1

Introduction

1.1 Purpose

This comparative dictionary provides an initial bottom-up reconstruction of one low-level Austronesian subgroup of the linguistic area of Wallacea: the Rote-Meto subgroup. This work forms one part of the larger bottom-up reconstruction of languages in this region, which is needed to fully understand the history of these languages.

This dictionary contains 1,174 reconstructions to Proto-Rote-Meto (PRM) or one of its lower branches (Chapter 3) along with the reflexes in modern languages that support these reconstructions. Proto-Rote-Meto is the hypothesised shared common ancestor of the Rote languages and the Meto language/dialect cluster spoken on the western part of Timor (Chapter 2).

This dictionary is *not* a reconstruction of any putative node between Proto-Malayo-Polynesian (PMP) and Proto-Rote-Meto (PRM). Nodes between these two levels are debated (Blust 1993; Donohue and Grimes 2008; Blust 2009b). Although reconstructions from putative intermediate nodes such as Proto-Central-Eastern Malayo-Polynesian (PCEMP) or Proto-Central Malayo-Polynesian (PCMP) are sometimes included as etyma for PRM forms, this should not be taken as a claim that these nodes exist. I simply follow the labels given by others to their reconstructions without passing judgement on the validity of these labels. For the purposes of this comparative dictionary I am agnostic regarding the validity of Central-Eastern Malayo-Polynesian (CEMP) and Central Malayo-Polynesian (CMP). The most fruitful way of assessing the evidence for such intermediate nodes is to continue and expand the kind of bottom-up work carried out in this dictionary.

1.2 Limitations

The reconstruction of Proto-Rote-Meto in this dictionary has a number of limitations. There are a number of minor limitations imposed by the sources on which this dictionary is based, as discussed further in §1.3 and §1.4.

However, probably the most serious limitation on this dictionary is the availability of data. While we have a wealth of data on the Termanu variety of Rote due to the work of the Dutch linguist J. C. G. Jonker in the nineteenth and earlier twentieth centuries, we have comparatively far less data on other varieties of Rote and the Meto cluster. Based on my experience so far, I fully expect that, as more data become available, it will be possible to add more reconstructions to this dictionary, while some of the current reconstructions will need to be modified.

The kind of additional data that will probably be added to this dictionary can be illustrated with two examples of reconstructed bird terms. Firstly, there is PRM *sinaraʔe '*Todiramphus spp*. (kind of kingfisher)'. This is supported by Bilbaa *sulae* '*Todiramphus chloris*', Rikou *sirae* '*Todiramphus chloris*', Lole *sinlaʔe* 'kingfisher', Ro'is Amarasi *sanae* '*Todiramphus spp.*', Kotos Amarasi *saʔnaʔe|k* '*Todiramphus spp.*', and Amfo'an *senae-l* '*Todiramphus spp.*'. None of the data supporting this reconstruction occur in the works of Jonker. Apart from the Kotos Amarasi term, I collected all the other cognates only at the end of 2017.

Secondly, there is PRM *tadeŋgus 'kind of dove, possibly rose-crowned fruit-dove, *Ptilinopus regina*'. In this case Dengka *leŋgus* and Oenale *reŋgus* 'pigeon, dove' (Dutch *duif*) occur in Jonker (1908:772), but none of the other cognates do. These other cognates were also collected at the end of 2017: Ro'is Amarasi *kuum treukus* 'rose-crowned fruit-dove', Kusa-Manea *ra~rukis* 'wild doves', as well as Landu and Rikou *rekus* 'rose-crowned fruit-dove'.

Such examples indicate that, even though Jonker's works are, perhaps, the most detailed and voluminous works published on any language of eastern Indonesia, more data on the Rote-Meto languages will allow the identification of yet more cognate sets and associated reconstructions.

Some such cognate sets are probably present in Jonker's works and simply await the discovery of a Meto form to allow reconstruction to PRM. One such example is PRM *hida 'how many', which is reconstructed to PRM on the basis of the Rote reflexes combined with Kusa-Manea (Meto) *hian* 'how many' — a term not otherwise currently known in the Meto cluster.

There are also probably other cognate sets that are not represented in Jonker's works due to his focus on Termanu. One such example is *leu 'now, already', for which the Rote cognates (Bilbaa and Rikou *leu*) were collected during my own fieldwork.

The relative lack of available data on speech varieties apart from Termanu is partially mitigated by the fact that I am personally most familiar with the Meto cluster (particularly Kotos Amarasi). The language Jonker knew best is thus complemented by the one I know best. Nonetheless, I am quite certain that more detailed and comprehensive data on Rote languages apart from Termanu, as well as more data on the Meto cluster, will allow reconstruction of many more terms than are currently included in this dictionary.

1.3 Data sources

This dictionary is based on a number of different sources of data. This includes published sources, unpublished descriptions and dictionaries, as well as my own field notes. In this section I discuss the sources on which I have drawn, and the limitations associated with these sources. This includes detailed discussions of Jonker (1908), from which most of my Rote data is drawn, as well as Middelkoop (1972), from which most of my Molo (Meto) data is drawn. The transcriptions used by these works and the problems associated with them are discussed in §1.4.

Data from several languages is frequently cited without any explicit reference to the source from which such data comes. The languages for which this is the case and the sources from which data for them comes are given in Table 1.1 in alphabetical order. All data for these languages come from sources given here, unless otherwise cited. Sources for other languages are always indicated.

Table 1.1: Sources for languages when no source is specified

Language	Source	Notes
Amfo'an	Culhane, Jumetan and Ora (2018)	unpublished Toolbox file, 1,711 headwords, revision published as Grimes et al. (2021)
Dela (Rote)	Tamelan (2017)	unpublished Toolbox file, 1,911 headwords
Dhao	Grimes, Ranoh and Aplugi (2008)	
Ende	Aoki and Nakagawa (1993)	
Galolen	Cristo Rei and Donohue (2012)	unpublished spreadsheet, 1,128 headwords
Hawu	Grimes, Lado, Ly and Tari (2008)	
Helong, Funai	Edwards (2018b)	
Helong, Semau	Balle and Cameron (2014)	unpublished Toolbox file, 3,368 headwords
Helong, Bolok	Balle and Cameron (2014)	Semau is the main dialect of this source
Kemak	own field notes	archived with PARADISEC[1]
Kisar	Christensen (in progress)	unpublished Toolbox file, 2,518 headwords
Meto lects	own field notes	see §1.3.1.2
Rote lects	Jonker (1908)	abbreviated as 'J'
Sumba lects	Onvlee (1984)	includes Anakalang, Kodi, Lewa, Mamboru, and Weyewa
Tetun Fehan	van Klinken (1995)	unpublished Toolbox file, 3,444 headwords
Waima'a	Himmelmann et al. (2006)	archived Toolbox file, 3,890 headwords
Welaun	own field notes	on LexiRumah[2] and PARADISEC

1.3.1 Rote

In this section I discuss the sources of my data for the Rote languages. Four kinds of Rote data occur in this comparative dictionary: data from Jonker (1908), data from my own fieldwork, data from work by linguists associated with the Language and Culture Unit (UBB) in Kupang, and data from a draft ritual language dictionary by James Fox (Fox 2016b). See the front matter and §2.2 for a list of Rote language/dialects and corresponding ISO 693-3 codes and Glottocodes.

1 catalog.paradisec.org.au/collections/OE8.
2 lexirumah.model-ling.eu/languages/west2547-welaun.

1.3.1.1 Jonker (1908)

Jonker (1908) is the source for the vast bulk of my data on the Rote languages. Jonker (1908) is an 805-page 'Rote'-Dutch dictionary. Termanu is the primary variety around which this dictionary is organised. The main section of this dictionary is 672 pages long and lists Termanu headwords, often as underlying roots, with derivatives as sub-headings along with definitions and example sentences. Jonker also usually provides cognates/equivalents of the Termanu headwords in eight other Rote varieties: Korbafo, Bokai, Bilbaa, Rikou, Ba'a, Tii, Dengka and Oenale. He also occasionally gives cognates from Oepao and Keka.

The cognates from these varieties are listed after the Termanu headword and are not given full definitions. Additionally, the final 140 pages or so of Jonker's dictionary are devoted to 'Forms and words from the other dialects'. This section lists non-Termanu words that have already been given in the main body of the dictionary, as well as forms that are not given there. While some of the headwords in this second part have definitions and example sentences, many entries are simply cross-references to the Termanu form in the main body of the dictionary.

In the introduction, Jonker explains that the non-Termanu words were collected by means of the Termanu form (or a derivative) mostly in a sentence and combined with a Malay translation. The non-Termanu equivalents were then given in written form by various schoolteachers from Rote who were Jonker's consultants. He explicitly states that it should *not* be assumed that the non-Termanu words have the same shades of meaning as their Termanu cognate/equivalent and notes differences in meaning when he is aware of them.

This imposes one limitation on this comparative dictionary. The exact semantics of most words from varieties of Rote other than Termanu is unknown. When I do not give a definition for a particular Rote form, this is because Jonker does not specify any definition. In such cases, while we can assume that the semantics are similar to the Termanu form, we cannot assume that they are identical. A word from another Rote variety may well have additional meanings that are not present in the Termanu form and/or Termanu may have uses that another variety does not have.

Regarding the form of words from varieties other than Termanu, Jonker states that he has confidence in the accuracy of his data and he does not include words that he suspected were mistakes. Similarly, forms he judged to be dubious are given in brackets and/or with a question mark.

1.3.1.2 Other sources

Three other kinds of Rote data occur in this comparative dictionary: data from my own fieldwork, data from work by linguists associated with the Kupang based Language and Culture Unit (UBB), and data from Fox (2016b). These other sources are summarised in Table 1.2.

