



Te Reo

the Journal of the Linguistic
Society of New Zealand

Volume 66
Issue 1 (*Regular Issue*)

Research Article

2023
Pages 60–96

November 2023

Tomoip phonetics and phonology

Russell Barlow
Max Planck Institute for Evolutionary Anthropology

Don Killian
University of Helsinki

Tomoip phonetics and phonology

RUSSELL BARLOW

Max Planck Institute for Evolutionary Anthropology

DON KILLIAN

University of Helsinki

Abstract

This paper provides a sketch of the phonetics and phonology of Tomoip, an Austronesian language of New Britain, Papua New Guinea. Tomoip does not appear to be closely related to the other Austronesian languages of New Britain. Furthermore, it has been in contact with multiple non-Austronesian languages of the island. These two facts suggest that the language may have had a particularly interesting history. As the first publication dedicated to Tomoip, this paper is intended to serve as a starting point for further description, analysis, and classification of the language.

Keywords

Tomoip; Oceanic; Austronesian; New Britain languages; phonology

1 Introduction

Tomoip is an Austronesian language spoken on the island of New Britain in the western Pacific.¹ Most Tomoip speakers live in Pomio District, East New Britain Province, Papua New Guinea, although there is also a small population found in Talasea District, West New Britain Province. Neighboring community members speak Kol, Sulka, Mengen, and Meramera. Kol and Sulka are both non-Austronesian languages, currently classified as isolates. Mengen and Meramera are both Austronesian languages, but they are not closely related to Tomoip. Indeed, Tomoip does not appear to be closely related to any of the other Austronesian languages spoken on the island of New Britain (Chowning, 1969; Ross, 1988).

Tomoip was first mentioned by Parkinson (1907, pp. 176, 777–781), who recorded 109 Tomoip (*Tumuip*) words and presented them in comparison with about as many Sulka and Mengen words. Rascher (1909, p. 224) may have been referring to Tomoip when he wrote that the language of the “Paleava” people is different from that of both the “Nakanai” and the “Baininger” (i.e., Baining). According to Meier (1914, p. 349), these “Paleava” are the Tomoip (*Tumuip*) people, the term “Paleava” in fact not referring to a particular tribe but rather simply

¹ We wish to thank Simon Mangil for sharing his knowledge of Tomoip with us, as well as Ger Reesink and Malcolm Ross for providing very helpful comments. We also thank the Department of Linguistic and Cultural Evolution at the Max Planck Institute for Evolutionary Anthropology for funding the field research that made this paper possible.

meaning “bush dwellers”. Burger (1913, p. 5) provides a map of New Britain, in which the Tomoip (*Tumuip*) people are depicted as being located north and west of the Mengen people, south of the Sulka people, and east of the “Gaktai” people. Here, too, Meier (1914, p. 349) offers a correction: “Gaktai” is said to be a term used by the Sulka people to refer to their enemies, in this case the Southeast Baining (*Südost-Baininger*), who should not be located (as they are on Burger’s map) south of the Gazelle Peninsula.

Grace (1955a; 1955c) recorded over 500 words, phrases, short sentences, and pronominal paradigms in his unpublished fieldnotes of Tomoip (*Tomōip*). An inspection of Grace’s transcriptions makes it clear that Tomoip is an Austronesian language. However, these notes probably remained unknown to other researchers for several years. Loukotka (1957, p. 59) includes Tomoip (*Tumuip*) among a list of Papuan (i.e., non-Austronesian) languages of New Britain. Lanyon-Orgill (1960, p. 28), on the other hand, writes that the “little-known Tumuip [i.e., Tomoip] language is closely allied to Mengen and contains more Melanesian [i.e., Austronesian] than Papuan elements.” Salzner (1960, p. 27) includes Tomoip (*Tumuip*) in a “Neupommern-Gruppe” (i.e., New Britain Group), along with Kilenge (i.e., Maleu-Kilenge), Bariai, Idne (i.e., Maleu-Kilenge), Kombe (i.e., Kove), Ba’u (possibly Bebeli), O Mengen (i.e., Mengen), and T(a)ulil (i.e., Taulil).

Capell (1962a, p. 91), writing of the sparsity of knowledge of the languages of the area, notes that “population figures for 1940 are not helpful, in that they are based on different and non-linguistic divisions.” Indeed, Capell reports that a combined “Sulka-Tumuip” group had a population of 2,839. Capell’s (1962b, p. 374) map depicts a “Gaktai-Tumuip” language, identifying it as a non-Austronesian language spoken to the southeast of Sulka. However, Lanyon-Orgill (1962, p. 414) comments that “Taulil and Tumuip [i.e., Tomoip] are basically Austronesian languages with a Papuan element”.

Scheller (1966) made audio recordings of two Tomoip speakers in Karlai Plantation in the Wide Bay area. We know of no fieldnotes or other writing associated with these recordings. The recordings consist of elicited words and sentences (mostly recorded without the Tok Pisin prompts), as well as two short texts.

Dyen (1965), who did have access to Grace’s notes (p. 13), recognized Tomoip as Austronesian; however, following his lexicostatistical methods, he could only assign the language to the broad “Austronesian Linkage” or “Malayopolynesian Linkage” (1965, p. 37, 40, 42). Dyen (1965) apparently thought that Tomoip was spoken in southwestern New Guinea and thus interpreted Grace (1955d, p. 339) as having implicitly assigned it to the same subgroup as Nakanai, Bakovi (i.e., Bola), and possibly Kapore (i.e., Bebeli); of such “southwestern” languages, Dyen (p. 52) claims that Kilenge (i.e., Maleu-Kilenge), Peleata (i.e., Pele-Ata), Tomoip, and Pililo (i.e., Solong) “are apparently no more closely related with members of this linkage [i.e., the Willaumez Linkage] than with each other and other members of the Austronesian Linkage.”

Chowning (1969, pp. 21–23) writes that Tomoip is clearly Austronesian (AN), although influenced by the non-Austronesian (NAN) languages of New Britain. Chowning (1969, p. 34) further notes:

The vocabulary contains a large number of AN forms, and there seems to be no doubt about the classification of the language, but it seems quite unlike any other AN language in New Britain on the basis of lexicon alone. There are a number of suggestive but not strong resemblances to some languages located much farther west, in [Southwest New Britain]. ... Its AN component cannot be explained in terms of loans from Mengen or Meramera; the forms are quite different from those in either language. Geography suggests that its relatives should be sought in New Ireland, but so far I have been able to examine little New Ireland material, and none of it resembles Tomoip.

Chowning (1976, p. 373) later adds: “It does seem safe to say that the AN component in Tumuip [i.e., Tomoip] is not derived from any other languages now spoken in New Britain”, concluding that it “is likely to have been there a long time; otherwise it should be easy to locate its place of origin” (p. 379). Chowning (1976, pp. 375–377) also provides a comparative wordlist including 30 Tomoip words, based on the wordlists by Parkinson (1907) and Grace (1955a). After gaining access to Grace’s notes, Capell (1971, pp. 267–268) writes:

Tumuip (Tomoip) is now regarded as fairly closely related to Mungen, and the AN status of Mungen is accepted. Both, however, seem to the author to be marginal cases: the AN element is not great. Tomoip forms (kindly supplied to the writer by Prof. G.W. Grace) do not seem to be structurally AN, apart from the lexical content, which has a degree of AN. The verbal forms, however, are doubtful ... [and] seem to be constructed on principles quite NAN, but the possessive suffixes to nouns are AN ...

Capell (1971, pp. 270–272) also provides a comparative wordlist including 24 Tomoip words, presumably based on Grace (1955a), but apparently with some typographical errors introduced. Wurm (1971, p. 645) follows Capell (1971) in stating that both the vocabulary and the grammar of Tomoip (*Tumuip*) have been influenced by NAN languages.

Lindrud (1980, pp. 178–179) provides a wordlist of 50 items for Tomoip as recorded in Kavale village. He writes that further research into Tomoip “could be of high priority because so little is known of it” (p. 168). Rath (1980, p. 201) provides 10 Tomoip words, which mostly resemble forms found in Lindrud (1980) and perhaps derive from the same field research. Ross (1980) recorded Tomoip speakers from Kaukum village (cf. Ross, 1988, p. 402).

Pawley (1982, p. 34) treats Tomoip (*Tumuip*) as its own primary branch of Oceanic (cf. Pawley & Green, 1984, p. 136).

Ross (1988, pp. 26, 257, 262, 291–293) classifies the language as being most closely related to the languages of New Ireland; he presents sound correspondences with other languages of the Meso-Melanesian cluster (pp. 266–268); and he provides examples of Tomoip reflexes (pp. 281, 286, 287, 354), as well as some description of various aspects of Tomoip grammar (pp. 274–278, 296–298). Although classifying the language as belonging to the New Ireland group, Ross (1988, p. 292) notes:

The sound correspondences for Tomoip ... are the most irregular of any WM Oceanic [i.e., Western Melanesian Oceanic] language and imply that the lexicon comes from two, perhaps three, Oceanic sources as well as perhaps from neighbouring non-AN sources.

Chowning (1996, pp. 22–23) remains skeptical of Tomoip’s inclusion within a New Ireland grouping.

Reesink (2005) contains paradigms for Tomoip pronouns (p. 169) and possessive marking (p. 181), as well as some grammatical and lexical notes on the language (pp. 150, 154–155, 162, 164, 166–167, 171, 182–183), primarily for the purpose of comparing it to the non-Austronesian language Sulka (along with the fellow Austronesian language Mungen).

Brief mention of Tomoip is made at various points in the volumes of the *Lexicon of Proto Oceanic* series: Osmond (2007, pp. 182; 2011, p. 31); Osmond & Ross (2016a, p. 203; 2016b, p. 228); Ross (2007, p. 150; 2008a, pp. 263, 283–284; 2008b, p. 393).

The present description of Tomoip phonetics and phonology is based on elicitation sessions conducted with Simon Mangil, a Tomoip speaker who was born in Milim village, Pomio District in 1967. Recorded elicitation sessions were conducted in October and November 2022 in the provincial capital Kokopo, where Simon Mangil was temporarily

residing. A wordlist based on these recordings is provided in Barlow & Killian (2023). In addition to these recordings, we have also considered the available data from previous researchers in making our analyses. However, unless otherwise stated, all cited Tomoip forms come from recordings made with Simon Mangil in 2022.

The Tomoip phoneme inventory consists of 20 segments: 15 consonants and 5 vowels.

2 Consonants

The Tomoip consonant inventory consists of 15 phonemes: 3 voiceless stops, 3 voiced stops, 3 nasals, 4 fricatives, and 2 liquids (Table 1). The palatal fricative (or approximant) /j ~ j/ is relatively infrequent and perhaps not a full-fledged phoneme. It is treated here, however, as being distinct from the high front vowel /i/ (§2.5). Although the glottal stop [ʔ] also occurs phonetically in the language, it is not considered to have phonemic status (§2.7).

Table 1. Tomoip consonants

	labial	coronal	palatal	velar	glottal
voiceless stops	p	t		k	
voiced stops	b	d		g	
nasals	m	n		ŋ	
fricatives	β	s	j		h
trill		r			
lateral		l			

Reesink's (2005, *passim*) orthography points to this same set of 15 consonants, with his <ng> corresponding to our /ŋ/, his <w> corresponding to our /β/, and his <y> corresponding to our /j/. The consonant correspondences for the languages of the Meso-Melanesian Cluster given by Ross (1988, pp. 266–268) imply that there are at least 13 Tomoip consonants: no comparative information is provided for a putative Tomoip /g/ or Tomoip /j/. Ross's (1988) <v> corresponds to our /β/.

Overall, the consonant inventory very closely resembles the 15-consonant inventory of the nearby non-Austronesian language Sulka, as presented by Tharp (1997, p. 1). The only two differences are that the Sulka inventory is shown with a palatal approximant <j> instead of a palatal fricative <j> and with a velar fricative <y> instead of a velar stop <g>. However, Tharp (1997, p. 5) notes that the palatal consonant “can also be fricativized”, thereby suggesting an even closer affinity with the Tomoip arrangement. Furthermore, Reesink (2005, p. 148) analyzes the Sulka velar fricative [ɣ] as an allophone of a velar stop /g/, thereby also suggesting great similarity between the two inventories. At the same time, however, Reesink (2005, p. 148) analyzes Sulka [β] and [d] as being (only) allophones of /b/ and /r/, respectively, while also including a bilabial approximant <w>, thus bringing his consonant count to 14. Either way, the inventory of consonant phones between the two languages is rather similar.

Of the geographically closest Austronesian languages, Meramera also has a rather similar consonant inventory (cf. Johnston, 1980a, p. 112; Ohtsuba & Ohtsuba, 1992, p. 1). However, Meramera lacks both a rhotic consonant and a palatal consonant. Furthermore, instead of a glottal fricative /h/, Meramera has a glottal stop /ʔ/. Mengen lacks the consonants /d, β, j, h/ (Rath, 1993, p. 72). Mamusi also seems to lack /d, j, h/, but it does have /β/ (Rath, 1980, p. 203). Lote lacks /b, d, g, β, j/, but it has a velar fricative /x/ (Pearson & van den Berg, 2008, p. 8). Nakanai lacks /h/ and – rather surprisingly – the nasals /n/ and /ŋ/ (Johnston, 1980b,

pp. 21, 252–253). According to Ross (1988, pp. 292–293), Tomoip may be more closely related to some geographically more distant languages like Minigir, Kuanua, and Ramoaaina, which, we note, have similar consonant inventories to that of Tomoip: Vinitiri (i.e., Minigir) only lacks /j/ and /h/ (Van Der Mark, 2007, p. 23);² Tolai (i.e., Kuanua) and the Ramoaaina dialects of Molot and Mioko lack /s/ as well as /j/ and /h/ (Mosel, 1980, p. 9).

