Title

Ninde as a Mixed Language: A Reconstruction of Proto North Central Malekula and Proto South West Bay

Author

Schaefer, Kevin

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Ninde as a Mixed Language:

A Reconstruction of Proto North Central Malekula and Proto South West Bay

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Linguistics

by

Kevin J. Schaefer

Committee in charge:

Professor Eric W. Campbell, Chair

Professor Bernard Comrie

Professor Marianne Mithun

Professor Russell Gray

December 2023

The dissertation of Kevin J. Schaefer is approved.

Bernard Comrie

Marianne Mithun

Russell Gray

Eric W. Campbell, Committee Chair

December 2023

ACKNOWLEDGEMENTS

The work of this dissertation would not have been possible without the colleagues, friends, and family who supported me.

I'm especially grateful to Eric W. Campbell, Bernard Comrie, Marianne Mithun, and Russell Gray for every draft review and piece of advice along the way.

To Eric, for always helping me find my bearings again and enduring the times in between.

To Bernard, whose feedback has always pinpointed what others have not seen.

To Marianne Mithun, you filled my mind with the wonders of language contact.

To Russell, for keeping me in check as a transparent scientist.

Tele numüngkut xo wut timan elipndu kitax maile, nda ras pesei tele naxa‘iyene a kamu. Kaitip Kami, tele nungk kuvaxvaxuwopa‘ kitax. Elder Letpen, nungk mbolowol wut mbolowol. Tele kaiwut xoxo: Jif Alben Ruben, Jif Edwel Kaiseng, Jif Ruben Daniel, Elda Luwi Saao Baal, Nemi Serao, ma‘ Jif Plas Kali. Deacon Raombong Serao raxma‘ Kelina Serao, kamox ruwo xapvitaxtax tatai ma‘ nayei tLawa‘.

I would also like to thank dear friends and colleagues for their wisdom and their feedback on this work and my role in it: Nate Sims, Megan Lukaniec, Chris Brendel, and Adrienne Tsikewa. Most of all, tuwa‘ Caroline Crouch, watawox a kine langlang - neispospo tlepe tox sene lele neruwoxiyene a kine tele taxpoxoxo noxoine.

And to my family, who patiently encouraged me to carry on.

VITA OF KEVIN J. SCHAEFER

December 2023

EDUCATION

Bachelor of Arts in Linguistics, Macalester College, May 2013

Master of Arts in Linguistics, University of California, Santa Barbara, June 2016

Doctor of Philosophy in Linguistics, University of California, Santa Barbara, December 2023 (expected)

ABSTRACT

Ninde as a Mixed Language:

A Reconstruction of Proto North Central Malekula and Proto South West Bay

by

Kevin J. Schaefer

The Ninde language (Oceanic, Vanuatu) has been placed in a Malekula Interior subgroup (Tryon 1976) and a Western Malekula Linkage (Lynch 2016), but oral tradition identifies two parent languages (Letpen 2018). These three proposals have different consequences for which languages should be included in a comparative study, but the two monophyletic approaches treat Ninde as a divergent branch within its subgroup or linkage. Based on new data and oral history collected as part of a community-led documentation project, this dissertation provides the first diachronic study of Ninde as a mixed language.

Computer-assisted applications of the comparative method were used to assess these competing proposals. A random forest model was fit to lexical data from northern Malekula, offering an alternative to the analysis by Lynch and Brotchie (2010) that linguolabial consonants represent an areally diffused trait. Diachronic sound change models, implemented algorithmically as sequences of exceptionless innovations, aided in identifying two sources of words that together

account for most of Ninde's lexicon. These novel tools do not replace the comparative method as the standard by which phylogenetic claims are assessed, but instead, complement them.

Ninde is not a divergent member of any subgroup or linkage; rather, it inherits lexicon and grammar from two parent languages. One of these is also an ancestor of Naati and Nahavaq, which are spoken in the South West Bay region where Ninde is spoken today. Descendants of the other include Avava and Neverver, which are spoken in North Central Malekula. These findings highlight the importance and precision of local traditional knowledge, which merely corroborated by the innovative computational tools used in this dissertation.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS

1	first person
2	second person
3	third person
ABL	ablative
ALL	allative
AN	alienable noun
ANA	anaphoric
APPL	applicative
BEN	benefactive
COMPL	completive
CONT	continuative
CPZR	complementizer
DEF	definite
DIST	distal
DU	dual

EMPH	emphatic
EXCL	exclusive
FOC	focus
FRUST	frustrative
IN	inalienable noun
INCL	inclusive
INST	instrumental
IRR	irrealis
LN	locative noun
LOC	locative
NCM	North Central Malekula
NEG	negative
NMLZ	nominalizer
NSG	non-singular
NSPEC	nonspecific subject
NTR	intransitive
OBJ	object

PAN	Proto Austronesian
PART	partitive
PERF	perfect
PFV	perfective
PL	plural
PMV	post-verbal modifier
PNCM	Proto North Central Malekula
PNCV	Proto North Central Vanuatu
POc	Proto Oceanic
POSS	possessive
PROX	proximal
PSWB	Proto South West Bay
R	realis
SG	singular
SUBJ	subject
SWB	South West Bay
TOP	topic

TR	transitive
TV	transitive verb
V2	non-initial serialized verb

1 Introduction

Languages of the Republic of Vanuatu, and in particular, the island of Malekula, have been classified according to two main schemes: one that divides the languages into branching language families, and another that groups them into linkages which are united by shared innovations that can be diffused across language boundaries. Contact has certainly made for complex diachronic work, but an Austronesian linguistic heritage is a point of pride for some language communities, who stand to gain from study of language prehistory. Linkages may represent language contact that could reflect aspects of cultural identity through the ages, but cannot generate falsifiable predictions about language forms if shared innovations are proposed to be fundamentally irregular.

This dissertation simultaneously offers a proposal of two reclassified clades of the Oceanic languages of Malekula and a reconstruction of a proto language for each one: North-Central Malekula (NCM) and South West Bay (SWB). At the intersection of these two clades is Ninde, a mixed language spoken by the Mewun people – this identity of Ninde as a mixed language has been known for generations by its speakers, but never acknowledged in written scholarship. This work also provides a formal description of the details of that mixing.

In demonstrating the layered relationships of shared inheritance and contact effects, this work serves as a revision of the linkages model that has previously identified Ninde as highly divergent from its nearest relatives. This work has imposed on synchronic lexicographic data the assumption that inherited innovations will reinforce older isoglosses, but contact-induced innovation is diffused across these stacked isoglosses, rendering some of them unusable as diagnostics of specifically shared *inheritance*. Using a novel proposal for the cladistic structure of Malekula subgroups, I then reconstruct the two clades (NCM and SWB) at the shallowest time depths based on all of the available lexical data using computational methods to reconstruct large parts of the lexicons. The feasibility of reconstruction is provided as a quantification of the viability of newly proposed clades, provided as the proportion of the lexicon that can be accounted for by regular sound change.

This diachronic work is a community-led endeavor that continues a tradition of indigenous thought about Ninde in diachrony. I am one of a group of linguists and community members documenting the Ninde language synchronically. The historical work here also reconstitutes the modern reflexes expected from modeled proto-languages where these are undocumented and flags forms that diverge in unexpected ways from the known linguistic innovations. The product of this work may provide community members of any of the seven languages investigated here with materials to recover unknown vocabulary or expedite documentation of words at the fringes.

In this chapter, I begin by contextualizing this work with respect to prior phylogenetic work and my own involvement with Ninde (§1.1). I then provide an overview of the sociocultural context of Malekula and describe in some detail the languages most relevant to the present study (§1.2). Then, I describe materials and methodologies I will employ, including a discussion of how these represent a novel approach (§1.3). Finally, I provide an outline for the remainder of the dissertation (§1.4). Note that throughout the work, I will refer to the (arbitrary and colonial) names of language families as NCM and SWB; the corresponding geographical regions will be identified as North-Central Malekula and South West Bay.

1.1 Context of the work

Vanuatu is at present the most linguistically diverse nation on the planet in terms of average number of speakers per language (Lynch & Crowley 2001:xii; François 2015). Phylogenetic proposals for the island of Malekula alone are at stark odds (Tryon 1976, Lynch 2016). François et al. (2015:11) note that, beginning with Tryon, these phylogenetic approaches have generally proposed overlapping units of phylogenetic structure in some form. Ross (1988) proposed the term *linkages* to designate groups of languages that share a critical number of substantial isoglosses, but cannot be strictly delineated as a language family or subgroup to the exclusion of other languages. Lynch's more recent proposal places more or less all of the languages in linkages rather than families, and is generally agnostic about inheritance vs. contact.

Such a model raises questions about the feasibility of reconstruction. Lynch has demonstrated for Malekula languages that "some sound changes can be truly irregular"

(Lynch 2009a:15) and Lynch and Brotchie (2010) attribute the messy status of the typologically rare linguolabial consonants in north Malekula to imperfect pattern copying – in other words, a borrowed sound change that irregularly affects native lexicon beyond borrowed words that may have replaced cognates. This has implications for the comparative method, which assumes that sound change is fundamentally regular. Bowern (2006) dismisses the notion that inherited language change cannot be differentiated from the effects of language contact, supporting traditional methods to identify patterns of inherited language change even if there are borrowed cognates and back-borrowing. The central question is whether sound change is too irregular for historical reconstruction of languages shaped by prolonged periods of stasis and multilingual contact.

For reasons of personal involvement, this project is centered on the Ninde language, the language spoken by the Mewun ethnic group in South West Bay; however, Ninde's linguistic relations suggest a history of migration that may make for a history shaped by influence from divergent linguistic sources. The comparative method reveals multiple sources that contribute to the lexicon, each in nearly exceptionless ways, even if the problem of identifying what those sources were remains problematic. In the present study, I propose that the oldest layer of lexical and grammatical influence is shared with the languages traditionally spoken in North Central Malekula (NCM) – this will be considered in the present work as a language family to which Ninde belongs. Following divergence from this language family, Ninde took on a number of lexical and phonological characteristics from languages of Malekula’s vast southern interior and especially from the South West Bay region, where Ninde is spoken today.

Oral history is critical to contextualizing Ninde’s linguistic history and records information ranging from historical changes that affected classes of words, phonologically conservative word forms actively produced in traditional songs, and narratives about the origins of Ninde. The modern language is considered to be an inextricable mix of two languages, now no longer spoken, but associated with living family lineages: Small Ninde and Big Ninde. Additionally, the form of Ninde used in songs typically represents the speech of spirits in folktales; however, it is not intelligible to Ninde speakers and without the context of matrix narratives, most of it would be unrecognizable. This is a common relationship between song and text in Vanuatu, and François (2002) considers that song may be from historical or borrowed (foreign) sources. For Ninde, it is clear that the forms of song would represent a close and conservative relative, were they conceptualized as independent languages.

Numerous challenges specific to this region make historical reconstruction difficult and a convincing historical account would be of theoretical importance. For one, relatively little is known about rarer sounds found in this region (like linguolabials and bilabial trills) in diachrony because they are extremely uncommon. The lexicon is perhaps prone to diversify as a result of multiple parameters of word taboo, namely in-law registers and spiritual word taboos (see Comrie 2000 on similar taboo-motivated borrowing in Papua New Guinea).

Genetic relatedness cannot easily be disentangled from layers of contact – oral history suggests that trade and exogamy likely contributed to extraordinary rates of multilingualism. This fact of the data leads Lynch (2016) to conclude that Malekula is best characterized by "criss-crossing linkages" rather than tidy family-tree divergence. If this is the case, it may be found that a "proto-linguistic-area" (following Dixon 1997) best accounts for the forms we see today – generations of contact in a long period of multilingual stasis would have obliterated patterns of regular linguistic innovation.

There is much to gain from advancement in historical reconstruction of Malekula languages. For various reasons, the documentation of most Vanuatu languages consists of little more than minimal word lists, but there is a great potential to reconstitute (Dench 2000), or forward-reconstruct, lexical forms based on models of language history. Many language communities have faced recent shift to other indigenous languages with more numerous users, which are in turn threatened by shift to Bislama, the national language of Vanuatu that is used most for cross-language communication (Crowley 1990). With new lifeways and traditions, and massive colonial-era displacement of peoples, important domains of the lexicon of even actively spoken languages are not routinely used (see for example Dimock 2009 on reported limitations on using Nahavaq in sermons). Contexts for relevant connected speech are difficult to manufacture for "naturalistic" documentation purposes and elicitation sometimes reveals words that are in a gray area of loss – speakers are aware of the existence of now uncommon words, but the forms are at best on the tip of the tongue.

Given the state of affairs on Malekula, the broader goals of this work are to deepen understanding of this culturally and linguistically diverse region in order to serve as a resource for community efforts to document local history, reclaim disrupted traditions, and restore intergenerational transmission of the Ninde language and associated customs – part of a greater trend of heritage-reclaiming movements by Indigenous Ni-Vanuatu (Love 2016). I aim to do this by interfacing with local traditions of scholarship and knowledge production. In service of these broader goals, I put to use new and expanded methods of hypothesis generation, historical modeling, and evaluation. This project is part of a larger-scale community-university collaboration to document the Ninde language; materials produced for model evaluation in the form of reconstituted vocabularies (expected word forms) can serve as an aid for word recall in lexical elicitation contexts or potentially as a means to approximate lost lexical domains for the language communities represented in the project, albeit not without risking data integrity.

1.2 Overview of languages and materials

Malekula is ethnolinguistically diverse, with typological features unusual for Oceanic languages that are of great interest from a historical perspective (Crowley and Lynch 2001). The reconstruction work addresses in particular the distribution of typologically unusual linguistic traits on and around the island of Malekula, Vanuatu. These include rare phonemes: bilabial trills throughout the island, phonemic and allophonic; linguolabial consonants in languages of the north (Lynch 2019a); vowel devoicing in languages of the interior; a lateral fricative with complex articulation in Ninde that patterns in the phonology with the uvular fricative/trill; and the presence of trills at three places of articulation for some Ninde speakers. The lexical systems of the languages also share some locally common traits: the languages of South West Bay (SWB) generally have clan-based taboo avoidance based on lineages traced to pre-human origins (e.g., sharks, mythological hominids); all of the languages in question have an in-law register marked by lexical replacement; terms for the days of the week identify days relative to a reference day, and this sequence extends four to seven days in either direction (past and future); as is typical for Austronesian languages, local and geographic-scale spatial relations employ an absolute frame of reference with three axes. Morphosyntactic traits include suppletive number marking in verb stems in SWB; possessive constructions encoding various possessor-possessum relationships (e.g., ownership of livestock vs. possession of meat for food); multiple innovations of complement-suppression strategies (e.g., a suffix in SWB that eliminates obligatory locative complements); and several case markings for lexical nouns in SWB.

1.2.1 Sociocultural background

Archeological evidence of human settlement and extinction of megafauna (a giant crocodile and flightless birds) points to an initial settlement associated with the Lapita archeological culture approximately 3,000 years ago possibly via the Bismarck Archipelago off the eastern coast of New Guinea (Bedford & Spriggs 2014:3-5). The people associated with this initial settlement brought with them yam, taro, banana, and domesticated pigs and chickens, but made extensive use of local food sources. The archeological record suggests that the first population quickly spread throughout the Vanuatu double archipelago and diversified culturally soon afterward. Lapita people probably brought the ancestor(s) of Oceanic languages and there is some evidence to suggest that trade networks connecting early Vanuatu to other parts of Oceania broke down or became specialized soon after initial settlement. Population genetics and ancient DNA suggest that the Lapita people who settled Vanuatu were genetically East Asian Austronesians (typical of the residents of Taiwan at the time) with little admixture with the Papuan populations native to the Bismarck Archipelago (Posth et al. 2018). Papuan genes subsequently "replaced" much of the East Asian genetic ancestry with subsequent waves of migration to Vanuatu. The presence of genetically Papuan individuals can be detected in ancient DNA in southern Vanuatu 2,500 years ago, but only 500-1,000 years later on Malekula.

In more recent times, contact with Europeans ushered in a period of dramatic population change. The island interiors were depopulated, largely toward the goals of missionization. European presence in Vanuatu brought disease and people were forcibly removed as fixed-term slaves, who were often returned to South Pacific islands that were not their original homes. Vanuatu was administered as the colony of the New Hebrides by the condominium of France and the United Kingdom. On 30 July, 1980, Vanuatu won independence after years of effort. In the early years of independence, the concept of *kastom* 'custom, tradition' crystallized as a legal and social concept that devolved power to community leadership.

The legacy of outright colonialism endures today. The Bislama language developed from the local varieties of Pacific Islands Pidgin, a product of the regional slave trade. The population is still largely concentrated along the coastal regions, causing strain for the environments that sustain the nation's rapidly growing population. Nevertheless, newer social movements promote a redistribution of population density toward their historic inland homes, with associated revitalization and revival of traditional languages and practices – notably those related to conservation and sustainability.

1.2.2 Sung speech

Archaic forms in traditional song represent in some ways Ninde's closest linguistic relative. The lexicon of songs is unintelligible to speakers of Ninde, yet the language employs grammatical structures in use in spoken Ninde. Prior reconstruction work on Malekula has largely ignored the text of traditional songs, some of which is contextualized monologue of spirits in traditional tales. Across northern Vanuatu, linguists documenting languages have reported that the text of these songs is not translatable (Crowley 2006a on Avava, Crowley 2006b on Nese, Brotchie 2006 on Tirax, Dodd 2014 on V'ënen Taut). François (2002) reports that Araki speakers believe they acquired their sung language (or perhaps only the sample included in that grammar) from a neighboring people.

In Ninde storytelling, however, song represents in some ways a conservative sister language to the spoken form, but in other ways an oral-historical record of an earlier stage of the language. The two main hypotheses about the language of song discussed in Ninde-speaking communities are similarly: (1) a foreign origin and (2) an ancient origin. In fact, the two hypotheses are not mutually exclusive. Song features in most stories and at traditional dances. According to Edwel Kaising (pers. comm.), the oral history of Mewun people reports that they encountered other peoples that had familiar tunes, but their lyrics were in their own languages. In some ways, the practice of maintaining archaic lyrics may continue in some form in hymn translations, which preferentially include known archaic word forms.

As a living practice, the forms in song are subject to at least some of the sound changes affecting the spoken language, but the unique constraints of rhythm and melody favor, for example, preservation of syllable counts (and conversely disfavor vowel deletion). To the extent that these lexical forms can be considered sister to the spoken language, I reconstruct Proto Ninde as the linguistic form ancestral to both speech (including regional variation) and traditional song. The treatment of song as oral documents of language diachrony sets this project apart from traditional reconstruction work.

1.2.3 Sources of data and analysis

This reconstruction relies in large part on data I helped collect during four fieldwork trips since 2015 and ongoing collaboration with Ninde speakers using online media; but the data importantly come from newer work done mainly by other community-outsider scholars around Malekula and other parts of central and northern Vanuatu. Research in recent years has generated reference grammars of various sizes and dictionaries (both for linguists and non-linguists) (Crowley 2006a, 2006b, 2006c, 2006d, Barbour 2012, Pearce 2015, Takau 2016), augmented documentation of described languages (Dodd 2014), and revamped phylogenetic proposals for the languages of Malekula (Lynch 2016, Gray et al. forthcoming).

In this dissertation, I will cite community experts as such, using where possible references to archive locations where audio and video containing traditional knowledge can be located[1]. I can attest to practices in the Mewun community of verifying the accuracy (or at least shared traditions) of oral histories and knowledge. Some community members have been identified as meriting especial respect in matters of oral tradition, and I have made an effort to privilege those sources. In many cases, Ninde speakers provided analysis or direction for investigation.

Ultimately, the primary analysis in this dissertation is based on the computational modeling of approximately 12,109 lexical forms distributed across 9,536 cognate sets (many

[1] Regrettably, the archive at present is located in a collection crediting the names of the community-outsider researchers who created and archived only the record of these events.

of them singletons) from ten languages or varieties. Initially, lexicons were also collected for four additional languages that were ultimately only included in one chapter (§4), but their inclusion nonetheless increases the number of cognate sets/singletons to 11,040. All lexical entries were painstakingly transliterated or transcribed and assembled into cognate sets by hand beginning in 2015, many of them from contexts in reference grammars with no separate lexicon.

Many types of data are particularly problematic for reconstruction. Of the data modeled, 178 are personal names, 243 local or regional toponyms, 1,174 low-level and generic taxonomic names for flora and fauna (i.e., glossed as a 'type of' tree, pig, yam, banana, etc.), and 241 identified as borrowings. There are thousands of lexical and grammatical constructions longer than a single word. These often appear to be calques; as such, they often increase the number of lexical sets, even though the forms are not cognates per se. These were often entered as a mix of word forms and the grammatical classes that they combine with. For reconstruction, this means that the phonetic context those forms appear in is variable. Within individual entries, there are 1,436 morpheme boundaries identified (with 251 of those marked as historical boundaries only). Morpheme boundaries were not consistently marked; if every member of a cognate set shared the same morphology and form, for example, the boundary was not marked. Boundaries that were represented in the data were eventually used by the algorithm to identify the conditioning environments and morphophonological changes specific to certain word-classes. For example, some sets could only be reconstructed with success if subversions of the model implemented morphological changes: reduplication, addition of a general noun prefix *nV-*, or a general verb prefix *i-*. These were cases that posed problems for mass reconstruction using a computational approach, but offer rich data for qualitative analysis as presented in Chapters 4, 5, and 7.

Despite these many challenges, the rate at which cognate sets could successfully yield proto forms was considerably high. For the NCM group, 97% of cognate sets successfully yielded common proto forms for Avava and Neverver, or a proto form was generated (i.e., reconstructed presumably for pre-NCM and then estimated for the later NCM) based on Naman or Neve'ei. Of apparent Ninde cognates in these sets, 16.8% could not be accounted for by regular sound change, a further 61.4% produced expected Ninde reflexes not found in the lexicon (because they are unattested, are not a part of the lexicon, or are too divergent from a real, attested Ninde word). For the SWB group, 90% of cognate sets successfully yielded common proto forms for Naati and Nahavaq. Of the Ninde cognates, 38.7% of the forms could not be accounted for by regular sound change and 46.8% produced expected Ninde reflexes that were not found in the lexicon. Of proto forms that failed to converge – meaning that the set of possible proto forms included candidates generated from Naati and Nahavaq with no intersection – 16.3% provided a match for Ninde with *either* Naati or Nahavaq, but not both. Generally, this occurred in partial cognate sets or where one form had unexplained fragments/morphemes. Approximately 5% of Ninde vocabulary could be attributed to both NCM and SWB, as the same reflex was predicted for either contributing source. While the failure rate for modelling attested Ninde forms is rather high when considering descent from one proto language, this is perhaps what can be expected for a mixed language.

Qualitative analysis of Ninde's lexical systems suggests details of the process of highly innovative language mixing. The pronoun system and kinship terms have what could be described as split semantic vs. lexical inheritance from NCM and SWB languages, as lexical material from the one source appears to have been used to calque the semantic categories encoded by the other. Numerals, on the other hand, appear to be influenced by contact with yet another language group. The lexicon for calendrical terms and much of the kinship system were greatly expanded in Ninde. Irregularity in the resulting mixed language is often characterized by creativity, spurred perhaps by competing doublets and conflicting taxonomies.

1.3 Roadmap for the dissertation

The chapters that follow have the following structure. Chapter 2 is an overview of the prior work on Malekula language phylogeny and synchronic description of individual languages. Chapter 3 is a proposal for a new clade structure for Malekula languages made on the basis of a small set of nested isoglosses. Chapter 4 attempts to account for the subgroupings in the previous chapter that run counter to claims in the literature of diffusion of linguolabials across clades – it also offers a basis for excluding the far-flung northwestern languages from the same clade as Ninde. Chapter 5 characterizes the common knowledge that Ninde is a mixed language, including some sound changes that cannot be resolved in Ninde without appealing to two major lexical sources. Chapter 6 is an account of the major sound changes modeled for the two clades, as well as the contribution rates for each source language family. Chapter 7 is a discussion of lexical systems, and what can be determined by reconstructing semantic oppositions within constrained domains. Finally, Chapter 8 offers some conclusions about the feasibility of carrying out reconstruction work in the face of deep language contact, and the potential benefits of forward reconstruction for future synchronic lexicography.