Table 1.2: Sources other than Jonker (1908) for Rote languages

Lect	Source	Notes
Dela	Tamelan (2017) Tamelan (2021)	Toolbox[3] file 1,911 headwords PhD thesis
Lole	Zacharias, Balukh Balle and Banamtuan (2014)	Toolbox file 2,688 headwords
Rikou	Nako, Nako, Balle and Banamtuan (2014)	Toolbox file 766 headwords
Tii	Grimes, Cheng, Hayer-Pah, Pandie, Mulosing and Banamtuan (2014)	Toolbox file 3,281 headwords
Bilbaa	own field notes	archived with PARADISEC
Landu	own field notes	archived with PARADISEC
Oepao	own field notes	archived with PARADISEC
Rikou	own field notes	archived with PARADISEC
Termanu	Fox (2016b)	draft ritual language dictionary

Note: PARADISEC can be found at catalog.paradisec.org.au/collections/OE9.

I carried out a week's worth of fieldwork at the beginning of November 2017 in eastern Rote on Bilbaa, Landu, Oepao and Rikou. The primary purpose of this trip was to gather preliminary data on Oepao and Landu, neither of which are well represented in Jonker (1908).[4] During this trip, I also worked through the Rikou forms that were then present in my comparative database, which Jonker gives as orthographically vowel initial to check whether they begin with an underlying glottal stop (see §1.4.1.1).

Another major source of Rote data comes from work carried out by linguists associated with the Kupang-based Language and Culture Unit (UBB). Particularly important is data on Dela collected and provided by Thersia Tamelan. This includes a draft Dela dictionary (Tamelan 2017), as well as her PhD thesis at The Australian National University (Tamelan 2021), which is a grammar of Dela. In addition to having lexical forms not found in Jonker (1908), Tamelan's work provides a clear, modern linguistic description of Dela, including information that cannot be easily extracted from Jonker's dense grammatical description of the Rote languages, focused on Termanu (Jonker 1915).[5] One example of the kind of information lacking in Jonker's work, but described in Tamelan's work, is the distinction between vowel initial and glottal stop initial words (§1.4.1.1).

1.3.2 Meto

Meto data mainly comes from two sources: Middelkoop (1972) and my own fieldwork. I also have data on Amfo'an compiled by Kirsten Culhane in the form of a toolbox file with 1,711 headwords (Culhane, Jumetan and Ora 2018), as well as Baikeno data provided by Charles E. Grimes. See the front matter and §2.3 for a list of varieties of Meto and corresponding ISO 693-3 codes and Glottocodes.

3 software.sil.org/toolbox/. Toolbox is software program for managing lexical and corpus data.

4 Particular thanks go to Pieter Sjioen and Yulius Iu who were my main consultants for Oepao and Landu respectively, as well as Paulus Nako who organised my trip and accompanied me on fieldwork in Rote.

5 Fox and Grimes (1995:611f) put it well regarding Jonker's works, stating: 'Neither Jonker's dictionary nor his grammar is, in any conventional sense, the study of a single language. Jonker used both of these works to advance comparative observations on an extensive array of other Austronesian languages. His desire to be comprehensive, exhaustive and at the same time comparative resulted in studies that present formidable obstacles to a simple comprehension of the basic structures of Rotinese [sic]'.

1.3.2.1 Middelkoop (1972)

One major source of Meto data is Middelkoop (1972), an unpublished draft 673-page Meto-Dutch dictionary, which was still under preparation when Middelkoop passed away. This dictionary is based on the Molo variety, though forms from other varieties of Meto are occasionally included.

Pieter Middelkoop was a missionary linguist posted to Timor in 1922, after which he began to learn Meto. The dictionary was initially compiled in card-file format, which Middelkoop states was mostly ready by about 1947. The copy that I possess is a photocopy onto (mostly) A4 pages of the typed version of these cards. This photocopy originally belonged to James Fox, who generously gave it to me when he rediscovered it when moving offices. I made a scan of the physical dictionary and had it OCR-ed to make it semi-searchable. References to pages of Middelkoop (1972) are to the pages of the PDF of this scanned copy.

Given that Middelkoop's dictionary is a draft, headwords usually have only a simple Dutch gloss and in many cases only an example sentence with translation occurs, from which the sense of a form can be worked out. There are also many handwritten corrections or emendations on parts of the dictionary. The transcription used in this dictionary is discussed in §1.4.2.

1.3.2.2 Fieldwork

My fieldwork on the Meto cluster has been focused on Amarasi, I have spent about eight months working on Kotos Amarasi (based in the *desa* of Nekmese') and, among other data, have compiled a lexical database of 2,509 headwords of which 2,110 are unique mono-morphemic roots. My work on Kotos Amarasi is the reason that many of the definitions for this language are more extensive than those of other varieties of Meto.

I have also conducted a month's worth of fieldwork on Ro'is Amarasi and 190 of the headwords in my Amarasi lexical database are marked as exclusively Ro'is and a further 234 Kotos Amarasi headwords given with non-identical cognate Ro'is variants. (Many Kotos and Ro'is Amarasi morphemes are also identical.) Ro'is Amarasi data is important from a comparative perspective as, according to the historical phonology, it forms a primary branch within the Meto cluster. All other Meto varieties, including Kotos Amarasi, form a single Nuclear Meto subgroup (§3.3.4).

In addition to all this Amarasi data, I have also collected data on the following varieties of Meto: Timaus (35 minutes of transcribed and translated texts and 72 minutes untranscribed texts, lexicon of 748 headwords), Kusa-Manea (4 hours of untranscribed texts, lexicon of 489 headwords), Amanuban (22 untranscribed texts, 8 wordlists), Ketun (1 transcribed text, 2 untranscribed texts, 3 wordlists), Kopas (1 transcribed text, 2 untranscribed texts, 5 wordlists), Fatule'u (2 wordlists), Amanatun (2 wordlists), Molo (2 wordlists) and Amfo'an (1 wordlist).

Finally, Jonker (1908) frequently provides forms from an unspecified variety of Meto in his etymological notes. When I include such forms, they are marked as *Meto* and a reference to Jonker (1908) is given. Jonker usually does not give definitions for these Meto forms.

1.3.3 Other languages

Forms from languages outside the Rote-Meto group are frequently given in the out-comparisons section of this comparative dictionary (§1.5.1). These are usually other languages from the greater Timor region, though data from languages of other regions are also occasionally given.

In most cases, a full citation for such data is given. The exceptions are languages for which data comes from electronic sources or unpublished sources: Dhao, Galolen, Hawu, Helong, Kemak, Kisar, Tetun Fehan, and Waima'a. The sources for these languages were given in Table 1.1.

Data from other languages are given full references, including page numbers. Languages that commonly occur in the out-comparisons that are drawn from published dictionaries are given in Table 1.3. Some of these sources are abbreviated in the references.

Table 1.3: Sources of other languages frequently cited

Language	Sources	Abbreviation
Bima	Ismail, Azis, Yakub, Taufih and Usman (1985), Jonker (1893)	—
East Tetun	Morris (1984)	Mo
Ili'uun	de Josselin de Jong (1947)	dJ
Kambera	Onvlee (1984)	On
Mambae	Grimes, Marçal and Fereira (2014), Fogaça (2017)	—
Sika	Pareira and Lewis (1998)	—

In addition to Kambera, Onvlee (1984) gives cognates/equivalents of his Kambera headwords for other languages of Sumba: Anakalang, Kodi, Lewa, Mamboru, and Weyewa. However, in a very similar way to Jonker (1908), these cognates/equivalents are usually given without any definition. When I cite such forms, I list them after the Kambera word without a definition. In such cases, these Sumba languages have the same reference as Kambera.

In addition to these sources, Jonker (1908) frequently gives putative cognates for his Termanu headwords from many other languages. Among these, the data for Meto, Helong, Ende, Hawu, Kambera and Bima come from his own work. Unfortunately, glosses are only given for languages other than Rote when they diverge significantly from that of the Rote forms.

Whenever possible I have sought independent verification for such cognates from other sources and give the form, definition and citation for such putative cognates according to these other sources. However, many terms from other languages remain known only from Jonker's etymological notes and thus must be given without definition. That Jonker does not include definitions for such forms indicates that their semantics are likely close to that of the Rote forms.

Jonker does not usually specify from which variety of Helong or Meto his data comes. Thus, when I give a form from one of these languages taken from Jonker (1908) I cannot usually specify the variety and each is given respectively as only 'Helong' or 'Meto'.

1.4 Transcription

PRM reconstructions and their reflexes are given in a phonemic transcription according to standard IPA conventions with the exception of the palatal approximant, which is transcribed <*y*> to avoid any confusion with the voiced palatal affricate <*ʤ*> [ʤ].

When the phonemic representation of a form is in doubt, it is enclosed in angled brackets < >. This is often the case for Molo forms taken from Middelkoop (1972), as discussed further in §1.4.2, as well as out-comparisons from Sumba taken from Onvlee (1984).

As discussed in §1.4.1.1, Jonker does not distinguish between vowel initial and glottal stop initial words in his works. Initial glottal stops in brackets (ʔ) are used to indicate that a word from one of the Rote languages may begin with a glottal stop, but that this has not been confirmed. Initial glottal stops without brackets in Dela, Oenale, Rikou, Oepao and Landu indicate that I have confirmation from other sources or my own fieldwork that this glottal stop is underlying.

A variety of different punctuation marks is used in these phonemic transcriptions to represent different kinds of morphology. These punctuation marks are summarised in the front matter and Table 1.4. Productive affixes are separated from their stem with the hyphen -. Reduplication is indicated by the tilde ~. Fossilised suffixes that are no longer productive are separated from the stem by a vertical bar |. Historic compounds for which neither member is known to exist as an independent word in the language are separated by a slash /. When one member of a (historic) compound does exist as an independent word, but the other does not, the two are separated by an underscore _.