2.1 Voiceless stops /p, t, k/

There are three voiceless stops in Tomoip: labial /p/, coronal /t/, and velar /k/. All three are common segments in the language and have fairly wide distributions: they each occur word-initially, word-medially, and word-finally. In word-final position, they often occur optionally unreleased (i.e., [p̚, t̚, k̚]). When immediately preceding vowels, they are often slightly aspirated (i.e., [p^h, t^h, k^h]). When immediately preceding front vowels, the coronal and velar stops /t, k/ are occasionally palatalized (i.e., [tʲ, kʲ]).³ We have not observed any labialization in Tomoip (i.e., there are no examples of †[p^w, t^w, k^w], nor of †[b^w, m^w, β^w], etc.). The voiceless velar /k/ is considerably more common than the other two voiceless stops; it is in fact one of the most frequently occurring consonants in the language.⁴ The voiceless stops occur more commonly in word-initial position than in word-final position.⁵ The voiceless coronal /t/ is especially uncommon in word-final position.⁶

Bilabial /p/ word-initially, word-medially (intervocalically), and word-finally

<i>pa</i>	‘to hunt’	<i>kapar</i>	‘white ant’	<i>lakap</i>	‘to run’
<i>pe</i>	‘fish’	<i>lapel</i>	‘to peel’	<i>uriep</i>	‘lizard’
<i>pi</i>	‘to come’	<i>hopiek</i>	‘Malay apple’	<i>morip</i>	‘dolphin’
<i>po</i>	‘to put’	<i>koponj</i>	‘snail sp.’	<i>muop</i>	‘navel’
<i>pu</i>	‘to go’	<i>lopu</i>	‘pigeon’	<i>ɲoup</i>	‘yam mound’

Coronal /t/ word-initially, word-medially (intervocalically), and word-finally

<i>ta</i>	‘PREP’	<i>ratamenej</i>	‘sleeping mat’	<i>ɲorat</i>	‘basket’
<i>tek</i>	‘excrement’	<i>korate</i>	‘child’	<i>get</i>	‘housefly’
<i>ti</i>	‘they [PL]’	<i>ɲoetiŋ</i>	‘side’	<i>rot</i>	‘string’
<i>to</i>	‘to make’	<i>tokitok</i>	‘adze’	<i>kindot</i>	‘scorpion’
<i>tu</i>	‘odor’	<i>kutuna</i>	‘turtle’	<i>kout</i>	‘narrow’

² Van Der Mark (2007, p. 25) additionally describes a marginal Vinitiri phoneme, a preglottalized alveolar nasal /ʎn/. However, this phone seems to derive (historically, at least) from /tn/ sequences, and we wonder whether it can still be analyzed underlyingly as /tn/, as is possibly the case for some speakers (Van Der Mark, 2007, p. 343).

³ The coronal stop /t/ is also optionally palatalized when it immediately precedes the lateral /l/ in a consonant cluster: /t/ → [tʲ] / _ l. As noted in §2.6, the lateral /l/ is generally pronounced as an alveolar [l] but is at times realized as a palatal [ʎ].

⁴ Of the set of 881 formally distinct unbound lexical entries in our wordlist, /k/ occurs in 237 (27%), /m/ in 237 (27%), /l/ in 236 (27%), /r/ in 212 (24%), /n/ in 178 (20%), /t/ in 139 (16%), /p/ in 135 (15%), /ŋ/ in 128 (15%), /b/ in 101 (11%), /h/ in 83 (9%), /β/ in 76 (9%), /s/ in 73 (8%), /d/ in 62 (7%), /g/ in 46 (5%), and /j/ in 25 (3%). These figures should only be taken as a rough guide.

⁵ Word-initial frequencies (among 881 lexical entries): /k/ 129 (15%), /m/ 104 (12%), /t/ 81 (9%), /p/ 78 (9%), /β/ 57 (6%), /l/ 56 (6%), /b/ 54 (6%), /h/ 50 (6%), /n/ 47 (5%), /r/ 44 (5%), /d/ 28 (3%), /s/ 28 (3%), /g/ 25 (3%), /ŋ/ 24 (3%), /j/ 8 (1%).

⁶ Word-final frequencies (among 881 lexical entries): /n/ 71 (8%), /k/ 70 (8%), /ŋ/ 69 (8%), /r/ 67 (8%), /m/ 60 (7%), /l/ 51 (6%), /p/ 25 (3%), /s/ 14 (2%), /t/ 11 (1%). The consonants /b, d, g, β, j, h/ do not occur word-finally.

Velar /k/ word-initially, word-medially (intervocalically), and word-finally

<i>ka</i>	‘fence’	<i>rakabuk</i>	‘animal’	<i>nak</i>	‘grave’
<i>ke</i>	‘you [SG]’	<i>dake</i>	‘oyster’	<i>horek</i>	‘men’
<i>ki</i>	‘bunch’	<i>lakilia</i>	‘knee’	<i>blik</i>	‘bad’
<i>ko</i>	‘crayfish sp.’	<i>makom</i>	‘eucalyptus’	<i>kok</i>	‘mucus’
<i>kur</i>	‘louse’	<i>baku</i>	‘bat sp.’	<i>tabluk</i>	‘eel’

In at least one verb there is an apparent change of /p/ to [β] when /p/ immediately precedes either /j/ or /r/. Thus, the stem-final consonant of the verb *lap* ‘to hit’, as found in *lap-ke* ‘hit you [SG]’ and in *lap a Palua* ‘hit Palua’, contrasts with that found in *lap-jo* ‘hit me’ [‘laβjo] and in *lap-rie* ‘hit him/her/it’ [‘laβrie]. Since /pr/ sequences are known to occur word-initially (cf. §6.1) but are unattested word-internally, we hypothesize that the preceding vowel is also relevant for conditioning this change. Thus, there perhaps exists a lenition rule affecting the voiceless bilabial stop /p/ when directly following a vowel and directly preceding a voiced (continuant) consonant.

/p/ → [β] / V _ C [+voice]

We do not have enough data to test the full extent of this rule. However, it does not seem that other stops (e.g., /k, b/) are lenited in this environment.

2.2 Voiced stops /b, d, g/

There is a complementary set of three voiced stops in Tomoip, matching the set of three voiceless stops in their place of articulation: labial /b/, coronal /d/, and velar /g/. All three occur word-initially and word-finally, but no voiced stop occurs in word-final position. The voiced velar stop /g/ is, after the fricative/approximant /j/, perhaps the least common consonant in Tomoip.⁷ All three voiced stops are pronounced at times with a degree of prenasalization (i.e., [ᵐb, ᵐd, ᵐg]).

The voiced stops are all phonemically contrastive with their voiceless counterparts, both word-initially and word-medially, but not word-finally, as voiced stops do not occur in word-final position.

Contrasts of /b/ versus /p/

<i>bale</i>	‘house’	<i>pale</i>	‘cockatoo’
<i>bla</i>	‘head’	<i>pla</i>	‘nail’
<i>bu</i>	‘to rain’	<i>pu</i>	‘to go’
<i>taba</i>	‘chest’	<i>kapar</i>	‘white ant’
<i>kabe</i>	‘outside’	<i>lapel</i>	‘to peel’
<i>bobo</i>	‘taro sp.’	<i>kopon</i>	‘snail sp.’

⁷ The voiced velar phoneme also has some unusual or unexpected realizations, some of which appear to be lexically conditioned. The realization of /g/ as a voiced velar fricative [ɣ] is more likely to occur intervocalically, but also occurs when /g/ precedes /l/ in consonant clusters, as well as perhaps in other environments. The function word *ga* ‘MOD’, possibly a modal particle, is frequently pronounced [ɣa], irrespective of the preceding segment. There are also occasional realizations of /g/ as the voiced velar approximant [ɰ]. We note that a phonemic velar approximant /ɰ/ is reported for the nearby Baining language Mali (Stebbins, 2011, pp. 22–23; although cf. Stebbins & Tayul, 2012, p. 16).

Contrasts of /d/ versus /t/

<i>denan</i>	‘one’	<i>tenan</i>	‘his/her mother’
<i>dij</i>	‘thorn’	<i>tij</i>	‘hiccup’
<i>dono</i>	‘men’s house’	<i>tono</i>	‘before’
<i>hoda</i>	‘cross-cousin’	<i>motam</i>	‘female pig’
<i>kaedi</i>	‘liver’	<i>ηoetij</i>	‘side’
<i>blodu</i>	‘nose’	<i>motu</i>	‘vulva’

Contrasts of /g/ versus /k/

<i>ga</i>	‘MOD’	<i>ka</i>	‘fence’
<i>ger</i>	‘blowfly’	<i>ker</i>	‘angry’
<i>gur</i>	‘Milky Way’	<i>kur</i>	‘louse’
<i>gom</i>	‘vegetables’	<i>kom</i>	‘testicle’
<i>mage</i>	‘yam sp.’	<i>dake</i>	‘oyster’
<i>agonj</i>	‘1SG.POSS.ED’	<i>makom</i>	‘eucalyptus’

Ross (1988, pp. 266–268) does not include Tomoip /g/ as a reflex of any reconstructed Proto-Oceanic (POc) phoneme. However, Ross’s (1980) wordlists include <g> in forms such as <gomil> ‘cloud’ (wordlist A, p. 3) and <negelem> ‘snake’⁸ (wordlist B, p. 3). The only Tomoip form exhibiting /g/ in Reesink (2005) is <göl> ‘stringbag’ (p. 182).

2.3 Nasals /m, n, η/

There is also a set of three nasal stops, again corresponding in place of articulation to the sets of oral stops: labial /m/, coronal /n/, and velar /η/. All three have a wide distribution and occur word-initially, word-medially, and word-finally. Furthermore, each of the three nasals functions as its own syllable, provided it occurs in word-initial position and immediately precedes another consonant with which it cannot form a consonant cluster (cf. §6.1).

Words exhibiting contrasts among /m/, /n/, and /η/

<i>ma</i>	‘dry’	<i>na</i>	‘to give’	<i>ηa</i>	‘to yawn’
<i>mi</i>	‘again’	= <i>ni</i>	‘PROX’	<i>ηi</i>	‘hibiscus’
<i>tomae</i>	‘to buy’	<i>tonae</i>	‘fish sp.’	<i>tona</i>	‘thence’
<i>βim</i>	‘yam sp.’	<i>βin</i>	‘to hide’	<i>riη</i>	‘sour’
<i>gom</i>	‘vegetables’	<i>gon</i>	‘cold place’	<i>goη</i>	‘1SG.POSS.ED’
<i>liem</i>	‘five’	<i>dien</i>	‘sweaty’	<i>lieη</i>	‘mushroom’
<i>mum</i>	‘chewed-up betel’	<i>bun</i>	‘song’	<i>muη</i>	‘to turn’

In word-initial position, immediately preceding consonants, nasals are usually syllabic (i.e., [m̩, n̩, η̩]). Often these word-initial syllabic nasals are realizations of the prenominal singular marker *ne=* ‘SG’, which is one of a set of three morphemes that index nominal number. Ross (1988, pp. 292–293) refers to these markers as “number-marking proclitics”, whereas Reesink (2005, p. 171) describes them as “articles”. Although it is not clear whether these morphemes indicate any specificity or definiteness (as expected of a class of articles), the marker *n=* ‘SG’ most likely derives from the POc article **na* (cf. Crowley, 1985). This singular marker *ne=*

⁸ This form includes the singular marker *ne=* ‘SG’.

‘SG’ has three single-segment nasal allomorphs that occur with some monosyllabic nouns. Whereas *ne=* ‘SG’ appears to be fully productive, these shorter allomorphs only occur with certain lexemes. The allomorphy can mostly be explained by the place of articulation of the following segment: *n=* occurs before alveolar or palatal consonants /t, d, s, j/; *m=* occurs before labial consonants /p, b, m/; and *ŋ=* occurs before velar or glottal consonants /k, g, h/. Although this allomorphy was doubtless (at some point) motivated by articulatory assimilation, there is no synchronic process of nasal assimilation in Tomoip.

Syllabic [n] deriving from the singular marker *ne=*

<i>n=diŋ</i>	‘thorn [SG]’	
<i>n=tek</i>	‘post of a house [SG]’	
<i>n=sal</i>	‘path [SG]’	
<i>n=le</i>	‘nit [SG]’	
<i>n=ju</i>	‘fire [SG]’	[niu] ⁹

Syllabic [m] deriving from the singular marker *ne=*

<i>m=buo</i>	‘pig [SG]’
<i>m=beŋ</i>	‘night [SG]’
<i>m=pe</i>	‘fish [SG]’
<i>m=per</i>	‘stone [SG]’

Syllabic [ŋ] deriving from the singular marker *ne=*

<i>ŋ=gom</i>	‘vegetable [SG]’
<i>ŋ=kin</i>	‘sand [SG]’
<i>ŋ=ha</i>	‘sun [SG]’ ¹⁰
<i>ŋ=hum</i>	‘earth oven [SG]’

In some cases, it seems best – synchronically, at least – to analyze the nasal as belonging to the nominal stem, since it is never dropped or replaced (e.g., by the plural marker *e=* ‘PL’). Whatever their origin, the nasals in such words are all pronounced syllabically and do not form a (syllable-internal) cluster with the following consonant.

Word-initial NC sequences with nasals that seem to be part of the nominal stem

<i>nde</i>	‘sea’	(cf. <i>ne=nde</i> ‘sea [SG]’, <i>e=nde</i> ‘sea [PL]’)
<i>ndin</i>	‘nest’	(cf. <i>e=ndin</i> ‘nest [PL]’)
<i>msik</i>	‘flying fox’	(cf. <i>ne=msik</i> ‘flying fox [SG]’, <i>e=msik</i> ‘flying fox [PL]’)
<i>mdip</i>	‘coral tree’	(cf. <i>ne=mdip</i> ‘coral tree [SG]’, <i>e=mdip</i> ‘coral tree [PL]’)
<i>ŋgol</i>	‘string bag’ ¹¹	(cf. <i>ne=ŋgol</i> ‘string bag [SG]’, <i>e=ŋgol</i> ‘string bag [PL]’)

⁹ This form creates trouble for the analysis of the palatal fricative [j] as a distinct consonantal phoneme (cf. §2.5), since it is pronounced as a single syllable, apparently with the diphthong /iu/, thereby lending support to the alternative analysis that [j] and [j̥] are both allophones of the high front vowel /i/.

¹⁰ Note that Parkinson (1907, p. 779), however, records <nega> ‘Sonne’ [‘sun’].

¹¹ Reesink (2005, p. 182) gives the forms <ng-göl> ‘SG-stringbag’ and <e-göl> ‘PL-stringbag’. If these forms represent an older pronunciation, then it does seem that our data for this word reflect a reanalysis of the singular marker *ŋ=* ‘SG’ as belonging to the nominal stem.

2.4 Voiceless fricatives /s, h/

There are two voiceless fricatives in Tomoip: a voiceless coronal sibilant fricative /s/ and a voiceless glottal fricative /h/. The sibilant fricative /s/ is variously realized as alveolar [s], as postalveolar [ʃ], or as alveolo-palatal [ç].¹² It occurs word-medially, word-initially, and word-finally, although word-final position is not particularly common. The glottal fricative /h/ occurs word-initially and word-medially, but does not occur word-finally. First, it may be shown that the two fricatives /s/ and /h/ are contrastive.

Words exhibiting contrasts between /s/ and /h/

<i>sa=</i>	‘DU/PL/1SG.IRR’	<i>ha</i>	‘sun’
<i>so</i>	‘to see’	<i>ho</i>	‘ashes’
<i>se=</i>	‘2SG.IRR’	<i>he</i>	‘stick’
<i>ser</i>	‘red snapper’	<i>her</i>	‘grasshopper’
<i>siam</i>	‘no’	<i>hiam</i>	‘sky’
<i>kases</i>	‘flat’	<i>plahe</i>	‘bark’
<i>βiso</i>	‘knife’	<i>tihō</i>	‘why?’