2 Prior Work

This reconstruction relies on the prior work of Ni-Vanuatu laypeople and mostly foreign academic linguists and it would not be possible without a body of literature reflecting a range of documentation and reconstruction work. This work can be located in two largely independent traditions, which occasionally interact: oral history and written documentation. Within each of these traditions, there is diachronic documentation of individual languages at various snapshots, and there is work identifying phylogenetic relationships.

Oral-historical sources are consulted to balance and complement the use of one-sided European colonial documentation of indigenous Ni-Vanuatu. In keeping with Fixico (1983), I consider that “oral history cannot stand alone as a single source for research, but it can be utilized in conjunction with written documents” (p. 11). In Vanuatu, it is also true that the oral tradition records events before written records and continues to the present, overlapping the earliest written records, such that the “oral tradition accounts for the same historical events as do written documents” (p. 10) – this parallel record of events makes for a richer understanding of documentation contexts and their impacts. Narratives and story-telling have long been a vehicle for traditional knowledge and also theory (Brayboy 2021:93). Mewun oral history contains records of the missionaries, anthropologists, and linguists who have worked with the Ninde language; this has provided key insight into the nature of the available data and guidelines for continued partnerships with community members.

Written documentation, too, needs to be evaluated in light of the oral record. Many of the earliest written records of languages were written by missionaries involved in transcription of scripture (Tryon 1976), yet the languages represented by these records were contact varieties, shaped grammatically (e.g., by borrowing English *if*) and lexically (borrowing and coining religious terms) by community outsiders (Thieberger 2017). Thieberger & Ballard (2008) document the many ambiguities in the nature of the pan-Efate language developed by Daniel MacDonald at the end of the 19th century; Arthur Capell and surely others believed that this constructed "compromise" language represented an indigenous language that was actually spoken. In fact, it was a blend of often moribund language varieties from around the island, developed to avoid the effort that multilingual translations would require as individual language communities faced dramatic population decline due to introduced disease.

The idea that socio-cultural research can be beneficial for both the nation and for rural communities, and the idea that such research by outsiders, especially non-Ni-Vanuatu, can be conducted in a non-exploitative way, are ideas that are still gaining ground in Vanuatu (Regenvanu 1999:98). During the collaborative documentation project I am involved with, Caroline Crouch and I (the two outsider linguists who have spent time with the community) initially met resistance from people who cited the extensive oral history of people who have visited and extracted linguistic knowledge, only to leave nothing of lasting use for the community. More broadly, research ethics and responsibilities constitute a major area of theorization by Indigenous communities, especially when standards of conduct are defined universally for (and not by) the community stakeholders without respect for cultural diversity (Smith 2021). For this research relationship, part of respecting the community as outsiders involves sharing research findings that can be used by the community and maintaining a long-term relationship.

Fortunately, there are models for developing better research practices that account for culturally-linked responsibilities and avoiding common shortcomings. Leonard (2018) proposes community-led efforts that center native language users' experiences and conceptions of their languages. This is rather obviously unlike the approach that has been traditionally taken to research in Vanuatu. Grinevald (2003) theorizes fieldwork as a learnable skill which includes close relationships with community members as people. This involves developing research *by* communities (possibly with expertise of outside experts as needed) and requires discretion as an outsider about whether to work with a community or not – that is, whether members of the community seek the help of outside experts. Dwyer (2006:36-7) lays out general principles of conduct that may be person-specific: the 'do no harm' principle encompasses unintentional harm that could be enacted simply by publishing a name of a community member (violating their privacy) or by not publishing a name (and withholding due credit).

With a very different lens, Berez-Kroeker et al. (2018) call for discipline-wide standards in the way that linguistic data are made available and referenced in linguistic writing. These approaches together can improve both our scientific responsibilities and our responsibilities to give materials back to language communities that produced them in the first place. The process of data collection that I have participated in is unavoidably extractive and appropriative. We have taken an approach that aims to mitigate this problem, but an awareness of the issues in data provenance is necessary to evaluate this dissertation and for any future researcher in this area to recalibrate the goals of their work.

The contribution of oral history and archival written records as analysis and data are explored further in §2.1 and §2.2, respectively. In §2.3, I describe the history of diachronic work and how it informs this dissertation. In §2.4, I characterize what is known in the academic literature about each of the languages in synchrony, clarifying the often multiple names used for the same languages over time; this section also characterizes the process of data collection on Ninde, much of which remains unpublished. Finally, I conclude in §2.5 with some of the lively areas of research and debate about the languages of this region.

2.1 Oral-historical records

Indigenous linguistic scholarship on Malekula languages is maintained by respective language communities in the oral-historical record. Speakers of Ninde – and this is no doubt the case for other speech communities in Vanuatu – keep track of language history, phonetic and semantic change to individual words, and some curated cognate sets. Some of this knowledge system appears to interact with the written record, but it is crucially an independent practice that is actively and skeptically discussed in Mewun communities. The section that follows is a brief overview of, and most likely the first written account of, the kinds of information that are actively maintained by Mewun people – the traditional owners of the Ninde language. For the other oral traditions, I regrettably resort to citing grammar writers and anthropologists, treating any information reported in those works as if they are *written* contributions to individual languages by scholars of a foreign tradition.

The oral history of Ninde starts with the emergence of the language from two formerly distinct source languages and continues on to the depopulation of the once heavily populated interior. Keepers of the oral record say that Ninde is the descendant of both Big Ninde and Small Ninde, which are no longer spoken but inextricably mixed, although they can be associated with traceable patrilineal lineages (Letpen 2018a,b). With the arrival of missionization, Mewun people – and with them the Ninde language – were relocated from villages in the interior highland to coastal South West Bay, where areas of flat land were designated to be church grounds fit for settlement by converts (Saao Baal 2018). This represented abrupt changes to the cultural contexts in which Ninde is used.

The written record of Ninde starts with what remains of written Christian hymn and gospel translations, and the advent of writing is documented in the oral tradition. There is basic knowledge about the identities of missionaries, anthropologists and their activities on Malekula. Generally, only the oldest speakers can read missionary texts, which are written in a defective orthography that obscures some phonological contrasts. It is possible that those translations used forms that were already archaic at the time of writing – the bound third-person singular subject pronoun *i-* is ubiquitous in Christian texts, but absent even in ostensibly conservative traditional songs. Observations about diachronic changes in Ninde could have been one way of facilitating reading these texts, but it is also likely that an archaic form was invented from comparative and oral-historical evidence. Setting aside the directionality of influence, it is clear that the now-obsolete features that enter into conversations about Ninde reflect the state of the language used in those written sources. Furthermore, those same forms are used in contemporary religious compositions (Letpen 2018c) and are audible in the data collected recently by Shimelman et al. (2019).

The products of linguistic and anthropological work have been considerably less accessible to most Mewun people; there is nonetheless a parallel oral history of noteworthy visits by outside scholars. In some cases, what these researchers likely intended to say off-the-record is on the contrary very much on record in the oral tradition. For example, it is said that Jean-Michel Charpentier, who conducted a lexical survey of southern Malekula in the 1970s, asked Mewun people why Ninde appeared to be highly divergent from its neighbors, and whether they might have African origins (Raobong Serao, p.c.). This story is frequently repeated in conversations about Mewun history and identity. Importantly, there is a pattern of missionaries and academics extracting data of cultural value, yet producing little of lasting use to the community. This was discussed at virtually every public event my colleagues and I organized in South West Bay, and at least one community member revealed he was only interested in collaborating with us after he saw print materials in circulation. A more detailed account of anthropological literature on Malekula is listed in §2.3 in the subsections for each language.

Conversations about Ninde's external relationships appear to be relatively common, but as linguists our presence is surely a factor in the frequency of these conversations. In these discussions, there seems to be a somewhat restricted inventory of cognate sets that people draw from to exemplify relatedness between languages. The recurring examples, along with the fact that they are not always invoked by multilingual people accessing their own linguistic repertoires, suggests that this cognate identification is a tradition of sorts. The languages most often compared to Ninde tend to be those spoken to the north, notably: Nowol, Nombotkote, Nesarian, and Dixon Reef. Unfortunately, none of these are currently documented, and they are among the most threatened languages of Malekula. Edwel Kaiseng, a *kastom jif* (designated expert on pre-colonial traditions), tells of a time in the past when Mewun people were surprised to find that their neighbors performed some of the same songs, except that the text was in their own languages (p.c.).

Historical knowledge about word forms concerns both obsolete word usage and archaic phonological forms of words. New usage is generally attributed to intergenerational change – younger people allegedly use *kaiwut* to mean 'man, husband' generically, but older people reserve it for 'elderly man'. In some cases, Bislama loans have all but replaced native forms, and the forms are known and discussed by older speakers, but not actively used. These include some rather important domains: highly common words like *was* ‘wash’, *huk* ‘go fishing’, and *ale* ‘so, and then’; most numbers; time-keeping terms like weekdays and months; most inalienable, and some vocative kinship terms; and introduced items like *waea* ‘wire’ and *wotakris* ‘watercress’. More challenging for the historical linguist, were it not for the oral-historical record, are recent loans from indigenous languages that native speakers know to be loans, like *kaitmys* ‘Albino’ and *naʔaivər* ‘thatch’ from Naati.

Knowledge of unconditioned sound changes arguably challenges notions that sound change is exceptionless because it is imperceptible. A diachronic sound change that is frequently discussed is deletion of root-final /ŋa/, affecting words like *l̪ˤo(nge)* ‘go’, *nda(nge)* ‘good’, *nombo(ngo)-* ‘mouth’, and even roots in reduplicated stems like *sa(nga)sa(nge)* ‘nearby, close’. This change is unlikely to be phonologically motivated, since word-final nasals are intact, including /ŋ/, and no other final syllables are deleted with other nasal onsets, like /ma/ or /na/. The deletion of these segments could be motivated by reanalysis, since *=nge* is a third-person object enclitic on the verb phrase and an anaphoric demonstrative enclitic on the noun phrase. It is crucial to understanding word-final vowel phenomena, since the vowels exposed by this deletion to the word edge resist word-final raising of the low vowels (/a/ → [e] or [o]). The nominalizing suffix *-iyene* /ijene/, which forms nouns from verbs, is reduced along with any final /n/ or /ʔ/ in the verb stem to [jne]:

Verb form	Nominalized form		Surface realization
/**ʁaʔ**/ ‘clear land, work’	/na-**ʁaʔ**-ijene/	→	[na**ʁa**<u>jne</u>] ‘garden, labor’
/(ka)**ʔan**/ ‘eat’	/na-**ʔan**-ijene/	→	[na**ʔa**<u>jne</u>] ‘food’
/nuwuta **ʁan**/ ‘become day’	/nuwuta-**ʁan**-ijene/	→	[nuta**ʁa**<u>jne</u>] ‘day’

The word [nuta**ʁa**<u>jne</u>] ‘day’, through reanalysis of its reduced form of /nuwuta-**ʁan**-ijene/, is the source of the temporal subordinator [t(a)ʁaj] ‘when’. These are some of the changes that people are very conscious of, even though they are complete and predate subsequent innovations on necessarily modified, already reduced forms.

Additionally, the linguistic forms used in song are suspected by some to represent an old form of the Ninde language. Regardless of their origin, the practice of transmitting song is a linguistic one, and the forms are clearly related in some way to spoken Ninde. All that is known is that the text of songs represents the speech of *kapat*, sometimes called devils. These figures are spirits active in the present day and the ancestral founders of *nasara*, traditional dancing grounds that are associated with human lineages, even though they are now seldom used for ceremony.

The oral tradition is subject to skepticism and few people are ratified to speak with authority about historical topics. When invoked, ideas like these are often asserted with a virtually neutral truth value, followed by evidence for and against the notion. Reputable individuals are generally the oldest men who attained high ranks in the age-graded *nalawan* society, now no longer practiced. Letpen (who only uses one name) is esteemed for her vast cultural knowledge and conservative manner of speaking; this is often attributed to her adolescence, which she spent largely caring for her father while her peers were in school. When a story is told or a claim made about the past by someone else, the provenance can at times be questioned. In this study, I draw hypotheses from oral traditions that have been shared or explained to me, and engage with the oral-historical record as a body of scholarship.

2.2 Written archival records

Generally, archival resources of various types are available for Malekula, but the vast majority are religious texts translated into native languages by uncredited individuals for missionary use and wordlists collected from broad surveys of several languages. The exact resources available are included in more detail in §2.3 on lexical surveys and in §2.4 in the coverage of individual languages.

A number of anthropological sources contain isolated words. Layard (1928) covers mainly ceremonial rights throughout Malekula, and Deacon and Wedgwood (1934a) include words from material culture, spiritual life, kinship, and social organization, also from languages around the island. Larcom (1984) more recently worked in South West Bay (where Ninde is spoken) to describe the impact of missionization on the region, but she also includes a small number of words. More recent data is found in the work of Charpentier (1982), who collected a wordlist of Ninde in the 1970's in his survey of southern Malekula. Although those data appear to confound front vowels followed by labialized stops with front rounded vowels (his transcriptions of sequences /m^{w}i/ and /my/ both appear as <mü>, neutralizing both the contrasts of /m^w/ vs. /m/ and /i/ vs. /y/, which continue to be contrastive), the transcriptions contain greater phonetic detail and a variety of cultural terms not contained in the biblical translations. Since this work was a series of word lists, it also provides valuable, if flawed, lexical forms from geographically close sister languages.

2.3 Prior diachronic work on Malekula languages

All of the indigenous languages of Vanuatu have long been identified as Austronesian languages within the Oceanic family (Ray 1986). The linguistic evidence of language in diachrony has long been considered alongside observable changes in material culture and population genetics throughout Island Southeast Asia and the Pacific. The Proto-Austronesian expansion is correlated with the spread of an archeological culture that introduced plants and animals to much of Oceania and perhaps drove some native species to extinction. It is difficult to determine whether this cultural and linguistic group expanded out of Taiwan or nearby mainland Fujian: this is in part because any linguistic relatives on the mainland would have been replaced by Min varieties and because some sources of the Pre-Austronesian population likely arrived on Taiwan via Fujian in the first place (see Jiao 2021 for an overview of points in favor of this view).

Genetic evidence beyond human DNA supports a Taiwanese origin. Helicobacter pylori (gut bacteria) found in human Austronesian-speaking hosts throughout the Pacific descend from clades found in a subset of indigenous Taiwanese people (Moodley 2009). The paper mulberry (*Broussonetia papyrifera*) that was introduced throughout greater Austronesia is descended from the southern Taiwanese lineages and not from the lineages introduced to Northern Taiwan from Fujian (Matisoo-Smith 2015).

Several wide-reaching studies of the languages of Malekula have been carried out and served as the basis for historical works. This dissertation has narrow focus, but will draw on work with broader scope. Tryon (1976) used lexicostatistics to obtain phylogenetic subgroupings for many Vanuatu languages. His classification remained the sole authoritative source for decades, but individual grammar writers and comparative linguists have rejected or clarified the specific inclusion or exclusion of certain languages (e.g., Crowley 2006c).

Sound change reconstruction work by Clark (2009) and Lynch (2016) has generated proto forms for comparison. The former provides a reconstruction of Proto Northern-Central Vanuatu (PNCV). The latter is an attempt to classify and subgroup Malekula languages, with an occasional reconstruction of Proto Malekula words (without strongly asserting a case for Proto Malekula).

Two sources – Lynch and Brotchie (2010) on the ordering of the innovation of labiodental obstruents relative to low vowel deletion in the languages of north Malekula, alongside Clark's (2009) higher-order reconstructions – serve as a starting point for Lynch's (2016) phylogenetic grouping of Malekula languages into three linkages. Of those linkages, he determines that Ninde belongs to the Western Malekula linkage. He does not ultimately claim a single Malekula grouping, but makes additional observations of sound changes toward a possible earlier Proto Malekula.

Clark's (2009) reconstruction of PNCV samples then-available lexicons from 15 Northern and Central languages, of which Ninde was one[2]. He concludes that the reconstructibility of that particular grouping of languages provides some evidence for a unified linguistic community on the basis of shared innovations found predominantly in these languages; however, no sound changes appear to pertain to all NCV languages and only to NCV languages. Clark considers the innovations shared with other Oceanic languages to be crosslinguistically common and not diagnostic. There are a number of doublets in his reconstruction, but Clark addresses them and dismisses potential concern about them, echoing an observation by Keesing and Fifi'i (1969) that Oceanic languages seem to tolerate

[2] The then-available lexical forms for Ninde did not reliably mark glottal stops, which are now known to be phonemic in Ninde. This led to an erroneous claim in that work that glottal stop was retained in only one NCV language.

high degrees of synonymy in the lexicon (for purposes of taboo avoidance and in-law registers). His suggestion is that daughter languages tended to lose one of many synonymous (and often derivationally or historically related) forms inherited from the proto language, and that the reconstruction of pairs of semantically equivalent proto words need not raise concern about the unity of the proto language.

Proto Oceanic (POc) is the next reliable level of reconstruction and several people have worked extensively on reconstruction. Among lexical reconstruction works, Ross, Pawley, and Osmond (1998, 2007, 2008, 2011, 2016) produced a five-volume series of lexical reconstructions of POc organized by domain. Analyses of specific sound changes have also greatly advanced the reconstruction of POc. That work has focused on specific reflexes or natural classes (Lynch 2002 on labiovelars, Lynch 2003 on low vowels, Lynch & Crowley 2003 on reflexes of glottal or dorsal stop *q, François 2011 on alveolar taps and trills). Lynch (2000a) also reconstructs patterns of stress assignment of POc.

Many have also addressed aspects of POc morphosyntactic structure. These include both the status of syntactic categories in the proto language (Crowley 1985 on nouns, Ross 1998 on adjectives), and diachronic questions about larger structures (Ross 2004 on the general morphosyntactic typology of Oceanic languages, François 2004 on spatial terms, Næss 2013 on the development of POc transitivity marking from PAn voice systems). Crowley (2002:160-168), after surveying serial verb constructions across Oceanic languages, argues for their existence in POc and begins to reconstruct them.

As for Proto Austronesian, there is plenty of work integrating data from many branches, including Oceanic. Dempwolff (1938, 1969a, 1969b) offers the earliest attempts to characterize the family using the comparative method. Further and more recent reconstruction work is collected and reproduced in online databases, such as the Austronesian Comparative Dictionary (Blust & Trussel 2013) and Austronesian Basic Vocabulary Database (Greenhill, Blust, & Gray 2008). Those sources have incorporated updated hypotheses about the internal structure of the family and reconstructed forms for intermediate proto languages representing subgroups in the phylogeny.

In addition to these reconstructions of inheritance, contact with Papuan languages has been proposed a number of times (among others: Ray 1926, Capell 1954, Lynch 1981, and Wurm 1982). Donohue and Denham (2008) include a summary of those claims, which propose among other influences a Papuan source in the development of contrastively rounded (or labialized) bilabials, linguolabials, and bilabial trills; quinary number systems and SVO word order. The development of quinary number suggests that at least some of the contact between Austronesian and Papuan languages occurred after the divergence of PNCV, since Clark reconstructs a decimal system, even though many of the languages have the quinary decimal (also called imperfect decimal where only values 6-9 are derived with a base of five) system.

Lynch's work on Malekula subgrouping serves as a starting point for comparison and reconstruction (Lynch & Brotchie 2010, Lynch 2016). These works propose a number of ordered sound changes and borrowing relationships, but I hypothesize a number of alternatives. Although advancements in the documentation of Malekula languages have been ongoing since most of this work was published, extensive work on higher-order proto languages and prehistory can nonetheless contextualize or supplement questions that appear to be inexplicable from bottom-up reconstruction alone.

2.4 Prior synchronic work on Malekula languages

Grammars and dictionaries are usually a result of long-term collaborations involving a linguist and members of a language community working together to document a language. Even where community members are not directly credited, I assume that the finished product generally represents to some degree native speakers' knowledge, insights, priorities, and analyses.

The subsections that follow provide an introduction to the languages from which data were drawn for the purposes of reconstruction only. A broader set of languages was considered in determining which languages to include in the reconstruction; those are listed in §3. Much prior synchronic work has been collected in the service of comparative diachronic study in the form of wordlists (§2.4.1), but some of the work is typological in nature (§2.4.2), and other studies center on the grammar of a single language (§2.4.3).

2.4.1 Lexical Surveys

In more recent times, much of the data available for historical work is in the form of multi-language word lists. Codrington (1885) surveyed 26 language varieties, and Capell (1962) also surveyed languages from throughout the archipelago, substantially increasing the literature on Vanuatu languages. The first large-scale list was published by Tryon (1976), who surveyed 177 languages from throughout Vanuatu and carried out lexicostatistics. Each of these scholars was interested in the structure of languages and worked out grammatical similarities across languages.

Charpentier (1982) represents a break from that tradition; possibly due to his untimely death, he only published a lexical survey. His lexical survey sampled over 1200 word sets (containing blanks for unknowns, duplicates for polysemous words, and a number of circumlocutions) in 19 languages of southern Malekula. Although it contains a number of likely errors and organizes lexical items by semantic equivalence (rather than in cognate sets), this work greatly advanced the study of linguistic relationships in the south of the island. More recently, Shimelman et al. (2019) have undertaken a large-scale survey of over 30 languages from the whole of Malekula and more from nearby Pentecost, Maewo, and Ambae islands, sampling the lexicons of some languages that were fully undescribed in linguistic scholarship; these were collected in order to perform Bayesian phylogenetic analysis of linguistic divergences on the island.

2.4.2 Areal typology

In this section, I discuss some of the typological properties typically *shared* by Malekula languages and data available from crosslinguistic lexical surveys. Following a general profile, I list languages that will have greater relevance to Ninde as they are implicated in genetic relationships proposed by Tryon (1976) and Lynch (2016). (See Table 1 for a Ninde-centric schematization of those two subgroupings, ordered by approximate relative distance of precolonial language communities from South West Bay.) For each language, I note additional reference grammars and lexicographic works that have been used.

The languages of Malekula are characterized by a number of shared traits, and many of the unique innovations can be demonstrably inherited from these. In the phonology, most of the languages have plain and prenasalized stops (as well as prenasalized and plain rhotic phonemes) but no voicing contrasts in the inventory; a labialized bilabial series that contrasts with plain (sometimes analyzed as velarized) bilabial consonants; and allophonic bilabial trills. In the morphology, most of the languages have bound possessor suffixes for most of the person and number combinations; fused *na- substantive prefixes in noun stems; a construct suffix *-n* (or *-ne*) used in the absence of a bound possessive form, which is homophonous with third-person possession; and nouns are generally divided between three types of alienability classes, with the languages of the north futher classifying two or three types of gastronomical types of possession[3]. Syntactically, the languages generally have basic

[3] These are forms differentiating, for example, a live pig owned as livestock vs. pork possessed for

SVO order; few prepositions that do not also serve as a verb or noun; multiple verb roots in a clause, via compounding and/or serial verb constructions; and tail-head linkage in storytelling.

2.4.3 The communities and their languages

Generally, through trade, exogamy, mixed heritage, and migration, older Ni-Vanuatu people on Malekula are highly multilingual. Malekula is in turn embedded in the society of Vanuatu, which is a nation with extraordinary linguistic diversity and density, with over a hundred indigenous Austronesian languages spoken by about 272,459 people (Crowley 2000, Vanuatu Census Data 2016). Deeper intra-national ties have brought people from more distant islands to villages of Malekula in recent generations, making for more linguistic diversity and a greater prominence of Bislama, the national language.

Virtually every grammar of a Malekula language makes reference to an oral history that locates the traditional homelands of the languages in the interior of the island. People relocated for churches, schools, copra plantations, and in many cases, formed larger settlements with speakers of other languages as populations collapsed in the aftermath of European disease and enslavement (see for example Crowley 2006b pp. 4-11 for a particularly thorough treatment of factors contributing to language loss in Naman). In each

consumption. Other categories may include possession for beverages and for chewable items.

subsection below, I provide the names of the ethnic group and their villages and/or territory, alternative names of the languages, approximate number of speakers, people who provided the data (if provided in the literature), grammar writers, and typologically uncommon characteristics of the language's structure and use.