Table 1.4: Punctuation representing morphological structure

Punctuation		Use
-	hyphen	productive affix
~	tilde	reduplication (including frozen reduplication)
|	vertical bar	fossilised affix
/	slash	historic compound; neither member independent
_	underscore	historic compound; one member no longer independent

1.4.1 Jonker's transcription

Jonker consistently represents the different contrastive sounds of the Rote languages, with the exception of initial glottal stops and some sequences of two identical vowels. Each of these is discussed in further detail in the next sections.

Jonker's representation of consonants (apart from the glottal stop) is mostly straightforward. Most consonant letters correspond to their modern IPA equivalents. The digraph <*ng*> represents the velar nasal [ŋ] and imploded stops are transcribed identically to plain voiced stops. Thus, for instance, Tii <*b*> = [ɓ] and <*d*> = [ɗ]. This is not a problem for comparative purposes as implosion is a non-contrastive feature of voiced stops in the Rote languages.

1.4.1.1 Glottal stop

Jonker consistently represents the glottal stop word medially and finally. Between two vowels the glottal stop is represented by a diaeresis on the second vowel; thus <V̈> = /ʔV/ Examples from Termanu are given in Table 1.5.

Table 1.5: Medial glottal stops in Termanu from Jonker (1908)

Medial glottal stop			**No medial glottal stop**		
Jonker	Phonemic	gloss	Jonker	Phonemic	gloss
<daï >	*daʔi*	‘dirt on body’	*<dai>*	*dai*	‘reach, arrive at’
<sòë>	*soʔe*	‘coconut spoon’	*<sòe>*	*soe*	‘disaster’
<leä>	*leʔa*	‘fathom’	*<leak>*	*lea-k*	‘cave’
<haö>	*haʔo*	‘mineral lime’	*<hao>*	*hao*	‘eat with hand’
<liü>	*liʔu*	‘hit with stick’	*<iu>*	*iu*	‘shark’

Apart from its use to represent a glottal stop, the diaeresis also represents a morpheme break between a reduplicant and its base when partial reduplication applies to a vowel initial word. Thus, for instance, Jonker (1908:10) transcribes Rikou *a~ana* [ʔa.ana] ‘small’, which has no underlying glottal stop as *<aäna>*.

Word finally the glottal stop is represented by a dot under the preceding vowel, thus: <Ṿ> = /Vʔ/. Final glottal stops occur most commonly in Korbafo, Bilbaa, Rikou, Dengka and Oenale, in which case they are usually suffixes. Examples from Dengka are given in Table 1.6.

Table 1.6: Final glottal stops in Dengka from Jonker (1908)

Final glottal stop			**No final glottal stop**		
Jonker	Phonemic	gloss	Jonker	Phonemic	gloss
<luị>	*lui-ʔ*	‘bone’	*<lui>*	*lui*	‘take off’
<atẹ>	*ate-ʔ*	‘liver’	*<ate>*	*ate*	‘slave’
<buna>	*buna-ʔ*	‘flower’	*<bina>*	*bina*	‘k.o. shellfish’
<kòlọ>	*kolo-ʔ*	‘hole’	*<lòlo>*	*lolo*	‘stretch out’
<seụ>	*seu-ʔ*	‘Alstonia villosa’	*<seu>*	*seu*	‘pick fruit’

Jonker states in the introduction to his dictionary that a final glottal stop in Bilbaa and Rikou was often not marked. Although regrettable, given that final glottal stops are almost always suffixes in these languages, this does not impact greatly on the application of the comparative method itself or the reconstruction of proto-forms in my comparative dictionary.

More worryingly, Jonker states that for Rikou (and to a lesser extent also Bokai) the difference between a sequence of two vowels without any intervening consonant and a sequence of two vowels with an intervening glottal stop was often not recorded. I have checked the Rikou forms against data from other sources (§1.3.1.2) wherever possible and the two nearly always agree; though in less than half a dozen cases, Jonker’s form has a medial glottal stop where my other sources do not. All such cases are noted.

Jonker does not write word initial glottal stops. All Rote varieties for which modern phonetic and phonological data is available have a contrast between vowel initial and glottal stop initial roots. The contrast is neutralised phrase initially (including in isolation) due to an automatic process of phrase initial glottal stop insertion. The difference emerges phrase medially.

Thus, for instance, Rikou vowel initial *ura-ʔ* 'scorpion' and glottal stop initial *ʔuse-ʔ* 'navel' are both realised with an initial glottal stop in isolation: [ˈʔuraʔ] 'scorpion' and [ˈʔusɛʔ] 'navel'. Phrase medially, however, 'scorpion' is vowel initial, as seen in the phrase *au ura=ka* [ˌʔauˈuraka] 'my scorpion', while 'navel' is glottal stop initial, as seen in *au ʔuse=ka* [ˌʔauˈʔusɛka] 'my navel'.

Jonker does not distinguish between such words and writes both with an initial vowel. A selection of examples of Rikou vowel initial and glottal stop initial roots (based on my own fieldwork), along with their representation in Jonker, as well as their realisations initially and medially, is given in Table 1.7 to further illustrate.

Table 1.7: Initial glottal stops in Rikou

Jonker	*<inde>*	*<ea>*	*<apa>*	*<ofạ>*	*<usẹ>*
Phonemic	*ʔinde*	*ʔea*	*ʔapa*	*ʔofa-ʔ*	*ʔuse-ʔ*
Initial	[ˈʔindɛ]	[ˈʔea]	[ˈʔapa]	[ˈʔɔfaʔ]	[ˈʔusɛʔ]
Medial	*ria ʔinde=na*	*au ʔea=ka*	*ria ʔapa=na*	*au ʔofa=ka*	*ria ʔuse=na*
Phonetic	[ˌriaˈʔindɛna]	[ˌʔauˈʔɛaka]	[ˌriaˈʔapana]	[ˌʔauˈʔɔfaka]	[ˌriaˈʔusɛna]
Gloss	'her/his spindle'	'my turtle'	'her/his buffalo'	'my canoe'	'her/his navel'
Jonker	*<ikọ>*	*<eị>*	*<ape>*	*<ò>*	*<urạ>*
Phonemic	*iko-ʔ*	*ei-ʔ*	*ape*	*oo*	*ura-ʔ*
Initial	[ˈʔikɔʔ]	[ˈʔeiʔ]	[ˈʔapɛ]	[ˈʔɔː]	[ˈʔuraʔ]
Medial	*ria iko=na*	*au ei=ka*	*ria ape=na*	*au oo=ka*	*ria ura=na*
Phonetic	[ˌriaˈikɔna]	[ˌʔauˈeika]	[ˌriaˈapena]	[ˌʔauˈɔːna]	[ˌriaˈurana]
Gloss	'its tail'	'my foot'	'her/his saliva'	'my bamboo'	'her/his scorpion'

That Jonker does not transcribe initial glottal stops is a problem for comparative purposes, as an initial glottal stop is often a reflex of an earlier consonant. Fortunately, however, I have access to data for both Dela (Tamelan 2017) and Rikou (own field notes), which allows me to distinguish between most vowel initial and glottal stop initial roots for these languages. I have thus transcribed all underlying glottal stops for Rikou, as well as Dela and Oenale as befits their phonemic status.[6]

Initial glottal stops that are suspected to be underlying in other varieties of Rote are transcribed in brackets to indicate their unconfirmed status. Thus, for instance, for 'buffalo' Jonker gives Dengka and Oenale *<amba>* while Tamelan (2017) gives Dela *ʔamba*. I thus transcribe the Dela and Oenale forms as *ʔamba* and the Dengka form *(ʔ)amba* to indicate that Dengka probably has an initial, but unconfirmed, underlying initial glottal stop.

6 Dela and Oenale are nearly identical and can be treated to some extent as a single lect.

The status of some initial glottal stops in Dela and Rikou remains ambiguous. This is either because the form in Jonker (1908) is not present in Tamelan (2017) or was not known by my consultants. Additionally, a number of Rikou forms with an ambiguous glottal stop were added to my database after my fieldwork in Rote. Such ambiguous glottal stops are given in brackets (ʔ) in the same way as all unconfirmed initial glottal stops in the Rote languages.

Meto varieties also have a process of glottal stop insertion before vowel initial words. However, unlike the Rote languages this glottal stop insertion affects vowel initial words in all phrase positions — not just phrase initially. The relevance of this process in Meto for comparative purposes is discussed in §2.4.2.

1.4.1.2 Vowels

Jonker represents the five phonemic vowels of the Rote languages with and without various accents according to a combination of the placement of stress and phonetic vowel quality. His transcription of vowels is not phonemic, but is consistent. A summary of Jonker's transcription of vowels is given in Table 1.8. Observe in particular that some letters represent both single and double vowels.

Table 1.8: Jonker's transcription of vowels

Jonker	Phonetic	Phonemic	Use by Jonker
e	[e]	*e*	all environments
o	[o]	*o*	all environments
è	[ɛː]	*ee*	final/only vowel of word
	[ɛ]	*e*	elsewhere
ò	[ɔː]	*oo*	final/only vowel of word
	[ɔ]	*o*	elsewhere
i	[iː]	*ii*	only vowel of word
	[i]	*i*	elsewhere
a	[aː]	*aa*	only vowel of word
	[a]	*a*	elsewhere
u	[uː]	*uu*	only vowel of word
	[u]	*u*	elsewhere
í	[iː]	*ii*	final vowel of word
á	[aː]	*aa*	final vowel of word
ú	[uː]	*uu*	final vowel of word

Before this transcription is discussed in detail, several facts concerning the phonology of vowels in Termanu must be summarised. A fuller overview of Rote phonology is given in §2.4.