Although /h/ does not occur word-finally, there are occasional instances of /s/ occurring in final position.

Words exhibiting /s/ word-finally

<i>lumbas</i>	‘sponge’
<i>kekes</i>	‘fish sp.’
<i>menes</i>	‘centipede’
<i>sos</i>	‘yam sp.’
<i>rabus</i>	‘coconut bark fiber’

There is occasional free variation between word-initial [h] and [Ø] in Tomoip – that is, an initial /h/ is at times deleted; likewise, an initial [h] is at times inserted before vowels. Similar variability in the use of [h], possibly determined in part by dialect, has been reported for nearby Lote (Pearson & van den Berg, 2008, p. 10) and Nakanai (Johnston, 1980b, pp. 16–18).¹³ Nevertheless, /h/ can be shown to be phonemic, since it contrasts with /Ø/ in minimal pairs.

Minimal pairs of /h/ versus /Ø/

<i>ha</i>	‘sun’	<i>a</i>	‘PERS’
<i>he</i>	‘stick’	<i>e=</i>	‘PL’
<i>hoha</i>	‘bird sp.’	<i>hoa</i>	‘vine, rope’
<i>puha</i>	‘behavior’	<i>pua</i>	‘fish sp.’

¹² The alveolo-palatal [ç] pronunciation is actually the most common realization in our data, particularly when immediately preceding the high front vowel /i/. However, based on Ross’s (1980) recordings, and on his and other researchers’ transcriptions, we do not find any indication that alveolo-palatal fricatives are particularly common in Tomoip. What our consultant pronounces as [ç] is perhaps more commonly pronounced by others (or in other dialects) as [s]. However, we note here a remark made in Ohtsuba & Ohtsuba’s (1992, p. 7) description of nearby Melamela (i.e., Meramera): “/s/ has more or less the quality of [ʃ], but not quite.”

¹³ We also note some variability of initial /h/ in Tok Pisin (cf. Hall, 1943, p. 14; Mihalic, 1971, p. 5; Laycock, 1985, p. 302).

Still, there are instances in which the free variation for a particular lexeme is so great that it is difficult to recover a single underlying form. For example, the word for ‘some, something’ can be pronounced either as [‘hasi] or as [‘asi] (as well as simply as [si]).

In Grace’s (1955a) phonetic transcriptions, /h/ is sometimes written as <g> or <y>, as in <bugε> ‘belly’ (p. 84) for our /buhe/, <burugo?> ‘meat’ (p. 85) for our /burho/, <bayo> ‘count’ (p. 87) for our /βaho/, and <gaya> ‘swim’ (p. 84) for our /haha/. In one case, there is variability between <g> and <h> recorded between two meanings of the same lexeme: <ha> ‘sun’ (p. 86) alongside <ga> ‘day’ (p. 86). Similarly, in Lindrud’s (1980, pp. 178–179) phonetic transcriptions, we find <g> for our /h/, as in <‘gətɛl> ‘egg’ for our /hotel/ or <‘bulhε (`bulgε)> ‘root’ for our /bulhe/. In Parkinson’s (1907) wordlist, we find instances both of <g> and of <gh>, as in <nega> ‘Sonne’ [‘sun’] (p. 779) and <búghelon> ‘sein Bauch’ [‘his belly’] (p. 779). In our data, however, /h/ is never pronounced as a velar consonant. However, as mentioned in §2.3, /h/ does pattern with the velar consonants in terms of selecting η= as the allomorph of the prenominal singular marker. Furthermore, we note an alternation between /g/ and /h/ in the paradigm of edible possessive classifiers (e.g., *agoη* ‘1SG.POSS.ED’ vs. *ahan* ‘3SG.POSS.ED’). There is perhaps dialectal as well as synchronic variation at play here.

2.5 Voiced fricatives /β, j/

There are two voiced fricatives in Tomoip: a voiced bilabial fricative /β/ and a voiced palatal fricative /j/. Like the voiced stops /b, d, g/, the voiced fricatives /β, j/ occur word-initially and word-medially but not word-finally. Although there is good evidence to support the existence of an underlying voiced bilabial fricative /β/, the phonemic status of the other voiced fricative, the voiced palatal fricative /j/, is less secure. This relatively infrequent phone might in fact prove to be (only) a consonantal realization of the high front vowel /i/. Nevertheless, in part based on phonetic clues and in part based on native speaker intuitions, this segment is treated here as a distinct phoneme, albeit a relatively marginal one.

Although the voiced bilabial fricative /β/ is realized in some environments as a labial-velar approximant [w] (and even in one word as a labiodental fricative [v]), it is never realized as a bilabial stop. Indeed, the bilabial fricative contrasts with the bilabial stop [b], which, it should be noted, is never pronounced as [β], even intervocally.¹⁴

Words exhibiting contrasts between /β/ and /b/

<i>βa</i>	‘to plant’	<i>ba</i>	‘bean’
<i>βer</i>	‘cod’	<i>ber</i>	‘lungs’
<i>βaβen</i>	‘plant sp.’	<i>kabe</i>	‘outside’
<i>koba</i>	‘ironwood tree’	<i>loban</i>	‘pain’
<i>nobek</i>	‘Jew’s harp’	<i>nobεη</i>	‘lime’
<i>koboε</i>	‘spleen’	<i>bobo</i>	‘taro sp.’

When immediately preceding a rounded vowel /u, o/, the voiced bilabial fricative /β/ is optionally realized as a labial-velar approximant [w]. Although this is almost always the case when /β/ is followed by /u/, there is more variability when /β/ is followed by /o/. For example, there are a number of words in which the underlying sequence /βu/ is pronounced as [wu].

¹⁴ The behavior of the bilabial plosive [b] may be contrasted with that of the velar plosive [g], which has the common intervocalic allophone [ɣ]. It may also be noted that the coronal plosive [d] is never realized as a fricative †[ð].

/β/ → [w] / _ u

/βulβul/	['wɔlwɔl]	'elbow'
/βuŋglie/	['wɔŋglie]	'vine sp.'
/βuβen/	['wɔβen]	'fish sp.'

This conditioning is less consistent when the following vowel is the mid back rounded vowel [o].

/β/ → [w] / _ o (optional)

/βoŋ/	[wɔŋ ~ βɔŋ]	'wasp'
/koβoe/	[kə'βoe]	'spleen'

Also, when immediately following a word-initial high back vowel /u/, the voiced bilabial fricative /β/ is realized as a labial-velar approximant, coalescing with the preceding vowel to produce a long consonantal segment [w:]. However, this environment seems to occur in only one word, whose underlying form is not entirely certain.¹⁵

/uβ/ → [w:] / # _

/uβe/	[w:e]	'tree'
-------	-------	--------

Despite their apparent affinity, the consonant /β/ and the vowel /u/ may nevertheless be contrasted, including by means of at least one minimal pair.

Minimal pair of /β/ versus /u/

βa	'to plant'	ua	'where?'
----	------------	----	----------

In at least one word, the bilabial fricative /β/ is sometimes realized as a labiodental fricative [v], presumably a lexically conditioned pronunciation. The word in question is *βiso* 'knife' ['viso ~ 'βiso]. The peculiar phonetic behavior of this word may be due to its status as an areal loanword.¹⁶

In Parkinson (1907), /β/ is transcribed as <v>, as in <vatōno> 'Gehöft' ['compound', i.e., 'village'] (p. 778) for our /βatono/. In Grace (1955a), /β/ is transcribed variously as <v> or , as in <tɛlva> 'sling' (p. 91) for our /tɛlβa/, or <bi-ri> 'wash' (p. 84) for our /βiri/. For the word 'to boil', both <bulbulj̄e> and <vulvulj̄e> are written (p. 89). Grace (1955a, p. 87) probably noticed the bilabial quality of the consonant, as he remarks at some point: "b or β throughout". Grace (1955a, p. 90) also probably noticed the contrast between /β/ and /b/, remarking "not v" after his transcription of <buq> 'pig'. Ross (1988, pp. 266–268) writes this

¹⁵ We note the variability in how this word has been recorded. Grace (1955a, p. 85) writes <ue> (also <'ue> 'mast', p. 91); Lindrud (1980, pp. 178–179) writes <'uwe>; Ross (1980, wordlist A, p. 4) writes <u-e>; Ross (1988, p. 281) writes <ue>. For 'Baum' ['tree'], Parkinson (1907, p. 778) has only a question mark, presumably having been unable to ascertain a lexical form for this concept.

¹⁶ Cf. Parkinson (1907, p. 782): "... die Wörter *a uamba* (die Axt) und *a visso* (das Messer), die wir bei unserem ersten Besuch bei den O Mengen an der Wasserfallbucht hörten, scheinen ihren Weg quer durch Neupommern gefunden zu haben." ["The words *a uamba* 'the axe' and *a visso* 'the knife', which we heard during our first visit to the Mengen in Waterfall Bay, seem to have found their way across New Britain.]. See Osmond & Ross (1998, p. 91) for similar forms in other Oceanic languages, including Nakanai *viso* '(steel) knife', which is flagged there as a possible loan. We also note Kol <'biso> 'knife' (Lindrud, 1980, p. 180).

consonant as <v>, whereas Reesink (2005, p. 164) writes it as <w> (as found in the causative prefix <wa->).

Tomoip speakers seem to write the voiced bilabial fricative as <v>. At least this is true for Simon Mangil, as well as for Malcolm Ross's consultant Martin Marlow.

The voiced palatal fricative [j] is the least frequently occurring phoneme. It is possible that this phone is simply a consonantal allophone of the high front vowel /i/. However, phonetically, it is quite different from the vowel, since it has a high degree of constriction. Also, like the voiced bilabial fricative /β/, it can be pronounced long (as [j:]). It is also at times realized as a palatal approximant [j̞].

Word-initial voiced palatal fricative /j/

<i>jaboŋ</i>	'sea snail sp.'
<i>jame</i>	'crocodile'
<i>jamleŋ</i>	'to dream'
<i>jo</i>	'I, me'
<i>ju</i>	'hair'

Intervocalic voiced palatal fricative /j/

<i>kujan</i>	'turmeric'
<i>majau</i>	'gecko'
<i>nuje</i>	'water'
<i>lojok</i>	'afternoon'
<i>donoju</i>	'trash pile'

One argument in favor of treating the palatal fricative [j] as phonemically distinct from the high front vowel [i] is that failing to do so would imply the existence of three-vowel sequences (ViV) in the language. Although Tomoip makes abundant use of diphthongs (cf. §4), three-vowel sequences are otherwise unattested in the language.

Another piece of evidence in favor of this distinction is the fact that the presumed consonantal phoneme /j/ conditions the lenition of a preceding bilabial stop /p/ (cf. §2.1), whereas the vowel /i/ does not, as in the words *kapietom* 'rainbow' or *hopiek* 'Malay apple', which are not pronounced with the sequence [βi].

Although Ross (1988, pp. 266–268) does not include Tomoip /j/ as a reflex of any reconstructed POc phoneme, his (1980) transcriptions made with Martin Marlow reflect this phone in forms such as <nuje> 'water'¹⁷ (wordlist A, p. 4) and <tanyali> 'upwards' (wordlist D, p. 5). The phoneme /j/ does not occur frequently in Reesink (2005), but it seems to be represented in forms such as <yo> '1SG' (p. 169) and <yu> 'hair' (p. 183). Parkinson (1907, p. 779) writes it as <j> in <nuje> 'Wasser' ['water']. Grace (1955a) generally transcribes it as <ĵ>, as in <nuĵe> 'water' (p. 86) and <eĵu> 'hair'¹⁸ (p. 88), whereas Lindrud (1980) transcribes it as <y>, as in <kayu> 'fire'¹⁹ (pp. 178–179).

A palatal phoneme with the basic realization [j] is not attested for neighboring languages. However, in his fieldnotes for the nearby non-Austronesian language Sulka, Grace (1955b, p. 122) notes "y quite fricative" (at least for the form <yir> 'hair'); Tharp (1997, p. 5)

¹⁷ Cf., however, <nuie> 'river' (Ross, 1980, wordlist B, p. 2), in which this same lexeme is written with <i> instead of <y>.

¹⁸ This form includes the plural marker *e* = 'PL'.

¹⁹ This form includes the class prefix *ka-*, which often precedes terms for trees, long things, and plants; the item in the wordlist probably refers to 'firewood'.

likewise remarks that this consonant “can also be fricativized”. Furthermore, Hellwig (2019, p. 31) reports that the voiced velar fricative /ɣ/ has the allophone [j] in the Baining language Qaqet.

Tomoip speakers seem to write the voiced palatal fricative as <y>. Again, this is based on writing samples we have from Simon Mangil and from Martin Marlow.

2.6 Liquids /r, l/

There are two liquids in Tomoip: a rhotic /r/ and a lateral /l/. The phoneme /r/ is most commonly pronounced as a voiced alveolar trill [r], but it is alternatively pronounced as a voiced alveolar tap [ɾ]. The lateral /l/ is generally pronounced as a voiced alveolar lateral approximant [l], but it is at times pronounced as a palatal lateral approximant [ɭ]. This realization is especially common when the lateral immediately precedes the high front vowel /i/. Both /r/ and /l/ occur word-initially, word-medially, and word-finally. It may be shown that the two sounds /r/ and /l/ contrast with each other.

Minimal pairs of /r/ versus /l/

<i>rie</i>	‘to dance’	<i>lie</i>	‘ginger’
<i>ro</i>	‘to carry’	<i>lo</i>	‘PREP’
<i>para</i>	‘pandanus sp.’	<i>pala</i>	‘bridge’
<i>sir</i>	‘straight’	<i>sil</i>	‘to sing’
<i>utor</i>	‘to talk’	<i>utol</i>	‘to be hungry’

A phonemic contrast between /r/ and /l/ is sometimes considered to be more of an Austronesian than a “Papuan” typological feature (cf. Foley, 1986, p. 55; Reesink, 2005, p. 147). However, such a contrast is common in the non-Austronesian languages of New Britain.²⁰

2.7 The glottal stop [ʔ]

There is no phonemic glottal stop in Tomoip. This phone, however, occurs quite regularly at the end of vowel-final utterances, at least in the speech of some speakers.²¹ There are no minimal pairs between forms with glottal stops and forms with any other segment (nor between those with and those without glottal stops). In connected speech, words are pronounced without any intervening glottal stops, even when one word ends with a vowel and the following word begins with a vowel. Thus, we characterize this phonetic phenomenon as being utterance-final rather than word-final. However, since our recordings contain many elicited words and phrases spoken in isolation, we have many examples of glottal stops.

