consistently syllabified this word as /ka-sto-mu/, preferring an obstruent plus obstruent cluster and epenthetic final vowel to any codas. All younger speakers tested, however, displayed considerable variability, giving both the coda syllabification /kas-tom/ and the onset cluster syllabification /ka-stom/ on the same occasion. A similar situation pertained with /dokta/. The older speaker

[6] This apparently odd loan is a relatively high frequency word, being a common term used to refer to the former Speaker of the Provincial Government, an important and high profile resident of Goveo village.
[7] Presumably from 'stink beetle'.

syllabified this as /do-ki-ta/, resisting both a coda and a /kt/ cluster (despite the fact that he produces the item in casual speech without the epenthetic vowel as /dokta/). Younger speakers, however, again gave the varied syllabifications /dok-ta/ and /do-kta/. For both of these items there is clearly confusion among speakers as to whether to permit a normally unacceptable obstruent plus obstruent cluster, or a normally unacceptable coda.

The same confusion appears to apply to a further item, /kaspotu/ 'clam sp.'. However the existence of this CC sequence and its consequent prosodic dilemma seems bizarre, given that it is a loan from the neighbouring Zabana, which does not permit clusters. The Zabana form is /kasipotu/. In Kokota older speakers seem to have that as the underlying form, though the same older speaker, who regards the word as Zabana not Kokota, syllabified it variably as /ka-si-po-tu/ and /kas-po-tu/. Younger speakers, however, regard the word as Kokota, with the underlying form /kaspotu/ (rejecting /kasipotu/ and syllabifying the form as a trisyllable). These speakers display the same syllabification variability on this item as with /dokta/ and /kastom/, giving the syllabifications /ka-spo-tu/ and /kas-po-tu/. On the basis of regular Kokota stress assignment and vowel syncope, an underlying form /kasipotu/ would give the surface form /kaspotu/ in casual speech. What seems inexplicable is that the synchronic syncope should become diachronic vowel loss, when this creates an unacceptable CC sequence in the new underlying form.

2.3.1.2.2 Absence of /h/ and voiceless sonorants in clusters

One of the striking features of Kokota onset clusters is the constraint on voiceless sonorants as C2. This, combined with the absence of /h/ as C1, and the relatively unusual presence of voiceless sonorants in the language, suggests the possibility that the full set of voiceless sonorant phonemes discussed above are not underlyingly phonemes. Instead, the possibility must be considered that they are in fact clusters of /h/ plus a voiced sonorant, which regularly undergo synchronic coalescence to form surface voiceless sonorants. This would resolve the unusual presence of voiceless sonorant phonemes, and explain the absence of /h/ as a cluster C1. However, while the voiceless sonorants exist as a result of the <u>diachronic</u> coalescence of /h/ and adjacent voiced sonorants, brought about as a result of widespread syncope of unstressed vowels, there is evidence that this is not the synchronic situation, and that the voiceless sonorants are in fact underlying phonemes.

Crosslinguistically the presence of /h/ in consonant clusters is rare. However, while that to some extent weighs against the likelihood that the underlying cluster hypothesis is correct, it does not in itself constitute evidence of anything. Kokota may be one of the rare languages that has such clusters.

More suggestive is the additional absence of /ɣ/ as C1. As discussed above, all labial and coronal fricatives occur as C1, but neither of the nonlabial noncoronal fricatives does so. The absence of /h/ as C1 can therefore be seen as a feature of the absence of fricatives in an entire place class. This is more strongly suggestive that clusters with /h/ do not occur, even underlyingly. However, it again does not constitute actual evidence.

Evidence of one kind comes from speaker judgments. A series of forms containing voiceless sonorants were tested on speakers, with careful enunciation of a CC sequence involving /h/ and the relevant sonorant in both possible orders. For example /koɾ̥o/ 'pull' was tested as both /koɾho/ and /kohɾo/, and /nam̥ari/ was tested as /namhari/ and /nahmaɾi/. If these were underlyingly consonant sequences it is likely one of the tested forms would have been acceptable. It will be recalled from the discussion of /ɣ/ deletion in 2.1.2.3.3.2 that when forms in which the /ɣ/ always deletes in certain environments were tested with the /ɣ/ present they were accepted by speakers as being good Kokota, with the note that they are not pronounced that way. If the voiceless sonorants were underlyingly clusters a similar judgment of acceptability would be expected. Instead speakers rejected all tested forms, and in subsequent discussion expressed the judgment that the voiceless sonorants are not 'really' /ɾh/ or /hɾ/, /mh/ or /hm/ etc but are /ɾ̥/, /m̥/ and so on.