- Mid-vowels are lax [ɛ] and [ɔ] when stressed and before a syllable containing a mid-vowel.
- Mid-vowels are tense [e] and [o] before a syllable containing a high vowel.
- Content words contain at least two vowels.
- Nearly all vowel sequences occur, including double vowels /ii/, /ee/, /aa/, /oo/ and /uu/.

- Double vowels are realised as phonetically long vowels, i.e. /$V_{\alpha}V_{\alpha}$/ → [V:].
- Stress is penultimate. Secondary stress is assigned to every second syllable to the left.

Given this knowledge of Rote phonology, the following generalisations account for Jonker's transcription. Jonker transcribes the lax allophones of the mid vowels [ɛ] and [ɔ] with a grave accent <*è*> and <*ò*> whenever these allophones are stressed. Unstressed lax vowels, as well as stressed tense allophones are transcribed as <*e*> and <*o*> without any accent.

Given that stress falls on the penultimate vowel of a word, a double vowel will always contain a vowel that is the locus of stress and the whole double vowel will thus bear stress (either primary or secondary). This — combined with the fact that the first vowel of a sequence of two mid vowels is by nature followed by another mid vowel — means that a double mid-vowel will always be lax. Thus, the double vowels /ee/ and /oo/ are always stressed lax [ɛ:] and [ɔ:].

Consequently, Jonker's <*è*> and <*ò*> represent double vowels /ee/ and /oo/ when they are the final or only orthographic vowel of a word. Jonker's <*è*> and <*ò*> do not represent double vowels when they are penultimate before another syllable or vowel. Examples of Jonker's transcription of the stressed lax allophones of the mid vowels in Termanu are given in Table 1.9.

Table 1.9: Stressed lax allophones of mid vowels

Jonker	Phonemic	Phonetic	gloss	Jonker	Phonemic	Phonetic	gloss
<*nè*>	*nee*	[ˈnɛ:]	'six'	<*dène*>	*dene*	[ˈdɛnɛ]	'kapok'
<*na-tanè*>	*na-tanee*	[ˌnataˈnɛ:]	'contain'	<*dèlo*>	*delo*	[ˈdɛlɔ]	'citrus'
<*nò*>	*noo*	[ˈnɔ:]	'coconut'	<*bòle*>	*bole*	[ˈbɔlɛ]	'arenga palm'
<*na-sakò*>	*na-sa-koo*	[ˌnasaˈkɔ:]	'sip'	<*bòö*>	*boʔo*	[ˈbɔʔɔ]	'cough'

For the non-mid vowels /i/, /a/, and /u/, Jonker transcribes a final double vowel with an acute accent, to indicate phonetic stress. Examples from Termanu include: <*nakapí*> = *na-ka-pii* 'tense', <*bitiná*> = *bitinaa* '*Kleinhovia hospita*', and <*na-tafú*> = *na-tafuu* 'inflated'.

In other situations, double /ii/, /aa/, and /uu/ are not distinguished orthographically from single vowels. However, given that Rote content words minimally contain two vowels, any content word that contains only a single orthographic vowel in Jonker (1908) must have a double vowel underlyingly. Termanu examples are given in Table 1.10. That such words do indeed contain a double vowel is confirmed by the fact that when they occur with a prefix or as final members of a compound, Jonker transcribes them with an acute accent, as shown in the right-hand side of Table 1.10.

Table 1.10: Double vowels in Termanu

Jonker	Phonemic	Phonetic	gloss	Jonker	Phonemic	Phonetic	gloss
<*ki*>	*kii*	[ˈki:]	'left, north'	<*alu-kík*>	*alu kii-k*	[ˌaluˈki:k]	'left shoulder'
<*dak*>	*daa-k*	[ˈda:k]	'blood'	<*nadá*>	*na-daa*	[naˈda:]	'bleed'
<*huk*>	*huu-k*	[ˈhu:k]	'trunk, source'	<*ai-húk*>	*ai huu-k*	[ˌʔaiˈhu:k]	'tree trunk'

In summary, Jonker's use of accents to mark stress, combined with the minimal divocalic word requirement, means that there are almost no cases in which it is unclear whether a particular orthographic vowel represents a single vowel or a sequence of two identical vowels.

The only words that remain ambiguous are functors with single <*i*>, <*a*> or <*u*>, such as <*ma*> 'and'. This is because functors can be monosyllabic in the Rote languages. In these cases, I have referred to other sources to determine whether such functors contain a single or double vowel. In the case of *ma* 'and', Tamelan (2017) gives Dela *ma* 'and' with a single vowel.

1.4.2 Middelkoop's transcription

The transcription used in Middelkoop (1972) is not phonemic and under-representation, particularly of the glottal stop and double vowels, is common. Middelkoop's orthography can be used to some extent by those who already know the language, but it cannot be used as a reliable, consistent representation of the phonological structures of Meto. The main issues with Middelkoop's transcription are summarised in the next sections.

- Double vowels /ii/, /ee/, /aa/, /oo/ and /uu/ are written with a single letter.
- The glottal stop is transcribed < ' > between two vowels.
- Word final < ' > indicates either doubling of the previous vowel or doubling of the previous vowel followed by a glottal stop; i.e. <V'> = /VV/ ~ /VVʔ/.
- Word finally and before consonants the glottal stop is usually not transcribed.
- The vowel sequences /ao/ and /au/ are both written <*au*>.
- The vowel sequence /ae/ appears to be written <*ai*> when non-final.
- Prefixes which are a single consonant are often written with a previous vowel final word.
- The final vowel of the pronouns *ina* 3SG, *sina* 3PL and *hita* 1PL.INCL is usually written with a following inflected verb.

Even with experience of working on Meto and having used Middelkoop (1972) for years, it is still sometimes unclear to me what the exact form of a particular word is. Such forms are indicated in this comparative dictionary with angle brackets < >. All words not given in angle brackets from Middelkoop (1972) have been re-transcribed according to their phonemic form.

Thus, Middelkoop's <*fule*> 'foam' is probably either *fuleʔ* or *ʔfuleʔ*, but I have no way of knowing whether the Molo form has an initial glottal stop or not; Kotos Amarasi has *ʔfuri-f*, indicating that an initial glottal stop might be present in the Molo form, while Kusa-Manea has *fa~fura-f* indicating that a glottal stop might not be present.

Similarly, the word for 'turtle' is given with the forms <*ke'*>, <*ke'a*>, <*ke*> and <*kel*> with no obvious way of knowing what the orthographic variation means. Indeed, there are probably variant forms of which the final two are certainly *kee* and *kee-l*, but whether <*ke'*> is *kee* or *keeʔ* is unclear and whether <*ke'a*> is *keʔa, kea, keaʔ*, *kee=aa* and/or *keeʔ=aa* is also unclear.

1.4.2.1 Glottal stop

The glottal stop phoneme /ʔ/ is usually only written when it occurs between two vowels. Between two vowels /ʔ/ is written with an apostrophe < ' >. An example is <*ma'eki*> = *maʔekiʔ* 'slippery'.

Word finally and before consonants the glottal stop is not usually written. Hence Middelkoop's <*mafena*> = *maʔfenaʔ* 'heavy' has two glottal stops in all known varieties of Meto, including my own Molo data collected on the basis of fieldwork.

Initial glottal stops before other consonants can sometimes be detected by the presence of variants with and without initial epenthetic /a/, which often occurs before consonant clusters. Thus, <*akalen*> and <*kale*> 'fraenulum' indicate a form with initial /ʔk/, probably *a-ʔkale-n* and *ʔkaleʔ*, respectively. However, such glottal stop detection is a very inexact science. Here my confidence that this word does indeed begin with a glottal stop comes mainly from my own Amarasi data where I have *ʔkare-f* 'palate'.

There are sporadic instances of the glottal stop being written < ' > when it is stem/root initial before another consonant, but this is rare and inconsistent. Sometimes it even appears to be written after the initial consonant, such as <*nak'ai*> = *na-ʔkai(ʔ)* 'hook (v.)'.

1.4.2.2 Double vowels

Like Rote, Meto has a minimum word requirement whereby every content word must contain at least two vowels, with double vowels realised phonetically as a single long vowel, e.g. *oo* → [ʔɔː] 'bamboo', *n-iit* → [niːt̪] 'sees'. The only words that can be exempt from this requirement are functors. But even for functors I found evidence that most also have a minimum of two vowels, at least historically.

Word final sequences of two identical vowels can be marked with the apostrophe < ' > in Middelkoop (1972). An example is < *o'* > = *oo* 'kind of bamboo'. But use of the apostrophe to represent double vowels is not a rule. Thus, for <*ne*> = *nee(ʔ)* 'six' there is no indication of the double vowel. Sometimes a word written with a final < ' > also ends in a glottal stop, such as in <*na'*> = *naaʔ* 'blood'. Word final < ' > indicates either doubling of the previous vowel or doubling of the previous vowel followed by a glottal stop, i.e. <V'> = /VV/ ~ /VVʔ/.

Before a consonant (other than glottal stop), sequences of two identical vowels are normally written with a single letter. Examples include <*bifel*> = *bifee-l* 'woman', <*hun*> = *huun* 'grass', <*lus*> = *luus* 'deer', and <*ek fui*> = *eek fui* 'kind of agave', and so on.[7]

1.4.2.3 Other vowel sequences

The vowel sequence /ao/ is written identically to /au/ as <*au*>. As an example, both *au* '1SG' and *ao-f* 'body' are given under a single <*au*> headword. Another example is <*nau*> = *nao* 'go'.