²⁰ Examples of non-Austronesian languages in New Britain with two (or more) contrastive liquid consonants include Anêm (i.e., Anem) (Thurston, 1976, p. 15), Kol (Lindrud, 1992, p. 1), Sulka (Reesink, 2005, p. 148), Ura (Stanton, 2007, p. 21), Mali (Stebbins, 2011, p. 18), Qaqet (Hellwig, 2019, p. 20), and Tulil (i.e., Taulil) (Meng, 2018, p. 16). On the other hand, the neighboring Austronesian language Meramera lacks such a distinction, having a lateral consonant but no rhotic one (Johnston, 1980b, p. 112; Ohtsuba & Ohtsuba, 1992, p. 1).

²¹ Although we observe many word-final glottal stops in our recordings, we notice them somewhat less often in those of Scheller (1966) and considerably less often in those of Ross (1980).

3 Vowels

The Tomoip vowel inventory consists of 5 phonemes: 2 high vowels, 2 mid vowels, and 1 low vowel: /i, u, e, o, a/ (Table 2). All five vowels have a fairly broad distribution, with /a/ being the most common and /u/ being the least common.²² All vowels occur word-finally.²³ However, it is fairly less common for vowels to occur word-initially, with the mid vowels /e, o/ essentially never occurring in word-initial position.²⁴ Although a schwa-like vowel [ə] optionally occurs in reduced syllables, it does not have phonemic status in the language (§3.6). Also, although vowels in stressed syllables are often pronounced with greater duration, there is no evidence of phonemic vowel length (§5).

Table 2. Tomoip vowels

	front	central	back
high	i		u
mid	e		o
low		a	

The forms provided in Ross (1988, *passim*) suggest that Tomoip has (at least) these same five vowel phonemes. Reesink's (2005, *passim*) orthography, however, suggests a seven-vowel system, consisting of these five vowels plus a low-mid front vowel /ɛ/ <ë> and a low-mid back vowel /ɔ/ <ö>. Although we note the presence of the phones [ɛ] and [ɔ], we analyze them here as allophones of /e/ and /o/, respectively (see §3.3 and §3.4).

Although Tomoip's consonant inventory resembles that of neighboring Sulka (cf. §2), its vowel inventory is notably smaller. Reesink (2005, pp. 149–150, 153) analyzes Sulka as having seven vowel phonemes /i, u, e, o, ɛ, ɔ, a/, as well as exhibiting phonemic vowel length for all vowels. Kol is reported to have a similar seven-vowel arrangement, with length distinctions present for most but not all vowels (Lindrud, 1992, p. 2; Reesink, 2005, p. 153).²⁵

On the other hand, the nearest Austronesian languages, Mengen (cf. Rath, 1993, p. 74) and Melamela (i.e., Meramera) (cf. Ohtsuba & Ohtsuba, 1992, p. 2), also have five-vowel systems consisting of /i, u, e, o, a/. This is also the system typically reconstructed for Proto-Oceanic (cf. Ross, 1988, p. 93; Lynch et al., 2002, p. 65). Lote, however, exhibits a six-vowel system of /i, u, e, ɛ, ɔ, a/ (Pearson & van den Berg, 2008, pp. 8–9). Minigir, Kuanua, and Ramoaina all exhibit six-vowel systems of /i, u, e, ə, o, a/ (cf. Van Der Mark, 2007, p. 26; Mosel, 1980, p. 20; Davies & Fritzell, 1992, p. 5).²⁶ Notably, Tomoip lacks a phonemic mid central vowel †/ə/, although a reduced schwa-like [ə] occurs phonetically in some environments (§3.6).

²² Of the set of 881 formally distinct unbound lexical entries in our wordlist, /a/ occurs in 409 (46%), /e/ in 307 (35%), /o/ in 281 (32%), /i/ in 267 (30%), and /u/ in 235 (27%).

²³ Word-final frequencies (among 881 lexical entries): /a/ 125 (14%), /e/ 117 (13%), /o/ 88 (10%), /i/ 67 (8%), /u/ 46 (5%).

²⁴ Word-initial frequencies (among 881 lexical entries): /a/ 41 (5%), /u/ 13 (1%), /i/ 11 (1%), /o/ 2 (0.2%), /e/ 1 (0.1%).

²⁵ We also note that Kol appears to have phonemically contrastive phonation types (e.g., modal voice vs. breathy voice), although more research is needed (authors' fieldnotes).

²⁶ Mosel (1980, pp. 20–21) also describes phonemic vowel length for at least the Rakunai and Raluana dialects of Tolai (i.e., Kuanua).

3.1 *The high front vowel /i/*

The high front unrounded vowel /i/ is generally pronounced like the cardinal vowel [i] in open syllables, but it is alternatively pronounced more centralized as [ɪ], especially when in closed syllables. The high front vowel /i/ may be contrasted with the only other high vowel in the language, the high back vowel /u/.

Minimal pairs of /i/ versus /u/

<i>i</i>	‘he, she, it’	<i>u</i>	‘smelly’
<i>in</i>	‘to eat’	<i>un</i>	‘here’
<i>pi</i>	‘to come’	<i>pu</i>	‘to go’
<i>kik</i>	‘kingfisher’	<i>kuk</i>	‘finger’
<i>him</i>	‘to kill’	<i>hum</i>	‘earth oven’

The high front vowel /i/ may also be contrasted with the only other front vowel in the language, the mid front vowel /e/.

Minimal pairs of /i/ versus /e/

<i>i</i>	‘he, she, it’	<i>e=</i>	‘PL’
<i>pi</i>	‘to come’	<i>pe</i>	‘fish’
<i>kuri</i>	‘belch’	<i>kure</i>	‘seed’
<i>him</i>	‘to kill’	<i>hem</i>	‘vein’
<i>sir</i>	‘straight’	<i>ser</i>	‘red snapper’

3.2 *The high back vowel /u/*

The high back rounded vowel /u/ is generally pronounced like the cardinal vowel [u] in open syllables, but it is alternatively pronounced more centralized as [ʊ], especially when in closed syllables. Contrasts between the high back vowel /u/ and the high front vowel /i/ are given in §3.1. The high back vowel /u/ may also be contrasted with the only other back vowel in the language, the mid back vowel /o/.

Minimal pairs of /u/ versus /o/

<i>pu</i>	‘to go’	<i>po</i>	‘to put’
<i>ju</i>	‘hair’	<i>jo</i>	‘I, me’
<i>kum</i>	‘crab sp.’	<i>kom</i>	‘testicle’
<i>kure</i>	‘seed’	<i>kore</i>	‘mountain ridge’
<i>mum</i>	‘chewed-up betel’	<i>mom</i>	‘to burn’

3.3 *The mid front vowel /e/*

The mid front unrounded vowel /e/ is usually pronounced higher than the cardinal vowel [e] when in open syllables – that is, as [e̞]. Word-finally, /e/ is typically pronounced as a diphthong [ei] or [eɪ], provided it is not immediately followed by an utterance-final glottal stop, in which case it often pronounced as [eʔ]. (In closed syllables – including those closed by a phonetic glottal stop – the mid front vowel is generally pronounced lower, as [ɛ].) This “high” quality of the vowel /e/ can also be found in the two geographically closest non-Austronesian

languages, Sulka and Kol. For Sulka, Reesink (2005, p. 149) describes the mid front vowel /e/ as fluctuating “somewhat between the more central close vowel [ɪ] and the close-mid [e]”. Similarly, Grace (1955b, p. 124) comments in his fieldnotes for Sulka: “e higher than cardinal e”. Müller (1915–1916, p. 75) attributes this variability to different dialects across the Sulka-speaking villages. As for Kol, we have observed the common quality [ɛ̞] of the mid front vowel /e/ (authors’ fieldnotes). This vowel quality thus contrasts, for example, with that of nearby Austronesian Nakanai, in which the mid front vowel /e/ is described phonetically as a mid-open front vowel [ɛ] (Johnston, 1980b, p. 253).

The mid front vowel /e/ may be contrasted with the only other mid vowel in the language, the mid back vowel /o/.

Minimal pairs of /e/ versus /o/

<i>le</i>	‘nit’	<i>lo</i>	‘PREP’
<i>he</i>	‘stick’	<i>ho</i>	‘ashes’
<i>kare</i>	‘to fry’	<i>karo</i>	‘tall ginger’
<i>ler</i>	‘to sleep’	<i>lor</i>	‘boil, blister’
<i>nek</i>	‘trash’	<i>nok</i>	‘coconut frond midrib’

Contrasts between the mid front vowel /e/ and the high front vowel /i/ are given in §3.1.

The mid front vowel /e/ almost never occurs word-initially. Indeed, it is only attested in this position as a plural marker *e* = ‘PL’ and, it seems, as a possessive preposition *e* ‘POSS’.²⁷

Although Reesink’s (2005) orthography suggests a possible phonemic contrast between [e] (written as <e>) and [ɛ̞] (written as <ë>), we consider these two phones to be allophones of the same phoneme, their alternation generally being conditioned by syllable shape. For example, Reesink (2005, p. 183) gives the form <buhë> ‘belly’. In our own recordings, however, we have found that the 3SG possessed form of this noun is generally pronounced [‘buheˈlɔn], whereas the unpossessed form is variously pronounced as [‘buhe], [‘buhei], or [‘buheʔ]. In other words, the low-mid front vowel [ɛ̞] occurs in closed syllables (including those closed by a glottal stop), whereas the mid front vowel [e] occurs in open syllables. Similarly, Reesink (2005, p. 171) gives the form <bu-lmë> ‘a coconut fruit’. Indeed, in our own data, the most common citation form for this item is [‘bɔlmɛʔ] – that is, with an utterance-final glottal stop. However, when not utterance-final, the item is pronounced [‘bɔlmei], as in [‘bɔlmei manˈmanaʔ] ‘the coconut floats’. We also note the apparent lexeme-internal variability (assuming no typographical error) between <kem> ‘1EXCL’ (Reesink, 2005, p. 164) and <këm> ‘1EXCL’ (p. 169), which also suggests the absence of a phonemic distinction between [e] and [ɛ̞].

Finally, we observe that Grace’s (1955a) phonetic transcriptions show alternations between <e> and <ɛ̞> for the same underlying vowel in different recordings of the same word: for example, <nuje> ‘milk’ (p. 84) versus <nujɛ̞> ‘water’ (p. 86). Sometimes two transcriptions are given in the same place, such as both <deltu> and <deltu> ‘standing’ (p. 84) and both <tiriɛ̞> and <tiriɛ̞> ‘dance’ (p. 85).

²⁷ Cf. Reesink’s (2005, p. 183) alternative analyses of the form *psie* ‘penis’ either as being monomorphemic or as containing a possessive classifier (or adposition): *psi-e* ‘penis-POSS’. We provisionally treat *e* ‘POSS’ as a possessive preposition, while acknowledging that the forms appearing to contain it are difficult to analyze.

3.4 The mid back vowel /o/

The mid back rounded vowel /o/ is usually pronounced higher than the cardinal vowel [o] when in open syllables – that is, as [ɔ]. Word-finally, /o/ is typically pronounced as a diphthong [ou] or [ɔu], provided it is not immediately followed by an utterance-final glottal stop, in which case it is often pronounced as [ɔʔ]. (In closed syllables – including those closed by a phonetic glottal stop – the mid back vowel is generally pronounced lower, as [ɔ].) Like the “high” quality of the mid front vowel /e/, the “high” quality of the mid back vowel /o/ also appears to be a feature of the non-Austronesian languages Sulka and Kol. In Sulka, there is reported fluctuation between [o] and [u] in the pronunciation of /o/ (Müller, 1915–1916, p. 75; Schneider, 1962, p. 12; Reesink, 2005, p. 150). Similarly, in Kol we have observed the common quality [ɔ] for the mid back rounded vowel /o/ (authors’ fieldnotes).

Contrasts between the mid back vowel /o/ and the mid front vowel /e/ are given in §3.3. Contrasts between the mid back vowel /o/ and the high back vowel /u/ are given in §3.2.

Like the other mid vowel, the mid back vowel /o/ essentially never occurs word-initially. There are only two words that seem to exhibit word-initial /o/. If not loans, their atypical phonotactic behavior is difficult to account for. The two words in question are *olsak* ‘hardwood tree sp.’ (*garamut* in Tok Pisin) and *otβie* ‘strap of a string bag’. There is also, possibly, a possessive preposition of the form *o* ‘POSS’, which, if it does occur, is attested with only one noun.²⁸ Finally, there are two other questionable [o]-initial forms, both of which occur as alternative versions of other forms: [orir] ‘fuzz’ (alongside *rir* ‘fuzz’) and [oβrik] ‘dew’ (alongside *βrik* ‘dew’).

Reesink’s (2005) orthography distinguishes between [o] (written as <o>) and [ɔ] (written as <ö>). However, we do not find any evidence of a phonemic contrast between the two forms. Rather, much like the alternations between [e] and [ɛ], the alternation between [o] and [ɔ] is generally conditioned by syllable shape: the mid back vowel [o] occurs in open syllables, whereas the low-mid back vowel [ɔ] occurs in closed syllables. It could be the case that we have conflicting data, however. For example, Reesink (2005, p. 183) provides the forms <mötu> ‘vulva’ and <ngom> ‘eye’, whereas we have recorded [ˈmotuʔ] ‘vulva’ and [ŋɔm] ‘eye’.

Furthermore, we observe that Grace’s (1955a) phonetic transcriptions show alternations between <o> and <ɔ> for the same underlying vowel in different recordings of the same word: for example, <nɔ·lɔ> ‘woman’ (p. 85) versus <nɔ·lo> ‘girl’ (p. 90). Sometimes two transcriptions are given in the same place, such as both <toni> and <tɔni> ‘sew’ (p. 85) and both <ko·lun> and <kɔlun> ‘hunt’ (p. 85).

3.5 The low central vowel /a/

The low central unrounded vowel /a/ is generally pronounced as the crosslinguistically common low central unrounded vowel – that is, as [ä]. It may be contrasted with the four other vowels in the language.

²⁸ Cf. Reesink’s (2005, p. 183) alternative analyses of the form *kao* ‘chin’ either as being monomorphemic or as containing a possessive classifier (or adposition): *ka-o* ‘chin-POSS’. If the latter analysis is correct, then this morpheme might be the same as the preposition *o* in the example sentence <Ti ta nging o yo> ‘They laughed because of me’ (p. 164). However, we have not found any evidence in our data of a morpheme *o* (aside from, perhaps, the form used with this one body-part term).