The languages of Malekula and their proposed relation are tabulated by relative aerial distance from South West Bay (Ninde) in Table 1. The details of the proposal put forth by this dissertation can be found in §3, the status of Ninde as a mixed language is addressed in §4, and the alternatives to Lynch's Malekula Peripheral languages are presented in §5.

Language	**Malekula Interior (Tryon 1976)**	**Western Malekula Linkage (Lynch 2016)**	**This proposal**
Ninde	A	Peripheral > Ninde	North Central Malekula + South West Bay
Naati		Peripheral > Naati	South West Bay
Aveteian	B	Peripheral > Southwestern	?
Lendemboi	B	Peripheral > Southwestern	*
Nahavaq		Peripheral > Southwestern	South West Bay

Language	Malekula Interior (Tryon 1976)	Western Malekula Linkage (Lynch 2016)	This proposal
Navwien	?	Peripheral > Southwestern	
Novol			
Nasarian	B	Peripheral > Southwestern	*
Na’ahai		Peripheral > Southwestern	
Nombotkote			
Nivat			
Axamb			
Nasvang			
Sörsörian	?		*
Avava	C	Peripheral > Southwestern	North-Central Malekula
Neve’ei	C	Central	North-Central Malekula
Naman	C	Central	North-Central Malekula
Neverver	C	Central	North-Central Malekula
Tape		Peripheral > Northwestern	
V’ënen Taut	C	Peripheral > Northwestern	

Language	Malekula Interior (Tryon 1976)	Western Malekula Linkage (Lynch 2016)	This proposal
Tirax	C	Peripheral > Northwestern	
Nese			

Table 1: Phylogenetic proposals for Ninde and related languages A, B, and C represent unnamed branches of Tryon (1976)'s Proto Malekula Interior (PMI). Asterisks in my proposal represent languages with significant contact. Question marks represent languages addressed but not placed in a clade.

Ninde (NCM and SWB)

Ninde is the language of Mewun people and is additionally spoken by familial relatives and regional neighbors of Mewun, but many children and especially Mewun living in Port Vila and abroad have shifted to Bislama or English. Several language groups spoken by people relocated to the coasts (notably Nasarian, Lendamboi, and Nowol speakers) have assimilated into Ninde-speaking social contexts and reflect varying degrees of shift to Ninde. Data for Ninde have come primarily from field notes and recordings. The data are primarily from documentation work with Elder Letpen, Chief Kaitipbuas (a.k.a. Luwi Saao Baal), Chief Alpen Reuben, Chief Edwel Kaiseng, Chief Robben, Elder David Kaiar, Chief Raobong Serao, Nemi Serao, Kelina Serao, and Elder Shemson Thompson, among many others.

Some forms have been recorded in Charpentier and in the Sound Comparisons project (Shimelman et al. 2019). Oral history of the language describes two forms of Ninde, an interior Small Ninde and a coastal Big Ninde, which became a relatively homogenous lect with shared heritage, each making inextricable contributions to the lexicon (Letpen, Raobong Serao, p.c.). Common examples provided to illustrate the difference between the ancestral dialects appear only to involve grammatical differences, but it is not apparent what each dialect contributed to the contemporary language. In one example from Raobong Serao, Small Ninde *tu-saxa* vs. Big Ninde *ti-taxse* 'drop off (a person)', the former is an extended sense from the usage of the same verb with an inanimate object, and the latter appears to be a causative form of *taxse* 'stand', either with *ti* 'tell' or partial reduplication functioning as a causative.

Among the written records are the four New Testament gospels in translation (clearly in turn from an English translation), approximately 20 hymns, and wordlists from missionaries' literacy materials, all dating from the 1890's. Though much information is absent (both lateral phonemes are written with <L> and glottal stops are absent), the words have overall more syllables, virtually only simple onsets (whereas today, complex onsets are common), and some apparent differences in vowel quality. Hymns are a valuable source of lexical data as well – they tend to be verbatim translations of the text of English hymns, with the lyrics of verses sometimes stretched across two verses.

Naati (SWB)

Crowley (1998), working with Aiar Rantes, provides a grammar sketch and collected a 750-item vocabulary of Naati. The language is spoken by a handful of people in Windua, where Ninde is also traditionally spoken. Crowley and Lynch (2001:69) report that this language was thought to be no longer spoken by the 1980's, but Crowley found that there were a few speakers in more recent years.

In addition to Crowley's (1998) grammar, there is also audio for the word list at the Vanuatu Voices project (Shimelman et al. 2019). Naati's phoneme inventory appears to be characterized by a partial merger of POc *s before front vowels with *j (probably a prenasalized coronal affricate or fricative) followed by a change of *s > h everywhere else, and the development of two front rounded vowels /y ø/ to the inherited five-vowel system (Lynch 2016). It also has allophonic bilabial trills before /u/. One of the few Ninde words identified as an acceptable borrowing is most likely from Naati: *naʔaivər* 'thatch'[4]. This suggests layered contact between Naati and Ninde.

[4] Letpen identified the source language by an ambiguous toponym, but Crowley (1998) records *naʔaivør* for Naati and Dimock (2006) has quite a different form for Nahavaq: *nuwur-jet* 'thatch, (lit. sago leaf)'.

Nahavaq (SWB)

Dimock (2009) provides a grammatical description of Nahavaq, the language of Sinesip people. The data were collected with many different speakers – more than 50 are implicated in the corpus metadata. Some of the most prolific individual contributors are Kalmar Jacobus, Alick Rongsin, Louise Aisig Mbuwas, Emile Massing Nambuas, and Aisul Manwei.

Many Ninde speakers in the southern part of South West Bay (mostly Windua) are bilingual speakers of Nahavaq, or Nahavaq is a heritage language for them. People have reported that there is some degree of lexical borrowing in Windua, particularly in kinship terms. Nahavaq is characterized by a merger of *dr (some prenasalized rhotic, or a prenasalized coronal stop with rhotic release) with *d; a realization of POc *j > h before front vowels and POc *s > h everywhere except before front vowels, followed by POc *j > s elsewhere[5]. It is likely that Nahavaq was in a contact relationship with Ninde, since Sinesip and Mewun people once shared a distinct style of sand drawing (Deacon & Wedgwood 1934b).

In addition to Dimock's grammar and accompanying corpus and dictionary, there is also audio for a wordlist by the Vanuatu Voices project (Shimelman et al. 2019), and a 1905 translation of some New Testament books and an undated primer/reader, and 11 glossed texts provided by Deacon and Wedgwood (1934a).

[5] Most likely, *j and *s merged in the direction of *j before front vowels, and in the direction of *s elsewhere before *s > h.

Lendamboi

At least one variety of Lendamboi (or Lendemboi) is recorded in Charpentier (1982)'s atlas of southern Malekula languages. Crowley and Lynch (2001) summarize reports of four mutually intelligible varieties traditionally spoken in the Lendamboi region: Ayiauleián$_a$, Mbotkót$_e$, Natanggan (which was moribund at the time), and Nioleien (Repanbitipmbangir and Niolenien/Repanbitip in Tryon 1979 and 1996, respectively). Charpentier describes the subscript vowels as post-tonic vowels that are "so destressed as to become unrecognizable (1982:63)." Recordings we made with a man from Lendamboi, David Kaiar, suggest that these could be voiceless vowels. In his speech, word-final vowels occasionally appear fully voiceless on spectrograms; there is also devoicing word-medially before voiceless consonants, and vowel-initial words have a non-phonemic prothetic [h].

We have approximately 180 minutes of elicited word forms and basic phrases (and about five minutes of connected speech) recorded from Kaiar, who identifies as the last speaker of a language or variety he calls Nevitangiene and wished to record words and phrases with us. His variety was spoken in the region of Lendamboi, but it is not yet clear why he does not identify his language as Lendamboi. At any rate, Kaiar seems to identify his variety as distinct, and his self-identification as the sole remaining speaker suggests, at the very least, that he has knowledge of otherwise reduced domains of lexical knowledge. Nonetheless, I will refer to his variety as Nevitangiene (Lendamboi).

The Vanuatu Voices archive (Shimelman et al. 2019) has the following varieties of Lendamboi represented: Natingatlang, Aingelemolesa, Newotenyene, Nevatanyene. None of these varieties are transcribed with voiceless vowels or prothetic [h] (although other languages in the database have this prothetic element transcribed); the audio does not suggest either of these exist in the speech of those speakers; and none of the final vowels seem meaningfully "destressed". This may mean that the phenomenon attested in historical sources is only present in Kaiar's variety. His speech therefore contributes a critical piece of the historical picture: the phonetics of post-tonic vowels, but also the pharyngealization of *l in Ninde (§5.4).

Avava (NCM)

Crowley and Lynch (2006a) provide a grammatical description with two sample texts for Avava. Unusually for Malekula, Avava features contrastive vowel length, optional diphthongization of word-final vowels via the addition of a short /i/, and paucal pronoun forms. Additionally, there is free variation among older speakers of noun-initial /nV/ alternating with a reduced or reanalyzed form /V/ (both apparently inherited from Proto North Central Vanuatu *na), and among the younger population between bare nouns and nouns with the reduced initial vowel. Crowley's grammar has one of the most complex accounts of the range of nominal compounding of the grammars of Malekula languages.

It is probable that there is an oral-historical record of a lect akin to Ninde's traditional song sampled in texts collected by Crowley and Lynch (2006a). A repetitive portion of one text (pp. 205-6) is italicized, with a note reading "Untranslatable" in place of a free translation. In the narrative context, the text is a bird's speech, suggesting that this portion is also performed as song. The untranslatable and sung nature of those forms resemble the traditional song forms embedded in narratives of Ninde (and other languages).

Neverver (NCM)

Neverver (sometimes spelled Nevwervwer to reflect pronunciation of the labialized bilabial fricatives in the Neve'ei language allonym) is described by Barbour (2012). It is traditionally spoken in the interior of the Dog's Neck region according to oral tradition, but is today spoken in the villages of Limap and Lingarakh on the eastern coast of that region. Barbour's reference grammar was written with data from Jacob Naus (who developed the community orthography and requested a linguist), James Bangsukh, Lerakhsil Moti, and transcription by Emma Vatdal, Nellie Vatdal, and Peter Vatdal, among many other community collaborators.

Neverver is set apart from all other Malekula languages for two major idiosyncrasies in its phonology and morphosyntax. It is the only language to have geminate consonants (voiceless stops and some of the continuants). In terms of grammar, it has innovated gendered singular third-person pronouns for human referents alongside the gender-neutral pronoun. Like Avava, prenasalized bilabial trills are phonemic in Neverver.

Neve'ei (NCM)

Neve'ei is spoken in the village of Vinmavis in the northwest of Malekula, but the traditional area of this language extends much further inland and nearly reaches the eastern coast (Musgrave 2007). The language has been documented in the form of individual words in the 1920's, when Deacon traveled to the interior of this region before Christianization and associated coastal removal. At the time of Musgrave's grammar, it was estimated that there were 750 speakers of Neve'ei. Crowley and Lynch (2006a) also includes isolated words as translations of Avava words (mostly botanical terms). Since this grammar was developed with data from a young student in New Zealand, the semantic domains represented contain less traditional knowledge than comparable reference grammars of other Malekula languages.

Comparatively little sets Neve'ei apart from its relatives; however, it has been challenging for diachronic study. Neve'ei is areally unique for having glottal stops, but these are shared with Ninde and more distant relatives. Like Avava, the singular third-person *i- has become fused to monosyllabic verb stems for all subject persons and numbers. Neve'ei is apparently the only language that can have objects dislocated or extracted and preposed from the typical SVO pattern.

Naman (NCM)

Naman is spoken in a large area called Lëngalëng in the Dog's Neck, stretching from coast to coast (Crowley and Lynch 2006a). The primary village is Vinmavis, but Naman speakers live in villages throughout the area, including the traditional area where Neverver was spoken – Neverver is now the dominant language of heritage Naman speakers.

Crowley and Lynch (2006a) provide a short reference grammar of the Naman language – this grammar was edited by John Lynch. The acknowledgments include the sole mention of the project collaborators in Vinmavis, but it is unclear what their role was: "Roy Jeremiah, Kalteri Jeremiah, Setoko Andy, Manu Simeon, Jeff Simeon, Andrew Tony, Simon Johneti, Edmond Johneti, Philip David, Pierre Harry, Daniel Harry, Thomson, Williamson, Kalmatak, Kalorongo, John Morrison, Sano, Marsden, Dudley and Aleris Fathley, Kevin Jack, Dansen Uren, Turan Uren, Malon, Indian and Ian" (Crowley & Lynch 2006b:xvi). Lynch (2019a) considers the diachronic phonology of Naman.

Typologically unique characteristics of Naman are generally found in the possessive constructions. Like most Malekula languages, Naman has alienable and inalienable possession and two constructions reflecting closeness of possession. Like Ninde, Naman has two constructions for close possession: one formed with bound possessor suffixes for personal pronouns and another construction that uses the singular third-person bound suffix -*n* and an independent pronoun of any person and number. Most languages generally have the latter. Moreover, Naman allows for the possessor to both precede and follow the possessum when the independent pronoun is used; lexical possessors are pronominalized in the position following the possessum. Finally, though Naman lacks the edible ('X for Y to eat') and potable ('X for Y to drink') possessors found in the languages to its north, it has an edible benefactive preposition *(de)nakha-* that can add the role of an eater to intransitive verbs like *tabëkh* 'cook food' (Crowley 2006b:75).

2.4.4 Community-led approaches

The history of engagement between linguistic scholarship (plus related disciplines) on one side and the members of language communities on the other is a complicated one. There is hope for the future in the growing number of Ni-Vanuatu linguists and specialists who work on their own or other Vanuatu languages; however, it is necessary to outline how the approach taken for this dissertation is an attempt to break with the troubling parts of that tradition. This concerns both the documentary fieldwork that was used to collect Ninde data and the approach to historical reconstruction.

The nature of the data that were recorded for this project is largely determined by community members. This includes the people who are represented and the topics of archived events. As a result, the corpus represents traditional knowledge in many domains, including language history, new religious traditions, and informal domains of language use. Most of the efforts in this project were devoted to community goals; the outcomes of the project are a community orthography and a fledgling literacy program, learning materials that are in circulation in the community, a hymnbook, and other media that support learning to speak and write Ninde.

Resources collected in more recent decades have disproportionately sampled word lists for diachronic aims, but depth has been sacrificed in recent times in favor of breadth. Diachronic work can support deep synchronic work, but the words and constructions that are of interest for community goals like language revitalization and recovery of traditional ecological knowledge extend beyond basic word lists. This dissertation has sampled virtually *all* of the known vocabulary for the languages that are its primary focus. In addition to modeling language history, I have also developed the tools for estimating undocumented corners of modern lexicons – these will be made available to language communities as possible.

2.5 Topics in Malekula languages

There are many issues that are of particular interest on Malekula. This section outlines some unresolved issues and areas of ongoing research. Many of these are topics that are of broad interest, but with high local relevance.

What are the nature and extent of contact-induced Papuan influence on Malekula languages? Specific features, like the quinary system and phonology, have been addressed, but little has been found in terms of lexical contributions. The linkage model of linguistic relationships has been framed in terms of mutual influence between varieties in a continuum, and less in terms of contact with the languages that Papuan people would have brought to Vanuatu in the period of population replacement.

Was there a Proto Malekula? Since Tryon (1972), the idea that Northern Malekula languages are more closely related to those of Espiritu Santo, a large island across a straight to the north, has largely been abandoned. Clark (2009) reaches no firm conclusions on the viability of a single proto language for the whole of Malekula, but he raises the issue in his reconstruction of Proto North Central Vanuatu, which includes both islands. Lynch's prolific works assume unity in the Malekula languages and contact with the languages of nearby Espiritu Santo island (i.e., diffusion of the linguolabials by contact), but no serious consideration has been given to inter-insular migration after divergence.

Can the nature of language change reveal the nature of human migrations? Many have proposed that Pulse-Pause models of settlement are supported by Oceanic phylogenies (Gray et al. 2009). Lynch (2019) lends support to this notion, characterizing the diachrony at the smallest scale in Vanuatu in terms of language-specific innovations, some of which crossed language boundaries. Population genetics and archaelogy correspondingly suggest that populations quickly became differentiated and localized soon after settlement in Vanuatu; the Lapita trade networks that kept far-flung populations connected seem to have collapsed within three centuries (Posth et al. 2018). These approaches suggest that diachronic typology could shed light on cultural history.

Indigenous people in Malekula are looking to tradition for its potential to mitigate environmental strain. McCarter et al. (2014) document some of the challenges specifically in the Mewun community and elsewhere in maintaining traditional ecological knowledge in the face of formal education. Love (2016) characterizes a popular movement supporting the reestablishment of depopulated villages. These efforts go hand-in-hand with language revitalization, not only because language has offered the tools for instruction of valuable skills in ecological maintenance, but for the value of ecological taxonomies and timekeeping practices that were integral to tracking and modulating resource consumption and health. These efforts parallel other projects that have reevaluated conceptions of indigenous practices and historical knowledge as beneficial to the study and maintenance of ecological health (c.f. Kimmerer 2013).

3 A New Approach to Malekula Linguistic Phylogeny

3.1. Introduction

The linguistic situation on the island of Malekula (also spelled Malakula), in the Republic of Vanuatu, has been characterized as one which can only be modeled in terms of genetic and areal *linkages* – groups of languages that share features in a family-resemblance fashion (Lynch 2016). Unlike language families, they are not intended to be *defined by* inclusion in innovations, but are characterized by incomplete participation in non-inherited sound changes. Lynch (2009:14) rejects factors like homophony avoidance and analogy as satisfying accounts of irregular sound change, arguing instead that a high degree of contact makes sound change in such a linguistic landscape fundamentally irregular. He has repeatedly claimed that irregular sound changes are at play in: the development of linguolabials (2005), changes affecting post-velar consonants (2009), and word-final vowel deletion (2014), among many other, more specific changes. Lynch's classification has proven to be a great improvement over the one put forward by Tryon (1976), but linkages offer limited insight into language prehistory without some traceable relation between modern languages and their ancestors.

In order to reassess the issues of competing phylogenetic subgroupings analyzed for Malekula (discussed in §2), I used data almost exclusively from Shimelman et al. (2019) to find evidence of innovations that are shared between languages in embedded and generally contiguous isogloss patterns – something that would indicate that some changes can be traced diachronically via inheritance. Following Bowern (2006), I consider any language change to be a matter of innovation and diffusion throughout a language community; whether the source of the innovation is internal or external to the community, it alters languages that have some semblance of intergenerational continuity, even if that is complex. Blust (1996:153) – like Lynch – allows for some exceptions in phonetic change, but does not reject the comparative method. Ultimately, this chapter is motivated by a need for phylogenetic subgrouping that guides language reconstruction and produces falsifiable expected reflexes for individual daughters, even if permeable boundaries between language communities mean that many innovations are induced by contact.

In this chapter, I present a preliminary reworked phylogenetic grouping of Malekula languages as an alternative to Tryon (1976) and the linkage-based model proposed by Lynch (2016), albeit one with much overlap. This was achieved by eliminating virtually all of the isoglosses that include idiosyncratic sets of languages, while selecting isoglosses that are fully contained within broader ones. Unlike previous proposals, this one uses a greater number of language varieties – albeit a smaller number of words per language – than has been previously available until the publication of the data used here. It also favors lexical isoglosses that involve derivation of new forms by the addition of new material and assessing a number of ordered sound changes. This working phylogenetic model has informed the reconstruction proposed in the chapters that follow; it allows for mixed languages that have dual linguistic parentage.

In the sections that follow, I introduce the geography (§3.1.1) and orthography (§3.1.2) that will be in use throughout this chapter. I then describe the methods used, including the data and procedure for identifying nested innovations (§3.2); the results in terms of both ordered innovations and the subgroups that emerge from them (§3.3); and some conclusions for the approach taken in this dissertation (§3.4).

3.1.1. Geography

Occasionally, I make reference to the present-day location of a language's speakers, so a rudimentary familiarity with the geography of Malekula is necessary. The shape of Malekula as viewed aerially has been described as resembling the silhouette of a sitting dog facing west. It has become conventional to refer to the northern third of the island as the dog's head. South of the dog's head is a wide isthmus that gradually widens as one moves southward: this is correspondingly the dog's neck. South West Bay is found on the west-facing coast in the southern half; the relatively much larger gulf containing smaller bays spans most of the southern coast; and the southeastern coast is marked by an inlet and smaller islands, including the Maskelynes archipelago. The interior of the southern portion of the island is characterized by rugged and nearly inaccessible highlands. This was the site of the

Lendemboi (or Letemboi) region, which was depopulated during the missionization era. A number of streams originate here, but they have limited navigability.

The northern part of the island is bounded by the Strait of Bougainville, beyond which lie the islands of Malo and Espiritu Santo, the largest island of Vanuatu. Though the focus is on Malekula languages, several languages share features with the languages of Espiritu Santo. In general practice and in the text that follows, it will be referred to simply as *Santo*[6].

Figure 1: Map of Malekula created with data from OpenStreetMap. All maps in this chapter are the author's own work.

[6] Colloquially, the urban area of Luganville (the second largest town in Vanuatu) is also called *Santo.*

The main island is surrounded by much smaller islands, and some of these are associated with languages and ethnicities that are also established on the mainland. These include (visible in Figure 1), from north to south, the islands of Vao, Atchin, Wala, Rano, and Uripiv, which is due east of the Malampa provincial capital, Lakatoro (marked by a star). The southeastern region has a large inlet and town, both named Lamap (formerly Port Sandwich). The farthest of the small islands in the south east are collectively called the Maskelynes. Along the South Coast is a large bay. In the center of that bay is the island of Axamb (or Akhamb), home of the Axamb language.

Each low-level node in this working phylogeny is presented with a map of the member languages in this chapter. These maps have estimated traditional areas inhabited by speakers of those languages, with each language community identified by a unique color – except where that range is limited to small islands. A black triangle indicates the location associated with a speaker who provided the data. In most cases, a triangle identifies multiple speakers in clustered villages. The triangles are by no means a comprehensive representation of settlement patterns for the language communities. When a language was sampled outside of its traditional range for any reason, it is linked by a thin straight line to the language label or to the colored field representing the range. These maps should be taken as approximations only; the authority on traditional and modern settlement patterns lies exclusively with the communities.

3.1.2. Orthography

Any low-level reconstructions provided here are cursory and primarily intended to highlight specific changes to the sound systems reflected in individual cognate sets. High-level reconstructions represent Proto North-Central Vanuatu (PNCV) (Clark 2009), or occasionally when specified, Proto Oceanic (POc) (Blust & Trussel 2020). At lower nodes, the precise reconstruction can eventually be refined in further work, but this is beyond the scope of this dissertation. Especially for lexical isoglosses, they are intended as approximations of word forms. The relevant characteristic in these cases is binary: whether a particular language inherited or borrowed a word of the general shape. In cases where the reconstructed forms are clearly divergent reflexes of a single etymon, the reconstruction is more specific for the relevant sounds (generally consonants). These qualifications are also highlighted in the text as appropriate.

The possibility of the presence of linguolabial consonants in the proto language receives special attention in this dissertation. Considerable work has reiterated that these are a diffused areal feature and not diagnostic of subgrouping (Lynch and Brotchie 2010, Lynch 2016, Lynch 2019). Accordingly, linguolabials were not included in the isoglosses identified here; nonetheless, it is possible to reconstruct linguolabials for at least one node: the languages that have *vovora 'star' (1.3). The status of V'ënen Taut as an intermediary for linguolabials borrowed is not supported by this work – V'ënen Taut could instead be the sole borrower of linguolabials, and all the languages with irregular distribution could have inherited the feature (along with languages of Santo). The conclusions of prior work that linguolabials were primarily diffused across language boundaries will be reassessed in the following chapter (§4).