Further evidence comes from the behaviour of onset clusters in reduplication. Kokota has productive partial reduplication which is discussed in detail in 2.4.

This reduplication formally involves an initial echo syllable. Roots which have an initial syllable with an onset cluster reduplicate that syllable in its entirety, generating an echo syllable with the same underlying CCV structure. However, on the surface only the initial consonant of the cluster is realised. Echo syllables undergo a process of reduplicated syllable cluster reduction which deletes the onset C2 of the echo syllable.

(2.24) /kɾisu/ 'scoop liquid' → /kɾikɾisu/ → [kikɾisu] 'be scooping liquid'

If Kokota voiceless sonorants were underlyingly clusters, then in one possible scenario the same C2 deletion should apply: a root with an initial voiceless sonorant would reduplicate to an underlying CCV echo syllable, which would then be realised on the surface with the C2 deleted. Assuming the standard obstruent plus sonorant structure for the underlying cluster, C2 deletion should prevent the sonorant from appearing on the surface. Alternatively, if the underlying structure reflects the order of sonorant plus obstruent (at odds with other clusters in the language), C2 deletion would generate a surface form with only the sonorant:

(2.25) a. */hɾuta/ → */hɾuhɾuta/ → *[huɾ̥uta]
b. */ɾhuta/ → */ɾhuɾhuta/ → *[ɾuɾ̥uta]

However, neither of these forms occur. Instead the voiceless sonorant occurs on the surface, corresponding to the pattern for the reduplication of single consonant onsets.

(2.26) a. /ɾ̥uta/ 'untangle' → [ɾ̥uɾ̥uta] 'be untangling'
b. /m̥aɣu/ 'be afraid' → [m̥am̥aɣu] 'be habitually fearful'
c. /ŋ̊au/ 'eat' → [ŋ̊aŋ̊au] 'be biting (of fish)'

This is not terribly strong evidence, because it is possible that /h/ plus sonorant coalescence occurs prior to the operation of the C2 deletion rule. Had the reduplication of voiceless sonorant onsets resulted in one of the surface forms in (2.25) it would have provided strong evidence of an underlying cluster. However this potential counter evidence to underlying voiceless sonorant phonemes does not present itself. That in itself is merely negative evidence. It has positive value in its contribution to a body of phenomena consistent with underlying phonemes and in the absence of evidence inconsistent with that hypothesis.

Stronger evidence involves the order of segments in proto forms. Kokota onset clusters are the result of a regular diachronic process in which certain unstressed vowels syncopate. Thus for example Kokota /knaha/ 'cough' is cognate with Zabana /kanaha/ 'cough'. The proto form can be assumed to have been identical with that in the phonologically conservative Zabana. As a result of frequent syncope of the vowel of the first syllable the two consonants normally adjacent on the surface were reanalysed as underlyingly adjacent. This diachronic reanalysis occurred with sequences of obstruent plus sonorant. However, other consonant sequences (sonorant plus obstruent, sonorant plus sonorant, and obstruent plus obstruent) were not reanalysed, and remain underlyingly separated by a vowel, even if they often appear as adjacent on the surface in casual speech. The order of the segments in a cluster thus reflects their order in the proto form.

Given that clusters were historically formed only by consonants in the sequence obstruent + sonorant, if the voiceless sonorants were underlyingly clusters with /h/ as the obstruent, they would have to reflect clusters with the order /h/ + sonorant. These in turn would have to reflect proto forms in which that segment order pertained. However the opposite is true. Voiceless sonorants in Kokota are the result of the historical coalescence of /h/ plus a sonorant. But the segment order in that diachronic process is the reverse of that which generated underlying clusters. For example if the voiceless sonorants are underlyingly clusters, then Kokota [kom̥u] 'year' must reflect a synchronic underlying */kohmu/, which in turn must reflect a proto form **/kohVmu/. However, the Zabana cognate, reflecting directly the proto form for Kokota, is /komuhu/, with the order sonorant + /h/. This segment order did not undergo diachronic reanalysis as clusters in Kokota. Six cognates of Kokota roots containing voiceless sonorants have been found in Zabana.