Contrasts between /a/ and the other four vowels /i, u, e, o/

<i>a</i>	‘PERS’	<i>i</i>	‘he, she, it’
<i>ka</i>	‘fence’	<i>ki</i>	‘bunch’
<i>sal</i>	‘path’	<i>sil</i>	‘to sing’
<i>alum</i>	‘child’	<i>ulum</i>	‘ilimo tree’
<i>dan</i>	‘brown ant’	<i>dun</i>	‘to sit’
<i>ba</i>	‘bean’	<i>be</i>	‘armpit’
<i>lap</i>	‘to hit’	<i>lep</i>	‘jewelry’
<i>pa</i>	‘to hunt’	<i>po</i>	‘to put’
<i>kare</i>	‘to fry’	<i>kore</i>	‘mountain ridge’
<i>βaha</i>	‘wide’	<i>βaho</i>	‘to count’

3.6 Reduced vowels [ə]

In certain environments, the non-high vowels /a, o, e/ are optionally centralized, losing their distinguishing features, neutralizing as it were to an indistinct schwa-like vowel [ə]. First, the low vowel /a/ is generally raised to [ə] when immediately preceding a spirantized realization of the voiced velar stop (i.e., [ɣ]). While the intervocalic position of the voiced velar stop /g/ conditions its realization as a fricative, this spirantization in turn conditions the raising and reduction of /a/ to [ə].

Reduction of /a/ to [ə] before [ɣ]

/pagas/	[pə'ɣas]	‘coconut frond’ ²⁹
/βagam/	[βə'ɣam]	‘flattened palm stem’
/mage/	[mə'ɣei]	‘yam sp.’

Second, the mid back vowel /o/ is generally raised and reduced to [ə] when immediately following the high back vowel /u/ (i.e., the diphthong /uo/ is often realized as [uə]).

Reduction of /o/ to [ə] after [u]

/buop/	[buəp]	‘REFL’ ³⁰
/muop/	[muəp]	‘navel’
/imuo/	[i'muə]	‘enemy’
/buo/	[buə]	‘pig’

This same phonological rule is described for Sulka (Reesink, 2005, p. 150).

Finally, the mid front vowel /e/ exhibits some lexically determined centralized allophones that occur in grammatical morphemes – namely, in the singular number-marking proclitic *ne* = ‘SG’, and in two prefixes, *me-* and *te-*, which commonly precede adjectives. When not deleted entirely, the vowel in these morphemes would be expected to be pronounced as [ɛ],

²⁹ Müller (1907, p. 81) notes a similar change in the neighboring Mengen language: “In unbetonten Silben ist der richtige Laut eines Vokales oft unklar, so hört man z. B. *pägūs tua* und auch *pügūs tua Brotbaum-Blüte ...*” [“In unstressed syllables, the correct pronunciation of a vowel is often unclear; thus, one hears, e.g., *pägūs tua* and also *pügūs tua* ‘breadfruit tree blossom’ ...”]. This Mengen word (cf. Madden, 2006, p. 33: *pagasi* ‘coconut branch’) could very well be a loan into Tomoip.

³⁰ See also Reesink (2005, p. 166): “The form *buop* [buəp] I obtained for reciprocal actions is the same as the one given to express reflexivity ...”.

as an unstressed mid front vowel followed by a consonant. Indeed, the forms [nɛ], [mɛ], and [tɛ] are common. However, it is also common for this vowel to be realized variously in these morphemes, closer to [ɪ], [ə], or [ɐ]; or, possibly, to be deleted entirely.

4 Diphthongs

Tomoip is a language rich in diphthongs. Since there are five vowels (and geminate vowels do not occur), there are 20 logically possible V_1V_2 sequences. It seems that potentially all of these vowel sequences are possible in the language, although some are relatively uncommon. Also, although the sequence /eo/ is attested, it is only known to occur across a morpheme boundary; it is not pronounced as a diphthong but rather as two syllables. The other 19 vowel sequences, however, have been observed within individual syllables. These 19 diphthongs are presented in Table 3, along with the number of words attested as exhibiting them (from a list of 881 formally distinct unbound lexemes).

Table 3. Tomoip diphthongs: number of occurrences (among 881 words)

	a	e	i	o	u
a	—	ae 19	ai 11	ao 5	au 12
e	ea 2	—	ei 2	eo *	eu 3
i	ia 21	ie 42	—	io 11	iu 8
o	oa 4	oe 9	oi 2	—	ou 5
u	ua 19	ue 2	ui 2	uo 12	—

* The sequence /eo/ is attested as occurring, but not as a diphthong.

Thus, the most commonly occurring initial vowels of diphthongs are /i/ (82 words), /a/ (47 words), and /u/ (35 words). The mid vowels are less common in this position: /o/ (20 words) and /e/ (7 words). As the second member of a diphthong, the most commonly occurring vowels are /e/ (72 words) and /a/ (46), with /o/ (28 words), /u/ (28 words), and /i/ (17 words) occurring somewhat less frequently.

The diphthong /ie/ is particularly common, occurring both in closed syllables and in open syllables.

Examples of words containing the diphthong /ie/

<i>bien</i>	‘worm sp.’
<i>kie</i>	‘dry’
<i>kusier</i>	‘we [PL.INCL]’
<i>lie</i>	‘ginger’
<i>piel</i>	‘cordyline’

Examples of other relatively common diphthongs in open syllables

/ia/	<i>biria</i>	‘breadfruit’	/ae/	<i>kae</i>	‘to choose’
/ua/	<i>nua</i>	‘digging stick’	/au/	<i>malau</i>	‘wild fowl’
/uo/	<i>buo</i>	‘pig’	/io/	<i>lio</i>	‘side of the body’
/ai/	<i>aidi</i>	‘3PL.POSS.ED’ [ai.di]	/oe/	<i>paβoe</i>	‘red emperor’

Examples of other relatively common diphthongs in closed syllables

/ia/	<i>mian</i>	‘ripe’	/æ/	<i>kotael</i>	‘fish sp.’
/ua/	<i>kuam</i>	‘to fear’	/au/	<i>kaur</i>	‘bamboo’
/uo/	<i>muop</i>	‘navel’	/io/	<i>mamiok</i>	‘papaya’
/ai/	<i>ŋainker</i>	‘swamp’	/oe/	<i>tomoep</i>	‘Tomoip’

Examples of less common diphthongs in open syllables

/iu/	<i>mabiu</i>	‘island lychee’
/ao/	<i>mao</i>	‘banana’
/oa/	<i>hoa</i>	‘vine, rope’

Examples of less common diphthongs in closed syllables

/iu/	<i>piuk</i>	‘tail’
/ao/	<i>kaon</i>	‘his/her chin’ ³¹
/oa/	<i>hoan</i>	‘fish sp.’

Sequences of /ei/ and /ou/ also occur, but it is not possible to assess whether or not they occur underlyingly in open syllables, since underlying /e/ and /o/ in open syllables are typically pronounced as [ei] and [ou], respectively (cf. §3.3. and §3.4). These sequences are known to occur in closed syllables, however, as in *rein* ‘handle of a knife’ and *bour* ‘trap’.

The remaining possible V₁V₂ sequences are: /eu, ea, oi, ue, ui, eo/. Although attested, these are all very uncommon, and some examples of them are perhaps spurious. The sequence /eu/ is only found in closed syllables: *beul* ‘betel pepper’, *meutu* ‘many’, and *Teul* (a man’s name), the first two of which are almost certainly polymorphemic: *ul* alone also means ‘betel pepper’; and *me-* is a prefix used in adjectives. However, /eu/ appears to be pronounced monosyllabically. The sequence /ea/ only occurs in two forms, both of which are alternative pronunciations of forms containing /ia/ – namely, *heam* ‘sky’ (alongside *hiam* ‘sky’) and *hean* ‘yam’ (alongside *hian* ‘yam’). Thus, [ea] possibly only exists as an allophone of /ia/. Next, /oi/ is only attested in our data in *βoi* ‘mango’, whose underlying form is uncertain, and in the greeting *moike* ‘hello’, which we suspect is composed of *mo* ‘and’ and *ike* ‘you [SG]’.³² The sequence /ue/ is only found in *tonkue* ‘lime spatula’ and in a man’s name, *Lelehuen*. Finally, the sequence /ui/ is only found in *simui* ‘dust’ and in the discourse word *kemui* ‘please’ (which we suspect could be polymorphemic as well).

The sequence /eo/ is only attested in our data across a morpheme boundary: the (highly unusual) [o]-initial word *olsak* ‘hardwood tree sp.’ has the (regular) plural form *e=olsak* ‘hardwood tree sp. [PL]’. It is pronounced as [e'olsak].³³

³¹ The only attestation of /ao/ in a closed syllable is across a morpheme boundary: /kao-n/ (with the 3SG possessive suffix /-n/). For consideration of an alternative morphological analysis, see Reesink (2005, p. 183).

³² Grace (1955a, pp. 89, 92) additionally records the forms <hoi> ‘wrap up’ and <tatōiluk> ‘top’. He also transcribes the name of the language as <tomoip> (p. 84). Although we maintain the English spelling of the language as *Tomoip*, we note that the pronunciation in our data is [ˈtomoep], with /oe/ as opposed to /oi/.

³³ Grace (1955a, p. 89) additionally records a form <teq> ‘sink, drown’. In our data, however, we have recorded *tio* ‘go into, go under’; cf. Ross (1980, wordlist D, p. 4), in which <tio> ‘(sun) sets’ likely contains the same lexeme.

5 Stress and suprasegmental features

There is no phonemic tone in the language, nor are there other phonemic suprasegmental distinctions.

Stress is realized in isolated words by a slight increase in loudness, vowel duration, and pitch. However, lexical stress does not carry a high functional load in the language. There are no known instances in which stress is contrastive. Generally, words tend to receive penultimate stress, although there are exceptions to this tendency. Thus we tentatively hypothesize that Tomoip exhibits non-phonemic penultimate stress.

Whereas prefixes such as class prefixes, the causative prefix *βa-*, and the prefixes *me-* and *te-* carry stress when affixed to monosyllabic bases, proclitics such as number markers and subject markers do not. Similarly to prefixes, suffixes such as transitive markers and direct possession markers count as part of the phonological word for purposes of stress assignment. Thus, a shift in stress can be observed, for example, between *delkua-n* ‘his/her neck’ [ˈdɛl.kuan] and *delkua-jo* ‘my neck’ [dɛlˈkua.jo].

Although we are unable at this point to describe Tomoip stress in great detail, we note that this general pattern of penultimate stress (with exceptions) is not uncommon among the languages of the area. For similar stress patterns in nearby Austronesian languages, see Johnston (1980a, pp. 125, 143) for Mangsing (i.e., Mangseng), Bebeli, Bali (i.e., Uneapa), Bulu, Bola, and Meramera; Johnston (1980b, pp. 256–259) for Nakanai; Van Der Mark (2007, pp. 33–35) for Vinitiri (i.e., Minigir); and Pearson & van den Berg (2008, p. 14) for Lote. Indeed, penultimate stress is a common characteristic of languages of the Oceanic subgroup in general and is sometimes reconstructed for the protolanguage as well (Lynch et al., 2002, p. 35). For similar stress patterns in nearby non-Austronesian language, see Lindrud (1992, p. 2) for Kol; Stanton (2007, pp. 81–85) for Ura; and Stebbins (2011, pp. 33–34) for Mali.³⁴

Furthermore, we note here that we have not found any evidence of contrastive vowel length. Reesink’s (2005) orthography suggests that (at least) the low vowel [a] has a long-vowel counterpart [a:] (written as <aa>), at least for the morpheme <taa>, which is described as one of several adpositions used to link the possessum to the possessor in body-part nouns that employ inalienable possession constructions; this morpheme occurs, for example, with <ngom> ‘eye’, which precedes forms such as <taa-n> ‘at-3SG’, <taa-yo> ‘at-1SG’, and <taa-ke> ‘at-2SG’ (pp. 182–183). In our own recordings, however, we have not observed a long vowel here: for example, we transcribe *ŋom tan* ‘his/her eye’ phonetically as [ˈŋomˈtan].³⁵ Similarly, we transcribe *lem tan* ‘his/her forehead/face’ as [ˈlɛmˈtan], also with a short vowel.

Based on his fieldnotes, Reesink (p.c.) suggests some possible indications of contrastive vowel length among other vowel qualities as well, as in <m-pe> ‘SG-fish’ versus <peemu> ‘axe’ and in <m-mën> ‘SG-bird’ versus <nëër> ‘rain’. However, at various places in our own recordings, we observe both phonetically short [e] and phonetically long [e:] both in *pe* ‘fish’ and in *pemu* ‘axe’.³⁶ Although we only have examples of short [ɛ] occurring in *men* ‘bird’ (the expected allophone of /e/ in a closed syllable), we find both short [ɛ] and long [ɛ:] in *ner* ‘rain’. Ross’s (1980) fieldnotes do not indicate any vowel length. Grace (1955a) does

³⁴ Although Tharp (1997, p. 4) also reports a similar stress pattern for the nearby non-Austronesian language Sulka, this is not borne out by the analysis in Reesink (2005, p. 150).

³⁵ We also note some irregularities in the possessive marking for *ŋom* ‘eye’, which, in our data, has /o/ instead of /a/ in the possessive marking for 2SG: *ŋom tom* ‘your [SG] eye’; cf. Grace (1955a, p. 84): <ŋomtɔm> ‘your eye’, <ŋomtɔŋ> ‘my eye’. This noun seems to have an alternative stem *ŋoma* ‘eye’, as in *ŋoma-jo* ‘my eye’ and *ŋoma-ke* ‘your [SG] eye’; cf. Grace (1955a, p. 84): <ŋomake> ‘your eye’.

³⁶ We also note that *pemu* ‘axe’ is possibly a loan from Kol <ˈpɛmu> ‘axe’ (Lindrud, 1980, p. 180), perhaps ultimately coming from Kuanua. Meyer (1961, p. 311) records <pemu> ‘Beil, Axt’ [‘hatchet, axe’] for the Vuatam (i.e., Watom) dialect of Tuna (i.e., Kuanua).

transcribe different vowel lengths, but his transcriptions – like ours – indicate variability within individual forms: for example, he has both <kiake> and <kia·ke> ‘his leg’ (p. 84);³⁷ both <ma·na> and <mana> ‘he thinks’ (p. 84); both <ti·ni> and <tini> ‘I eat’ (p. 84); both <ko·die> and <kōdie> ‘shoots (bow)’ (p. 85); and both <bugē·> ‘belly’ (p. 84) and <bugē> ‘stomach’ (p. 88).

6 Phonotactics and syllable structure

Various Tomoip syllable shapes are attested: syllables with onsets, syllables with codas, syllables with both onsets and codas, and syllables with neither onsets nor codas. Complex codas are attested, including a variety of CC combinations. However, there are no onsets that consist of more than two consonants. Complex codas, on the other hand, are not attested at all.