7 There are rare exceptions in which a double vowel is written with two letters such as <*eem*> '2PL come'.

When it is non-final the vowel sequence /ae/ appears to written as <*ai*>. Thus, Amarasi has *na-ʔaekaʔ* 'soak', which is cognate with the Molo forms <*naik*> *n-(ʔ)aek* and <*u aikat*> *u-ʔaeka-t* in Middelkoop (1972). Word finally Middelkoop writes /ae/ as <*aè*>. Two examples are <*saè*> = *sae* 'go up' and <*maè*> = *mae* 'ashamed'.

1.4.2.4 Morpheme boundaries

Verb agreement is usually obligatory in Meto, and one set of agreement prefixes consists of only a consonant: *ʔ-* 1SG, *m-* 2SG/2PL/1PL.EXCL, *n-* 3SG/3PL, *t-* 1PL.INCL (§2.6.5). When the agreement prefix *m-* occurs after any of its corresponding pronouns — *hoo* 2SG, *hii* 2PL or *hai* 1PL.EXCL — the prefix is often written as the final consonant of this pronoun. Examples include:

- <*ho**m** pau fe*> = *hoo **m-**pao feʔ* 'keep waiting for me'
- <*hi**m** tok mitloe*> = *hii **m-**took mi-tloe* 'you sit parallel'
- <*hai**m** hek manu*> = *hai **m-**heek manu* 'we caught a chicken'

Consistent with the fact that Middelkoop (1972) does not usually write initial glottal stops, the 1SG prefix *ʔ-* is not usually written. Thus, <*au tebi kukis*> = *au ʔ-tebi kukis* 'I crumble bread'.

The 3SG, 3PL and 1PL.INCL pronouns are *in, sin* and *hit*, respectively. These pronouns have vowel final forms with are used before consonant clusters: *ina, sina* and *hita*, respectively. When these pronouns occur before an agreeing verb whose root begins with a consonant, the final vowel of these pronouns is written with the verb in Middelkoop (1972). Examples include:

- <*i**n a**nlo'*> = *in**a** n-looʔ* 's/he throws up' (p. 280)
- <*si**n a**note hau neki fani*> = *sin**a** **n-**ʔote hau n-eki fani* 'they cut a tree with an axe'
- <*hi**t a**tlaksaè talali noel*> = *hi**ta** t-laak sae ta-lali noe-l* 'we stepped over the river'

In such instances, a case can be made for writing these forms orthographically with the /a/ attached to the verb as this aids morpheme recognition. While this may be helpful for native speakers, it is unhelpful for identifying the underlying structures for the purposes of linguistic analysis and reconstruction.

Occasionally, enclitics are written attached to their host, but given the draft state of Middelkoop's dictionary these could just be typographical errors. One example is <*hom nau man**kit** ho feto*> = *hoo m-nao m-aan **=kiit** hoo fetoʔ* 'go and get your sister for us'.

1.5 Structure of the dictionary

This comparative dictionary is structured around reconstructions. Headwords are reconstructions to Proto-Rote-Meto, or to one of its sub-nodes (see §3.3). Headwords are arranged alphabetically with IPA symbols placed after their nearest equivalent.

Prenasalised stops are treated as separate letters. The complete order of all letters which occur in PRM reconstructions is as follows: *a, *b, *ɓ, *d, *ɗ, *ʤ, *e, *ə, *f, *h, *i, *k, *ʔ, *l, *m, *mb, *n, *nd, *ŋ, *ŋg, *o, *p, *r, *s, *t, *u, *w.

Each reconstruction is defined and a variety of other information is also given. Below each reconstruction are the forms in the Rote-Meto languages, which provide evidence for the reconstruction. A simple example entry is exemplified in Example 1.1.

Example 1.1: Simple PRM reconstruction

***ika** *Morph:* ***ika-k**. *PRM.* fish. *Etym:* *hikan. *Pattern:* k-10.
iʔa-k *Termanu.* fish. (J:200)
iʔa-ʔ *Korbafo.*
iʔa-k *Bokai.*
ika-ʔ *Bilbaa.*
ika-ʔ *Landu.* fish. (own field notes)
ika-ʔ *Rikou.*
ika-ʔ *Oepao.* fish. (own field notes)
iʔa-k *Ba'a.*
iʔa-k *Tii.*
ia-ʔ *Dengka.*
ʔuʔu_ia-ʔ *Dela.* all kinds of fish. *[Form:* **ʔuʔu** is the normal word for 'fish' in Dela.*]*
ika|ʔ *Ro'is Amarasi.* fish.
ika|ʔ *Kotos Amarasi.* fish.
ika|ʔ *Molo.* fish. *(*M:159)
ika|ʔ *Kusa-Manea.* fish.
Out-comparisons:
ikan *Semau Helong.* fish.
ikan *East Tetun.* fish. *(*Mo:90)
iʔa *Dhao.* fish.

Each entry is headed by a reconstruction in boldface, in this case ***ika**. Directly after the root is given any morphology with which this root occurred (§1.5.3.1); in this case, *ika took the nominal suffix *-k (§2.6.1). The level at which this reconstruction is made (§3.4) is given after this in italics, in this case Proto-Rote-Meto (PRM). This is followed by the reconstructed meaning ('fish'), which in turn is followed by any etymon at a higher level, usually PMP (§1.5.3.3). After the etymology, notes on any issues or problems with this reconstruction are given (§1.5.3.7).

Below the reconstruction are given its reflexes in the Rote-Meto languages. This begins with the varieties of Rote in the following order: Termanu, (Keka), Korbafo, Bokai, Bilbaa, (Landu), Rikou, (Oepao), Ba'a, (Lole), Tii, Dengka, Oenale and Dela.[8] Lects in brackets in this list are not always given, while forms from the other lects are given whenever they are known to be cognate.

8 The ordering in which the Rote languages are given is a compromise between geographic position and relatedness. Termanu is given first as this is the variety on which Jonker (1908) is based, followed by Korbafo, which is extremely similar. Bokai is given next with an eastwards progression from there to Bilbaa, Landu, Rikou and Oepao. After this Ba'a is given with a mostly westwards progression to Lole, Tii, Dengka, Oenale and Dela.

In line with the nature of Jonker (1908) (see §1.3.1), from which most Rote data comes, Termanu is often the only Rote variety defined. The definition is a free translation of Jonker's Dutch definition. When the source of a form does not come from Jonker (1908), it is usually glossed and the source is given. Thus, in the case of *ika 'fish' in Example 1.1 the Landu and Oepao forms *ika-ʔ* 'fish' come from my own field notes. When no definition is given for a variety of Rote, the definition must be assumed to be similar to that of Termanu. Thus, the Korbafo, Bokai, Bilbaa, Rikou, Ba'a, Tii and Dengka forms in Example 1.1 can be assumed to also mean 'fish'.

After the definition, any necessary notes are given on issues concerning this reflex. The final part of the entry for a reflex is the source of this form. In Example 1.1, Termanu *iʔa-k* comes from page 200 of Jonker (1908). When no alternate source is given for another Rote variety, it comes from the same page of Jonker (1908) as the Termanu form. The only exception is Dela, data for which always comes from Tamelan (2017). For more discussion on the sources on which this dictionary is based see §1.3.

After Rote, forms from different varieties of Meto are given. Often only a Kotos Amarasi and Molo form are given, as these are the varieties for which the most data is available. When other lects are given they are ordered roughly from east to west.[9] These entries follow the same model as the Rote entries except that glosses for all varieties are given.

1.5.1 Out-comparisons

The final part of an entry consists of putative cognates in languages outside of the Rote-Meto subgroup. These are preceded by *Out-comparisons* and further indented. These out-comparisons are forms from languages outside of the Rote-Meto group that are formally and semantically similar to the Rote-Meto languages. In most cases, these out-comparisons are from Austronesian languages in the greater Timor region, though occasionally out-comparisons from non-Austronesian or more distant languages are also given.

All languages that occur more than three times in the out-comparisons are listed in the front matter, along with their geographic location, ISO 693-3 codes, and Glottocodes (Hammarström et al. 2020). Equivalent information for languages that occur three or less times in the out-comparisons is included each time a form from that language occurs.

These out-comparisons have *not* been thoroughly vetted for whether they show regular sound correspondences with the PRM forms and thus cannot be taken unquestioningly as cognate. While it is likely that many of the out-comparisons are indeed cognate with the PRM forms, this cannot be assumed to be the case. A more thorough investigation of the sound correspondences may show that certain forms in certain lects are borrowings and not the result of shared inheritance.

9 The complete ordering of Meto lects is as follows: Ro'is Amarasi, Kotos Amarasi, Amanuban, Amanatun, Ketun, Kopas, Fatule'u, Molo, Amfo'an, Timaus, Baikeno, Kusa-Manea. With the exceptions of Ketun, Kopas and Timaus (which probably originate elsewhere), this is roughly east to west along the south coast, then east to west among remaining varieties, with Timaus after Amfo'an, from which it originated.

This is particularly the case for Helong out-comparisons. Helong and the Rote-Meto languages appear to have been in contact with one another since the time of PRM and have remained in contact ever since (see §3.3.4.2). It remains to be determined to what extent it may be possible to detect contact at different time depths between Helong and Rote-Meto.[10]

While I do give preliminary notes on what appear to be irregular sound correspondences between the PRM reconstruction and out-comparisons when I am aware of them, the nature of such irregularities remains to be properly investigated.

In general, I have searched fairly thoroughly for putative cognates in Tetun, Ili'uun, Helong and Kisar based on available sources. I have not made a thorough search for cognates in other regional languages, but have included similarities that I opportunistically stumbled upon.