In all six the order is sonorant + /h/, the opposite of that which diachronically generated Kokota clusters, while no cognates have been found reflecting the expected /h/ + sonorant order.[8]

A final piece of evidence supporting the view that voiceless sonorants are not underlyingly clusters arises from the constraint that all cluster C2 sonorants must be coronal. Since noncoronal voiceless sonorants occur, if these did represent coalesced underlying clusters those clusters would violate that constraint. Moreover, the constraint on noncoronal C2s has developed since the reanalysis of adjacent consonants as clusters. As this constraint developed, clusters which violated it underwent a phonological change in which noncoronal C2s became coronal. Thus Kokota /knisu/ 'spit' is cognate with Zabana /kakamisu/ 'spit', suggesting a Proto Isabel form */kamisu/[9]. This then underwent diachronic syncope of the first syllable vowel, resulting in a Proto Central Eastern Isabel form *kmisu. The constraint on noncoronal C2s developed subsequently, causing a shift of the /m/ to /n/.[10] The intermediate stage is in fact synchronically present in Cheke Holo. Holo also reflects cluster formation, indicating a shared ancestor with Kokota which was a sister of pre-Zabana. However, unlike Kokota, Holo did not go on to develop a constraint on noncoronal clusters. Thus the Holo reflex of Proto-Isabel */kamisu/ is /kmisu/. These three languages closely reflect the stages of diachronic change:

(2.27)	Proto-Isabel	Zabana	Cheke Holo	Kokota
	*/kamisu/	/kakamisu/	/kmisu/	/knisu/

If the Kokota voiceless sonorants were underlyingly clusters, this change would have turned all the noncoronal C2s into coronals, leaving only coronal surface voiceless sonorants. However, noncoronal voiceless sonorants do occur, indicating that they are not the result of the same diachronic process that generated clusters.

Thus the root [nam̥aɾi] cannot underlyingly be *[namhaɾi] because the cluster is sonorant + obstruent, and the process of cluster formation applied only to obstruent + sonorant sequences. Equally, it cannot be [nahmaɾi] because the C2 is labial and thus does not accord with the development of the constraint on noncoronal C2s. The same is true of [ŋ̊au] 'eat', where the sonorant is velar.

[8] The cognates are:

Kokota	Zabana
/kom̥u/ 'year, fruit'	/komuhu/ 'season from April - Dec.'
/koɾ̥o/ 'pull'	/koɾoho/ 'pull'
/mam̥i/ 'sakale sp.' ~ /mamam̥i/ 'tree sp.'	/mamihi/ 'tree sp. Spondias dulcis'
/nam̥a/ 'love, be friendly'	/namaha/ 'love'
/naɾ̥o/ 'creeper, rope'	/naɾoho/ 'rope'
/nom̥o/ 'hear'	/nonomoho/ 'hear'

[9] Supported by cognates outside Isabel in the New Georgia chain - Vangunu /kamisu/ 'spit', and slightly further afield in Mono-Alu /amisu/ 'spit'.

[10] The development of clusters and subsequent constraint on noncoronal C2s provides subgrouping evidence for the languages of Isabel. Apart from Bughotu, spoken on the eastern tip of the island, all Isabel languages are descended from a common ancestor which can be referred to as Proto Isabel. All these languages except for the phonologically conservative Zabana, spoken at the western end, have obstruent + sonorant cluster onsets, suggesting that development of these clusters took place in a common ancestor, a sister of pre-Zabana that may be called Proto Central Eastern Isabel. The central members of this subgroup, Kokota, Laghu and Blablanga all have a constraint on noncoronal C2s not found in Cheke Holo, and perhaps Gao, suggesting that this constraint developed in an ancestor to these central languages that may be called Proto Central Isabel. Cheke Holo and Gao may be daughters of a Proto Eastern Isabel, a sister of PCI which did not develop the no noncoronal C2 constraint. However since at this stage no shared innovation has been found for both Cheke Holo and Gao it seems safer to suggest that PCI was a sister of pre-Holo and pre-Gao.

This evidence demonstrates that the voiceless sonorants cannot be surface realisations of underlying clusters, as the clusters they would reflect cannot be generated by the processes that generated other existing clusters.

2.3.2 Nuclei

Kokota allows a maximum of two morae per syllable. In light syllables the nucleus consists of a single vowel. In heavy syllables the nucleus consists of a diphthong, except with a small number of recent loans where a coda is present, where the nucleus consists of a single vowel (the additional mora being contributed by the coda).