For the purposes of consistency with prior scholarship, the reconstructions are represented partially in an informal, emerging phonetic alphabet that resembles English-based orthographies adopted for many South Pacific languages. This system will be used for all levels of reconstruction – where lower-level nodes contain sounds specific to that language group, IPA values are used. The main divergences from IPA are the value of *j, which is a prenasalized sibilant (often with language-internal free variation between alveolar vs. postalveolar and fricative vs. affricate articulation), and the voiced symbols *b, *d, and *ɡ, which represent prenasalized segments with corresponding IPA place features. (Note that several popular orthographies use ‹g› for /ŋ/ and ‹q› for [ᵑɡ] or [ʔ]) The grapheme *y represents IPA [j], and there are some unique vowel conventions: diaresis/umlaut represents front vowels on rounded-vowel base letters, but *ë represents [ə]. The phonetic values of reconstructed consonants (Table 2) and vowels (Table 3) are schematized below, but do not represent the phonetic inventory of any single synchronic linguistic snapshot of Malekula.

	Linguolabial	**Bilabial**	**Labiovelar**	**Coronal**	**Dorsal**	**Glottal**
Stop (plain)	*p’ [t̼]	*p [p~b]	*p^{w} [p^{w} ~ b^{w}]	*t [t~d~t̪~d̪]	*k [k~g]	*ʔ [ʔ]
Stop (prenasalized)	*b’ [^{n}d̼]	*b [^{m}b]	*b^{w} [^{m}b^{w}]	*d [^{n}d~^{n}d̪]	*g [^{ŋ}g]	
Affricate (plain)				*č [ʦ~ʧ]		
Affricate (prenasalized)				*z [^{n}ʤ~^{n}ʥ]		
Fricative (voiceless)		*f [ɸ]	*f^{w} [ɸ^{w}]	*s [s]	*x [x]	*h [h]
Fricative (voiced)	*v’ [ð̼]	v [β]	v^{w} [β^{w}]		*ɣ [ɣ]	
Plain trill		*P [ʙ̥ ~ ʙ]				
Tap/Trill (prenasalized)		*B [^{m}ʙ]		*D [^{n}d^{r}~^{n}ɖ^{ɽ}]		
Nasal Stop	*m’ [n̼]	*m [m]	*m^{w} [m^{w}]	*n [n]	*ŋ [ŋ]	
Approximant			*w [w]	*l [l]	*y [j]	

Table 2: Consonant representation in reconstructed forms.

	Front	Central	Back
High	*i [i] – *ü [y]		*u [u]
Mid	*e [e] – *ö [ø]	*ë [ə]	*o [o]
Low-Mid/Low	*ɛ [ɛ]	*a [a]	*ɔ [ɔ]

Table 3: Vowel representations in synchronic and reconstructed forms. Contrastive rounded vowels are located to the right of en-dashes.

3.2 Methods

The approach taken in this chapter seeks evidence for types of diachronic change that are patterned in nested groups of languages, instead of innovations that criss-cross subgroups. These innovations include shared general sound changes and unique cognate lexical forms. In what follows, the description of the data (and its coverage of the languages of Malekula) and the selection process for diagnostic changes reveal the great amount of noise and variability that one must contend with.

3.2.1 Data

For the sake of comparability, data were restricted to Shimelman et al. (2019) as the primary source for lexical data. That work represents a large-scale lexical database documenting basic vocabulary across Malekula and some small offshore and large nearby islands. The database includes audio files from speakers and IPA transcriptions.

Since that primary source does not yet include a lexical survey of the languages of Santo at the time of writing, I refer to Tryon (1976) for characterizations of Santo languages. This source similarly provides an original lexical survey alongside older compiled word lists, but for much more of the Vanuatu archipelago. Due to technological limitations of its time, only transcriptions are available, and the word forms may be abstracted across speakers, varieties, and possibly obligatory morphosyntactic context. For the Ninde language, I mainly used field notes and lexical forms I personally helped to collect on Malekula as well as insights from prior work on Ninde's linguistic history. This includes occasional reference to sung Ninde (discussed in more detail in §2), which has a frozen lexicon that is semantically opaque (and therefore resistant to elicitation). An ancillary goal of this work is also to contextualize lexical data I helped to collect from a language identified by a man named Elder David Kaiar as Nevitangiene. Kaiar identifies as the last full speaker of this language. Finally, for languages which have more extensive documentation available, I incorporated insights from work with those data; often, these provided characterizations of language diachrony that could not otherwise be determined from the data – these external sources are cited as relevant in the text that follows.

3.2.2 Languages

The languages considered in this chapter as sources of primary data are all indigenous, spoken languages of the Oceanic branch of the Austronesian language family, but the linguistic makeup of Malekula is more complex. Public institutions like schools and medical facilities have continued to use English and French, often alongside Bislama. There may be indigenous sign languages on Malekula and since 2006, New Zealand Sign Language has been introduced as a language for Deaf education – prior efforts have brought Melanesian Sign from Papua New Guinea, among others (Iseli 2018). Ni-Vanuatu from other islands frequently relocate to Malekula, usually as refugees from vulcanism, as public workers, or as spouses of Malekula islanders. Foreign missionaries, aid workers, Peace Corps volunteers, and business people also bring languages, if only temporarily. Ni-Vanuatu who work abroad also sometimes acquire additional languages, like Chinese.

Ultimately, this work privileges languages of Malekula that have been recorded with modern equipment and phonetics training. Many of these languages face language endangerment and some are barely remembered by descendants of speakers. A better understanding of the phylogeny of this area may eventually inform work that is limited to elicitation and reconstitute languages from messy and scant data. This work has the potential to promote the acquisition and study of spoken languages, both synchronically and diachronically – but there is likely more work needed to document all of Malekula's languages more thoroughly. The languages that were considered in this work are tabulated in Table 4.

Language	Variety/Toponym	Ethnonym	Alternative names	Island
Alavas	Alavas			Malekula
Alavas	Wowo			Malekula
Angavae	Dixon Reef			Malekula
Aore				Santo
Atchin	Orap		(Northeast Malakula)	Malekula
Atchin	Small Island			Malekula
Aulua	Loxse-Asolokh		Aulua Bay	Malekula
Aulua	Loxse-Lanfitfit			Malekula
Avava	Khatbol		Khatbol, Katbol	Malekula
Avava	Tisvel			Malekula
Avok	Hokai		(Axamb)	Malekula
Avok	Small Island			Malekula
Axamb	Small Island	Axamb	Akamb, Ahamb, Akhamb	Malekula
Axamb	Small Island-Maliambor			Malekula
Bangaasak	Taremb		Umbbuul	Malekula
Batarxopu	Benut			Malekula
Batarxopu	Gonwar			Malekula

Language	Variety/Toponym	Ethnonym	Alternative names	Island
Batarxopu	Lembilmbil			Malekula
Batarxopu	Lipitav			Malekula
Batarxopu	Lumsak			Malekula
Burmbar	Burmbar		Banam Bay, Vartavo	Malekula
Burmbar	Fartavo			Malekula
Burmbar	Tengan			Malekula
Espiegel's Bay	Litebral			Malekula
Espiegel's Bay	Solwota			Malekula
Espiegel's Bay	Unspecified			Malekula
Fifti	Khatbol		Vivti	Malekula
Fifti	Malmbor			Malekula
Kerepua				Santo
Larevat	Fwishile		Larëvat	Malekula
Larevat	Larevat			Malekula
Larevat	Mosox			Malekula
Letemboi	Aingelemolesa		Lendamboi, Lendemboi, Repanbitip/	Malekula
Letemboi	Natingatlang			Malekula

Language	Variety/Toponym	Ethnonym	Alternative names	Island
Letemboi	Nevatanyene		Niolean (variety)	Malekula
Letemboi	Newotenyene			Malekula
Malua Bay	Coast		Middle Nambas	Malekula
Malua Bay	Marasup			Malekula
Malua Bay	Petarmul			Malekula
Malua Bay	Tiptir			Malekula
Mandri	Faru			Malekula
Mandri	Farun			Malekula
Na'ahai	Mbatvanui		Malfaxal, Malvaxal	Malekula
Na'ahai	Toman			Malekula
Naati	Windua	Wilemp	Nāti, Nahati	Malekula
Naha'ai	Malfaxal		Malfaxal, Malvaxal	Malekula
Naha'ai	Mbatmbang			Malekula
Nahavaq	Banour	Sinesip	South-West Bay, Sinesip	Malekula
Nahavaq	Caroline Bay			Malekula
Nahavaq	Lembinwin			Malekula
Nahavaq	Loorndavo			Malekula

Language	Variety/Toponym	Ethnonym	Alternative names	Island
Najit	Tanmial			Malekula
Naman			Litzlitz	Malekula
Nasarian	Ngava			Malekula
Nasarian	Wileven			Malekula
Nasvang	Big Island			Malekula
Nasvang	Farun			Malekula
Natangan	Mbonvor			Malekula
Navwien	Mbonvor			Malekula
Nese	Sanboise			Malekula
Nese	Tanmial			Malekula
Netimb	Bangir		Natimbo	Malekula
Neve'ei	Coast		Vinmavis	Malekula
Neve'ei	Vinmavis			Malekula
Neverver	Limap	Mindu	Lingarak, Nevwervwer	Malekula
Neverver	Mindu			Malekula
Neverver	Sakan			Malekula
Ngata	Taremb			Malekula

Language	Variety/Toponym	Ethnonym	Alternative names	Island
Ninde	Labo	Mewun	Labo, Meaun	Malekula
Ninde	Lawa			Malekula
Ninde	Opmobamba			Malekula
Ninde	Southwest Bay			Malekula
Ninde	Windua			Malekula
Nisvai	Blacksands			Malekula
Nisvai	Ronevie			Malekula
Nitita	Khatbol			Malekula
Nombotkote	Blacksands			Malekula
Nombotkote	Melkin			Malekula
Nombotkote	Neioleian			Malekula
Novol	Bangir			Malekula
Novol	Dixon Reef			Malekula
Novol	Tavendrua			Malekula
Pangkumu	Datisman		Rerep, Tisman, (Unua-Pangkumu)	Malekula
Pangkumu	Ndarmif			Malekula

Language	Variety/Toponym	Ethnonym	Alternative names	Island
Port Sandwich	Lamap		Lamap	Malekula
Port Sandwich	Merivar			Malekula
Rano	Chinambong		(Northeast Malakula)	Malekula
Rano	Potora			Malekula
Rano	Twalung			Malekula
Salang	Dixon Reef			Malekula
Siviti	Gonwar		Njav	Malekula
Siviti	Jericho			Malekula
Siviti	Tanmililib			Malekula
Siviti	Womol			Malekula
Tape	Mae	Tape	Maragus, Maraxus	Malekula
Tape	Tautu			Malekula
Tesmbol	Melaklak			Malekula
Tesmbol	Usus			Malekula
Tirax	Bakru	Batarvxalin	Dirax, Dirak, Mae	Malekula
Tirax	Bethel			Malekula
Tirax	Botmelesnial			Malekula

Language	Variety/Toponym	Ethnonym	Alternative names	Island
Tirax	Mae			Malekula
Tirax	Olamb			Malekula
Tirax	Orap			Malekula
Tutuba				Santo
Uliveo	Lurtes		Uluveu, Maskelynes	Malekula
Uliveo	Pelongk			Malekula
Uliveo	Peskaros			Malekula
Unua	Benuamor		(Unua-Pangkumu)	Malekula
Unua	Mbatumbo			Malekula
Unua	Noferai			Malekula
Uripiv	Durpif		(Northeast Malakula)	Malekula
Uripiv	Potnambe			Malekula
Uripiv	Tautu			Malekula
Uripiv	Tautu Village			Malekula
V'ënen Taute	Alpalak	Taut	V'ënen Taut, Big Nambas	Malekula
V'ënen Taute	Benenavet			Malekula
V'ënen Taute	Benmara			Malekula

Language	Variety/Toponym	Ethnonym	Alternative names	Island
V'ënen Taute	Hamax			Malekula
V'ënen Taute	Unmet			Malekula
Vao	Beetehul	Vao		Malekula
Vao	Canada			Malekula
Vao	Peteri			Malekula
Vao	Torono			Malekula
Wala	Botku	Wala	(Northeast Malakula)	Malekula
Wala	Petreros			Malekula
Wala	Preng			Malekula
Wala	Worprev			Malekula
Wowo	Lesarmalas		Mpotovro, Botovro, Vovo, Bieria	Malekula
Wowo	Wowo			Malekula
Xoli			Mbwenelang, Bwenelang	Malekula

Table 4: The doculects considered in this work with corresponding ethnonyms and glossonyms used in other work.

Intra-language contact

For those language communities that are the subject of a reference grammar, recent history of population contraction and consolidation of villages paints a picture of intense contact between once distinct varieties. The languages of Avava, Salang, and Bangaasak represent a likely scenario of interdialectal mixing. Crowley and Lynch (2006a:8-9) identifies three originally distinct varieties of a single language that formed a linguistic amalgam (Bislama: *kampani lanwis*) of Modern Avava as populations declined in colonial times: (Original) Avava spoken in Vovoka, Umbbuul[7] spoken in Bangasak, and Gara spoken on the northern bank of the Nurumbat River (just north of Vovoka). Avava speakers report that Fifti (whose remaining speakers live among Avava speakers in Khatbol village) and Nitita are not mutually intelligible with Avava, but they share 73.75% of vocabulary in Crowley's samples (Crowley 2006a:11-3). Similarly, the Neverver speaking community also underwent relocation and consolidation after depopulation (Barbour 2012:2). While modern Neverver arose from the Mindu dialect, some people still identify with the no longer spoken Sakhan dialect.

Massive population reconfiguration affected all communities, but not necessarily the languages per se. Other language communities also underwent removal from the interior, but do not have a documented history of interdialectal shift, like Neve'ei (Musgrave 2007:2-3).

[7] The Bangaasak doculect in this chapter.

Some languages communities were affected by migration from other regions of Vanuatu in the face of economic growth, like Nese, which has undergone a near-complete shift to Bislama (Takau 2016:16). These could represent languages that were more homogenous across the geographical range, or a history of shift and mixing is simply not felt to be a key part of the narrative.

Proto languages

The vast majority of the sources of innovation have been identified in reconstructed lexicon, but a handful still remain uncertain. A number of these are semantic shifts from near-synonyms. In many cases, these have already been identified by Clark (2009), who also reconstructed the PNCV forms. In other cases, a word is formed by compounding or affixation, but still a large number of words remain that have no identified origin in Oceanic languages.

With no viable sources for some lexical innovations within the documented parts of Austronesian lexicons, perhaps some of the new lexical forms were introduced via non-Austronesian languages that are thought to have been brought in antiquity with individuals from New Guinea who adopted Austronesian languages. As discussed further in §7.4, the decimal numeral system of Tirax is known to be older (more widespread in Oceanic), and the quinary systems in Malekula's other languages are thought to have developed in contact with speakers of Papuan languages who arrived after Vanuatu's initial wave of settlement by Austronesian-speaking people (Lynch 2009). Much of the Bismarck Archipelago, where the Lapita culture probably originated (and the closest part of mainland New Guinea from Vanuatu), and other parts of eastern Papua New Guinea are now predominantly Austronesian-speaking regions as well.

Some superficial similarities can be found between Malekula language and members of the Trans New Guinea languages. For example, I reconstruct regional *xar-sa 'ear' for parts of south and east Malekula, and this likely includes a regular reflex of the possessive *-sa; Shafer (1965:316)[8] reconstructs Proto Kiwai *gaɾe 'ear' from Proto Trans New Guinea *ka(nd,t)(i,e)C (Pawley 2005). Several languages also innovated *kapʷi 'egg', which resembles Proto Kiwai *kikopu 'egg, fruit' (Shafer 1965:318) (cf. Proto Trans New Guinea *maŋgV [Pawley 2005] or *man(a,u)ka 'egg' [Ross 2014]). The Kiwaian languages are spoken over 2.7 hundred kilometers away from Malekula and are by no means the closest part of New Guinea, but these two forms in this branch are particularly similar. Like Malekula languages, and unlike other Austronesian or Trans New Guinea languages, Kiwaian languages also have singular, dual, (crucially) trial, and plural pronoun sets (Ross 2005).

[8] As cited in Usher and Suter (2022) – but the original source and citation could not be located.

3.2.3 Structure of the data

I use the glossonyms for dialects as represented in Shimelman et al. (2019) or Tryon (1976) – these are the ones presented in Table 3 above. Word lists are identified by language and locale, and most combinations have one representative speaker (but some have two). In the case where there are two speakers representing one village, they are numbered with an Arabic numeral 1 or 2. When it is necessary to discuss isoglosses that cross-cut these language communities, I refer to the variety by the name of the locale with the glossonym in parentheses.

Generally, because the data come from a word list, the available information is somewhat limited by the sampling method. In many cases, lexical substitution has obfuscated a language's participation or non-participation in a widespread sound change. This has meant that only some sound changes were meaningful to consider with the data, and a comparatively small set of words is generally sufficient to demonstrate whether that sound change or lexical innovation took place in a particular language.

3.2.3 Procedure

An innovation is an innovation, whether diffused within a language or across languages (Bowern & Koch 2004). If an innovation is not diffused throughout a language community, yet it becomes established enough to become part of a basic word list elicited from a community member, then it can be assumed to be in use in a community, whether individuals actively produce it or have only passive knowledge of it. Initially, an incompletely diffused form may reflect language divergence that is in process; in this case, it can create the appearance of greater similarity between a lending and borrowing language. With time, even borrowings are subject to innovations affecting the lexicon. These changes to shared inheritance and established loans reveal genealogical transfer.

To disentangle contact-facilitated cross-linguistic diffusion from patterns of inheritance, I identified lexical and phonetic reflexes that can be gleaned from the data and categorized them into three sets: (1) innovations that have crucial ordering relative to each other; (2) areally linked innovations that show diffusion over a contiguous land area but do not overlap with or interact with other changes; and (3) innovations that are shared by unique combinations of languages. Only innovations which are shared by increasingly exclusive groups of languages met the criteria for (1), but it is possible that late retention required for a subsequent change was occasionally treated as an innovation – a neutral approach to innovation vs. retention was taken if the direction of change at shallow clades was not obvious. In some cases, however, innovations are shared between *neighboring* languages with apparently distant genealogical relationships (based on ordered innovations established as group 1). In this case, depending on the nature of the shared change, it was noted as an areal feature. When isogloss boundaries proved to be too messy for categorization in groups (1) or (2), they were not included (group 3). The remaining isoglosses most likely represent retentions made rare by more recent innovations, areal influence obscured by later population shifts, or parallel developments.

There are three ways that sound changes were considered to form a natural sequence based on formal properties alone: (1) sound changes were considered incremental changes in sonority, (2) they could be ordered in terms of feeding and bleeding relationships, and (3) they appear to propagate through increasingly broad contexts in increasingly restricted sets of languages. In conjunction with distribution across languages, these were taken as strong evidence of sequential innovation.

Of sonority-related changes, spiranticization and debuccalization of stops and fricatives, on the other hand, could precede deletion. This means that if some language shares all its innovations with another and has deleted stops (based on higher-order reconstruction) where others have fricatives, this is not counted against the grouping. This mirrors some characterizations of lenition as any change that is likely to be found on a diachronic pathway toward deletion (Hock 1991).

Some changes were systematically ignored. The numeral system is reconstructed as a decimal pattern in Proto Oceanic, yet the vast majority of Malekula languages have a quinary pattern deriving numerals 6-9 from words for 1-4. Every previous account of the phylogeny of Malekula has grouped Tirax, which has a decimal system, with only some of the languages that have the quinary pattern. This suggests that numerals 6-9 may have been restructured following high-level divergence, perhaps on analogy with patterns found in Papuan languages. Similarly, linguolabial consonants were considered to be not diagnostic of genealogical relationships in previous work (Lynch & Brotchie 2010). For these reasons, these innovations did not inform the analysis in this chapter.

3.3 Results

This process uncovered several innovations that pattern together. There are two main groups of languages on Malekula: one northern group with reflexes of *xarsa ‘ear’, *vivora ‘star’, and retention of PNCV*mai ‘come’ and *m^{w}adu ‘back’; the languages spoken south of the dog’s head have reflexes of *v^{w}ale ‘come’, but either innovated forms or retentions of PNCV forms *daliŋa ‘ear’, *taku-Ru ‘back’, and *vituu ‘star’. Several lexical items show sequential innovations that build on prior ones, and more accumulated changes are consistent with other nested patterns of shared innovations.

3.3.1 Sequential sound change

Two changes showed clearly increasing contexts for change. One of the two changes is the rhoticization of *d to *D, which more broadly occurs in the context of a following back vowel, but is extended to the context before *l in some languages. The other is the neutralization of liquids *r and *l, most broadly before high vowels but expanding into full neutralization in all contexts. Each of these changes can also be ordered relative to other changes: they are in a feeding or bleeding relationship with each other in languages that take part in both innovations.

A basic set of words was used to assess whether a language has taken part in one of these sound changes. If a form was clearly not a cognate, other forms were used, if possible. These changes were:

<u>For rhoticization of *d:</u>

1. ‘Back(side)’ as a diagnostic of whether PNCV*d is rhoticized before *u in languages with reflexes of *du- or *m^{w}adu-.
2. ‘Ear’ as a diagnostic of whether PNCV*d is rhoticized before *(V)l for languages with reflexes of PNCV*daliŋa.

<u>For liquid neutralization:</u>

3. ‘Five’ for phonological isoglosses relating to rhoticization of *l in the environment before *i for reflexes of PNCV *lima.
4. ‘Tooth’ for additional phonological isoglosses relating to rhoticization of *l in the environment before *i for reflexes of PNCV *livo; additionally for rounding of *i in the environment before a syllable with a rounded vowel, and the relative ordering of these two innovations, as relevant.

5. ‘Egg’ for rhoticization of *l in the environment before *u for reflexes of *katulu ~ *kadulu ~ *tulu (< PNCV*ʔatolu). However, *u in this environment may have been fronted allophonically and often yields /i/ as a reflex.
6. ‘Heavy’ for languages with *r deletion in reflexes of PNCV*marazi and the ordering relative to *l-rhoticization.

Debuccalization and deletion:

7. ‘Name’ for reflexes of intervocalic *k and *s of PNCV*nakisa.

In the case of each of these sound changes, there is no direct evidence for a language participating in only the less common change (more environments for change) but not the more common one. This is consistent with the more common environment representing the initial stages of the sound change and the second environment another step in the change.

Rhoticization of *d is not only restricted to northern Malekula, but also the nearby island of Santo. Of the 41 Santo languages in Tryon’s sample, only the data from the languages Kerepua and Nonona suggest a fully maintained distinction between PNCV *d and *D; all other Santo languages have merged the two in at least some environments. A number of them additionally have [k] as a reflex of PNCV*D and of PNCV*d in rhoticizing environments – this change probably happened after PNCV*d was rhoticized to *D (in the environment preceding *l with or without an intervening vowel), as [k] appears as a conditioned reflex of *d in these environments. The range of possibilities here are parallel innovation, partial "reversal" in some languages, or a change that expanded to a new context, but in any case, the change affecting both possible contexts preceded a change *D > k.

While rhoticization forms clear isogloss boundaries circumscribing contiguous linguistic regions on Malekula, there is one case of cross-cutting that can be explained by chance convergence. Several languages have prenasalized reflexes of PNCV*t (> *d) after a syllable with a nasal onset. If PNCV *t was prenasalized in a Malekula language after a prefixed nominal *na-, then the resulting *d was also rhoticized in these environments. For example, in the PNCV*(na)-tulu- 'egg', the PNCV *t has become /D/ in several languages. The prenasalization here could be a newer innovation – instead of introducing a prenasalized coronal contrast before /u/ (i.e., /du/ vs. /Du/), the subsequent change could have been a direct change from PNCV *t > D / nV _ {u, o} in keeping with the phonotactic constraints established by former changes. Unlike the proposed expansion of context from only *u to also *l, this change would maintain the distribution of *d and *D shared with other languages.