The voiced obstruents /b, d, g, β, j/ and the glottal fricative /h/ do not serve as codas. Codas thus are known to consist of any nasal /m, n, ŋ/, liquid /l, r/, or non-glottal voiceless consonant /p, t, k, s/. This is the exact same phonotactic distribution found in Sulka (Reesink, 2005, p. 153).³⁸ The Tomoip/Sulka pattern is rather different from that found in the nearby Austronesian languages Mengen, Meramera, and Nakanai. In Mengen, only one consonant occurs word-finally: the velar nasal /ŋ/ (Reesink, 2005, p. 153). No consonants occur word-finally in Melamela (i.e., Meramera) (Ohtsuba & Ohtsuba, 1992, p. 6) or in Nakanai (Johnston, 1980b, p. 254). In the non-Austronesian language Kol, the only consonants to occur word-finally are the velar nasal /ŋ/ and the lateral /l/ (Reesink, 2005, p. 153). On the other hand, the somewhat more geographically distant Austronesian language Kuanua (or at least some of its dialects) resembles Tomoip in terms of its attested word-final consonants: /m, n, ŋ, l, r, p, t, k/ occur word-finally, whereas /b, d, g, β/ do not (Kuanua lacks /s, j, h/) (cf. Meyer, 1961; Franklin et al., 1974). In Lote, all consonants except /h/ occur word-finally (authors’ fieldnotes), although it should be noted that Lote lacks phonemic voiced obstruents (Pearson & van den Berg, 2008, p. 8).³⁹

Diphthongs are common (cf. §4). In word-initial position immediately preceding another consonant, nasal consonants can also serve as syllables. Although Tomoip syllables can employ codas, diphthongs, and complex onsets, there is no syllable known to exhibit all three (i.e., †CCVVC is not attested). The attested syllable shapes are as follows:

V (or N), VV, CV, CVV, VC, VVC, CVC, CVVC, CCV, CCVV, CCVC

The simplest syllables thus consist of a nucleus (typically a vowel), without any onset or coda.

Syllables without onset or coda (V)

<i>a</i>	‘PERS’	<i>i</i>	‘he, she, it’
<i>u</i>	‘smelly’	<i>e=bale</i>	‘house [PL]’ [e‘ba.lei]

The vowels /a, i, u/ all occur in syllables without onsets. The mid front vowel /e/ is only attested as a syllable nucleus without an onset in words marked with the plural marker *e*= ‘PL’. The

³⁷ These forms should perhaps be glossed instead as ‘your [2SG] leg’.

³⁸ The pattern in the Baining languages also seems fairly similar: Qaquet has word-final /m, n, ŋ, l, r, p, t, k, s/ and does not have word-final /^mb, ⁿd, ^ŋg, β, γ ~ j, r/ (Hellwig, 2019, p. 46); Mali has word-final /m, n, ŋ, l, r, p, t, k, s, β/ (with /β/ only marginally in this position) and does not have word-final /^mb, ⁿd, ^ŋg, w, ɿ, j, ɰ/ (Stebbins, 2011, pp. 18–23).

³⁹ Pearson & van den Berg (2008, p. 11) mention that all consonants in Lote occur in coda position, but we can find no evidence of /h/ serving as a coda, with the possible exception of the exclamation <Éh!> ‘hmm!’ (p. 143).

mid back vowel /o/ generally does not occur as a syllable nucleus without an onset, although there are apparently two exceptions: *olsak* ‘hardwood tree sp.’ and *otβie* ‘strap of a string bag’. We do not know of any reports of nearby languages with similarly limited mid-vowel-initial syllables. However, based on Stebbins & Tayul’s (2012) dictionary of Mali (Baining), it seems that both initial /e/ (p. 86) and initial /o/ (p. 143) occur in only a handful of words.

Any of the three nasals /m, n, ŋ/ can serve as an entire syllable provided it is at the beginning of a word and it immediately precedes another consonant with which it does not form a consonant cluster (cf. §2.3, §6.1).

Word-initial syllabic nasals (N.C-)

<i>msik</i>	‘flying fox’	[mˈsɪk]	<i>mdip</i>	‘coral tree’	[mˈdɪp]
<i>ŋgol</i>	‘string bag’	[ŋˈgɔl]	<i>n=sal</i>	‘path [SG]’	[nˈsal]

Syllables of the shape VV – that is, syllables containing a diphthong as nucleus, but no onset or coda – are rare. The only diphthongs that seem to occur in syllables without onsets are /ai/ and /au/.

Syllables without onset or coda – diphthong as nucleus (VV)

<i>au</i>	‘PREP’	<i>aidi</i>	‘3PL.POSS.ED’ [ˈai.di]
-----------	--------	-------------	------------------------

It is much more common for syllables to contain onsets. There are thus many examples of words containing syllables of the shape CV. The clearest illustrations of this syllable shape are found in monosyllabic words.

Syllables with simple onset but no coda (CV)

<i>ha</i>	‘sun’	<i>pa</i>	‘to hunt’
<i>le</i>	‘nit’	<i>ki</i>	‘bunch’
<i>mo</i>	‘and’	<i>bu</i>	‘to rain’

Likewise, there are a number of monosyllabic words consisting of an onset and a diphthong.

Syllables with simple onset but no coda – diphthong as nucleus (CVV)

<i>hoa</i>	‘vine, rope’	<i>tia</i>	‘clam’
<i>kae</i>	‘to choose’	<i>kie</i>	‘dry’
<i>mao</i>	‘banana’	<i>lio</i>	‘side of the body’

Codas also occur. Since there are relatively few syllables that lack onsets, there are not many syllables of the form VC (i.e., CVC is much more common). However, some examples are to be found.

Syllables with coda but no onset (VC)

<i>in</i>	‘to eat’	<i>ul</i>	‘betel pepper’
-----------	----------	-----------	----------------

Syllables with coda but no onset – diphthong as nucleus (VVC)

aun ‘here’

There are a number of monosyllabic words consisting of a single onset and a single coda (i.e., CVC).

Syllables with simple onset and coda (CVC)

<i>sal</i>	‘path’	<i>pan</i>	‘ray’
<i>lep</i>	‘jewelry’	<i>diŋ</i>	‘thorn’
<i>nok</i>	‘coconut frond midrib’	<i>ŋuk</i>	‘lie, falsehood’

Syllables with simple onset and coda – diphthong as nucleus (CVVC)

<i>biam</i>	‘plain’	<i>buol</i>	‘root’
<i>huan</i>	‘cucumber’	<i>kaur</i>	‘bamboo’
<i>liep</i>	‘lake’	<i>piuk</i>	‘tail’

Tomoip exhibits a variety of consonant clusters (cf. §6.1). Therefore, it is not uncommon for there to be syllables of the shape CCV, CCVV, or CCVC. However, there are no clear examples of syllables of the shape CCVVC.

Syllables with complex onset but no coda (CCV)

<i>bla</i>	‘head’	<i>pla</i>	‘skin’
<i>ple</i>	‘spirit type’	<i>tli</i>	‘sea almond tree’
<i>pso</i>	‘rack above hearth’	<i>pto</i>	‘buttress root sp.’

Syllables with complex onset but no coda – diphthong as nucleus (CCVV)

<i>glie</i>	‘shield’	<i>krie</i>	‘buttress root sp.’
<i>βria</i>	‘tilapia’	<i>mliu</i>	‘yam sp.’

Syllables with complex onset and (simple) coda (CCVC)

<i>ptak</i>	‘mud’	<i>glep</i>	‘arecoid palm’
<i>blik</i>	‘bad’	<i>kses</i>	‘flat’
<i>tluŋ</i>	‘bay’	<i>kner</i>	‘stinger of an insect’

6.1 Consonant clusters

As seen in some of the examples already given, Tomoip exhibits a variety of syllable-internal consonant clusters. While some CC sequences only occur across morpheme boundaries, there are a number of CC clusters that occur within a single morpheme. In all instances these occur as onsets (never as codas).

All six oral stops are known to occur as the first element of a consonant cluster. Of these, the bilabial plosives /p, b/ are the most common in this position. The coronal plosives /t, d/ are much less common. Many words begin with a nasal followed by another consonant, but, in most of these, the nasal consonant is syllabic and thus these do not exemplify complex onsets

(cf. §2.3). However, the bilabial nasal /m/ occurs as the first consonant in complex onsets when the second element is a sonorant coronal (i.e., /ml-/ , /mn-/ , or /mr-/). In such cases, an epenthetic schwa is optionally inserted. The /m/ in such clusters, however, is never pronounced as a syllabic consonant. Similarly, the bilabial fricative /β/ occurs as the first consonant in a cluster when the second element is a sonorant coronal (i.e., /βl-/ or /βr-/; †/βn-/ is not attested).

As the second member of consonant clusters, the following consonants occur: the liquids /r, l/, the voiceless coronal fricative /s/, the voiceless alveolar stop /t/, and the alveolar nasal /n/. Or, categorizing these segments strictly by place of articulation, we can say that as second elements in consonant clusters, all coronals except /d/ occur – that is, /r, l, s, t, n/.

Given these eight initial consonants (/p, t, k, b, d, g, m, β/) and five following consonants (/r, l, s, t, n/), there would be 39 possible CC combinations (excluding the theoretically possibly geminate /tt-/ , which would be formed by the only consonant known to occur in either position). However, only about half of these are attested.

The attested consonant clusters are presented in Table 4, along with the number of words known to exhibit them word-initially (from a list of 881 formally distinct unbound lexemes). The sequences marked with a single asterisk (*) are not attested in word-initial position but they occur as clusters within individual syllables. The sequences marked with a double asterisk (**) are attested in word-initial position but not within an individual syllable (i.e., the first C is syllabic).

Table 4. Tomoip consonant clusters: number of word-initial occurrences

	r	l	s	t	n
p	pr 2	pl 5	ps 2	pt 2	—
t	—	tl 2	—	—	—
k	kr 1	kl 1	ks 1	—	kn 3
b	br 1	bl 4	—	—	—
d	dr 1	—	—	—	—
g	gr *	gl 2	—	—	—
m	mr 1	ml 3	ms **	mt **	mn 2
β	βr 2	βl *	—	—	—

As onsets, these combinations of consonants generally align with the Sonority Sequencing Principle, although sonority plateaus like /ml/ and /pt/ are also possible.

Examples of consonant clusters

/pr/	<i>praŋpa</i>	‘adze’	/pl/	<i>pla</i>	‘nail’
/ps/	<i>psi</i>	‘to flatulate’	/pt/	<i>ptak</i>	‘mud’
/tl/	<i>tluŋ</i>	‘bay’	/kr/	<i>krie</i>	‘buttress root sp.’
/kl/	<i>kloy</i>	‘spirit type’	/ks/	<i>kses</i>	‘flat’
/kn/	<i>kner</i>	‘stinger of an insect’	/br/	<i>breβoe</i>	‘earwax’
/bl/	<i>bla</i>	‘head’	/dr/	<i>dran</i>	‘swamp’
/gr/	<i>βaŋgrem</i>	‘lime container’	/gl/	<i>glep</i>	‘arecoid palm’
/mr/	<i>mrek</i>	‘to be thirsty’	/ml/	<i>mliu</i>	‘yam sp.’
/mn/	<i>mnor</i>	‘dog flea’	/βr/	<i>βria</i>	‘tilapia’
/βl/	<i>siβlo</i>	‘outrigger’			

Some of these forms deserve comment. First, *kseš* ‘flat’ has an alternative pronunciation that lacks the cluster: [‘kasəs]. However, considering the existence of the cluster /ps/ – another heterorganic voiceless stop and alveolar fricative combination – the cluster /ks/ should not raise too much suspicion. Second, the noun *dran* ‘swamp’ is pronounced either as [dran] or as [ʰdran]. Although prenasalization can occur with any voiced stop, it seems to be especially prominent in this word, occurring also in the plural form *e=dran* ‘swamp [PL]’ [eʰndran] (perhaps even [ɛnʰdran]).⁴⁰ Although we have no examples of word-initial /gr/, the form *βangrem* ‘lime container’ must contain this cluster, since otherwise it would have an unattested complex coda: the word syllabifies as [‘βaŋ.grem].⁴¹ Similarly, although we have no examples of word-initial /βl/, the form *siβlo* ‘outrigger’ must contain such a cluster, since voiced obstruents do not occur as codas: the word syllabifies as [‘ei.βlou].⁴²

Many of the permissible clusters presented above are nevertheless only minimally represented in our wordlist. It could be the case that more clusters than these also occur in the language, only that we have no examples of them.

Furthermore, there are apparently a number of lexical roots that begin with sequences of consonants that are never pronounced without an immediately preceding vowel. In practice, this means that these forms never occur without prefixes or proclitics.

CC-initial roots that require prefixes or proclitics

<i>lko</i>	‘to fall’	(e.g., <i>ta=lko</i> , <i>se=lko</i>)
<i>rpek</i>	‘to cry’	(e.g., <i>ti=rpek</i> , <i>ne=rpek</i>)
<i>rβio</i>	‘to fight’	(e.g., <i>ti=rβio</i> , <i>ta=rβio</i>)
<i>rle</i>	‘smooth’	(e.g., <i>me-rle</i> , <i>te-rle</i>)
<i>lme</i>	‘coconut’	(e.g., <i>ne=lme</i> , <i>e=lme</i> , <i>bu-lme</i> , <i>ka-lme</i> , <i>la-lme</i> , <i>ra-lme</i>)

Tomoip proclitics found in verb phrases include portmanteau forms that indicate modality (realis or irrealis) while also indexing the subject. They include *ta=* ‘DU/PL/1SG.REAL’, *ti=* ~ *t=* ‘2SG/3SG.REAL’, *sa=* ‘DU/PL/1SG.IRR’, *se=* ‘2SG.IRR’, and *so=* ‘3SG.IRR’ (cf. Reesink, 2005, p. 169 for a slightly different analysis).⁴³ Proclitics in noun phrases include the number-marking proclitics *ne=* ‘SG’, *ro=* ‘DU’, and *e=* ‘PL’ (cf. Ross, 1988, pp. 292–293; Reesink, 2005, p. 171). This singular form *ne=* ‘SG’ (or a homophone of it) also seems to function as a nominalizer, for example, in forms such as *ne=rpek* ‘worry, problem, concern’ (from *rpek* ‘to cry’). Some nouns also take class prefixes, which can specify the semantics of the referent; these include *buŋ-* ~ *bu-* (for fruits, round things, bunches), *ka-* (for trees, long things, plants), *la-* (for baskets, open things, pieces), and *ra-* (for leaves). These have been identified as “quantity markers” by Ross (1988, pp. 293–298) and Reesink (2005, p. 171; cf. p. 154); our data, however, have not revealed any quantity-marking function of these morphemes. The prefixes *me-* and *te-* occur with adjectives. They are possibly in origin stative markers, derived

⁴⁰ The fact that the Tok Pisin word *dram* ‘drum’ has been borrowed into Tomoip as *daram* ‘drum’ may suggest that /dr/ clusters are avoided in the language. However, it could be that the loaning dialect of Tok Pisin also lacked this cluster.