In the small number of recent loans involving a coda, none have a coda following a diphthong. These thus do not violate the maximal bimoraic syllable constraint. The possible syllable structures are thus as follows:

```
(2.28)  a.         σ           b.         σ             c.         σ
              /    |                  /   |    \               /   |    \
            /      μ                /     μ    μ             /     μ    μ
          /        |              /       |    |           /       |    |
        ((C)C)     V            ((C)C)    V    V         ((C)C)    V    C
```

The structure shown in (2.28)c. is extremely marginal. Of the other two, representing monophthong and diphthong nuclei, the former occurs far more frequently. In the word list in Appendix 2, out of 746 syllables, 697 (93.4%) had a monophthong nucleus, while only 49 (6.6%) had a diphthong.

2.3.3 Codas

Other than in a handful of Pijin loans, codas occur only as a consequence of widespread vowel syncope in casual speech (discussed in detail in 3.2). This syncope frequently brings together two consonants which are not adjacent underlyingly. This occurs morpheme internally, across morpheme boundaries, and even across word boundaries.

(2.29) a. /banesokeo/ 'a place name' → [bansokeo]
b. /mane dou/ 'big man' → [man dou]

Syncope of this kind generates codas in which the first of the affected consonants, which would otherwise be syllabified as an onset, attracts the mora associated with the syncopated vowel. It is then syllabified with the vowel of the preceding syllable to form a new, heavy syllable. So in (2.29) a. the surface form has four syllables rather than five, the first of which has a coda. In (2.29) b. the first word is realised as a monosyllable with a coda.

Other than that, codas only occur in a small number of recent loans. Sometimes these are word final, and thus unequivocally codas, as in /kastom/, for example, though it is worth noting that in careful speech older speakers often insert an epenthetic final vowel, giving the surface form [kastomu]. However for younger speakers the form is always realised with the final coda. In other instances the possible coda occurs medially. In these instances there is considerable variation within single speakers. In examples like /kastom/ the first consonant in the medial CC sequence could potentially be a coda of the first syllable, or the C1 of a second syllable onset cluster. As discussed in 2.3.1.2.1 in relation to cluster constraint violations, speakers typically exhibit considerable variation in this situation, with individual speakers giving both syllabifications /ka-stom/ and /kas-tom/ in a single elicitation.

2.3.4 Word level syllable structure tendencies

The distribution of onset types varies depending on where the syllable falls in the word. The data for this is as follows:

Table 2.16: Distribution of onset types across syllables.

	syllable 1	syllable 2	syllable 3	syllable 4	syllable 5
no onset	23 (6.9%)	16 (5.1%)	6 (7.2%)	0	0
C	294 (87.8%)	274 (87.6%)	74 (89.2%)	14 (100%)	1 (100%)
CC	18 (5.4%)	23 (7.3%)	3 (3.6%)	0	0
total	335[11]	313	83	14	1

The smaller proportion of onsetless second syllables is understandable given that in any VV sequence some will form diphthongs, thus only a proportion will result in onsetless syllables. This is not the case with word initial vowels.

The figures for cluster onsets are more striking: there are more syllable 2 cluster onsets than for syllable 1, despite there being less second syllables overall. This can be explained by the nature of the diachronic development of clusters. Kokota clusters are the result of a process of vowel loss. Widespread unstressed vowel syncope occurs in synchronic Kokota. It appears that at an earlier stage, apparently at the stage of a Proto Central Isabel, historical syncope became diachronic vowel loss in obstruent + vowel + sonorant sequences. In disyllables no such loss is possible because the only consonants must be separated by a stressed vowel. In trisyllables regular synchronic stress assignment assigns stress to the antepenultimate syllable, meaning that syncope would bring together non-initial consonants, leading through vowel loss to medial clusters. Only a minority of synchronic trisyllables are assigned stress on the penultimate syllable, however, historically in such roots the first syllable vowel, being unstressed, could syncopate, bringing together the first and second syllable onsets, and through vowel loss, generate initial clusters. The stress regime of synchronic Kokota, if assumed for Proto Central Isabel, would explain the preponderance of medial rather than initial clusters. This preponderance provides evidence that the existing stress regime applied at that earlier stage. This also removes the possibility of interpreting the disproportionate number of medial clusters as indicating a preference to be able to make cluster C1s codas. Medial cluster C1s are not syllabified as codas by speakers and do not appear to occur as codas in the prosodic structures found in normal speech.