A viable sequence of innovations can be formalized as follows:

1. **Santo and Northern Malekula:**

   $*d > *D / \_ \{u,o\}$

   Rhoticization of *d before back vowels

2. **Southern Santo and Northern Malekula:**

   $*d > D / \_ V_0 l$

   *d is rhoticized before *l, with or without intervening vowels

3. **Some Santo languages:**

   $*D > k$

   *D becomes /k/ and the previous change (1) feeds this one

4. **Malekula languages in the dog's head and neck**

   4.1. $*t > d / n V \_$

   *t is prenasalized after a nasal

   4.2. $*d > D / \_ V$ [+back]

   The resulting *d is rhoticized before back vowels

Changes affecting the liquid contrast (*l vs. *r) reconstructed for PNCV are also in feeding and bleeding order with other sound changes. The contrast may be partially or completely neutralized, with /r/ preceding high front vowels and /l/ preceding other vowels. Rhotic reflexes are conditioned by *i and sometimes by fronted reflexes of *u, but some languages seem to have re-phonologized /l/ vs. /r/ by merging (allophonically front) *u and *i. In the southeast of Malekula, the *o and *u of *ʔatolu ‘egg’ sometimes become *i, yielding various reflexes for *l. Additionally, some languages have deleted *r in postvocalic positions, and a change of *t > r has taken the place of *r. At shallow levels, these changes make it virtually impossible to group languages of the southeast, with each language undergoing an ostensibly unique ordering of relevant changes. Nonetheless, higher-level changes exclude the languages of the southern and southeastern coasts from any grouping with Ninde.

3.3.3 Morphological isoglosses

Most of the morphological structure of the data cannot be assessed from a wordlist, but some word forms show layers of derivation. The following three forms were used to assess whether languages had undergone the same degree of morpheme accretion, assuming that a morpheme in the citation form of a word has generally been affixed or compounded with the stem before outer layers of morphemes/affixes.

8. 'Back(side)' for isoglosses reflecting PNCV *(m^w)adu vs. apparently unprefixed PNCV*adu.

9. 'Sleep' for morphological isoglosses for languages that have a form derived from *p^wač with a prefix of the shape *ga-.

10. 'Star' for morphological isoglosses involving reduplication or derivational affixes, reflecting PNCV*vituu and PNCV*m^wazoe, for which I propose a single origin in at least some reflexes.

11. 'Egg' reflecting PNCV *ʔatolu- with prefixed * m^wa- of unidentified origin.

'Star'

Malekula languages can be divided into groups on the basis of their reflex for 'star', for which I propose a single etymon, deviating from Clark's two reconstructions: PNCV*vituu and PNCV*m^wazoe[9] . Innovations affecting this lexical form reflect broader *sound changes* (not a mere lexical difference) and morphological developments. More importantly, the forms have a necessary relative chronology and subgrouping on the basis of this lexemes

[9] This is reconstructed as a form meaning 'disc', 'planet', with common extensions to 'star', but also occasionally contrasting with 'star'. Perhaps Clark's forms represented two different lexical items, that have simply mutually influenced each other or become homophonous in some daughters.

correspondence to other innovations. The forms reflecting *m^{w}azoe are crucially found in languages that have progressive nasalization, such that an intermediate partially reduplicated *v^{w}e<u>n-v^{w}</u>edeu could regularly surface as *v^{w}e<u>nm^{w}</u>edeu. Clade protos are offered in tables 5-13, and lower-level data is cross-referenced where the clades are mentioned. I suggest that POc *pituqin underwent the following sequences of changes in various Northern Vanuatu languages (starting from Santo and moving southward):

1. **In all Northern Vanuatu languages:**

*pituqin > *vitui

Regular spiranticization of the initial *p

Regular loss of intervocalic glottal *q

1.1. Some Santo languages and Vao:

*pituqin > *vitui > *vitiu

Metathesis of final vowels (no data on regularity)

This change is potentially shared by all Malekula languages

1.2. Some Santo and Northern Malekula languages develop a compound:

*pituqin > *vitui (?) > *vitiu (?) > *vitiu + *vovora

Only Aore and Tutuba (both spoken on Santo) retain the full compound:

Tutuba: *βitu βoβora* ‘star’

Aore: *ð̼itu βaβora* ‘star’

1.3. **All other languages in this group** retain the second word only:

> *vitiu + *vavora > *vavora

(Semantic) displacement

2. In almost all coastal southern and eastern Malekula languages:

> *vitiu > *na-v^wutiu > *na-v^wučiu > *(na)-m^wučiu

Metathesized final vowels

Regularly palatalized *t > *č / _ i

Nasalized initial *v^w after *na-

3. **In the languages predominantly of the interior (Table 10):**

> *na-v^wutiu > *na-v^wen-v^wutiu

Partial reduplication of the root

Particularly conservative are Lendamboi languages:

Natingatlang: *navanvinde*

Nevitangiene: *nevenvinde*

3.1. In the southern interior:

> *na-v^{w}en-v^{w}udiu > *navenvut

Loss of final vowels

3.2. In the northern interior (and Ninde):

> *na-v^{w}en-v^{w}udiu > *na-v^{w}enm^{w}ezeu

Nasalization of *v^w / n_

This sequence of morpho-phonological changes emerged as one of the most reliable (in that it is reinforced by other isoglosses). Some languages have clear reduplication of the first syllable (3), resulting in a nasal coda [n] in the reduplicand as a realization of prenasalized *d or *z. The oral *v^w in both reduplicand and root is retained in languages of the southern interior and central west coast (3.1), but nasalized in the root in languages of the northern interior (3.2). Such nasal realizations of *d and *z are relatively common word-internally, as in Neve'ei, and occasionally also in the reduplicand (Crowley 2006b:52).

I have analyzed these forms as derived from one proto form, but Clark's reconstruction of two forms was motivated also by semantics. Clark identifies languages throughout central and northern Vanuatu that have *vituu reflexes alongside *m^{w}azoe reflexes, with the latter meaning 'Venus', a planet in general, or a specific star; these languages are: Mota, Raga, Sakao, Vara Kiai, Tamabo, and Paamese. Examples of star terms in stage (2) could in fact be reflexes of another root altogether, as he has reconstructed. Clark (2009) also reconstructs *m^{w}aza 'spear', a reflex of which forms the first part of a compound in Nahavaq *nam^{w}as-nal* 'shooting star' with a literal meaning of 'sun spear' (Dimock 2009:94). Whatever the origin of these terms, they could have exerted mutual influence on each other, at least in Nahavaq; or what appears to be reduplication could even be a compound of both roots.

Since entries for 'star' are well recorded in word lists and reinforce other shared innovations (reduplication behavior, conditioned nasalization of *v), languages are also categorized here by the "stage" of morphological complexification in this word. The processes shaping the complex forms of this word likely escaped detection in higher-order reconstructions like Clark's PNCV because of the substantial phonetic change in the languages that are best documented in available literature. Of the languages that have many speakers and some amount of grammatical description, Avava *v^{w}inmese* has lost the initial *na-, Neverver *nivinnimdžo* has apparently metathesized the reflex of *m with the nasal portion of palatalized *z (yielding a geminate /nː/), and Ninde *nanmysi* ~ *nanm^{w}əsi* has regularly lost the intervocalic *v. In a song performed by Plas Kali (2018) intended to evoke the moon and stars on a cloudy day, a form *nawanmasi* can be heard; it is likely that this word is a parallel reflex of 'star' that preserves the reduplication more faithfully.

Generally, Santo has the greatest diversity in 'star' isoglosses (exhibiting *vitiu, *m^{w}učiu, and *vovora), which suggests it as the location of some early linguistic innovations found on Malekula, or a history of bidirectional linguistic interchange between Santo and Malekula. This lends some support to a shared origin of linguolabials – another rare feature shared across many of the languages that have two of the three innovations in 'star' – and against a Proto Malekula ancestral to all Malekula languages.

'Sleep'

The most widespread form for 'sleep' is minimally changed from PNCV *maturu, but another form for 'sleep' can be reconstructed as *p^wač (with a voiceless bilabial stop, low vowel, and a coronal stop or affricate). This form appears to represent layered innovation as well. Some forms have reflexes of the stative prefix *ma- (*ma-p^wač), others have velar initial consonants consistent with initial material of the shape *ga- (*ga-p^wač). These latter complex reflexes fit different points along a sequence of changes to the internal *p^w, which is spiranticized or deleted: *ga-p^wač > *ga-v^wač > *gač ~ *gos. It is not clear if these are derived forms or inflected stems used as citation forms, but the isoglosses are comparatively straightforward. The bare roots are found in coastal languages of southeastern Malekula, the transparently prefixed forms are found in a subgroup of those languages, and the coalesced forms are found in the interior highlands. The languages of the interior highlands have been classified differently in this chapter from the bare and analyzable forms.

Two further lexical alternatives meaning ‘sleep’ pattern geographically with the coalesced forms of *gač and *gos; they are provisionally reconstructed as *m^{w}one and *iɛn. Reflexes of these forms replace PNCV*maturu and are probably derived from PNCV*eno ‘lie down’. These are not implicated in any morphological complexification, but are also localized to the languages of the interior highlands of Malekula. Each of these innovations are neatly nested within the subgroups proposed in this chapter, but in contrast to layered morphology of words like ‘star’, the forms identified as most advanced in terms of morphophonological change – the coalesced syllables of *gač and *gos – do not reinforce the patterns of nested innovations of the simpler states *ga-p^{w}ač, *ga-v^{w}ač, and *p^{w}ač.

‘Egg’

Some forms take on what appears to be a prefix *m^{w}a- with uncertain semantics. Reflexes of PNCV*ʔatolu ‘egg’ have an added initial element *m^{w}a- in several languages, and this could be related to the alternation in PNCV*adu vs. PNCV*m^{w}adu ‘back of body’ or ‘spine’. It resembles the prefix PNCV *v^{w}a- that Clark (2009:70) identifies as a fruit classifier (nasalized after the nominal prefix *na-, perhaps). Some element *m^{w}av^{w}i- in other languages looks tantalizingly like words for ‘white’ across the southern Malekula languages (PNCV *ovo=ovo or *voge, perhaps with the stative prefix *ma-). It could also reflect a reduced form of *kap^{w}i ‘egg’ that I reconstruct for group C in §3.3.4. Although the forms are rather different after phonetic reduction and prefixation, I consider it likely that these forms contain the morphological reflexes of the older words for ‘egg’.

Importantly, the forms are also layered and associated with specific sound changes. Reflexes with *m^wa- have either a root *kadulu or *tulu (reflexes in part or full of PNCV*ʔatolu). Forms that have *m^wav^wi- can either have a final element *tul or simply *l, and in an even more restricted set of reflexes, the *l and *v^w reflexes have become metathesized in words of the shape *m^wav^wi-l, such that a bilabial consonant is in word-final position. Except in low-level clades, the layered morphology of ‘egg’ does not support subgrouping.

3.3.3 Lexical isoglosses

Finally, these phonological and derivational changes were strongly associated with restricted sets of lexical isoglosses:

12. ‘Mouth’ for a high-level isogloss, PNCV*ziŋo ‘snout’ replacing PNCV*b^waŋo.
13. ‘Ear’ for a high-level isogloss, PNCV*b^wero ‘outer ear’ (also associated with animals) replacing PNCV*taliŋa-.
14. ‘Come’ for a high-level isogloss, *v^wale (of unknown origin) replacing PNCV *ma(i).

15. ‘Back(side)’ for lexical isoglosses at several levels: PNCV*du, PNCV*m^{w}adu, or PNCV*taku-Ru-.

16. ‘Sleep’ for low-level lexical isoglosses *iɛn and *p^{w}ač (of unknown origin), replacing PNCV*maturu, and for morphological isoglosses for languages that have a form derived from *p^{w}ač.

17. ‘Heavy’ for numerous lexical isoglosses: inherited PNCV*mava or PNCV*marazi, or forms with unidentifiable origin *dVv^{w}, *nVb, *lVb.

18. ‘Egg’ for lexical isoglosses at many levels.

3.3.4 Subgroups

Three main subgroups emerge from this set that specifically include Ninde, and they are all implicated in a higher-order subgroup with each other (excluding most Malekula languages). These are identified on the tree diagram in Figure 2 by letters A-F. The letters in these diagrams correspond to the lettered subsections that follow, as well as accompanying maps (for low-level groupings) and tables. The text progresses generally starting from more numerous language subgroups to more restricted sets posited as descendants, but this structure means that all intermediate proto forms are first presented without accompanying data, then finally repeated alongside the data.

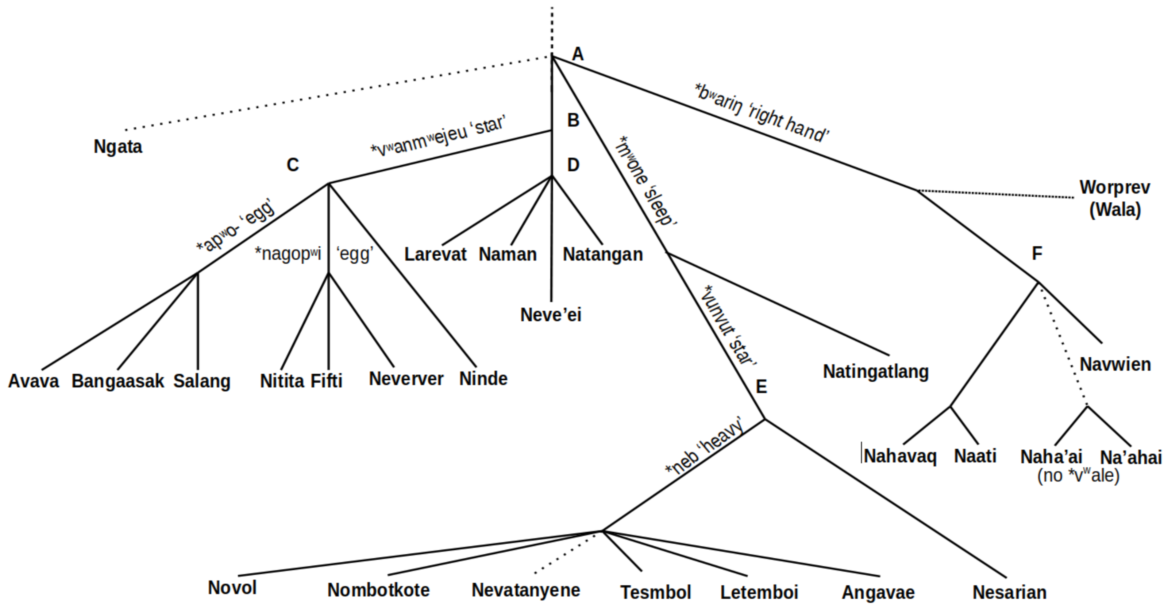

Figure 2: Proposed phylogenetic subgrouping for some of the Malekula languages.

A.

The larger subgroup can be defined by innovations represented by reflexes of: *vʷale 'come' (vs. PNCV*mai), *livukat 'night' (containing PNCV*livuka- 'middle'), and probably *vʷVdu 'star', with reflexes that have a rounded bilabial fricative or nasal and likely a prenasalized coronal stop or affricate (reconstructed as PNCV*vituu and PNCV*mʷazoe). This group does not take part in any of the rhoticization isoglosses – reflexes of *d and *D are distinct (but there is a documented tendency for younger generations across languages to produce /D/ as /d/), and so are the reflexes of liquids *l and *r. None of these languages retain *mai 'come', most of them retain *boŋi 'night' in secondary senses only ('evening' or ‘past day’). None of them participate in innovations like *xarsa ‘ear’ or the semantic shift of PNCV*bʷero ‘outer ear’ > ‘ear’ (all retain PNCV*daliŋa), the expanded form *mʷadu ‘back’ (all have PNCV*taku-Ru reflexes except for one group), or *pʷač ‘sleep’ (though several have the expanded form *ga-pʷač).

In terms of morphological isoglosses, all of the languages are consistent with a prenasalized coronal in *vituu > *vidui or *vidiu ‘star’, regardless of the reflex. The evidence for an intermediate *d shared by all of the languages in this group is the fact that the partial reduplicand is *v(w)en- in all the languages that have a reduplicated form, and this *n could have been the expected reduplicand coda allophone of *d as it is in Neve’ei (Crowley 2006b:52). In the environment of the following front vowel, the second *d could have merged with *z, as well as instances of *t that palatalized to *č, and in some languages ultimately with *s (at least in the same environment of palatalization). Forms with a final plain oral coronal /t/ (not followed by any vowel) nonetheless have a number of possible pathways from an intermediate *d: *t could have been a word-final reflex of *d after vowel loss, or even of *č or *z. Any analysis, even reconstructing both *vituu and *mʷazoe, supports the same isogloss patterns, as any explanation must account for the unexpected distribution of nasals in apparently reduplicated forms.

Furthermore, the languages with these innovations fully include all the languages with the innovations listed below (B-F). This subgrouping is reinforced by reflexes of a form /bʷar{a,i}{g,d,ŋ}i{n,ŋ}/ ‘right (side)’ (cf. PNCV*matuʔa). This form seems to be idiosyncratically affected by feature metathesis and/or assimilation in the final two consonants, though it is not clear whether this affected place of articulation or nasality; a reconstruction at this level of *b^waragin will be explained at the lower-level nodes.

B.

Languages: Avava-Bangaasak-Salang, Neverver-Fifti-Nitita, Ninde, Neve'ei, Naman, Larevat, Natangan

Marginally: Novol, Nombotkote, Angavae, and some varieties of Letemboi.

The next subset is spoken along the western coast of Malekula and the interior. It includes languages that have all of the following: reflexes of a form *nV-mVkut 'person, man' and *{v^w,m^w}idiu 'star', and most have a partially or fully reduplicated root *tn 'cook' (cf. PNCV*taʔo-ni 'cook in an earth oven') – those of this group that do not, have no attested form for 'cook' at all. This group fully contains groups C and D.

The word for 'man' or 'person' could reflect a compound of PNCV*m^w{a,e}ra 'person of' and PNCV*ʔuta 'place', but the root meaning 'person of' where attested has an idiosyncratic velar reflex of *r, if it does not have a different origin altogether. This is evident in Neverver *nemax* and Naman *mox-* 'denizen, person of' and it is also suggested by the first syllable of Ninde *Mewun* (the name of the general ethnic group that speaks Ninde). As a compound, this would also parallel Bislama *manples* ‘indigenous person’ (from *man* and *place*) – these are rare noun-noun compounds in Bislama, which otherwise forms complex noun phrases with the possessive prefix *blo(ng)*.

The reflexes of *tn ‘cook’ all have a syllabic nasal, which is also unique to the phonologies of these languages. These languages are joined in this lexical isogloss, however, by Novol, Nombotkote, Angavae, and the Aingelemolesa variety of Letemboi. With one exception – Neverver has *titn* ‘cook’ – all of the languages have a fully reduplicated form /tntn/. All of the languages also have syllabic /l/ in a reflex of similar shape *tl ‘three’ (from PNCV*tolu), making this a shared phonological development as well.

In terms of sound changes, all of the languages have either /č/ or /s/ as reflexes of PNCV*t before *i in non-final position. This affects words like PNCV*vati ‘four’ and PNCV*tika-i ‘not exist’. Since many of the languages underwent debuccalization of PNCV*s but not of palatalized PNCV*t, this would have to be reconstructed as group proto *č, which subsequently merged with *s in group C, probably after divergence of that group.

C.

Languages: Avava, Bangaasak, Salang, Neverver, Fifti, Nitita

Marginal members: Ninde, Neve'ei

This group of languages (Table 5) more narrowly shares a morphological isogloss representing *na-v^{w}in-m^{w}ezeu 'star', as well as lexical innovations including a form with a shape /tele/ 'another' (though this is also shared with Neve'ei) and a root *čog 'grass' with a strident coronal consonant (vs. PNCV *m^{w}anayu or *valisi) that is reduplicated in all the languages that have it except for Ninde *nesoŋgonei* (which is a compound with *nei* 'tree').

The reflexes of what is tentatively reconstructed as *na-v^{w}in-m^{w}ezeu 'star' are opaque, but all of the languages crucially have a likely reduplicated form, albeit one with further changes affecting the root: *v^{w} has a nasal reflex *m^{w} and *t (c.f. PNCV*vituu) has a sibilant reflex, both in the root only. Neverver, which provides the only overt evidence of a prenasalized sibilant, would have metathesized only the place of articulation in the final two (subsegmental) nasal gestures, yielding a geminate form in /nivinnimzo/. The reduplication in Ninde *na-n-müsi* 'star' is obscured by the general loss of intervocalic *v^{w} after *a, but a sung form /nawanmasi/ has the expected reflex /w/ of *v^{w} in other environments. This word was not identified with any meaning, but it appears in corpus item 134 (Kali 2018), which is a song sung to coax a star to reveal itself on a cloudy night. These forms support at least a shared origin in this lexical form, rather than a separate etymon.

Overall, Ninde is only a marginal fit, since several other forms that reinforce this grouping are not found in Ninde, which instead has many of these in common with various languages of groups Avava-Bangaasak-Salang and Neverver-Fifti-Nitita. Excluding Ninde, all of the languages have a form ‘egg’ compatible with reflexes of a velar consonant (Avava typically deletes root-initial *k in nouns along with the historical *nV- nominal prefix), followed by a low vowel, a (possible rounded) bilabial stop, and a high vowel: *kapʷi (which incidentally bears a resemblance to Proto Kiwai *kikopu). The exception of Ninde [namal̪ˤap] ‘egg’ has little weight here, because it is most similar to the Natingatlang variety of Letemboi [namʷalaɸ], which which it shares few other features.

Also excluding Ninde, a reconstruction of *bʷarakin ‘right (side)’ is possible for this group. The final portion of the reflexes in the Avava-Bangaasak-Salang group reflects [i(ː)n] (with a likely long vowel associated with a deleted velar) and Neverver-Fifti-Nitita reflects [t(i)n] (with regressive assimilation in place of articulation to the following nasal stop). These differ from all the other languages in having completely oral stops (or expected reflexes) in the onset of this final portion.

Since all of the languages are spoken in the Northern part of Central Malekula, that is, just south of the dog’s neck, I classify the languages as North-Central Malekula (NCM) languages – an arbitrary name based on current human geography. Ninde’s membership in this family is complicated by its status as a mixed language that, as recorded in oral tradition, inherits grammar and lexicon from Big Ninde and Small Ninde. The latter is likely part of this NCM group for reasons detailed in §5.

Language:	Proto	Avava Bangaasak Salang	Neverver Fifti Nitita	Ninde					
Doculect:		Proto	Proto	Proto	Windua	Lawa	Opmo bamba	Labo	Southwest Bay
‘come’	*vʷel	*vʷel	*vul(em)/ *ʙulem	*wul	–	-pul	-pul	-βul	-pul
‘mouth’	*noboŋo-	*boŋo-	*noboŋo-	*noboŋo	–	no-ᵐboŋo-n	no-ᵐboŋo-ne	no-ᵐbo-re	no-ᵐboŋo-ne
‘ear’	*nidiliŋa	*diliŋa-	*nidiliŋa	*nidiliŋa-	ⁿdlɪŋa-	nɪ-ⁿdɪliŋa-	ⁿdɪliŋa-	ⁿdɪlɪŋa-	nɪ-ⁿdlɪɪŋa-
‘back’	–	*Du-	*gale-	*nituko-	ni-tuku-	ni-tuku-	nɪ-tuku-	ni-tuko-	tuko-
‘star’	*vʷinmʷezeu	*vʷinmʷese	*nivʷinmʷizo	*nanmʷisi	na-nmʷisi	na-nmisi	na-nmʷisi	na-nmusi	na-nmʷusie
‘grass’	*čog-	(*sogosoŋ)	*čog(u)čog	*nesogo-nei	ne-soⁿgone	nɛ-soⁿgone	nɛ-soⁿgone	nɛ-soⁿgoni	nɛ-soŋone
‘right (side)’	*barVkin	*bari:n	*baratin	*des-motine*	*des-motne*	*ⁿdɛs-modnə*	*des-mʷotɪne*	*des-modne*	*desmotɪnɛh*
‘person’	*mukut	*mu:t	*numuɣut	*numügut	nu-mugʊt	nu-mugut	–	nu-mugʊt	aiwut
‘four’	*vač	*-vat	*-vas	*ves	-βɛs	-vɛs	-vɛs	-ves	i-vɛs
‘five’	*-lim	*-lim	*-lim	*selme	sɛlmɛ	-sɛlme	sɛlme	selme	sɛlɪmɛ
‘sleep’	*mVtur	*-matur	*matur	*mitoʁ	-mitox	-mitox	-mitox	-mɪⁿdox	-mɪtox
‘egg’	*kapʷi	*avʷV-ton	*nogopʷi	*namal̪ˤap*	*na-maðap*	*na-maðap*	*na-malap*	*na-maðap*	*na-maðap*
‘heavy’	*map	*map	*map	*lub*	*-luᵐʙ*	*-luᵐʙ*	*-luᵐʙ*	*-luᵐʙ*	*-luᵐʙ*
‘tooth’	*nilivʷV-/*bVrVx	*borox	*nilivu-*	*nibaʁi-a-	–	niᵐʙaɣi-a	–	–	–
‘another’	*tele(k)	*tele-	*dle	*tlaʔ	tlalis	tlaʔ	tlaʔ	–	–

Table 5: Comparative reference data for the C group, with reconstructed forms for the smaller subroups drawn from following tables.