When the out-comparison field contains forms that can be identified as loans from a Rote-Meto language, or a Rote-Meto entry can be identified as a loan from one of the out-comparisons, the form that is a loan is preceded by the degree sign ° and the source of the loan is given. An example is *ɗ̀ele 'Job's tears, Coix lachryma-jobi', from PMP *zəlay, which is inherited regularly in the Rote languages. Meto, however, has irregular °*sone*, which is probably a loan from Welaun *sole* (also from *zəlay). When all the forms under a single headword can be identified as loans, they are given in the Loan Distribution section (§1.5.4).

1.5.2 Multiple reflexes

In many cases, a language has more than one reflex of a single reconstruction; usually morphologically related forms. In such cases each form except the first is numbered sequentially, with these numbers corresponding to the equivalent numbered section of the definition. An example is given in Example 1.2.

Example 1.2: Multiple reflexes

***lasi**$_2$ *Morph:* ***ma-lasi-k, *na-ma-lasi**. *PRM.* old, aged.
lasi-k (2) na-ma-lasi *Termanu.* 1) old (especially of people and animals). 2) to be or become old, of people and animals. (J:283)
lasi-ʔ *Korbafo.*
lasi-k *Bokai.*
lasi-k *Bilbaa.*
lasi-ʔ *Rikou.*
lasi-k *Ba'a.*
lasi-k *Tii.*
lasi-ʔ *Dengka.*
lasi-ʔ *Oenale.*

10 As an additional complicating factor, it is not unlikely that Helong and the Rote-Meto languages both had contact with the same non-Austronesian language(s) and/or families that have since been lost due to shift to Helong and/or Rote-Meto. Again, it has not yet been determined to what extent it may be possible to detect which equivalents/cognates shared between Helong and the Rote-Meto languages are the result of mutual borrowing and/or borrowing from a third source.

m|nasi|ʔ *Ro'is Amarasi.* old, aged.
m|nasi|ʔ (2) na-m|mnasi *Kotos Amarasi.* 1) old, aged. 2) be or become old/aged.
m|nasi|ʔ *Molo.* old. (M:325)
m|nasi|ʔ (2) m|nasi-k *Kusa-Manea.* 1) old (e.g. of fruit). 2) old, aged (of people).

In this example, Termanu has two different forms. The definitions of these forms are given after the forms. Form (1) *lasi-k* means 'old (especially of people and animals)', and form (2) *na-ma-lasi* means 'be or become old, of people and animals'. When other varieties of Rote taken from Jonker (1908) also have multiple forms, these are given the same numbers as the parts of the Termanu entry to which they correspond.

Example 1.3: Example with multiple reflexes in all Rote varieties

***beli** *Rote.* price, bride price. *Etym:* *bəli 'buy, value, price; marriage prestations, bride price; purchase'.
beli (2) belis *Termanu.* 1) cost, price, value. 2) that which must be paid for a girl when taken for marriage, either paid with goods or money, the purchase price of a woman. (J:41)
beli (2) belis *Korbafo.*
beli (2) belis *Bokai.*
beli (2) belis *Bilbaa.*
beli (2) belis *Rikou.*
beli (2) belis *Ba'a.*
ɓeli (2) ɓelis *Tii.*
feli (2) felis *Dengka.*
feli (2) felis *Oenale.*
Out-comparisons:
foli-n *East Tetun.* price, cost, value; objects for barter. (Mo:35)
heli *Ili'uun.* property, valuable things. (dJ:117)
weli *Kisar.* buy.

Example 1.3 shows how the first form in all the Rote lects corresponds to Termanu *beli* 'cost, price, value' and the second form corresponds to Termanu *belis* 'bride-price'. Often it is only the Termanu entry that has multiple forms, as in Example 1.2. While equivalent numbered forms across the Rote lects are given the same numerals, such equivalency does not necessarily correspond to any forms among the Meto cluster.

I generally only give multiple reflexes of a single reconstruction when these additional forms contribute to understanding the history or development of the reconstruction. Thus, in Example 1.2, the Termanu and Meto forms attest that *lasi$_2$ probably had two forms in PRM: nominal *ma-lasi-k and verbal *na-ma-lasi.

In some cases multiple reflexes occur for a single language with no indication of a semantic difference. In such cases, the two forms are separated by a comma and receive a single gloss. Thus, for instance, two Molo reflexes of *mbeɗa 'put down' occur with no know semantic difference. They are given as **na-pela**, **na-bela** *Molo*. put down. (M:56, xxxix). Note also that in such cases the ordering of page numbers in the source

follows the order in which the reflexes are given in this dictionary, rather than their actual ordering in the source — that is, *na-pela* occurs on page 56 of Middelkoop (1972) and *na-bela* occurs on page xxxix.

1.5.3 Fields/parts of entries

Apart from the forms and definitions, many entries have additional fields that give additional information on the entry. These fields include a morphology field marking any morphology with which the reconstruction occurred (§1.5.3.2), an etymology field indicating a higher level etymon (§1.5.3.3), a field for indicating doublets (§1.5.3.4), fields tracking irregular sound changes (§1.5.3.5), a field tracking patterns or correspondences among proto-phonemes that undergo unconditioned splits (§1.5.3.6), and four fields for different kinds of notes (§1.5.3.7).

1.5.3.1 Definitions and glosses

Glosses of reconstructions are always given, but glosses of the forms in daughter languages are not always given. All glosses of reconstructions are my own proposed semantics for the PRM form, but glosses for the reflexes in daughter languages follow verbatim the sources from which data is drawn (see §1.3). This includes not giving any gloss when no gloss is given in the source, as is usually the case for non-Termanu forms taken from Jonker (1908).[11] In such cases, we must assume that the meaning of the non-Termanu terms is similar to that of Termanu (see §1.1.1.1 for more discussion). Thus, most Rote forms are not glossed and the Termanu gloss can be taken as a proxy for the meaning.

The convention is to supply glosses as in the sources verbatim, including a lack of gloss, not to leave out glosses that are identical to the first language listed. Thus, nearly all data not drawn from Jonker (1908) is glossed. This is exemplified in Example 1.1 where the forms in Landu, Oepao, Dela, Meto and the out-comparisons are all glossed according to the sources from which they are drawn, even though theses glosses are usually identical. The forms without glosses are taken from Jonker (1908) where they listed as cognates of the Termanu headword.

In some cases, Jonker (1908) explicitly gives the meaning of one or more Rote forms as differing from their Termanu cognate. Likewise, word sets taken from the final section of his dictionary, which is devoted to ‘Forms and words from the other dialects’, often have no Termanu cognate. In this case, the gloss for the first language listed is given and the glosses for subsequent varieties of Rote in Jonker (1908) are not glossed and must be assumed to be equivalent to the language for which the gloss is given. When a cognate set is restricted to Dengka and Oenale the gloss is usually repeated for both forms.

The glossing of terms referring to biota follows the practice used elsewhere in the dictionary of citing glosses verbatim from the sources. Many biota terms only receive a vague definition in Jonker (1908), as ‘kind of X’, though they are often accompanied

11 The statements here also apply to terms from the Sumba languages taken from Onvlee (1984), in which case we must assume that the unglossed forms are similar to the Kambera forms. These statements also apply to some forms from Maluku, in which case the unglossed forms must be assumed to be similar to the first language listed.

by a description and/or their name in Kupang Malay, which can be used to identify the species referred to. Whenever a biota term is defined precisely in one source (often with a scientific name) and defined only vaguely in other sources, I usually assign the more precise definition to the reconstructed PRM form.

Some discussion of introduced species of biota is needed. There are a number of reconstructed terms in this dictionary that refer to recently introduced species of biota in daughter languages. When the sound correspondences of such terms are regular, this indicates that semantic shift has taken place. In an insightful article on the history of traditional agriculture, Fox (1991) has shown how newly introduced crops were assimilated to pre-existing categories in the Rote languages. Thus, for instance, based on a cursory examination of the present-day semantics, we could assign the meaning 'maize' to PRM *mbela. However, to quote Fox (1991:250):

> On Roti [sic], it is clear that maize when it was introduced was culturally assimilated to the category of 'sorghum'. It is also conceivable that at an earlier period when sorghum was introduced, it was assimilated to the category of Job's tears. Thus this category, *pela* [from PRM *mbela], may subsume three stages of an agricultural progression.

See the discussion in the entry for *mbela for more details on this particular form.

There is evidence that similar patterns of assimilation and subsequent semantic shift have occurred in many other cases with biota terms in the Rote-Meto languages. Thus, for instance, the usual designation of Kotos Amarasi *rinah*, a reflex of *ɗĩlah, is 'pomegranate', which is an introduced species in Timor. However, there is another term *riin fui* 'wild *rinah*' that refers to another fruit tree, probably Aegle marmelos, which appears to be native to the region.

In some cases, reflexes of a reconstructed term only refer to introduced species. Thus, for instance, all reflexes of *uas in daughter languages refer to 'jicama Pachyrhizus erosus', which is a tuber native to central America probably introduced into Southeast Asia during the 16th century. If the PRM form is valid, and did not spread by borrowing, this meaning cannot have been the original sense. Instead, *uas must have undergone semantic shift in all daughter languages. However, in the absence of further information as to what other tuber this term may have originally referred to, I have chosen to assign this meaning to the PRM term rather than a vague meaning such as 'kind of tuber'. I have done this on the assumption that any native designation was likely similar to the introduced term. Importantly, however, no conclusions as to the age of PRM should be drawn on the basis of terms that apparently refer to introduced biota in this dictionary.