Indeed, far from there being a preference for medial clusters, there appears to actually be a preference for initial clusters, despite the figures in Table 2.16. This preference is evident in a diachronic shift currently underway in the language in which the C2 of a second syllable cluster is transferred to the initial syllable onset. Thus with certain lexical items younger speakers have the item with an initial cluster, while older speakers either use a medial cluster, or at least regard the medial cluster as correct Kokota and the initial cluster form as incorrect:

(2.30) a. /bakɾu/ → /bɾaku/ 'liquid'
b. /baknoa/ → /bnakoa/ 'be slow'
c. /faklano/ → /flakano/ 'rock outcrop on beach'

It is not clear at this stage whether the phonetic similarities between the first syllables and the second syllable C1s in these examples are significant.

The examples in (2.30) are presumably representative of a general shift underway, being special only to the extent that they are in mid change, the shift thus being visible. It therefore seems likely that while some words with initial syllable clusters reflect loss of a former first syllable vowel (as in (2.31) a.), others reflect the loss of a second syllable vowel, with the C2 of the consequent second syllable cluster subsequently shifting to the first syllable onset (as in (2.31) b.).

(2.31) a. CVCVCV → CCVCV
b. CVCVCV → CVCCV → CCVCV

[11] The discrepancy in the totals for this column is the result of rounding.

This means that the number of medial clusters currently present in synchronic Kokota may be assumed to be only a proportion of those originally present. Consequently at least some of the words which currently have an initial syllable cluster must reflect a proto-form which underwent loss of the second syllable vowel, with the implication that that vowel was unstressed. This further strengthens the argument that in Proto Central Isabel most trisyllables were assigned antepenultimate stress: not only do all current CVCCV roots reflect that, but a proportion of the current CCVCV do as well.

2.3.5 Word minimality

A word may consist of a single syllable. This may be a heavy syllable with a diphthong nucleus, such as /mai/ 'come', /dou/ 'be big' or /ao/ 'this (touching)'. A number of roots such as /su/ 'breast', /do/ 'mosquito' and /fro/ 'squeeze' are underlyingly monomoraic, meaning that the nucleus consists of a single short vowel (see 3.1). However, as discussed in 3.1, stress cannot be assigned to a monomoraic form, thus a stress bearing word may not consist of a single light syllable. Monomoraic roots are lengthened to create a heavy syllable to allow stress assignment. Underlyingly monomoraic grammatical particles are not always assigned stress, in which case they remain unlengthened.

Thus, the minimum Kokota word is a single syllable, and the minimum stress bearing word is a single heavy syllable.

2.4 Reduplication

Kokota displays considerable evidence of historical partial reduplication and some examples of historical full reduplication of disyllabic roots. Only partial reduplication appears to occur synchronically.

2.4.1 Partial reduplication

2.4.1.1 Function of partial reduplication

Derivational reduplication involving an initial echo syllable occurs very extensively in Kokota. The extent to which it is productive is not entirely clear. It has a general derivation function, the nature of the derivation varying widely. Some reduplicative derivations involve change of word class. In some instances a verb is derived from a noun.

(2.32)	a.	/fiolo/ 'penis'	→	/fifiolo/ 'masturbate (of male)'
	b.	/piha/ 'small parcel'	→	/pipiha/ 'make a *piha* parcel'
	c.	/puki/ 'round lump of s.th.'	→	/pupuki/ 'be round'

Much more commonly, however, word class altering derivation involves the derivation of nouns from verbs, either transitive or intransitive.

(2.33)	a.	/siko/ 'steal'	→	/sisiko/ 'thief'
	b.	/lase/ 'know'	→	/lalase/ 'knowledge, cleverness'
	c.	/kamo/ 'transfer'	→	/kakamo/ 'smoldering stick for transferring fire'
	d.	/maku/ 'be hard'	→	/mamaku/ 'leatherjacket (fish w. tough skin)'
	e.	/gufu/ 'smoke s.th.'	→	/gugufu/ 's.th. for smoking (ie. tobacco)'

In many instances reduplicative derivation functions to derive a new lexeme from a root without any change to word class. Often there is little semantic predictability in the meaning of the new form. Noun from noun derivations, for example, include:

(2.34)	a.	/baɣi/ 'wing'	→	/babaɣi/ 'side roof of porch'
	b.	/bul̥i/ 'cowrie'	→	/bubul̥i/ 'clam sp.'
	c.	/tahi/ 'sea'	→	/tatahi/ 'stingray'
	d.	/pɾotu/ 'distant object'	→	/popɾotu/ 'small lump on body'

Some verb from verb derivations display similar semantic unpredictability:

(2.35) a. /ŋau/ 'eat' → /ŋaŋau/ 'be biting (of fish)'
b. /pɾosa/ 'slap self w. flipper (of turtles)' → /popɾosa/ 'wash clothes'
c. /maɾ̥i/ 'be in pain' → /mamaɾ̥a/ 'be in labour'
d. /vahe/ 'carve' → /vavahe/ 'operate surgically'

Some verb from verb derivations show a more predictable semantic relationship, with the derived form coding an habitual, ongoing or diminutive event:

(2.36) a. /m̥aɣu/ 'be afraid' → /m̥am̥aɣu/ 'be habitually fearful'
b. /safɾa/ 'miss' → /sasafɾa/ 'always miss'
c. /seha/ 'climb' → /seseha/ 'climb all about'
d. /fogɾa/ 'be sick' → /fofogɾa/ 'be a little bit sick'

However, a substantial subregularity involves the reduplication of transitive verb roots. In some instances the derived form is an unaccusative stative verb.

(2.37) a. /lage/ 'castrate' → /lalage/ 'be castrated'
b. /sito/ 'make hot' → /sisito/ 'be hot to the touch'
c. /hoti/ 'sting (tr)' → /hohoti/ 'be very sore and tender'

Usually, however, the derived verb is unergative. This appears to be productive to the extent that unless the reduplicated form of a transitive verb has some unpredictable (and hence lexically specified) meaning, it will be an unergative version of the root.

(2.38) a. /piɾi/ 'tie' → /pipiɾi/ 'be tying'
b. /kɾisu/ 'scoop liquid' → /kikɾisu/ 'be scooping liquid'
c. /sofo/ 'grab' → /sosofo/ 'be grabbing'
d. /taho/ 'count' → /tataho/ 'be counting'

In some instances the reduplicated form of a root can have more than one of the possible derived meanings. For example a number of transitive verb roots reduplicate to derive both an unergative verb and a noun.

(2.39) a. /tuɾi/ 'tell s.th.' → /tutuɾi/ 'chat; a story'
b. /ɣato/ 'think (tr)' → /ɣaɣato/ 'think (itr); thoughts'
c. /keɾe/ 'sting (tr)' → /kekeɾe/ 'sting (itr); thorns'

2.4.1.2 Formal characteristics of partial reduplication

Partial reduplication involves the presence of an initial echo syllable which is underlyingly identical to the initial syllable of the root. As syllables are underlyingly open in Kokota, the effect is initial onset and nucleus reduplication. As indicated in 2.3, the majority of Kokota syllables have a single consonant onset. Reduplication of roots with such an initial syllable formally involves reduplication of the initial CV. This generates all but three of the examples given in 2.4.1.1. However, reduplication also occurs with some roots which have an initial onsetless syllable. In the absence of an onset only the nucleus is reduplicated.[12]

(2.40) /ipi/ 'wear (clothes)' → /iipi/ 'be wearing clothes'

[12] Unlike some other closely related languages. For example in Simbo (New Georgia group) in the absence of an initial syllable onset the second syllable onset reduplicates, so /opere/ 'spear' becomes /opopere/ 'one who spears', perhaps suggesting reduplicative infixation rather than prefixation in Simbo.

As well as roots with onsetless initial syllables, reduplication also occurs with roots with an underlying first syllable onset cluster. In this situation the echo syllable underlyingly also has the cluster onset. However, on the surface only the initial consonant of the cluster is realised, exemplified by (2.34) d., (2.35) b. and (2.38) b. Once the entire first syllable is reduplicated, a process of reduplicated syllable cluster reduction deletes the onset C2 of the echo syllable.