Avava, Bangaasak, and Salang

The group made up of Avava, Bangaasak, and Salang (also identified by Crowley 2006a as a subgroup) (Table 6) has shared innovations for almost every change identified. All of these languages have no reflex of nominal *na- in multisyllabic nouns, but may retain the vowel of that prefix in monosyllabic forms. Reflexes of 'star' are nearly identical, with unrounded final vowels suggesting deletion of *u after *e or *i. Forms meaning 'egg' share a general shape *av^{w}o-to- (a low vowel, bilabial stop, non-low rounded vowel, a stop, and sometimes nasal stop representing the third-person singular possessor), but this has been greatly reduced by vowel deletion and subsequent fortition of *v in the coda, and/or place assimilation. The presence of an initial vowel in 'egg' is typical of words with an initial velar stop in cognates when the root is preceded by the nominal prefix *na-, making the form for 'egg' much like that of the following Nitita-Fifti-Neverver group. All of the languages have compensatory lengthening in vowels before deleted *k, as in the word for 'person'. Given this change, the long vowel recorded by Crowley and Lynch (2006a), and the higher-order reconstruction, I reconstruct such a long vowel for *b^{w}ari:n 'right (side)'. These shared changes strongly support this low-level group.

All of these languages have retained [t] reflexes of *t before word-final *i, whereas the other group C (NCM) languages have strident consonants in this position. This could suggest a bleeding order with deletion of word-final *i, which would place this group well outside of subgroup C (NCM) with a unique retention. Given evidence that Ninde – proposed as a sister to this group – formerly had *č reflexes of PNCV*ti (§6.3.1) restricted to word-final positions, it is more likely that *č can be reconstructed for group C (NCM) and yielded *t in Avava-Bangaasak-Salang ultimately *s in Ninde. A similar change would have affected the prenasalized affricate *z, which surfaces instead as a nasal stop word-finally as in 'banana' in Avava; this form is not included in the data used here, but *apm* in Avava (Crowley 2006a), and *nys* in Ninde (reconstructible as *navuz for both). This alternative analysis makes for a much simpler subgrouping.

Language:	Proto	Avava		Salang	Bangaasak
Locale:		Khatbol	Tisvel		Taremb
‘come’	*v^{w}el	-vɛl	-wɛl	-vɛl	-β^{w}el
‘mouth’	*boŋo-	^{m}boŋɔ-	^{m}boŋo-	^{m}boŋo-	^{m}boŋo-
‘ear’	*diliŋa-	^{n}deliŋa-	^{n}dɪliŋa-	^{n}diliŋa-	^{n}dɪliŋa-
‘back’	*Du-	^{n}druː-	^{n}d^{r}u-	^{n}d^{r}u-	^{n}dru-
‘star’	*v^{w}inm^{w}ese	finməse	f^{w}inm^{w}ɔze	f^{w}ilmɛse	f^{w}inmɛse
‘grass’	(*soɡosoŋ)	sokosoŋ	soŋosoŋ	–	*mom^{w}os*
‘right (side)’	*bariːn	bərin	^{m}baɾin	barlin	–
‘person’	*muːt	muːt	mut	muːt	muːt
‘four’	*-vat	i-vat	i-vat	i-fat*	i-vat
‘five’	*-lim	i-lim	i-lim	i-lim	i-lim
‘sleep’	*-matur	-matur	-matur	-matur	-matur
‘egg’	*av^{w}V-to-n	ɔpm	ovoto	autn	ɔbm
‘heavy’	*map	-maː-p̄	-map	-map	-map
‘tooth’	*borox	^{m}boroh	^{m}borox	*lit-*	borohðei
‘another’	*tele-	telenan	telenan	telemai	telenan

Table 6: Comparative reference data for Avava, Salang, and Bangaasak.

Ninde is either a close relative to this group or has been in extensive contact with these languages. A number of innovations are shared between subsets of the languages. Notably, Avava, Bangaasak, and Ninde are the only doculects on Malekula that could share a form that can be reconstructed as *borVx ‘tooth’. The form for ‘star’ suggests a shared loss of the final vowel of word-final diphthong *{i,e}u. Ninde was clearly affected by contact, but were it not for loanwords from a variety of sources, it may be classified as more closely related to Avava than Neverver.

Nitita, Fifti, and Neverver

The remaining languages in this subgroup also form a strong low-level group: Nitita, Fifti, and Neverver (Table 7). They all have generally disyllabic forms of ‘come’, usually with a fully nasal consonant that other languages included in group C (NCM) do not have; coronal place of articulation in the final two stops of ‘right (side)’; otherwise fricative reflexes of *k as in ‘man’; and unique lexical forms of the shape *gal- ‘back’.

There is evidence both for and against shared lenition of velar stops with the Neverver-Fifti-Nitita subgroup and Ninde. Since those languages all delete *k in many post-vocalic environments, it is possible that all shared an intermediate velar fricative. On the other hand, non-lenited forms are in complementary environments across the larger group C (NCM): forms for ‘egg’ begin with *g in this group and this could be prenasalized after the nominal *na-; together with the forms (*av^{w}o-to) in Avava-Bangaasak-Salang, this suggests a uniquely shared lexical innovation *kav^{w}V ‘egg’. Ninde shares this post-nasal fortition of *k in other cognates, but additionally in the root-medial environment after *m (as in ‘person’). Finally, the set representing reflexes of *baratin ‘right (side)’ in this group would most likely have been inherited as *b^{w}arVkin, with the penultimate consonant *k assimilating in place to *t with the final consonant *n (and/or the preceding *r). This leaves only the context of (allophonically oral) low vowels where *k could be lenited to *x or *ɣ in group C (NCM). Tentatively, *x is reconstructed for PNCM, but as an allophone of *k in the context of low vowels.

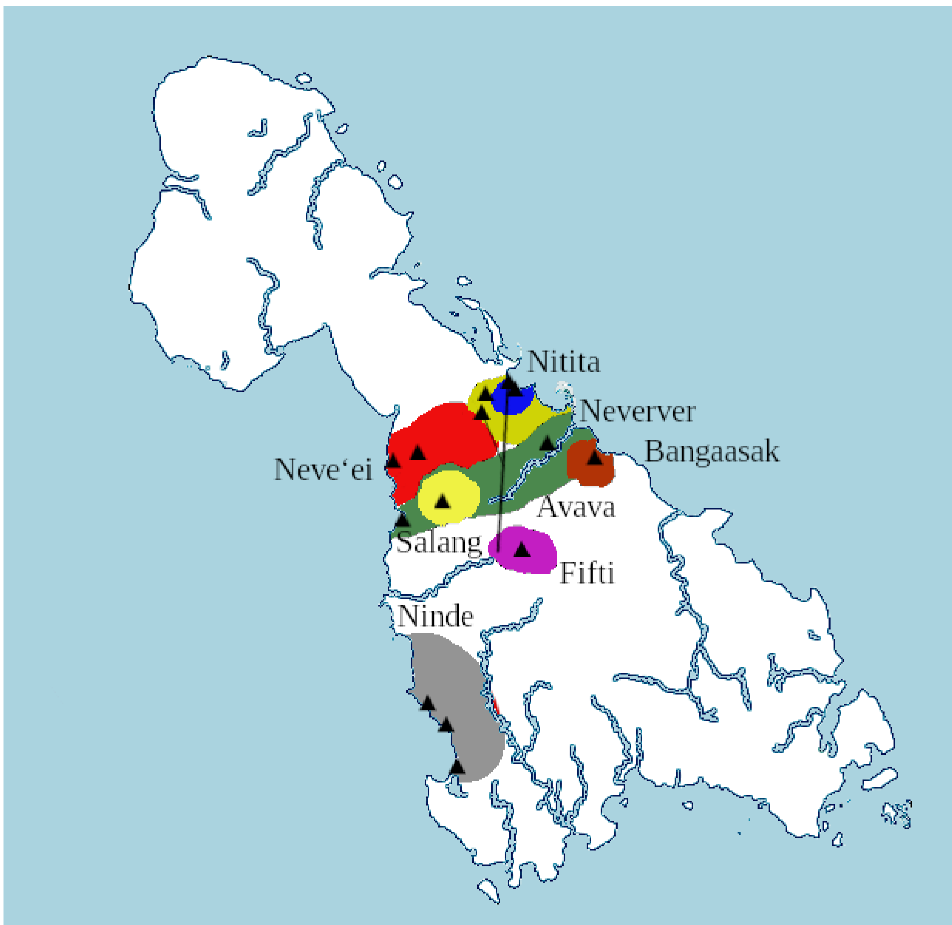

Figure 3: Approximate geography of group C, the North Central Malekula languages.

	Proto NFN	Neverver			Fifti		Nitita
		Sakan	Limap	Mindu	Khatbol	Malmbor	Khatbol
‘come’	*vul(em)/ *ʙulem	^{m}ʙulɛm	-^{m}ʙulɛm	^{m}ʙulɛn*	-v^{w}ɛl	^{m}ʙuel	-vunɛl
‘mouth’	*noboŋo-	no^{m}boŋo-	nomoŋo-	nomuŋo-	nimoŋo-	nimoŋo-	nʊ^{m}boŋo-
‘ear’	*nidiliŋa	nɪ^{n}dliŋa-	nɪ^{n}dliŋa-	nɪ^{n}dliŋa-	^{n}dɪlɪŋa-	nɪ^{n}dɪlɪŋa-	^{n}dɪlɪŋ-
‘back’	*gale-	$^{\text{ŋ}}$galɛ-	ni$^{\text{ŋ}}$gallɛ-	$^{\text{ŋ}}$galɛ-	$^{\text{ŋ}}$galda-	$^{\text{ŋ}}$galɛ-	ne-$^{\text{ŋ}}$kaðɛ-
‘star’	*niv^{w}inm^{w}izo	nɪvinəmʦɔ	nɪv^{w}inmʧɔ	–	nu$^{\text{ŋ}}$gunm^{w}uzo	f^{w}ɪnm^{w}ɪsɔ	nɪv^{w}ɪnimɪsor
‘grass’	*čog(u)čog	nu-soŋsoŋ	nʊ-^{n}t͡sɔŋt͡sɔŋ	–	ni-so$^{\text{ŋ}}$kuso$^{\text{ŋ}}$k	nu-soŋusoŋ	ni-so$^{\text{ŋ}}$kuso$^{\text{ŋ}}$k
‘right (side)’	*baratin	barat-n	barat-n	barat-n	^{m}barati-nia	^{m}baratɪn	barat-n
‘man’	*numuɣut	nu-muɣʊt	nu-muɣman	nu-muɣʊt	nu-mʊɣʊt	nu-muɣʊt	nə-mʊɣʊte
‘four’	*-vas	i-vas	i-vas	i-vas	i-vas*	i-vas	i-vas
‘five’	*-lim	i-lim	i-lim	i-lim	ɪ-ldin	i-lim	i-lin
‘sleep’	*matur	-matur	-mætur	-matur	-matur	-matur	-matur
‘egg’	*nogop^{w}i	nu$^{\text{ŋ}}$govi-	no$^{\text{ŋ}}$govi-	–	nu$^{\text{ŋ}}$koβ^{w}ɪ-	no$^{\text{ŋ}}$govi-	no$^{\text{ŋ}}$gop^{w}i-
‘heavy’	*map	-maː-p̚	-map	–	-map	-map	-maɸ
‘tooth’	*nilivu-	nuluvu-	noluvu-	nɛluvu-	nilbu-	nilɪβu-	niliβu-

	Proto NFN	Neverver			Fifti		Nitita
'another'	*dle	a^{n}dlean	–	–	–	–	–

Table 7: Comparative reference data for Neverver, Nitita, and Fifti.

D.

Languages: Larevat, Naman, Natangan, Neve'ei, Wala (Worprev)

The languages excluded by the main subgroups – those that do not have forms of the shape *b^{w}arV{g,k,t}i{ŋ,n} 'right hand' and that have unreduplicated forms of the shape *m^{w}učiu 'star' – are Larevat, Naman, Natangan, Neve'ei, and the Wala doculect identified as Worprev (but none of the others identified as Wala) (Table 8). They are unlikely to be a unified group based on the evidence collected here, since these are examples of non-participation in these changes. Natangan and Worprev (Wala) may be closely related: both have a fricative reflex /ɣ/ for the velar stops (including the nasal stop *ŋ) and /s/ for the affricates *z and probably *č. Additionally, the form [^{ŋ}gos] 'sleep' in Natangan may be (derived from) a prefixed form of *p^{w}ač, which Worprev also inherited or borrowed as *pas*.

This group, although not asserted in this chapter as a clade, is nonetheless consistent with prior work. Tryon (1976:87) places Larevat, Naman, and Neve'ei in his Malekula

Central subgroup of Malekula Interior and Lynch (2016) in his Central Western linkage within the greater Western Malakula linkage. There is no absence of shared innovations with other doculects in this group, but the patterns do not support subgrouping; one or all of the languages may be shaped by extensive contact.

Language:	Proto E	Larevat		Naman	Natangan	Wala	Neve'ei
Doculect:		Fwishile	Larevat/ Mosox		Mbonvor	Worprev	Coast/Vinmavis
'come'	*vʷel/*bʷel	-fɛl	-ᵐbɛl	-vale	-val	-ᵐbʷɛl	-{fʷ,vʷ}ɛlɛm
'mouth'	*boŋo-	ᵐboŋor	–	ᵐboŋo-	noᵐborɣɛ-	nɪᵐboɣo-o-	noᵐboŋo
'ear'	*daliŋa	ⁿdɛliŋa-	–/ⁿdʳɪliŋa-	ⁿdɛliŋa-	ⁿdalɪŋa-	nɪⁿdaliŋa-	naⁿd(ɪ)laŋa-
'back'	*bile-/ *taɣV-	ᵐbəlɛ-	ᵐbɪlɛ-	ⁿtɛru-	nɛtaɣə-	nɪtaɣo-	n{e,ɛ,ɪ}taa-
'star'	*mʷaze	manʧe	mnʧe	monʧe	namʷase	numʷosi	nɪmʷɪⁿʧi
'grass'	*velvelus/ *nimʷeney	vɛlvɪlɪs	–/n-vɛlvɛlɪs	ni-βɛlβɛlʉs	namanai	nɪ-mune	nɪ-mʷ{i,ɪ}ni
'right (side)'	*metu/ *bʷaradaŋ	ⁿdɪlɪmɛt	ⁿdʳɪlɪmɛt	xo-metu	natiⁿdaŋ	ti-ᵐbʷariŋin	b(ʷ)(ɛ)raⁿdaŋ
'man'	–	moɣman	–/moxɔman	moɣʊt	nɪmæɣət	nɪ-maɣiᵐbʷariŋin	nɛ-mʷɛn/ nu-mor~ nu-mʷur
'four'	*-vas	fes	vɪs/vɛz	i-bɪs	ivas	i-ᵐbas	{bɛ,i}-va(h)
'five'	*-lim	tlɪm	lɪm	i-ləm	i-lme	i-lim	{ba,i}-lim
'sleep'	*matur	-metr	–/-metr	-metur	ᵑgos	-pas	-m{a,æ}tu{r,ɾ}
'egg'	{*ma/ *ʔa}- *adile	naranⁿdʳɪz-	–/naranⁿdʳɪl-	naʔaⁿdl*	manlɛ-	nɪmaⁿdɛlʊ-	naʔaⁿd(a,ə)lɛ-*
'heavy'	*dov/*map	-ⁿdʳav	-ⁿdʳak	-ⁿdup	-nɛmb	-ⁿdoɸ	-map̄
'tooth'	*livʷo-	lɪv-	–/nlv-	nilβɪ-	nɛluvʷo-	niluvo-	nolo{f,v,β}u-
'banana'							
'another'		–	–	tiⁿ͡tʃɛᵐbənɛn	–	–	–/ telinɛn~ telimam

Table 8: Comparative reference data for the D group. Larevat (Larevat and Mosox doculects) share a column, and Neve'ei (Coast and two Vinmavis doculects) share columns language-internally, with abstract form representations that cover all the variation.

E.

Languages: Novol, Nombotkote, Letemboi, Tesmbol, Angavae

Marginal: Nesarian, Natingatlang variety of Letemboi

The second high-level subgrouping includes mainly languages of the southern interior and the western coast north of South West Bay (Dixon Reef) (Table 9). All of these languages have partially reduplicated forms of the shape [vVn-vV{t,d}(V)] 'star', and except for Natingatlang, all more narrowly have reflexes of a form that can be reconstructed as *vunvut 'star' in forms of the shape [v({ʊ,u,æ})n(ʊ)-v{u,ʊ}t]. All excluding Natingatlang and Nasarian have reflexes of *nVb 'heavy' in forms of the shape [{n,l}Vb], and all of the languages have reflexes of *m^{w}one 'sleep' in forms of the shape [m^{w}{ɔ,o,ə}n({ɛ,ɪn})] (but Natingatlang has [gas], associated with languages not discussed in this chapter). The final syllables of 'right (side)' is [ŋən{ə,ɛ}] in all of these languages except Novol (where it is [ŋaŋ]) and Angavae (where the word is [tuai]); this would reflect a fully nasal reflex of *g in the putative inherited form *b^{w}aragin of the proto language for the largest group (group A).

	Proto E	Novol			Nombotkote		Tesmbol		Angavae
		Tavendrua	Bangir	Dixon Reef	Blacksands/ Melkin	Neioleian	Melaklak	Usus	Dixon Reef
'come'	*v^{w}Vl	-βol	-β^{w}əl	-βol	-ɸul	-wul	-b^{w}alɛ	-βal	-^{m}bolə
'mouth'	*boŋoi	^{m}b^{w}oŋi	a^{m}b^{w}oŋ	^{m}boŋə-	na^{m}boŋ{ɔ,ə}-	no^{m}boŋoi	^{m}boŋi	^{m}buŋi	hə^{ŋ}garɪ-
'ear'	*daliŋa	hɛ^{n}dɛliŋi	-^{n}dɛliŋa-	-^{n}dɛliŋa-	ɛ-^{n}dɛli{ŋ,^{ŋ}ga}-	-^{n}dɛli^{ŋ}gai	^{n}daliŋa-	^{n}daliŋa-	hɪ-^{n}dɛliŋa-
'back'	*tak-	ɛta:-k-	ata:-g-	atax-	nɛta{k,xə}-	ataɣi	taxo-	tag-	hatag-
'star'	*vunvut	nɛvunvʊt	nɛvunvʊt	ɛvunvʊt	hanɛvnvʉt/ ɛvʊnvʊt	nɛvʊnvʊt	vɛlvit	ɛvɪnvɪt	hɛvʊlvʉt
'grass'	*m^{w}onai	*æ-^{n}d^{r}amb*	a-mɔnai	*a-^{n}d^{r}amp*	n{a,ə}-m^{w}o-nai	–	m^{w}əne	m^{w}ɛle	–
'right (side)'	*b^{w}araŋəne/ *tub^{w}ai	boraŋənɛ	boraŋaŋ	tetuboi	boraŋənə	tuwai	at^{m}b^{w}i	at^{m}b^{w}i	–
'man'	*mukut	a-mugʊt	na-muɣʊt	a-mgʊt	na-mugʊt/ a-mʊɣʊt^{h}	nu-mʊt	mut	na-mʊt	hɛ-mægʊt
'four'	*-vas	i-vas	i-ɛvaz	i-vas	i-va(ˑ)s	i-vas	-^{n}bɛs	-^{n}bɛs	i-vas
'five'	*-lime	i-ləmɛ	i-ɛləmɛ	i-ləme	i-l(ə)mɛ	i-lɪmɛ	-limɛ	-limɛ	*i-sɛlmɛ*
'sleep'	*m^{w}one	-m^{w}ɔnɛ	-m^{w}ɔnɪn	-m^{w}ɔnɛ	-m^{w}onɛ	-m^{w}ɔnɛ	-m^{w}ɔn	-m^{w}ɔnɛ	-m^{w}onɛ

	Proto E	Novol			Nombotkote		Tesmbol		Angavae
		Tavendrua	Bangir	Dixon Reef	Blacksands/ Melkin	Neioleian	Melaklak	Usus	Dixon Reef
‘egg’	*mʷovʷi	namʷovi	amɔvʷi	amovʷi	nomʷovi/ namʷop	–	lɛti	lɛti/moβi	hamovʉl
‘heavy’	*neb	-nɛmᵇ	nɛm	-nɛmp	-nɛmᵇ	-nɪmb	-nɪᵐʙ	-nɪᵐʙ	-nɛmᵇ
‘tooth’	*livʷi-	ɛluwi	hɛlui	ɛluvo-	nɛl(u)vʷ{o,ə}	-nɛlui	nɛlvʷi	nɛlwi	hɛluvo-
‘another’	–	–	–	–	–	–	anbi	atli	–

Table 9: Comparative reference data for the E group, with reconstructed forms taking into account only core group members. These may be compared with Proto Letemboi reconstructions from Table 11. Of Nombotkote forms, Blacksands and Melkin are combined into a single column.

This makes for a relatively higher level group (Table 10) including Letemboi, Nasarian, Natingatlang (a divergent doculect of Letemboi), and, based on the form for ‘star’, Nevitangiene (not identified with any language by Kaiar, whose traditional home is in Letemboi). It could be that Letemboi is an ethnogeographic identity more than it is a linguistic one. This group is represented in three historical layers, with Natingatlang sister to group E, but comparatively less evidence supports this than the other subgroups proposed in this chapter.

	LNN	Letemboi	Letemboi	Nasarian	
	Proto	Proto	Natingatlang	Ngava	Wileven
‘come’	*v^{w}ale	*wul/*bul	-v^{w}əle	-^{m}b^{w}ɛl	-volɛ^{h}
‘mouth’	*bo(ŋV)-	*b^{w}o(ŋV)-	*nə^{ŋ}gar*	^{m}boŋo-	*haŋaram*
‘ear’	*daliŋa	*deleŋV-	ɛ^{n}daliŋa-	^{n}dɛliŋa-	hɛ-^{n}dɛliŋa-
‘back’	*tak-	*takV-	nɛt^{h}a:-g-	atak-	hætag-
‘star’	*vVnvVdV	*vunvut	navanvɪ^{n}de	ɛvunv^{w}ut	havænʊvʊt
‘grass’	*m^{w}onai	*m^{w}onai	ɛ-vɛnm^{w}onai	a-m^{w}onai	ha-^{n}d^{r}am
‘right (side)’	*b^{w}araŋVnə	*b^{w}araŋənə	^{m}boraŋənə	^{m}boraŋanə	boraŋənɛ
‘person’	*makut	*mVkut	a-mugʊt̚	mo:t	ha-magʊt
‘four’	*-vaš	*-vas	i-vas	i-vas	i-va·ʃ
‘five’	*-lime	*ləme	i-ləme	i-ləme	*i-ʃɛlmɛh*
‘sleep’	*m^{w}one	*m^{w}one	*^{ŋ}gas*	-m^{w}on	-m^{w}onɛ
‘egg’	*m^{w}op^{w}i	*m^{w}op^{w}i	nam^{w}alaɸ	amoʙi	amoβə-
‘heavy’	*neb/ *(me)mave	*neb	-nam^{b}	-mɛmɛ	-mɛmaβɛ
‘tooth’	*liv^{w}o	*liv^{w}i-	nalupo-	nɛluv^{w}o-	hɛlɛβə-
‘another’	–	–	–	–	–

Table 10: Comparative reference data for Letemboi, Natingatlang (Letemboi) and Nasarian (LNN), with reconstructed forms for the smaller subroup of core Letemboi languages drawn from following Table 11.