⁴¹ In addition to *βangrem* ‘lime container’, the only examples we have of sequences of three consecutive consonants are *βunglie* [‘wɔŋ.glie] ‘vine sp.’ and *ŋ=glie* [ŋ.glie] ‘shield [SG]’, neither of which contains a true CCC onset. In all instances the sequence is velar nasal followed by voiced velar stop followed by liquid. It could be the case that the liquids in the clusters /gl/ and /gt/ condition an increase in prenasalization of the preceding velar stop.

⁴² However, the lenition rule postulated in §2.1 suggests that this word could possibly have the underlying form /siplo/. We note, at least, that Simon spelled this word as <sivlo>.

⁴³ Ross (1988, p. 292) also alludes to the realis forms: “Tomoip marks non-future, non-habitual verb phrases with *t-*.” In our data, the allomorph *t=* occurs when the 2SG/3SG.REAL marker is immediately followed by a vowel.

from POC *ma- and *ta- (see Evans, 2003); however, their function in Tomoip remains uncertain.⁴⁴

The prevalence of consonant clusters is perhaps the aspect of Tomoip phonology that is the most unusual for Austronesian languages of the area. As far as we know, no other New Britain languages assigned to Ross's (1988) Meso-Melanesian Cluster exhibit complex onsets.⁴⁵ However, similar consonant clusters are common in some of the non-Austronesian languages of New Britain, including Anêm (i.e., Anem) (Thurston, 1976, pp. 21–22) and Sulka (Reesink, 2005, pp. 147, 153). In particular, the Baining languages seems to exhibit similar clusters to those found in Tomoip, for example, Mali (Stebbins, 2011, pp. 27–28) and Ura (Stanton, 2007, pp. 39–41), in which coronal consonants in particular are found as second members of complex onsets.⁴⁶ However, complex onsets are also found in New Britain in some of the languages of the North New Guinea subgroup of Austronesian languages, for example, Bariai (Gallagher & Baehr, 2005, p. 19), Amara (Thurston, 1996, p. 205), and Mangseng (Milligan, 1992, p. 3). Chowning (1976, p. 378) refers to three somewhat non-Austronesian-looking features, “initial and medial consonant clusters, reduction to monosyllables, and vowel shifts”, noting that the combination is “characteristic of south-west New Britain” and that “Tumuip [i.e., Tomoip] shares these features, although the vowel shifts are different”.

6.2 Geminate consonants

The presence of the singular marker *ne*= ‘SG’ – or, more specifically, its allomorph *m*= – can lead to geminate nasals (cf. §2.3). Nominal roots that begin with /m/ can thus be pronounced with an initial /mm/ sequence (i.e., as [m:]).

Word-initial NN sequences with the first nasal derived from the singular marker *ne*=

<i>m</i> = <i>me</i>	‘rat [SG]’	(cf. <i>e</i> = <i>me</i> ‘rat [PL]’)
<i>m</i> = <i>men</i>	‘bird [SG]’	(cf. <i>e</i> = <i>men</i> ‘bird [PL]’)
<i>m</i> = <i>mana</i>	‘knowledge [SG]’	(cf. <i>e</i> = <i>mana</i> ‘knowledge [PL]’)

It is worth noting that, although the allomorphs *m*= and *ŋ*= of the singular marker *ne*= ‘SG’ probably arose due to regressive place assimilation, nasal assimilation is not a productive phonological process in the language. Indeed, there are lexemes that exhibit the crosslinguistically uncommon sequences of [nk] and [ng], as well as [mk] (†[mg] is not attested).

⁴⁴ It is possible that these prefixes derive adjectives from other parts of speech, but no clear patterns have been found. For example, both *pur* ‘big’ and *me-pur* ‘big’ are attested as adnominal modifiers (following nouns). When adjectives occur as predicates, the prefixes *me*- and *te*- are sometimes present, sometimes absent. In an example sentence with *me-kakae* ‘small’, Ross (1988, p. 292) glosses the morpheme *me*- simply as “L” (i.e., “ligature”).

⁴⁵ For example, syllables with consonant clusters are not found in Ramoaaina (Davies, 2001, p. 1), Tolai (i.e., Kuanua) (Mosel, 1980, p. 18), Vinitiri (i.e., Minigir) (Van Der Mark, 2007, pp. 27–28, 31), Melamela (i.e., Meramera) (Ohtsuba & Ohtsuba, 1992, p. 7), Nakanai (Johnston, 1980b, p. 254), Bola (van den Berg & Wiebe, 2019, p. 29), Bali (i.e., Uneapa) (Ross, 2002, p. 363), or Vitu (i.e., Muduapa) (van den Berg & Bachet, 2006, p. 12). Nor are they found in the nearby Mengenic subgroup of the North New Guinea linkage (for Lote, cf. Pearson & van den Berg, 2008, p. 11; for Mengen, cf. Rath, 1993, p. 97).

⁴⁶ However, not all non-Austronesian languages of New Britain exhibit consonant clusters. They do not seem to be found in Kol (Reesink, 2005, p. 153) or in Ata (i.e., Pele-Ata) (Hashimoto et al., 2008, p. 3). Although Meng (2018, pp. 25–26) reports CCV and CCVC syllable types for Tulil (i.e., Taulil), these appear to be restricted to syllables with semivowels as the second segment (at least based on the data presented).

Words exhibiting *lack* of nasal assimilation (i.e., $no \uparrow N \rightarrow [n] / _ [+velar]$)

<i>kunkun</i>	‘sugarcane’	<i>an-ke</i>	‘2SG.POSS.GEN’
<i>usingi</i>	‘octopus’	<i>tumka</i>	‘work’
<i>me-mke</i>	‘sweet’ ⁴⁷	<i>alum-ke</i>	‘your [SG] child’

We note, however, that Reesink (2005, p. 182) reports that the final nasal of the possessive classifier assimilates in place to the following consonant (e.g., in /an-ke/ ‘2SG.POSS.GEN’, as given above). However, this is not reflected in our data. We have, for example, recorded [‘anke] ‘2SG.POSS.GEN’, as opposed to †[‘aŋke].

Likewise, bilabial /m/ or velar /ŋ/ can occur immediately before alveolar consonants without assimilating in place to [n].⁴⁸

Words exhibiting *lack* of nasal assimilation (i.e., $no \uparrow N \rightarrow [n] / _ [+coronal]$)

<i>hemto</i>	‘cedar tree’	<i>jamleŋ</i>	‘to dream’
<i>mdip</i>	‘coral tree’	<i>hodiŋtoŋ</i>	‘charcoal’
<i>toŋlo</i>	‘channel’	<i>loŋnet</i>	‘sea worm sp.’

Finally, alveolar /n/ or velar /ŋ/ can occur immediately before bilabial consonants without assimilating in place to [m].

Words exhibiting *lack* of nasal assimilation (i.e., $no \uparrow N \rightarrow [m] / _ [+labial]$)

<i>manmie</i>	‘worm sp.’	<i>manβar</i>	‘bird sp.’
<i>praŋpa</i>	‘adze’	<i>buŋ-mao</i>	‘banana fruit’
<i>buŋ-biria</i>	‘fruit of the breadfruit tree’		

That said, there are no attested sequences of †/nb/ or of †/np/. Notably, the absence of nasal assimilation is also a feature of (at least some of) the Baining languages, including Ura (Stanton, 2007, p. 37) and Qaqet (Hellwig, 2019, pp. 34, 48).

Much like the singular marker *ne=* ‘SG’, the prefix *me-* can be a source of NC sequences in Tomoip, namely when the vowel of the affix is elided. This prefix occurs in many adjectives.

Word-initial NC sequences in adjectives

<i>m-rarek</i>	‘light (not heavy)’	
<i>m-buk</i>	‘deaf’	
<i>m-ma</i>	‘dry’	
<i>m-mia</i>	‘red’	(also <i>me-mia</i>)
<i>m-mian</i>	‘ripe’	(also <i>me-mian</i>)

Thus, whether resulting from the allomorph *m=* of the singular marker *ne=* ‘SG’ or an elided version of the prefix *me-*, geminate bilabial nasals can occur in Tomoip. In all instances these geminates occur across morpheme boundaries. However, there are at least three forms that exhibit geminate nasals ([m:] or [n:]) in word-medial position. It is not clear whether or not these straddle a morpheme boundary (whether synchronic or diachronic). All three terms refer

⁴⁷ This form probably never occurs without a prefix (cf. §6.1).

⁴⁸ Almost all nasal-plus-alveolar combinations are attested: the sole unattested sequence is †/ŋd/.

to parents: *amma* ‘father’ and *anno* ‘mother’ are both vocative forms and do not take possessive marking, whereas *mimma* ‘father’ refers to the inalienable kinship relation and receives suffixes that index the possessor. The inalienable kinship relation *tena* ‘mother’ (which does not contain any geminate nasals) similarly receives possessor suffixes.

Geminate nasals in terms for ‘mother’ and ‘father’

<i>anno</i>	‘mother (vocative)’	(also <i>no</i>)
<i>amma</i>	‘father (vocative)’	(also <i>ma</i>)
<i>mimma</i>	‘father’	

Finally, postnominal demonstratives often exhibit geminate nasals. It is not clear, however, whether these geminates belong to the underlying forms or they are only surface realizations. Notably, the velar nasal does not exhibit such gemination, despite occurring in some demonstrative forms (i.e., †[ŋ:] is not attested).

Gemination in postnominal demonstratives

[ani ~ an:i ~ =ni ~ =n:i]	‘PROX’
[ana ~ an:a ~ =na ~ =n:a]	‘MED’
[aŋa ~ =ŋa]	‘DIST’

The numeral *denan* ‘one’, which likely contains a demonstrative element, is similarly often pronounced with a geminate, as [den:an].

7 Conclusion

We introduced a basic overview of the phonology of Tomoip, building on earlier work by Ross (1988) and Reesink (2005). Although undoubtedly an Oceanic language, Tomoip exhibits a number of unexpected sound correspondences and other lexical differences that have so far made its precise classification challenging. This present paper thus hopefully serves as a step towards a better understanding of the language in a wider historical context. Since several phonetic features of Tomoip seem to reflect patterns also found in nearby non-Austronesian languages, this historical context likely includes contact and borrowing as well as inheritance.

That said, we acknowledge the limitations of our analysis, since our corpus of data is rather small and mostly limited to elicited words and sentences. Furthermore, as our primary data and recordings come from only one speaker, we cannot determine the degree to which forms or pronunciations reflect individual idiosyncrasies as opposed to language-wide patterns. However, comparisons with unpublished fieldnotes and earlier recordings from other researchers have largely corroborated our findings.

A number of questions about Tomoip remain unanswered and deserve further attention. For example, it is still unclear whether or not there is a phonemic palatal consonant in the language. Stress is also a topic in need of further investigation, ideally to be conducted with more naturalistic recordings of longer texts. It would also be worthwhile to undertake a more detailed analysis of vowels – both monophthongs and diphthongs – since the exact acoustic values of these segments is not always clear. We hope that this brief introduction to Tomoip will help other researchers in analyzing the language in greater detail, not only its phonetics and phonology, but also its morphology and syntax, about which we have only been able to provide occasional hints.

Abbreviations and symbols

1	1 st person	INCL	inclusive	PREP	preposition
2	2 nd person	IRR	irrealis	PROX	proximal
3	3 rd person	MED	medial	REAL	realis
AN	Austronesian	MOD	modal particle	REFL	reflexive
DIST	distal	NAN	non-Austronesian	SG	singular
DU	dual	PERS	personal marker	sp.	species
ED	edible classifier	PL	plural	*	reconstructed
EXCL	exclusive	POc	Proto-Oceanic	†	unattested
GEN	general classifier	POSS	possessive		

References

- Barlow, Russell & Killian, Don. 2023. Tomoip wordlist: Based on elicitation sessions with Simon Mangil from Milim village, recorded in Kokopo, East New Britain, October to November 2022. <https://doi.org/10.5281/zenodo.8437515>
- Burger, Friedrich. 1913. *Die Küsten- und Bergvölker der Gazellehalbinsel: Ein Beitrag zur Völkerkunde von Neuguinea unter besonderer Hervorhebung rechtlicher und sozialer Einrichtungen* (Studien und Forschungen zur Menschen- und Völkerkunde 12). Stuttgart: Strecker & Schröder.
- Capell, Arthur. 1962a. *A linguistic survey of the south-western Pacific: New and revised edition* (South Pacific Commission Technical Paper No. 136). Nouméa: South Pacific Commission.
- Capell, Arthur. 1962b. Oceanic linguistics today. *Current Anthropology* 3(4): 371–428. <https://doi.org/10.1086/200305>
- Capell, Arthur. 1971. The Austronesian languages of Australian New Guinea. In Thomas A. Sebeok (ed.), *Current trends in linguistics, volume 8: Linguistics in Oceania, part one: Indigenous languages*, 240–340. The Hague: Mouton. <https://doi.org/10.1515/9783111418827-010>
- Chowning, Ann. 1969. The Austronesian languages of New Britain. In Arthur Capell, Ann Chowning & Stephen A. Wurm (eds.), *Papers in linguistics of Melanesia, no. 2*, 17–46. Canberra: Research School of Pacific Studies, the Australian National University. <http://hdl.handle.net/1885/145143>
- Chowning, Ann. 1976. Austronesian languages: New Britain. In Stephen A. Wurm (ed.), *New Guinea area languages and language study, volume 2: Austronesian languages* (Pacific Linguistics: Series C–39), 365–386. Canberra: Research School of Pacific Studies, the Australian National University. <http://hdl.handle.net/1885/145151>
- Chowning, Ann. 1996. Relations among languages of West New Britain: An assessment of recent theories and evidence. In Malcolm D. Ross (ed.), *Studies in languages of New Britain and New Ireland, I: Austronesian languages of the North New Guinea Cluster in Northwestern New Britain* (Pacific Linguistics: Series C–135), 7–62. Canberra: Research School of Pacific and Asian Studies, the Australian National University. <http://hdl.handle.net/1885/145888>
- Crowley, Terry. 1985. Common noun phrase marking in Proto-Oceanic. *Oceanic Linguistics* 24(1/2): 135–193. <https://doi.org/10.2307/3623065>
- Davies, Robyn. 2001. *Participant reference in Ramoaaina narrative*. Ukarumpa: Summer Institute of Linguistics. Unpublished manuscript. <https://www.sil.org/resources/archives/43869>
- Davies, Robyn & Fritzell, Lisbeth. 1992. *Duke of York grammar essentials (Ramoaaina)*. Ukarumpa: Summer Institute of Linguistics. Unpublished manuscript. <https://www.sil.org/resources/archives/43873>