(2.41)	a. /pɾotu/ 'a distant object'	→ /pɾopɾotu/	→	[popɾotu] 'small lump on the body'
	b. /pɾosa/ 'slap self w. flipper (of turtles)'	→ /pɾopɾosa/	→	[popɾosa] 'wash clothes'
	c. /kɾisu/ 'scoop liquid'	→ /kɾikɾisu/	→	[kikɾisu] 'be scooping liquid'

On the surface, even in normal careful speech, the C2 is never realised. However it is present underlyingly. In syllabifications speakers invariably give the echo syllable with the same form as the entire first syllable of the root. In evaluating emphasised tested forms speakers invariably judge forms with the C2 realised as correct while those with it unrealised as unacceptable (then may immediately give the unrealised C2 form in elicitation). It is apparent that speakers are unaware of, and do not 'hear', the C2 surface deletion.

Similarly, where the first syllable consists of a diphthong, only the first vowel of the VV sequence is realised on the surface.

(2.42)	a. /toi/ 'cook'	→	/totoi/ 'fire'
	b. /saeko/ 'mango'	→	/sasaeko/ 'liver'
	c. /ŋau/ 'eat'	→	/ŋaŋau/ 'be biting (of fish)'

2.4.1.3 Non-synchronic echo syllables

A substantial number of lexemes have identical first and second syllables, indicating historical partial reduplication, but with no corresponding root present in synchronic Kokota. Thus, for example, /mimido/ 'penis', /mamaɾa/ 'be deep', and /ɾeɾeɣi/ 'look after' are synchronically monomorphemic, there being no corresponding forms */mido/, */maɾa/ and *ɾeɣi/.

This historical reduplication was formally identical to the synchronic regime described above. Onsetless echo syllables occur (/ooe/ 'say, talk; word, language', with no synchronic */oe/), as does cluster reduction:

(2.43)	a. */blata/	→	/blablata/	→	[bablata] 'bat sp.'
	b. */preku/	→	/prepreku/	→	[pepreku] 'lip'

As with synchronic reduplicative cluster reduction, speakers have the full cluster in the underlying form of the echo syllable, with deletion occurring on the surface. As with synchronic reduplication, non-synchronic echo syllables also display loss of the second vowel in a VV sequence.

(2.44)	a. */daegɾa/	→	/dadaegɾa/ 'shake with surprise'
	b. */kau/	→	/kakau/ 'crab'

2.4.2 Full reduplication

A small number of quadrisyllabic roots reflect historical full reduplication of disyllabic roots. Of the thirteen attested examples of echo disyllables, two may be onomatopoeic and therefore not reflect reduplication[13]. Of the remaining 11, six have no corresponding synchronic unreduplicated form: /kilekile/ 'k.o. custom axe' but no */kile/, /kulikuli/ 'seaweed sp.' but no */kuli/, /bulobulo/ 'tree sp.'[14] but no */bulo/,

[13] These are /ŋiɾeŋiɾe/ 'cicada' and /ŋuɾuŋuɾu/ 'make k.o. sound (like water running over rocks or wake of a boat)'.

[14] Possibly cognate with /bubulo/ 'early morning ground mist' and /buloma/ 'inedible betel sp.'.

/maramara/ 'lagoon'[15] but no */mara/, /ɣiliɣili/ 'tickle' but no */ɣili/, and /fa ɣonoɣono/ 'be ready' but no */ɣono/.

The remaining 5 lexemes have semantically related unreduplicated counterparts. For four, the relationship is idiosyncratic:

(2.45)	a.	/sekuseku/ 'black trevally'	/seku/ 'tail'
	b.	/fa manemane/ 'be dressed up (of man or woman)'	/mane/ 'man, male'
	c.	/fa ɣaseɣase/ 'be dressed up to show off (woman only)'	/ɣase/ 'girl, female'
	d.	/fa ɣanoɣano/ 'be very good'	/ɣano/ 'smell/taste good; edible root'

However, one example appears to conform to the two subregular uses of partial reduplication: deriving unergative verbs from transitive roots, and nouns from verbs:

(2.46) /izu/ 'read s.th.' → /izuizu/ 'be reading; a reading'

Despite that one apparent subregularity, there does not seem to be any grounds for claiming synchronic full reduplication. The instances of fully reduplicated lexemes with corresponding unreduplicated forms are so few as to warrant regarding them as marginal. Of those five, three ((2.45) b., c. and d.) also display further derivational complexity with the particle /fa/ (with no corresponding */manemane/, */ɣaseɣase/ or */ɣanoɣano/). Even the apparently functionally subregular (2.46) does not appear to suggest a synchronic process, as there is no reason why full reduplication should occur rather than the usual partial reduplication. While the vowel initial nature of the root could suggest a motivation, it is not seen in other similar examples (such as /iipi/ 'be wearing clothes', from /ipi/ 'wear s.th.'; and the historical /ooe/ 'say, talk; word, language').