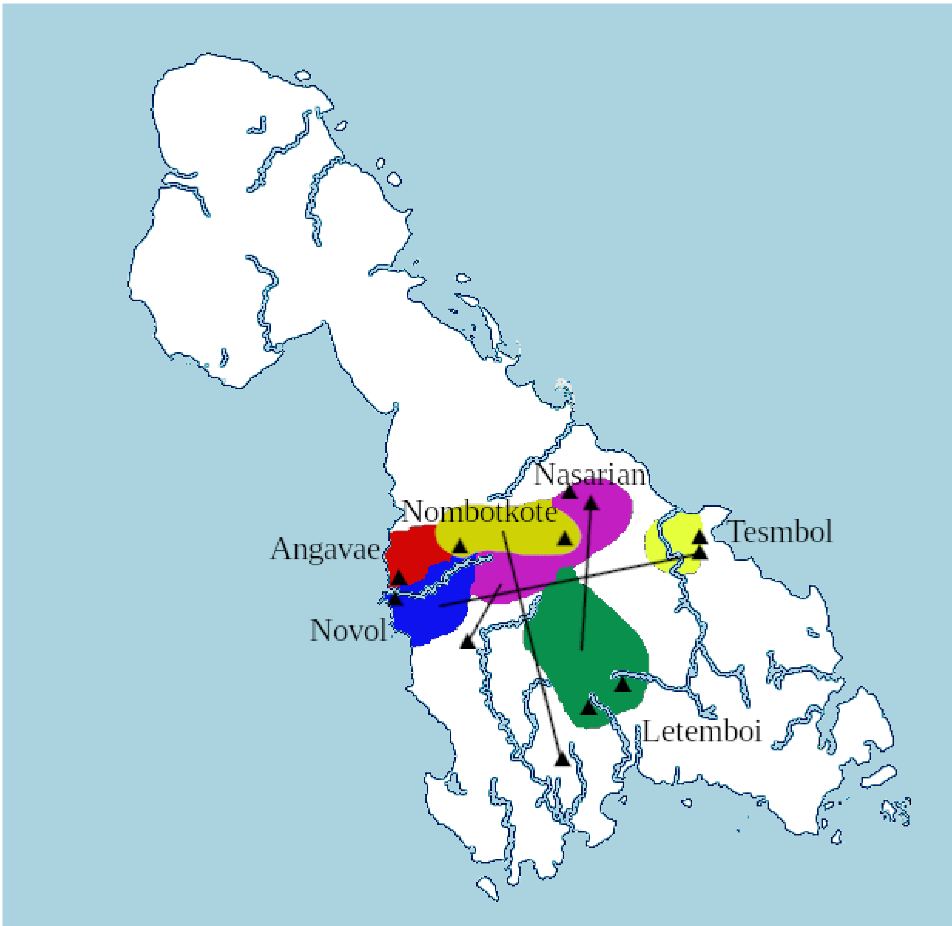

Figure 4: Approximate geography of group E.

A Proto Letemboi lexical set is tentatively reconstructed in Table 11 from just the doculects of Aingelemolesa, Newotenyene, and Nevatanyene. This excludes Natingatlang, which patterns more closely with Nesarian and shares the form for 'star' that we recorded from Nevitangiene. The motivation for reconstructing this node for the remainder of Letemboi doculects is mainly the shared linguistic identity; the starting assumption was that

Letemboi represented a single language, but this was not supported for varieties outside of these three.

	Proto Letemboi	Letemboi		
		Aingelemolesa	Newotenyene	Nevatanyene
‘come’	*wul/*bul	-wul	-bul	–
‘mouth’	$*b^{w}o(ŋV)$-	$na^{m}b^{w}ɔŋi$	$ɪ^{m}b^{w}ə$-	$a^{m}ʙura^{ŋ}ga$-
‘ear’	*deleŋV-	$ɛ^{n}dɛlɛŋi$	$^{n}dɛliŋa$-	$ɛ^{n}dɛlɛgɛ$-
‘back’	*takV-	nata:-ki	nɛtag-	atægə-
‘star’	*vunvut	nɛvʊnvʊt	ɛvnvʊt	ɛvʊnəvʊt
‘grass’	$*m^{w}onai$	$nɛ\text{-}m^{w}onai$	$nɛ\text{-}m^{w}onai$	–
‘right (side)’	$*b^{w}araŋənə$	baraŋənə	boraŋənə	boraŋənə
‘man’	*mVkut	ha-mogʊt	a-mugʊt	ɛ-mægʊt
‘four’	*-vas	-fas	i-vas	i-vas
‘five’	*ləme	i-ləmɛ	i-ləmɛ	i-ləmɛ
‘sleep’	$*m^{w}one$	$\text{-}m^{w}onɛ$	$\text{-}m^{w}ənɛ$	^{ŋ}gas
‘egg’	$*m^{w}op^{w}i$	namobi	amɔwi	$nam^{w}ɔvi$
‘heavy’	*neb*	$\text{-}nɛm^{b}$	$\text{-}nɛm^{b}$	$\text{-}la^{m}b$
‘tooth’	$*liv^{w}i$-	$nɛləβ^{w}i$	ɛluə-	ɛluvu-
‘another’	–	iaŋasua	–	–

Table 11: Proto Letemboi reconstruction using data from the varieties Aingelemolesa, Newotenyene, and Nevatanyene.

There are sporadic shared innovations with Ninde in this larger group, but they are not structured. Ninde *wul* ‘come’ is in line with the monosyllabic cognate found in many of these languages. The word *namal̪ˤap* ‘egg’ is uniquely like the form [namʷalaɸ] found in Natingatlang, and *selme* ‘five’ (bearing the erstwhile prefix forming numerals 6-9 from 1-4) is shared with only the Wileven variety of Nesarian. Nonetheless, many of the other unique lexical and phonological features of this set cannot be found in Ninde. For this reason, Ninde is not included here even as a marginal member, even though many speakers identify Dixon Reef languages as similar. A possible reason for the identification with these languages is in former similarities in suprasegmental phonology – pervasive vowel devoicing – that could explain several dramtic changes affecting Ninde (§6.3).

F.

Languages: Nahavaq, Naati

Marginal: Ninde, Naha’ai/Na’ahai, Navwien

The final group of languages includes Nahavaq and Naati as a small South West Bay clade defined by debuccalization of oral velar PNCV *k > [ʔ] everywhere, of prenasalized velar *ɡ > [ʔ] in the coda (excluding the first-person possessive suffix /-ɡ/), and of PNCV *s > [h] where it was not palatalized; Navwien and Na'ahai/Naha'ai share with Nahavaq and Naati (Table 12) a merger of oral and prenasalized velars that could have preceded such a change, but would have been extended to intervocalic environments after divergence, if these languages form a clade. Both these latter languages have also merged the nasal velar *ŋ with the oral *k and prenasalized *ɡ (at least intervocalically). The inclusion of Naha'ai/Na'ahai in this group is complicated by the retention of PNCV *ma(i) 'come' as [mi], but otherwise it has many affinities with Navwien. Navwien – unusually for this group – does not have glottal stops, but *does* participate in the innovation of *vʷale 'come' with subgroup A. All of the languages in this group (except Ninde) have a final syllable in 'right (side)' of the shape [{ŋ,ɡ}{i,ɪ}n], reflecting a generally conservative word of the shape *bʷaraŋin. All of the languages except for two doculects of Nahavaq, have a fully nasal reflex of *ɡ.

Ninde unexpectedly shares some changes with this set. These include a form for 'back' of the shape *-taku or *-taʔu. Even though the glottal stop of those languages corresponds to a velar stop in Ninde, an inherited velar from the common ancestor of group C (North Central Malekula) is expected to be deleted. Multiple sources of velars in Ninde also help to explain /k/ that has resisted prenasalization, as in *niki-a* 'name' (c.f. Naati *neʔia* and Nahava *neʔ(e)he*). Elsewhere, particularly in environments with no high vowels, Ninde also has glottal stops. Many of the reflexes of PNCV*s have been deleted in Ninde, but there is evidence that they were once intermediately *h based on changes to the lateral approximant discussed in §6. Ninde also has a raised first vowel in *mitoʁ 'sleep' and *-diliŋa- 'ear', which can be more easily reconciled with the forms *metur 'sleep' and *deliŋa 'ear' of this group than with widespread retained low vowels of PNCV*matur and PNCV*daliŋa.

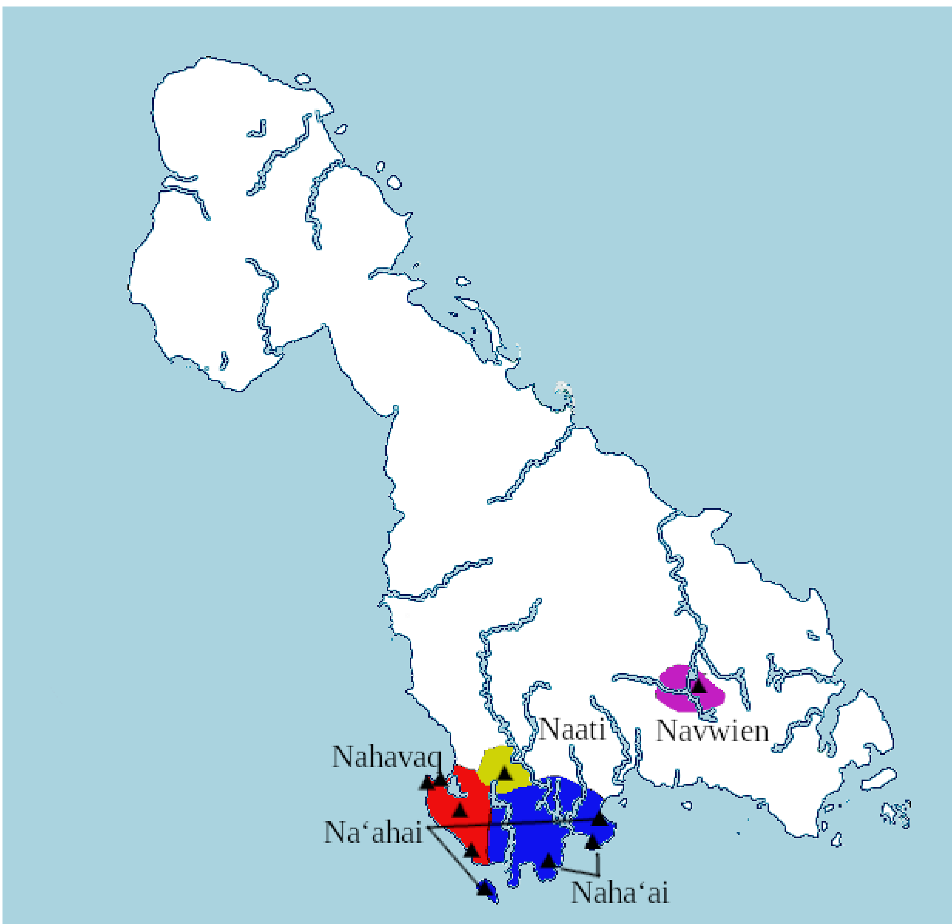

Figure 5: Approximate geography of group F.

	Proto F	Navwien	Naha’ai		Na’ahai	
		Mbonvor	Mbatmbang	Malfaxal	Toman	Mbatvanui
‘come’	–	-^{m}b^{w}ɛl	-mi	-mi	-mi	-mi
‘mouth’	*boŋo-	nɪ^{m}boɣo-	no^{m}boʔo	no^{m}boɣo-	no^{m}boʔo-	no^{m}boɣo-
‘ear’	*daliŋa-	nɪ^{n}daliŋa-	nɛ^{n}dalɪŋã-	na^{n}dalɪŋa-	^{n}dalɪŋa-	^{n}dalɪŋa-
‘back’	*tako-	nɪtaɣo-	nɛtaʔo-	nataɣo-	tao-	nataɣu-
‘star’	*m^{w}ošoi	num^{w}osi	nom^{w}osoi	nam^{w}osoi	nɪm^{w}osoi	nom^{w}ɔsoi
‘grass’	*m^{w}VnV	nɪmune	ni-m^{w}ini	ni-m^{w}ini	nim^{w}ini	nɪ^{m}ʙun_m^{w}ini
‘right (side)’	*baragin	ti^{m}b^{w}ariŋin	^{m}baraŋin	^{m}barəŋin	^{m}baraŋin	^{m}barəŋin
‘man’	*ba-b^{w}aragin	nɪ-maɣi^{m}b^{w}ariŋin	na-^{m}ba^{m}bariŋin	na-^{m}ba^{m}bariŋin	nɛ-^{m}bo^{m}bariŋin	^{m}ba^{m}bariŋin
‘four’	*-vas	i^{m}bas	i-væs	i-vas	i-væs	i-vas
‘five’	*-lim	i-lim	i-lim	i-lim	i-lim	i-lim
‘sleep’	*pas	pas	-pas	-pas	-pas	-pas
‘egg’	*(ma)-adelu	nɪma^{n}dɛlʊ-	namanlʉ-	namanlu-	manlɛ-	namanlɛ-
‘heavy’	*dip^{w}	-^{n}doɸ	-^{n}dop̚	-^{n}dop	-^{n}dop	-^{n}dop
‘tooth’	*liv^{w}o-	niluvo-	nɪliv^{w}o-	nilivo-	niliv^{w}o-	niliwo-
‘another’	–	–	ahɛn	–	–	–

Table 12: Comparative reference data for the F group, with reconstructed forms for the smaller subroups drawn from following tables.

The core of this group is Nahavaq and Naati, which can be reconstructed with few problems for the data here (Table 13). The only exceptions are reflexes for ‘egg’ and ‘another’ – Nahavaq *tišile~tisile* ‘another’ resembles many of the other languages’ forms in group A, but Naati is clearly not cognate. Naati m^{w}a^{n}dılı- ‘egg’ is the only form to have a preceding element [m^{w}a-] rather than [m^{w}av^{w}i-]. These two languages are very likely to form a group, even though Lynch (2016) had considered Naati to be its own divergent branch within the Western Malekula Linkage.

	Proto NN	Nahavaq				Naati
		Loorndavo	Caroline Bay	Banour	Lembinwen	Windua
‘come’	*v^{w}eleŋ	-β^{w}ɛlɛn	-wılɛŋ	-β^{w}ɛlɛn	-β^{w}alɛŋ	-^{m}b^{w}iliŋ
‘mouth’	*boŋo-	^{m}boŋo-	^{m}boŋo-	^{m}boŋo-	boŋo-	^{m}boŋo-
‘ear’	*deliŋa-	^{n}diliŋɛ-	^{n}diliŋɛ-	^{n}diliŋɛ-	^{n}dɛliŋɛ-	^{n}deliŋa-
‘back’	*taʔu-	taʔo-	taʔu-	taʔo-	taʔo-	taʔu-
‘star’	*m^{w}oši	musi	wuši	nimuši	nimuʃi	num^{w}osi
‘grass’	*m^{w}ane	m^{w}ane	m^{w}əne	nu-m^{w}ɛne	–	nı-muni
‘right (side)’	*baragin	^{m}baraŋin	^{m}baraŋın	^{m}bara^{ŋ}gin	^{m}bara^{ŋ}gin	^{m}baraŋin
‘person’	*morot	morot˺	m^{w}orot	morot	morot̚	nɛ-murut
‘four’	*-ves	i-vɛs	i-vɛs	i-vɛs	i-vɛs	i-vɛs
‘five’	*-lim	i-lim	i-lim	i-lim	i-lim	i-lim
‘sleep’	*metur	-mɛtur	-mɛtur	-mɛtur	-mɛtur	-mɛtur
‘egg’	?-*adili	nɔʔo-ɔnli-	noʔonli-	nɔʔo-ɔnli-	nɔʔo-ɔnli-	m^{w}a^{n}dılı-
‘heavy’	*dip	-^{n}dip	-^{n}dip	-^{n}dip	-^{n}dip	-^{n}dip
‘tooth’	*liv^{w}o-	liv^{w}o-	livo-	livo-	luwo-	neliv^{w}u-
‘another’	–	–	tisile	tiʃile	–	amataŋ

Table 13: Comparative reference data for the F group core languages: Nahavaq and Naati.

3.4 Discussion

3.4.1 Implications given data

The subgroups proposed in this chapter were those that emerged without knowledge reported in reference grammars, since this could have biased the analysis toward a phylogeny that groups documented languages. Expanding the view to consider what is known from documentation raises some additional points of similarity and points of difference. For example, a shared trait in this larger area is accretion in verb roots for all persons of a third-person singular subject prefix *i- including Naman (Crowley 2006b:66) and Ninde; or the existence of one inflectional class that is differentiated from another by the addition of /i/ to the pronominal prefixes in Avava (Crowley 2006a:68-73); Neverver has a similar alternation, but it can be explained in terms of the sonority of the verb-initial consonant (Barbour 2012:167-173). In this data set, this change can be inferred from unexpected palatal /j/ reflexes of initial consonants or those that have an inserted *i following them.

Neverver is unique in that it has geminate consonants (Barbour 2012:42), but they may be an old feature of the NCM languages. If that is the case, Neverver alone has retained, if not expanded them in its phoneme inventory. They can be identified where partially reduplicated verb forms (of Neve'ei or Naman) correspond to idiosyncratic stop reflexes in Avava and Ninde (where a lenited form is expected), but it is unlikely that the data exist to demonstrate whether geminates can be reconstructed to the whole North Central Malekula group.

3.4.2 Continuity with prior work

The outcome of this work is at odds with previous findings in some important ways, but many of the conclusions of prior work are supported by the findings presented here. Most of the distinct sound changes and lexical innovations of Malekula languages do not form clear isoglosses, except those that link varieties already established as languages or dialect chains. This supports the observation made by Lynch (2010), that most of the easily observable linguistic innovation in this region has taken place after languages already diverged. Comparatively few changes seems to be inherited by two daughters from a deeper clade. Nevertheless, the implication that the number of languages has been stable since the initial settlement of Vanuatu, or that diachrony has been shaped by contact over language divergence, is not supported for all languages of Malekula.

The proposal in this work is aligned in many ways with both Tryon's original phylogenetic grouping based on lexicostatistics and Lynch's linkages. Like Tryon, I propose that the languages of a vast part of northern Malekula are genetically related to languages of Santo more closely than they are with other Malekula languages. There are a number of languages that I have declined to place into any single group – these languages appear to be particularly shaped by contact. Ninde is exceptional in this regard: the oral history describing it as a mixed language, my familiarity with the language, and the appearance of layers of influence along a clear geographic trajectory support grouping Ninde into a clade with a set of languages it shares little in common with.

3.5 Conclusions

Considering the patterns of nested isoglosses and shared innovations, I have reconstructed a proto language ancestral to Avava, Bangaasak, Salang, Neverver, Nitita, Fifti, and (Small) Ninde as Proto North Central Malekula and Nahavaq, Naati, and (Big) Ninde as Proto South West Bay. This is the most fruitful choice of languages for reconstruction, since:

1. there is a compelling chronology of lexical and phonetic innovations for each of these groups without Ninde, yet otherwise unique innovations in both groups are shared with Ninde;
2. the languages that share unique changes with Ninde, but do not belong to either of these groups, are languages that today are spoken in the intervening regions, as would be expected from an associated history of migration;
3. the languages' isoglosses all appear to be nested within broader isoglosses, and are likely related at a deeper level, even if the subgroupings proposed here are inaccurate; and
4. these clades would be ancestral to some of the more extensively documented languages of Malekula (Neverver, Avava, Naati, and Nahavaq).

The contemporary geographic distribution of these languages hints at an old layer of inheritence from the North Central Malekula languages in Small Ninde. Assuming that proto homelands were located at the site of greatest diversity, the speakers of Small Ninde would have passed on their way to South West Bay through the area around Dixon Reef, where languages have unique isogloss bundles (including *m^{w}one 'sleep' and *nVb 'heavy') and sporadic innovations shared between especially Ninde, Natingatlang, Nevitangiene, Nasarian, Angavae, and Novol (*ja-lima 'five' and *na-movul 'egg'). This putative migration would end in South West Bay, where Ninde's new neighbors (Big Ninde) restructured its phonology, lexicon, and grammar. This appearance of a trajectory of migration for Small Ninde from a North Central Malekula region to South West Bay is demonstrated in Figure 6.

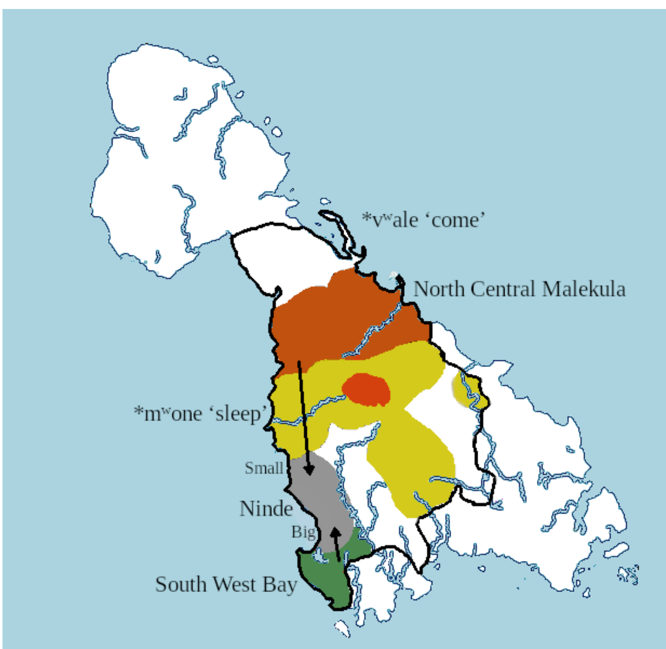

*Figure 6: Geography of the contiguous *v^{w}ale isogloss (outlined in black) with the main contributors to Ninde's mixed grammar and lexicon indicated by the colored fields.*

The characterization of Malekula languages as defying phylogenic subgrouping that is crucially *defined* by morphophonological innovation holds to a limited extent; the isogloss bundles that have emerged are indeed marked by criss-crossing isoglosses, but it is still possible, with great qualification, to trace inherited innovations. One of these difficulties is the number of exceptions that must be tolerated in order to do this. Most groupings are exceptionless for the diagnostic innovations that are highlighted here, but otherwise were limited to one exception per group. This cursory phylogeny is put to the test by reconstructing the lexicon with models of sound changes in §6. This phylogeny offers a starting point for reconstruction of Ninde's shared diachrony with other Malekula languages.

Compared to previous work, this approach to the data that were used yielded similar groupings where previous work by Tryon and Lynch had already agreed. This suggests that the morphological approach taken here largely reinforces the lexicostatistical approach and sound change isoglosses that have been used. Where the prior approaches are at odds, the approach taken in this chapter often failed to yield any deeper subgrouping at all; often, these were doculects that could be characterized as generally conservative in the lexicon. The most divergent clade proposed in this chapter is the one that includes (Small) Ninde. There is some support in prior work for a grouping that includes Avava and Neverver, but the two previous approaches differ with respect to their closeness. Nonetheless, the goal of this dissertation, and more specifically §7, is to reconstruct all known lexical items for Ninde and its relatives. This goal can only be accomplished with the more robustly documented languages, and

Avava, Neverver, and Nahavaq are comparatively well documented. They are implicated in some subgrouping in prior work – even if they are not closely related, all of the other languages in their groups are underdocumented.