- Dyen, Isadore. 1965. *A lexicostatistic classification of the Austronesian languages* (Indiana University Publications in Anthropology and Linguistics, Memoir 19 of the International Journal of American Linguistics). Baltimore: Waverly Press.
- Evans, Bethwyn. 2003. *A study of valency-changing devices in Proto Oceanic* (Pacific Linguistics 539). Canberra: Research School of Pacific and Asian Studies, the Australian National University. <http://hdl.handle.net/1885/72242>
- Foley, William A. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- Franklin, Karl J. & Kerr, Harland B. & Beaumont, Clive H. 1974. *Tolai language course* (Language Data, Asian-Pacific Series 7). Ukarumpa: Summer Institute of Linguistics. <https://www.sil.org/resources/archives/3907>
- Gallagher, Steve & Baehr, Peirce. 2005. *Bariai grammar sketch* (Data Papers on Papua New Guinea Languages 49). Ukarumpa: Summer Institute of Linguistics. <https://www.sil.org/resources/archives/23630>
- Grace, George W. 1955a. The George Grace linguistic research and slides of Melanesia collection, 1955–1956: Online documents: 34. Columbia University Notebook, pt. 3. Unpublished manuscript. <https://digital.library.manoa.hawaii.edu/grace/docs>
- Grace, George W. 1955b. The George Grace linguistic research and slides of Melanesia collection, 1955–1956: Online documents: 37. Swains Exercise Book, number 1, pt. 3. Unpublished manuscript. <https://digital.library.manoa.hawaii.edu/grace/docs>
- Grace, George W. 1955c. The George Grace linguistic research and slides of Melanesia collection, 1955–1956: Online documents: 41. Yellow 5 inch x 8 inch spiral notebook. Unpublished manuscript. <https://digital.library.manoa.hawaii.edu/grace/docs>
- Grace, George W. 1955d. Subgrouping of Malayo-Polynesian: A report of tentative findings. *American Anthropologist* 57(2): 337–339. <https://doi.org/10.1525/aa.1955.57.2.02a00110>
- Hall, Robert A. 1943. *Melanesian Pidgin English: Grammar, texts, vocabulary* (Special Publications of the Linguistic Society of America). Baltimore: Linguistic Society of America at the Waverly Press.
- Hashimoto, Chiyoko & Hashimoto, Kazuo & Kume, Nozomi. 2008. Organised phonology data: Ata (Pele Ata, Wasi) language [ATA], Kimbe, West New Britain Province. Unpublished manuscript. <https://www.sil.org/resources/archives/42427>
- Hellwig, Birgit. 2019. *A grammar of Qaqet* (Mouton Grammar Library 79). Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110765793>
- Johnston, Raymond. 1980a. The languages and communities of the Kimbe Bay region. In Raymond L. Johnston (ed.), *Language, communication and development in New Britain*, 107–158. Ukarumpa: Summer Institute of Linguistics.

- Johnston, Raymond Leslie. 1980b. *Nakanai of New Britain: The grammar of an Oceanic language* (Pacific Linguistics: Series B–70). Canberra: Research School of Pacific Studies, the Australian National University. <http://hdl.handle.net/1885/148444>
- Lanyon-Orgill, Peter A. 1960. *A dictionary of the Raluana language (New Britain, S.W. Pacific): Raluana-English, with an introduction and English-Raluana index*. Victoria: Self-published.
- Lanyon-Orgill, Peter A. 1962. Comments in Capell, Arthur (1962b), 414–415. Oceanic linguistics today. *Current Anthropology* 3(4): 371–428. <https://doi.org/10.1086/200305>
- Laycock, Don. 1985. Phonology: Substratum elements in Tok Pisin phonology. In Stephen A. Wurm & Peter Mühlhäusler (eds.), *Handbook of Tok Pisin (New Guinea Pidgin)* (Pacific Linguistics: Series C–70), 295–307. Canberra: Research School of Pacific Studies, the Australian National University. <http://hdl.handle.net/1885/145234>
- Lindrud, Stellan. 1980. The languages and communities of the Open Bay, Wide Bay and Southern Gazelle Peninsula region. In Raymond L. Johnston (ed.), *Language, communication and development in New Britain*, 159–183. Ukarumpa: Summer Institute of Linguistics.
- Lindrud, Stellan. 1992. Organised phonology data: Kol language [KOL], East New Britain Province. Unpublished manuscript. <https://www.sil.org/resources/archives/42266>
- Loukotka, Čestmír. 1957. Classification des langues papoues. *Lingua posnaniensis* 6: 19–83.
- Lynch, John & Ross, Malcolm & Crowley, Terry. 2002. *The Oceanic languages*. Richmond: Curzon Press.
- Madden, Fred. 2006. Mengen dictionary. Unpublished manuscript. <https://www.sil.org/resources/archives/31158>
- Meier, Joseph. 1914. Dr. Friedrich Burger, *Die Küsten- und Bergvölker der Gazellehalbinsel* (review article). *Anthropos* 9(1/2): 348–352. <https://www.jstor.org/stable/40444289>
- Meng, Chenxi. 2018. *A grammar of Tulil*. Bundoora: La Trobe University PhD thesis. <https://doi.org/10.26181/21857724.v1>
- Meyer, Otto. 1961. *Wörterbuch der Tuna-Sprache* (Micro-Bibliotheca Anthropos 34). Posieux: Anthropos-Institut.
- Mihalic, Francis. 1971. *The Jacaranda dictionary and grammar of Melanesian Pidgin*. Milton: The Jacaranda Press.
- Milligan, Lloyd. 1992. Organised phonology data: Mangseng [Umua dialect] (Mangsing) language [MBH], Kimbe, East and West New Britain Province. Unpublished manuscript. <https://www.sil.org/resources/archives/42261>

- Mosel, Ulrike. 1980. *Tolai and Tok Pisin: The influence of the substratum on the development of New Guinea Pidgin* (Pacific Linguistics: Series B–73). Canberra: Research School of Pacific Studies, the Australian National University. <http://hdl.handle.net/1885/144532>
- Müller, Hermann. 1907. Grammatik der Mengen-Sprache. *Anthropos* 2(1): 80–99. <https://www.jstor.org/stable/40389626>
- Müller, Hermann. 1915–1916. Erster Versuch einer Grammatik der Sulka-Sprache, Neu-Pommern (Südsee). *Anthropos* 10/11(1/2): 75–97. <https://www.jstor.org/stable/40442797>
- Ohtsuba, Hideki & Ohtsuba, Masae. 1992. Organised phonology data for the Melamela language (WNBP). Unpublished manuscript. [An adapted version is available at <https://www.sil.org/resources/archives/42309>.]
- Osmond, Meredith. 2007. Navigation and the heavens. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, volume 2: The physical environment* (Pacific Linguistics 545), 155–191. Canberra: Research School of Pacific and Asian Studies, the Australian National University. <http://hdl.handle.net/1885/106908>
- Osmond, Meredith. 2011. Fish. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, volume 4: Animals* (Pacific Linguistics 621), 25–136. Canberra: School of Culture, History and Language, the Australian National University. <http://hdl.handle.net/1885/106908>
- Osmond, Meredith & Ross, Malcolm. 1998. Household artefacts. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, volume 1: Material culture* (Pacific Linguistics: Series C–152), 67–114. Canberra: Research School of Pacific and Asian Studies, the Australian National University. <http://hdl.handle.net/1885/146170>
- Osmond, Meredith & Ross, Malcolm. 2016a. The human body. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, volume 5: People: Body and mind* (Asia-Pacific Linguistics 28), 75–207. Canberra: College of Asia and the Pacific, the Australian National University. <http://hdl.handle.net/1885/106908>
- Osmond, Meredith & Ross, Malcolm. 2016b. Bodily conditions and activities. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, volume 5: People: Body and mind* (Asia-Pacific Linguistics 28), 209–333. Canberra: College of Asia and the Pacific, the Australian National University. <http://hdl.handle.net/1885/106908>
- Parkinson, Richard. 1907. *Dreißig Jahre in der Südsee: Land und Leute, Sitten und Gebräuche im Bismarckarchipel und auf den deutschen Salomoinseeln*. Stuttgart: Strecker & Schröder.

- Pawley, Andrew. 1982. Rubbish-man commoner, big-man chief? Linguistic evidence for hereditary chieftainship in Proto-Oceanic society. In Jukka Siikala (ed.), *Oceanic studies: Essays in honour of Aarne A. Koskinen* (Transactions of the Finnish Anthropological Society 11), 33–52. Helsinki: The Finnish Anthropological Society.
- Pawley, Andrew & Green, Roger C. 1984. The Proto-Oceanic language community. *The Journal of Pacific History* 19(3): 123–146. <https://doi.org/10.1080/00223348408572489>
- Pearson, Greg & van den Berg, René. 2008. *Lote grammar sketch* (Data Papers on Papua New Guinea Languages 54). Ukarumpa: SIL-PNG Academic Publications. <https://www.sil.org/resources/archives/23806>
- Rascher, Matthäus. 1909. *Aus der deutschen Sudsee: Mitteilungen der Missionare vom heiligsten Herzen Jesu, Band 1: Land und Leute*. Münster: Aschendorf.
- Rath, Daniel. 1980. The languages and communities of the Mengen region. In Raymond L. Johnston (ed.), *Language, communication and development in New Britain*, 197–223. Ukarumpa: Summer Institute of Linguistics.
- Rath, Daniel D. 1993. Mengen phonology essentials. In John M. Clifton (ed.), *Phonologies of Austronesian languages, no. 2* (Data Papers on Papua New Guinea Languages 40), 71–98. Ukarumpa: Summer Institute of Linguistics. <https://www.sil.org/resources/archives/23739>
- Reesink, Ger. 2005. Sulka of East New Britain: A mixture of Oceanic and Papuan traits. *Oceanic Linguistics* 44(1): 145–193. <https://doi.org/10.1353/ol.2005.0026>
- Ross, Malcolm. 1980. *Konomala Mamusi Uvol Tomoip*. X-WAV/MPEG/PDF/TIFF/JPEG. MR1-012 at catalog.paradisec.org.au. <https://dx.doi.org/10.4225/72/56FFE00776765> [The Tomoip recording begins around 07:13 in the file *MR1-012-B.wav*; the Tomoip wordlist is transcribed in the righthand column of the file *MR1-012-A2_Mamusi.pdf*.]
- Ross, Malcolm. 1988. Proto-Oceanic and the Austronesian languages of western Melanesia (Pacific Linguistics: Series C–98). Canberra: Research School of Pacific Studies, the Australian National University. <http://hdl.handle.net/1885/145428>
- Ross, Malcolm. 2002. Bali-Vitu. In John Lynch, Malcolm Ross & Terry Crowley, *The Oceanic languages*, 362–386. Richmond: Curzon Press.
- Ross, Malcolm. 2007. Meteorological phenomena. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, volume 2: The physical environment* (Pacific Linguistics 545), 119–153. Canberra: Research School of Pacific and Asian Studies, the Australian National University. <http://hdl.handle.net/1885/106908>

- Ross, Malcolm. 2008a. Staple foods: Root crops, bananas, breadfruit and sago. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, volume 3: Plants* (Pacific Linguistics 599), 255–292. Canberra: Research School of Pacific Studies, the Australian National University. <http://hdl.handle.net/1885/106908>
- Ross, Malcolm. 2008b. Other cultivated plants. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, volume 3: Plants* (Pacific Linguistics 599), 389–426. Canberra Research School of Pacific Studies, the Australian National University. <http://hdl.handle.net/1885/106908>
- Salzner, Richard. 1960. *Sprachenatlas des Indopazifischen Raumes*. Wiesbaden: Otto Harrassowitz.
- Scheller, Meinrad. 1966. *Scheller4*. MPEG/X-WAV. MS2-004 at catalog.paradisec.org.au. <https://dx.doi.org/10.4225/72/56F9484682969> [The Tomoip recording begins around 02:20 and ends around 20:30 in file MS2-004-B.wav.]
- Schneider, Joseph. 1962. *Grammatik der Sulka-Sprache (Neubritannien)* (Micro-Biblioteca Anthropos 36). Posieux: Anthropos Institut.
- Stanton, Lee. 2007. *Topics in Ura phonology and morphophonology, with lexicographic application*. Canterbury: University of Canterbury MA thesis. <http://dx.doi.org/10.26021/4044>
- Stebbins, Tonya N. 2011. *Mali (Baining) grammar: A language of the East New Britain Province, Papua New Guinea* (Pacific Linguistics 623). Canberra: College of Asia and the Pacific, the Australian National University. <http://hdl.handle.net/1885/146770>
- Stebbins, Tonya N. & Tayul, Julius. 2012. *Mali (Baining) dictionary: Mali-Baining amēthamon angētha thēvaik* (Asia Pacific Linguistics A-PL 001). Canberra: College of Asia and the Pacific, the Australian National University. <http://hdl.handle.net/1885/9273>
- Tharp, Douglas. 1997. Organised phonology data: Sulka (Moglon) language [SLK], Milim, East New Britain Province. Unpublished manuscript. <https://www.sil.org/resources/archives/42479>
- Thurston, William R. 1976. *The phonology of Anēm, a non-Austronesian language of West New Britain*. Hamilton: McMaster University MA thesis. <http://hdl.handle.net/11375/10322>
- Thurston, William R. 1996. Amara: An Austronesian language of northwestern New Britain. In Malcolm D. Ross (ed.), *Studies in languages of New Britain and New Ireland, I: Austronesian languages of the North New Guinea Cluster in Northwestern New Britain* (Pacific Linguistics: Series C–135), 197–248. Canberra: Research School of Pacific and Asian Studies, the Australian National University. <http://hdl.handle.net/1885/145888>

- van den Berg, René & Bachet, Peter. 2006. *Vitu grammar sketch* (Data Papers on Papua New Guinea Languages 51). Ukarumpa: Summer Institute of Linguistics. <https://www.sil.org/resources/archives/23711>
- van den Berg, René & Wiebe, Brent. 2019. *Bola grammar sketch* (Data Papers on Papua New Guinea Languages 63). Ukarumpa: SIL-PNG Academic Publications. <https://www.sil.org/resources/archives/80489>
- Van Der Mark, Sheena Chantal. 2007. *A grammar of Vinitiri: An Austronesian language of Papua New Guinea*. Bundoora: La Trobe. (University PhD thesis.)
- Wurm, Stephen A. 1971. The Papuan linguistic situation. In Thomas A. Sebeok (ed.), *Current trends in linguistics, volume 8: Linguistics in Oceania, part one: Indigenous languages*, 541–657. The Hague: Mouton. <https://doi.org/10.1515/9783111418827-017>