2.5 Local orthography

The long period of activity on Santa Isabel by the Anglican Melanesian Mission, more recently the Church of Melanesia, particularly given their emphasis on education and the provision of religious texts in local languages, has meant that there exists a limited but relatively long standing tradition of reading and writing in the languages of the area (as well as English). Emphasis has been primarily on Bughotu, Cheke Holo (Maringe) and Zabana (Kia), the languages with the largest numbers of speakers. Initial attempts (involving the Melanesian Mission conventions described in Codrington 1885) resulted in a certain amount of variation in the orthographies in use. More recent attempts by the Church to develop a standardised locally acceptable orthography resulted in Bosma's (1981) report. Although the proposed system was designed for the three target languages, it was regarded as suitable for the remaining languages, including Kokota, as none had phonemes not found in the target languages.

Bosma's report made a number of proposals which were ultimately not adopted. Subsequent work by the Solomon Islands Translation Advisory Group, primarily a Bible translation organisation, resulted in an orthography which has since been widely accepted by Cheke Holo (Maringe) speakers. That orthography is used in the Maringe Bible, and was adopted by White et al for their (1988) Cheke Holo dictionary. This orthography has also been adopted by Kokota speakers. As shown above, a number of phonemes exist in Holo that are not present in Kokota. Symbols for these phonemes are not used in the Kokota orthography, with the exception of the use of the apostrophe to represent a glottal stop. As discussed above, the glottal is not phonemic in Kokota, although it appears to be in Cheke Holo. Kokota speakers often use the apostrophe to represent the optional intervocalic epenthetic glottal, especially in words which are cognate with Holo words in which the glottal is present. A further anomaly in the Kokota orthography reflects the equally anomalous use of both 'j' and 'z' in Holo. Possibly as the manifestation of a sound change underway in Holo, (White et al 1988:x) both [dʒ] and [ʒ] reflect a single phoneme. Although this situation does not pertain in Kokota, and the affricate is not present even phonetically, under the influence of Holo cognate

[15] Presumably reflecting the same now lost root as /mamaɾa/ 'be deep', /maɾakasa/ 'rainbow runner (fish)', /maɾava/ 'bluefin trevally' and /maɾafa/ 'crayfish sp.'

forms that are spelt with a 'j' in Holo tend to be spelt that way in Kokota.[16] In addition to these non-phonemic symbols, in line with Holo orthography the letter 'w' is used in some Pijin loans, despite the absence of [w] from the Kokota phoneme inventory. The orthography in use among Kokota speakers is presented in Table 2.17.

Table 2.17: Kokota orthography.

phoneme	letter	phoneme	letter	phoneme	letter
p	p	m	m	a	a
b	b	m̥	mh (hm)	e	e
t	t	n	n	i	i
d	d	n̥	nh (hn)	o	o
k	k	ŋ	n̄	u	u
ɡ	ḡ	ŋ̊	n̄h (hn̄)		
f	f	l	l		
v	v	l̥	lh (hl)		
s	s	r	r		
z	z (j)	r̥	rh (hr)		
h	h	-	w		
ɣ	g	-	(')		

The bracketed representations of the voiceless sonorants reflect long standing orthographic variation but do not reflect the adopted standard used for Holo.

The only material that I am aware of that has been previously published in Kokota is a booklet containing the Anglican Liturgy (*Buka Nhau Blahi Ka O'oe Kokota* - 'Holy Communion Book in the Kokota Language'). This appears to predate recent standardisation moves. Consequently it reflects earlier mission practice, with /ɣ/ written as 'g', /ɡ/ as '*g*' (in italics), and /ŋ/ written as '*n*' (in italics) (though there are some inconsistencies - 'God' is written without italics, as are all instances of the voiceless velar nasal, which is simply written as 'nh').

Currently further religious translation work is proceeding using the orthography in Table 2.17, on my advice excluding the bracketed alternatives and the glottal.

[16] As indicated in 2.1.1.3, variation exists in Kokota with the corresponding phoneme to Holo [dʒ~ʒ]. This variation, [ʒ~z], is not reflected in the Kokota orthography.