1.5.3.2 Morphology

The reconstructed headwords are roots. Whenever there is evidence that such a root obligatorily took certain morphology in PRM — that is, the root probably did not occur as a bare stem — the form(s) in which it occurred are given in the morphology field. If the morphology field is filled, this means that I have no evidence that the reconstructed root occurred as a bare stem in PRM.

Thus, for instance, PRM *lasi ‘old, aged’ in Example 1.2 is given with the derivatives *ma-lasi-k and *na-ma-lasi. This means that all Rote-Meto reflexes are from either *ma-lasi-k or *na-ma-lasi. No modern-day language attests *lasi alone without morphology. Similarly, the PRM root *ŋgala ‘*Sesbania grandiflora*’ is given with the derivative *ŋga~ŋgala. This means that all reflexes in the PRM languages are inherited from reduplicated *ŋga~ŋgala. There is no evidence that *ŋgala alone occurred at the level of PRM.

When the only reflexes for a reconstruction are morphologically complex, but there is no agreement on a single complex form for PRM, the morphology field is left empty. An example is PRM *ɓua ‘gather’, for which no modern language attests a bare root: Termanu has *na-ka-bua, bua~bua*, *bu~bua-k* and *be-bua* (among other derivatives), while varieties of Meto have *na-bua*, *na-k|bua?*, *bua?~bua?* and *n-bua*. Based on these reflexes we can probably posit that PRM had derivatives *na-ka-ɓua and *ɓua~ɓua, but we cannot posit that these derivatives were the only forms of this root that occurred in PRM. Thus, the morphology field is left blank for *ɓua.

When the reconstructed root itself is a reflex of a morphologically complex PMP term, this is not indicated in the headword. An example is PRM *mea ‘red’, which is a reflex of PMP *ma-iRaq, without any indication of the earlier *ma- prefix in the PRM headword.

When the root contains frozen morphology, which is not clearly inherited from a reconstructed morphologically complex form, this is indicated in the headword with the vertical bar, which is used for frozen morphology (§1.4). An example is the PRM root *natu|n ‘hundred’, which is a reflex of PMP *sa-ŋa-Ratus (via intermediate **ŋatus) with a final consonant *n of unclear origin, which is plausibly a suffix.

When a verb is reconstructed as taking agreement prefixes in PRM, the morphology field is filled by the third person form of the prefix, either *na- or *n-. An example is *faɗa ‘say, tell’ for which all reflexes take person agreement. Thus, I propose that this verb obligatorily took agreement in PRM and give *na-faɗa in the morphology field to indicate this.

As discussed further in §2.6.1, the Rote languages have a distinction between nouns that take the nominal suffix *-k/-ʔ* in isolation and as the final member of an attributive phrase and nouns that are not eligible to take this suffix. This distinction is reconstructible to PRM. Thus, when the morphology field is filled by a form with final *-k, this means that I posit that this noun was eligible to take this suffix in the appropriate environments (not that it obligatorily occurred with this suffix in all environments). One example is *ika-k ‘fish’ in Example 1.1, for which I posit there was an NP final form *ika-k and an NP medial form *ika.

1.5.3.3 Etymology

The etymology field usually contains a reconstruction taken from the online Austronesian Comparative Dictionary (Blust and Trussel ongoing). This is the source of all reconstructions in the etymology field, unless otherwise indicated.

Most of these reconstructions are to Proto-Malayo-Polynesian (PMP). In a small number of cases no reconstruction at this level exists and I give a reconstruction from one of its putative daughter nodes: Proto-Western Malayo-Polynesian (PWMP), Proto-Central Eastern Malayo-Polynesian (PCEMP) and Proto-Central Malayo-Polynesian (PCMP). As discussed in §1.1, my use of these labels simply follows the use of others, in this case reflecting the labels given by Blust and Trussel (ongoing). For the purposes of this comparative dictionary, I am agnostic regarding all putative nodes between PMP and PRM, with the exception of Timor-Babar, for which evidence is presented in §3.5.3.1.

The transcription of PMP reconstructions follows the conventions of Blust and Trussel (ongoing), with the exception of PMP <e> [ə] which is transcribed <ə> to avoid confusion with *e [e], which occurs in putative sub-nodes of PMP, such as PCEMP.

Certain symbols within this transcription tradition are not standard IPA or have unexpected phonetic values. The symbols used for P(CE)MP reconstructions in this comparative dictionary, along with the phonetic values ascribed to them by Blust (2009a:547, 623) and Wolff (2010:31, 241) are given in Table 1.11.[12] Note particularly that the values of *z, *j, *r and *R may be unexpected.[13]

Table 1.11: Proto-Malayo-Polynesian consonants

P(CE)MP	*p	*t	*k	*q	*b	*d	*z	*j	*g	*m	*n	*ñ	*ŋ
Blust (2009a)	[p]	[t]	[k]	[q]	[b]	[d]	[ʤ]	[gʲ]	[g]	[m]	[n]	[ɲ]	[ŋ]
Wolff (2010)	[p]	[t]	[k]	[q]	[b]	[d]	[ɟ̟ʲ]/[d͡ʒ]	[g]	—	[m]	[n]	[ɲ]	[ŋ]

P(CE)MP	*h	*s	*l	*r	*R	*w	*y	*i	*e	*ə	*a	*o	*u
Blust (2009a)	[h]	[s]/[ʃ]	[l]	[ɾ]	[r]	[w]	[j]	[i]	[e]	[ə]	[a]	[o]	[u]
Wolff (2010)	[h]	[t̪͡s̪]/[t̪͡θ̪]	[l]	—	[ʁ]	[w]	[j]	[i]	—	[ə]	[a]	—	[u]

Several kinds of additional information regarding reconstructions can also occur in entries. Firstly, when the reconstruction does not come from Blust and Trussel (ongoing), the source is given after the reconstruction. Secondly, when the PMP and PRM glosses are substantially different, the gloss ascribed to the PMP form is given in quotation marks after the PMP reconstruction. Often this means that semantic shift has occurred between PMP and PRM.

Thirdly, any comments on the reconstruction are given in brackets. This includes when the reconstruction is not to PMP, but at a putative lower level such as PWMP, PCEMP or PCMP. Other comments relate to problems that the PMP reconstruction presents, or give formally and semantically similar reconstructions that have also been made.

Apart from reconstructions to PMP, I also occasionally give putative forms marked with a double asterisk and assigned to 'pre-Rote-Meto', a level before PRM, which is left deliberately vague. Such pseudo-reconstructions are given when out-comparisons (§1.5.1) indicate that the PRM form is probably inherited from a higher node, but no

12 Wolff (2010) uses different symbols for several of his proto-phonemes, and he does not accept all the proto-phonemes posited by Blust. Reconstructions from Wolff (2010) and other sources are (re)transcribed according to the conventions used by Blust and Trussel (ongoing).

13 Blust (2009a) also reconstructs PMP *c [ʧ] and *D [d], but these putative proto-phonemes do not occur in any of the reconstructions in this dictionary.

reconstruction has yet been made by another scholar. These pseudo-reconstructions should not be considered proper reconstructions, as we have only a preliminary understanding of the way in which the Rote-Meto languages relate to other language groups (§3.5) In the small number of cases in which I am confident of the level and form of my own reconstruction, I give it with a single asterisk and give *own reconstruction* as the source.

1.5.3.4 Doublets

When a form has an etymologically related counterpart, this counterpart is included preceded by *Doublet:*. In this comparative dictionary, a 'doublet' is only used for formally distinct reconstructions that are inherited from a single reconstructed etymon without being morphologically related. Thus, for instance, PRM *fai 'day, time' and *hoi 'dry in the sun' are both reflexes of PMP *waRi, but with different sound changes. Similarly, *mane 'man, male' and *mone 'man, male' are both from PMP *maRuqanay via slightly different pathways.

Formally and semantically similar reconstructions that may ultimately be cognate but cannot (yet) be identified as descending from a single etymon are not marked as doublets. Instead, such forms contain cross references to one another. This includes forms that are inherited from those given as doublets by Blust and Trussel (ongoing). An example of the latter is *lea-k 'cave' and *lua|t 'cave', which are inherited from PMP *liaŋ and *luaŋ, respectively.

1.5.3.5 Irregular sound changes

When the forms given do not show the expected correspondences (summarised in §3.2), this is recorded in one of three different fields. Based on our current understanding of the Rote-Meto languages, as well as our reconstruction of PMP, such comparisons require positing irregular sound changes (or irregular retention of certain sounds).

To allow the reader to quickly ascertain the strength of the reconstruction, the notes recording irregular sound changes are usually given in the entry for the reconstruction before all other notes. The only exception is when such notes are given for the out-comparisons, in which case they are given in the entry for the out-comparison.

Firstly, there is a field marked *'irr. from PMP'*. This field records irregular sound changes that must be posited to derive the reconstructed PRM form from a putative PMP etymon. Many such irregular sound changes only have one putative attestation.

Secondly, there is a field marked *'minority from PMP'*. This field records irregular sound changes between PMP and PRM that have multiple attestations. Thus, for instance, while the usual reflex of PMP *q is Ø in PRM, there are seven instances of PMP *q > PRM *h. These latter instances have the note: *[minority from PMP:* *q > *h*]*. See §3.5.1 for discussion of unconditioned splits between PMP and PRM.

The final field for irregular sound changes is marked *'Sporadic'*. This field records sporadic sound changes (both between PMP and PRM, and/or PRM and its daughters), such as consonant metathesis, as well as changes that are only partially complete in

certain lects, such as the raising of final *a > *e* in West Rote and Meto (§3.2.3). When a sporadic sound change only affects one language, it is often given in the entry for that language.

Under a strict neo-grammarian view in which sound change is completely regular, all comparisons that involve irregular correspondences would need to be excluded from this comparative dictionary. This is not the approach I take. Instead, I include such comparisons when the form and semantics of two morphemes are so similar that I feel uncomfortable excluding them from the dictionary. Furthermore, the inclusion of such forms means that no potentially valuable information is excluded from the dictionary.

Unfortunately this is an inherently subjective exercise. How similar is similar enough? This is the main reason I explicitly list irregular sound changes. Others may wish to exclude comparisons they judge too dissimilar, or further evidence may show that the identification of a certain cognate set is erroneous in some way. This is also why I give irregular sound changes before most other notes. The problems with such reconstructions should be front and centre in order to allow them to be subject to proper scrutiny.[14]

A second reason for explicitly listing putative irregular sound changes is that they have the potential to advance our understanding of Austronesian comparative linguistics. It may be that a higher level reconstruction is currently in error, or it may be that irregularities similar to those in PRM are found in other languages and can be adduced as evidence for a higher node.

Despite the subjectivity of this criterion, in my experience most cases are actually quite clear-cut. Thus, consider PRM *naa-k 'brother of a woman' (from PMP *ñaRa) with the following reflexes: Rote *naa-k/naa-ʔ* and Meto *naʔo/nao-f* all 'brother of a woman'. I do not think any comparativist would attribute the formal and semantic similarity between these forms to chance. However, final *o* in Meto languages is not a regular reflex of *a. The reconstruction *naa-k thus requires irregular *a > *o* in Meto.

Similarly, consider PRM *ka-nduna-k 'nest' with the following reflexes: Oenale, Dengka *nduna-ʔ,* Rikou, Oepao *runu-ʔ*, Bokai, Bilbaa *lunu-ʔ*, other Rote *ndunu-k/ndunu-ʔ* 'nest', Ro'is Amarasi *kuna|ʔ*, other Meto *ʔ|kuna|ʔ* 'nest'. Again, these forms are almost certainly related but final *a > *u* in Rote languages is not regular and *nd > *k* in Ro'is Amarasi is also irregular (all other sound correspondences are regular). Thus, these irregular sound changes are flagged.

Whenever the posited irregular sound change can be motivated, such as PRM *ka-nduna-k > Termanu *ndunu-k* 'nest' where *a > *u* is probably sporadic vowel assimilation, such explanations are given in brackets after the irregular sound change. In this case the sound change is given as: '*[irr. from PRM:* *a > *u* in nRote (sporadic assimilation)*]*'. Particularly speculative or ad-hoc explanations are further followed by a question mark '?'.

14 I welcome correction on any reconstructions in this dictionary, as well as putative connections between my reconstructions and higher levels that can be shown to be false. I only ask that if others debunk reconstructions or connections when I have explicitly flagged them as involving irregular sound changes, that they extend me the generosity of stating that I acknowledge the problems with such reconstructions.

1.5.3.6 Patterns

The pattern field is used to show patterns of correspondences among the reflexes of PRM *k and *d. Each of these proto-phonemes shows unconditioned splits in which the reflexes are not completely random. Instead, there are patterns of regularity in the correspondences. Four patterns can be identified for initial *k, six for medial *k, and two for *d in all word positions. The correspondences for each pattern are summarised in Table 1.12 for a select number of daughter languages. The full correspondences are given in §3.2.

Table 1.12: Patterns for Proto-Rote-Meto *k and *d

PRM	*k-					*-k-						*d	
env.	#_	#_		#_	#_	V_V	V_V	V_V	V_V	V_V	V_V	all	all
pattern	k-1	k-2		k-3	k-4	k-5	k-6	k-7	k-8	k-9	k-10		d-2
Dela	*k*	*ʔ*		*ʔ*	*h*	*k*	*ʔ*	*k*	*ʔ/Ø*	*ʔ*	*ʔ/Ø*	*r*	*r*
Dengka	*k*	*ʔ*		*ʔ*	*h*	*k*	*ʔ*	*k*	*ʔ/Ø*	*ʔ*	*ʔ/Ø*	*l*	*l*
Tii	*k*	*k*		*k*	Ø	*k*	*k*	*ʔ*	*ʔ*	*ʔ*	*ʔ*	*d*	*r*
Termanu	*k*	*k*		*k*	Ø	*k*	*k*	*ʔ*	*ʔ*	*ʔ*	*ʔ*	*d*	*l*
Bilbaa	*k*	*k*		*k*	*Ø/k*	*k*	*k*	*k*	*k*	*k*	*k*	*d*	*l*
Rikou	*k/ʔ/Ø*	*ʔ/Ø*		*ʔ*	Ø	*ʔ/k*	*ʔ/k*	*k/ʔ*	*ʔ/k*	*ʔ/k*	*k/ʔ*	*d*	r
		k-2a	k-2b										
Ro'is	*k*	*k*	*k*	*h*	*h*	*k*	*k/ʔ*	*ʔ*	Ø	*ʔ*	*k*	*n*	*r*
Kotos	*k*	*k*	*ʔ*	*h*	*h*	*k*	*k/ʔ*	*ʔ*	Ø	*ʔ*	*k*	*n*	*r*
Molo	*k*	*k*	*ʔ*	*h*	*h*	*k*	*k/ʔ*	*ʔ*	Ø	*ʔ*	*k*	*n*	*l*

Each of these patterns is tracked throughout this comparative dictionary in the field marked *Pattern*. Thus, for instance, a reconstruction with *k marked as '*Pattern:* k-1' indicates that this *k has reflexes according the k-1 pattern in Table 1.12, '*Pattern:* k-2' that it follows the k-2 pattern, and so on. When data from diagnostic languages is lacking, the possible patterns are given. Thus, k-2/3 indicates a *k, which could belong to either pattern k-2 or k-3.

Forms that mostly follow one of these patterns, but with one or two deviations in the reflexes, are marked with a prime symbol, with the deviations given afterwards. An example is *koro 'Rainbow Bee-eater' for which the reflexes follow pattern 1 with the exception of Meto which has *k > Ø rather than expected *k = *k*. Thus, this entry contains '*Pattern:* k-1′ (but *k > Ø in Meto; expect *k = *k*)'. Instances of *k or *d that do not follow any of these patterns are marked as k-irr. or d-irr. with the irregularities tracked in the *irr. from PRM* field, as discussed in §1.5.3.5.

1.5.3.7 Notes

Up to four different kinds of general notes are given for entries. Firstly, there is a field marked *Notes:*, which is used for all notes that do not clearly fit into any of the categories discussed. After this comes notes on irregular sound changes, discussed in more detail in §1.5.3.5.

After this is a note field flagged *Form*. This is used for notes on the form of an entry. Examples of the kinds of comments given here include giving the possible unmetathesised form(s) of a Meto word, which has only been attested metathesised, giving the gloss of part of a compound, or indicating regular sound changes that might not be immediately apparent. This field is also used for alternate forms from sources apart from the primary source used for a lect.

The note field flagged *History* is used for notes on the history of a form. This includes discussion of possible higher etyma for a PRM form, or comparisons with reconstructions that cannot straightforwardly be identified as the etymon for my reconstruction. Other miscellaneous notes on the history of a term are also given.

The note field flagged *Semantics* gives notes on the semantics of an entry. Examples of the kinds of notes given in this field include more detailed discussion of semantic shifts that certain comparisons involve, or indicating when a botanical term has only a vague gloss such as 'kind of tree' in my sources.

1.5.4 Loan distributions

In addition to the main part of this comparative dictionary, I have also included a number of loan distributions in the final part of the dictionary. That is, cognate sets that can be identified as being ultimately borrowed from another language. This section has the same organisation and structure as the main part of the dictionary, with the exception that the headwords in this section are preceded by a superscript hash $^{\#}$ indicating that these headwords are not proper reconstructions, but rather generalisations across the forms in daughter languages.

This section does not include every instance of a loan in the Rote-Meto languages, but is focused on sets that might be mistaken for cognates shared by common inheritance. Some of the cognate sets in this section may have been borrowed at the level of PRM or one of its daughter nodes, though this seems unlikely for most sets.

1.5.5 Finder lists

Two finder lists are also included in this dictionary. Firstly, there is an English finder list. The glosses by which this finder list is organised include glosses for the reconstruction, and glosses for the reflexes in daughter languages which have undergone semantic shift.

Thus, for instance, PRM *ndelat is reconstructed with the meaning 'lightning', but the Meto cluster has forms that have undergone semantic shift to include 'gun' (e.g. Ro'is Amarasi *renet* 'gun', other Meto *kenat* 'gun'). Thus *ndelat is included in the finder list after both 'lightning' and 'gun'. Doing this allows the reader to look up the etymology of a particular word from one of the modern Rote-Meto languages based on its gloss. It also makes it easier for the reader to find potential PRM cognates that may have undergone similar semantic shift in related languages.

The finder list only gives glosses for PRM reconstructions or reflexes of them. Glosses for PMP forms that are the etymon of a PRM form are not given in the finder list. Thus for instance PMP *kuhkuh ‘claw, talon, fingernail’ has undergone semantic shift to PRM *kuku ‘finger, toe’ and no reflexes of this form attest the PMP semantics. As a result, ‘fingernail’ in the finder list does not point to PRM *kuku.

In addition to the English finder list, a finder list organised by reconstructions to a node higher than PRM is also given. This finder list gives the forms in the etymology field (§1.5.3.3) followed by the PRM reflex(es) and the gloss of the PRM reconstruction. Not all PRM reconstructions occur in this finder list, as not all are known to be inherited from a higher node.

This text is taken from *Rote-Meto Comparative Dictionary*, by Owen Edwards, published 2021 by ANU Press, The Australian National University, Canberra, Australia.

doi.org/10.22459/RMCD.2021.01