4 Linguolabials as a diagnostic innovation: Evidence against the Western Malekula Linkage

Linguolabials are exceedingly rare speech sounds in the languages of the world, found at the highest density in Oceanic languages of Vanuatu. Those found in languages outside of the Oceanic family do not share this feature with any relatives: the Kajoko variety of Bijago (Niger-Congo, Guinea Bissau) (Olson et al. 2009) and Umotina (Bororoan, Brazil) (Ladefoged & Maddieson 1986). Like Kajoko, however, the linguolabials of Vanuatu were innovated from plain bilabials that contrasted with labialized bilabials (Maddieson 1987). Also like Kajoko, bilabials in loanwords tend to be realized as linguolabial, at least in V'ënen Taut (Fox 1979). Unlike Kajoko and Umatina, the Oceanic languages of Vanuatu are all related at some time depth or another and only they can give us a window onto how linguolabial consonants behave in diachrony subsequent to their innovation.

In Vanuatu, linguolabial consonants are found in languages spoken throughout Espiritu Santo island and in the north of Malekula island, which lies immediately to the south of Espiritu Santo (see Figure 7).

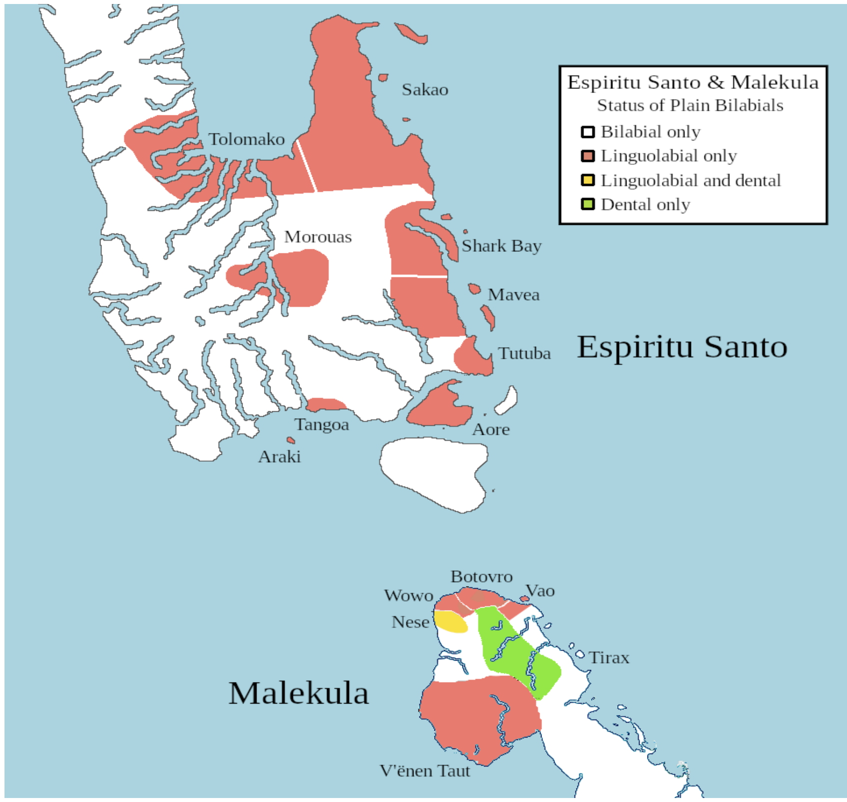

Figure 7: Geographic distribution of languages on Malekula and Espiritu Santo that have linguolabials or comparative evidence of past linguolabials from plain bilabials, represented in terms of the sound correspondences to plain bilabials in other Oceanic languages.

The Malekula languages are the focus of this chapter, because it is in these languages that the idea of shared inheritance of linguolabials has been rejected. These include:

- the closely related group made up of Vao, Wowo, and Botovro;
- Nese, which has some bilabial and dental reflexes where linguolabials are expected; and

- V’ënen Taut, which has linguolabials that correspond regularly to velarized (or plain) bilabials (as opposed to labiovelars, or labialized bilabials) in other Oceanic languages.

Of languages that do *not* have linguolabials, the consensus view based on sound correspondences is that Tirax has most likely had linguolabial stops prior to their merger with dental consonants (and possibly a fricative that merged with the bilabial fricative). Tape, which has no linguolabials, is the closest relative of V’ënen Taut, which does have them, so either it lost them or V’ënen Taut innovated them. Even in closely related languages, these speech sounds have complex distributions and messy reflexes.

Malekula languages have been classified according to two different schemes: Tryon’s (1976) phylogenetic model based on lexicostatistics and Lynch’s (2016) linkages model based on lexical and phonological change. Tryon had grouped all the languages of northern Malekula island with the languages of southern Espiritu Santo island (which lies just north of Malekula). Several of these languages across both islands have linguolabial consonants in their phoneme inventories, but Lynch and Brotchie (2010) reject the idea that linguolabials were inherited by all of these languages based on proposed bleeding relationships in sound change chronology. Lynch (2016) considers them to be a diffused areal feature and classifies the languages that have them in different linkages. The issue is particularly important for this dissertation, since his Western Malekula linkage includes Ninde – which does not have linguolabials – with a (phylogenetic) clade made up of Tape and V’ënen Taut. In the previous chapter (Chapter 3), a Tape-V'ënen Taut supgroup was reaffirmed as a clade, but not classified with any other subgroup represented on Malekula. This chapter seeks to justify the exclusion of Tape and V'ënen Taut in the reconstructions of Ninde's linguistic ancestors (Chapters 6-9) by providing an alternative to Lynch and Brotchie's (2010) explanation for areal diffusion of linguolabial consonants and reevaluating the criteria for linkage-hood used by Lynch (2016).

Importantly, Lynch's linkages model posits innovation-linked, rather than innovation-defined subgroups (Lynch 2000) associated with divergence and contact in periods of linguistic stasis. This work assumes that phonological correspondences are fundamentally irregular in this social context, because linguistic innovations are diffused incompletely across proto languages that were dialect continua (see Lynch 2006). A deeper analysis disentangling exceptionless, conditioned sound changes from irregular contact effects suggests that the sound correspondences seen in Northern Malekula languages are the result of mergers and loss of linguolabials and not irregular diffusion.

I refer to the sounds in question using the more familiar term *linguolabial*, but literature on the topic has frequently used the more specific term *apicolabial*. These sounds are produced with the tongue tip as the active articulator making full or partial constriction with the upper lip. In addition to being more articulatorily precise, the IPA transcriptions associated with *apicolabial* use the apical diacritic on bilabial bases <p̺ ^{m}b̺ β̺ m̺>, which facilitates visual assessment of cognate sets that include conservative bilabial consonants.

Lynch refers to the innovation of linguolabials from plain bilabials as the *apicolabial shift* throughout his writing. I use the coronal letter bases with the linguolabial diacritic <t̼ ⁿd̼ ð̼ n̼> instead – this is equally justified by regional diachrony, as linguolabials partially yield coronal reflexes in Tirax and Nese[10]. The convention that has emerged in Vanuatu for representing linguolabial consonants in practical orthography (at least by missionaries and linguists) is to use a corresponding labial base letter with an apostrophe: <p’ b’ v’ m’>. The diversity in orthographic representation arises, of course, because this place of articulation is relegated to a diacritic, rather than a set of base letters.

In this chapter, I make a case for shared inheritance of linguolabials in Nese and Tirax and identify subsequent conditioned change that creates the diachronic illusion of irregular distribution. The question of whether there is a Proto Malekula node is beyond the scope of this dissertation, and data from languages of Espiritu Santo have not been systematically studied. Since these two languages have been particularly challenging in prior work, the observation of regularity here opens up the possibility that linguolabial consonants can be reconstructed to a proto language with modern daughters spanning at least parts of Malekula and Espiritu Santo (and including V’ënen Taut and Tape).

The data for this study were painstakingly collected from reference grammars and assembled in cognate sets: Fox (1979a) for V’ënen Taut, both Crowley and Lynch (2006c)

[10] We have recorded an elderly speaker of Ninde who produced linguolabials in place of dentals, possibly using the lips as a substitute because he had lost his teeth – this is an example from the same region of labiodentals from dentals, albeit in “disordered” speech.

and Takau (2016) for Nese, Crowley and Lynch (2006d) for Tape, and Brotchie (2009) for Tirax. All of the reconstructions are drawn from Clark's (2009) Proto North Central Vanuatu (PNCV), because this is the lowest-level reconstructed language ancestral to all the languages discussed here and reflects changes at every level of linguistic structure since Proto Oceanic. These reconstructions have generally proven accurate in predicting phenomena in the daughter languages. Plain and labialized bilabials are reconstructed for PNCV as non-contrastive before rounded vowels. In at least one descendant, however, plain/velarized bilabials condition fronting of rounded vowels (Dimock 2007 on Nahavaq). Such an analysis works more effectively for northern Malekula linguolabials, as this chapter will later argue. When lexical data are taken from other sources, those are cited explicitly – but all descriptive claims are cited regardless of origin. This chapter also contains original claims based on data presented in those reference grammars.

To demonstrate the complexity of the sound correspondence sets involving at least one labial consonant between Nese, Tape, Tirax, and V'ënen Taut, they are shown in Table 14. There are 32 correspondence sets for the data, listed in order of frequency. In that table, the column bearing the heading "In set" reports the number of perfectly matched lexical cognate sets, which have the same gaps in lexical attestation (indicated by a question mark). Of these, only 20 have attested reflexes for each of the four languages in question (Nese, Tape, Tirax, and V'ënen Taut) and 25 of them (with any number of reflexes) only characterize one single cognate set each in the data. The "Broad" column offers a different count for the sets because it includes the sum of exact matches and partial correspondence sets that are not explicity different (because of missing data). The latter is redundantly counted wherever appropriate and includes singletons and cognate pairs.

An additional six full cognate sets can be presumed by combining complementary pairs of cognate sets – those that each have one missing reflex, but two overlapping corresponding reflexes. In Table 14, these are repesented as sets with 0 representatives (in the column with the heading “In set”) and are constructed in duplicate by the merger of two or three correspondence sets that have exactly three out of four languages represented. The highest number of cognate sets that fit a sound correspondence pattern is eight (with an additional maximum of seven partial sets that are compatible). In many cases, the unattested reflex is simply unattested because the semantic equivalent for a word-form does not appear to be cognate for a particular reflex.

Sound correspondence				Counts		Lexical example				
Nese	**Tape**	**Tirax**	**V'ënen Taut**	**In set**	**Broad**	**Nese**	**Tape**	**Tirax**	**V'ënen Taut**	**Gloss**
n	m	n	n̼	8	15	li**ne**	ilə**m**	-li**n**	lə**n̼**	'five'
ð̼	β	β	ð̼	3	17	**ð̼**anax	**β**ənax	-**β**naxɛ	**ð̼**ənaɣ	'steal'
m	m	m	m	3	12	na**m**at	nə**m**ot	n**m**at	nə**m**at	'snake'
n	m	m	n̼	2	16	na**n**ata-	**m**əte-	**m**ta-	**n̼**ata-	'eye'
^{m}b	^{m}b	^{m}b	p	2	5	nə**b**əŋ	nə**b**əŋ	**b**uŋ	na**p**(ə)n	'day'
n̼	m	n	n̼	2	3	nen̼e**n̼-**	mi**mi**-	ne**ne**-	nə**n̼-**	'tongue'
?	m	n	n̼	2	2	–	**m**eləx	-**n**elik	**n̼**alək	'dark'
β	β	β	ð̼	1	10	tas**β**e	sə**β**ərən	h**β**eliŋ	**ð̼**aran	'when?'
n	m	n	m	1	10	nai**ne**	nə**m**ax	nai**n**	nə**m**aɣ	'house'*
β	p	β	ð̼	1	7	le**β**	le**p**	le**β**	la**ð̼**	'take'
?	p	^{n}d	t̼	1	6	–	li**p**ax	li**d**ax	li**t̼**aɣ	'dog'

Sound correspondence: Nese	Tape	Tirax	V'ënen Taut	Counts: In set	Broad	Lexical example: Nese	Tape	Tirax	V'ënen Taut	Gloss
m	m	n	n̪̼	1	5	nu**m**	**m**ən	-**n**in	-**n̪̼**ən	'drink'†
m	m	n	n	1	4	ne**m**	nə**m**ax	ne**n**	nə**m**aɣ	'house of'*
β	Ø	β	Ø	1	4	nad̪̼ata**β**	pəte_	bta**β**	t̪̼ate_i	'breadfruit'
^{n}d̪̼	p	^{m}b	t̪̼	1	4	na**d̪̼**ataβ	**p**əte	**b**taβ	**t̪̼**atei	'breadfruit'
ð̪̼	p	β	?	1	4	nararа**ð̪̼**	darə**p**	dra**β**	–	'coral tree'
^{n}d̪̼	^{m}b	^{n}d	t̪̼	1	3	naxa**d̪̼**	ni**b**	na**d-**xan	na**t̪̼**	'fire(wood)'
m	m^{w}	m	m	1	2	**m**uloun	**m^{w}**əliun	**m**leun	**m**li(u)n	'chief'
^{n}d	p	^{m}b	t̪̼	1	2	**d**aɾaβ	**p**əre	**b**raβ	**t̪̼**arei	'long, tall'
n̪̼	m	n	m	1	1	-**n̪̼**ot	**m**it	-**n**ɛt	**m**ət	'black'
Ø	β	Ø	t̪̼	1	1	_tro	**β**ətir	_tur	-**t̪̼**tir	'stand (up)'

Sound correspondence				Counts		Lexical example				
Nese	Tape	Tirax	V'ënen Taut	In set	Broad	Nese	Tape	Tirax	V'ënen Taut	Gloss
^{n}d	β	^{n}d	t̼	1	1	xa**d**e	e**β**i	a**d**ɛ	i**t̼**i	'where?'
^{m}b	^{m}b	^{m}b	t̼	1	1	ne**b**ito-	(be)**b**ət	**b**ito-	**t̼**ət	'navel'
^{m}b	p	?	ð̼	1	1	**b**al**b**al	**p**əl**p**ol	–	**ð̼**əl**ð̼**əl	'fight'
?	m^{w}	m	m	1	1	–	**m**^{w}ili	**m**ul	**m**əl	'no longer'
^{m}b	β^{w}	?	p	1	0	na**b**ob	ni**β**^{w}ip	–	nə**p**ap	'sprouted coconut'
?	β	t	t̼	1	0	–	tə**β**aləx	buta**t**	ta**t̼-**	'grandmother'
Ø	β^{w}	?	?	1	0	nu_am	nə**β**^{w}ib	–	–	'fish-poison tree'
β	u	?	?	1	0	nana**β**	mama**u**	–	–	'yawn'
m	^{m}b^{w}	?	?	1	0	na**m**ʤo	nə**b**^{w}əd	–	–	'wild yam'

Sound correspondence				Counts		Lexical example				
Nese	Tape	Tirax	V’ënen Taut	In set	Broad	Nese	Tape	Tirax	V’ënen Taut	Gloss
m	^{m}b	?	?	1	0	nua**m**	nəβ^{w}i**b**	–	–	‘fish-poison tree’
?	^{m}b	?	ð̼	1	0	–	**b**eleβər	–	**ð̼**əl	‘(to) thunder’
m	m	m	n̼	0	14					
^{n}d̼	β	^{n}d	t̼	0	11					
^{m}b	^{m}b	^{n}d	t̼	0	6					
ð̼	β	β	Ø	0	5					
^{m}b	p	^{n}d	p	0	3					
β	?	β	β	0	2					

Table 14: Sound correspondences which include at least one labial reflex. “In set” refers to the number of full cognate sets characterized by the correspondence; “Broad” is the sum of “In set” plus partial sets (redundantly included) that could be characterized by the correspondence but are excluded for missing data. Voiced IPA obstruents are to be interpreted as prenasalized (and voiced) in lexical transcriptions.

** Note the potentially misleading redundant inclusion of Tape and V'ënen Taut 'house', which is alienable and inalienable and corresponds to both the alienable and phonetically reduced inalienable forms of Nese and Tirax.*

† This form is treated as the product of metathesis, but cf. Tape NP məne-X*, Tirax* NP nma-X*, and V'ënen Taut* NP n̼an̼a-X *'X's NP to drink' for the order of Tirax nasals in the potable possessive construction.*

The rest of this chapter proceeds as follows. In §4.1, I represent the consensus view of the origin of linguolabials in at least some of the languages of Vanuatu and how this account can be complexified to account for more of the data. In §4.2, I discuss assimilation patterns in plain bilabial and linguolabial consonants that operate synchronically in Nese and Nahavaq – these could reflect diachronic changes. In §4.3, I explain why the novel approach taken here is necessary given the available data. The methods used to detect patterns of interaction (random forest models) are described in §4.4. The findings are presented in §4.5: there is comparatively little evidence of contact and many patterns of conditioning by both labials and coronals. I ultimately conclude in §4.6 that the innovations that Tape and V'ënen Taut share with other Western Malekula languages are superficial, justifying the leaner set of languages chosen for the Ninde-centric proto language reconstructions offered by this dissertation[11].

[11] All the lexical data available from these two languages were initially considered in an earlier attempt (Schäfer 2018) to reconstruct either Proto Malekula Interior (Tryon 1976) or a Proto Western Malekula dialect continuum based on Lynch (2016).

4.1 Diachrony of linguolabials

4.1.1 Origin

The distribution of linguolabials among the languages of Malekula is complex and appears unpredictable. Lynch (2019b) argued that linguolabials are recent innovations by contact on Malekula, but might be reconstructed for all of Espiritu Santo languages that have them, but his analysis is restricted to Malekula. Based on qualitative evidence that linguolabials tend to dissimilate with each other and assimilate to both bilabials and coronals (presented in §4.2), this chapter presents an exploratory statistical approach (random forest models) to organizing the sound correspondences and features of their environments. Many patterns emerge from this type of model, and it is not adequate to characterize the distribution of linguolabials as simply unpredictable.

The consensus view is that plain bilabial consonants became linguolabial ones, but this alone does not explain the sound correspondences on its own. Tape either never underwent this change, or the change was reversed. In Tirax, a consonant that previously became linguolabial subsequently became dental (Lynch and Brotchie 2010). V'ënen Taut retains the crucial linguolabial stage, such that Tape:V'ënen Taut:Tirax correspondences of m:n̼:n in Tape /tə**m**es/, V'ënen Taut /ta**n̼**a/, Tirax /t**n**ah/ 'devil', for example, mimic a diachronic progression through stages in the order they are presented here.

Labialized bilabials (or labiovelars) are retained in Tape, but became plain bilabials in V'ënen Taut and Tirax – at least before non-front vowels. This yields a correspondence pattern of m^w:m:m, as in Tape /**mʷ**əliun/, V'ënen Taut /**m**liun/, Tirax /**m**leun/ 'chief'. Only V'ënen Taut retained labialization before front vowels, at least until the time that Fox (1979a) documented the language.

There is no consensus about – and in fact no discussion of – the patterns (as patterns) observed for Nese. Expected linguolabials can have bilabial, linguolabial, or dental reflexes in Nese. For this reason, Lynch (2019) suggests that the distribution in that language and the (less numerous) exceptions to his predictions for Tirax are not shaped by inheritance and sound change, but by contact.

4.1.2 Age of linguolabials

Generally, V'ënen Taut linguolabials are readily predictable, whereas those in Tirax and Nese are not; these facts have been used to support the notion that linguolabials are older in V'ënen Taut. Some instances of V'ënen Taut linguolabials do not reliably correspond to dentals in Tirax: e.g., Tape /nə**m**en/, V'ënen Taut /**n̼**anex/, Tirax /**m**anix/ 'bird' (expected Tirax: /**n**anix/). The fact of regularity in V'ënen Taut initially led Lynch and Brotchie (2010) to conclude that it was this language that initially innovated the sound and that the sound was "imitated, though imperfectly" (Lynch 2016) by other languages in the area. Lynch (2019b) later reframed the hypothesis of diffusion, suggesting that V'ënen Taut or Vao acquired linguolabials during a historical period of trade with speakers of Espiritu Santo languages before passing them along to other languages. This approach runs counter to general assumptions that linguistic complexity takes time to develop, and it is Nese that has the most phonemic labial contrasts.

As an alternative hypothesis, I propose that Tirax may have a comparatively *long history* since the innovation of linguolabials. The shift of PNCV *marani > Tirax *laran* 'tomorrow' represents a number of changes spurred by the innovation of linguolabials. There is an expected change of *m > *n̼ > n, but this particular word represents a third change: the intermediate form *naran could have been reanalyzed as having a common noun prefix *nV-* (see Lynch 2017 for a historical account of the noun prefix), which alternates with the adverbial prefix *lV-*, with both prefixes containing a vowel colored by the following vowel of the noun root. This adverbial prefix (which is shared throughout Vanuatu) is prototypically locative, but also forms temporal adverbs. The underlying combination of the nominal *na- and *rani 'be daylight, dawn' is not attested elsewhere on Malekula. Other nominalized forms of *rani incorporate the obligatory subject PNCV *ʔuta 'place' (Tirax has PNCV *na + *ʔuta + *rani > *notrɛn* 'day') and adverbs derived with the stative prefix *ma-. Over these alternatives, I propose that Tirax speakers reanalyzed *marani as a nominal *naran (after *m > n̼ > n) and derived a corresponding temporal *laran* 'tomorrow'. This process would represent two sound changes and a morphological reanalysis made more likely by a state of the language with no trace of the word's historical derivation. The absence of similarly layered change in highly regular V'ënen Taut suggests, if anything, that its linguolabials are *newer*.

While it is clear that Tape is the closest relative to V'ënen Taut (Tryon 1976, Lynch 2016) and that Tape does not have any sign of linguolabials (Crowley 2006:100), it is not impossible that Tape once had them. Since virtually all Oceanic plain bilabials correspond to linguolabials in V'ënen Taut, a change of linguolabials (back) to bilabials in Tape would leave no trace of the sound in Tape's phonology. This type of reversal has been attested in languages like Tutuba and the V'ënen Taut of youth today, and in many cases may make any conclusions impossible to reach (Lynch 2019b:310-311). This means that apparent retention of plain bilabials cannot indicate non-participation in the linguolabial shift (Lynch 2019b:229-230).

4.1.3 Ordering of sound changes

Lynch and Brotchie (2010) argue that unstressed low vowels were deleted before bilabials became linguolabial in Tirax. For the form reconstructed as PNCV ***b**aravu 'long', the following steps would have taken place:

(1) the unstressed low vowel of the initial syllable would have been deleted, whereas the high vowel of PNCV ***b**ura 'smash, break' would be spared. In the intermediate form ***b**ravu, the

labial contrast (plain vs. labialized) would have been neutralized before another consonant, and

(2) the high *u would have “irregularly” become *i. Only for ***b**ura (now ***b**ira) is the *b in the prevocalic environment to

(3) become linguolabial and subsequently

(4) dental. Finally,

(5) final high vowels are deleted and low vowels are raised for both forms in modern Tirax.

The different outcomes for PNCV *b are schematized below for the environments before an unstressed low vowel and an unstressed high vowel:

PNCV		1		2		3		4		5 **(Tirax)**
***b**aravu ‘long’	>	***b**ravu	>	***b**ravi	>					**b**rav
***b**ura ‘smash, break’			>	***b**ira	>	***d̼**ira	>	***d**ira	>	**d**re

The conditioning environment for bilabials to become linguolabial would be before /i e a/, but some instances of *u were fronted and unrounded. The irregular change of step (2) can be resolved if PNCV reconstructions were modified to have contrasting *b^{w} vs. *b before *u, and *b conditioned *regular* fronting of following *u. This would also be in line with Nahavaq, which has such a contrast (Dimock 2009). There is another explanation for the pattern above that better accounts for more of the data.

Lynch and Brotchie's account presents examples which all have one thing in common, without acknowledging the pattern: the proposed low-vowel deletion (Step 1) only blocked the shift of plain bilabials to linguolabials if it would have resulted in a consonant cluster with /t ^{n}d ^{n}d^{r} r l/ as the second consonant. They admit that their account does not explain some labial reflexes, but what they do not note is that these are all instances where low-vowel deletion would have produced conceivably better-formed complex onsets with linguolabials followed by /l h/:

PNCV		**Tirax**
***b**ala-ti 'wattled structure'	>	***d**las* 'be closed'
***m**a-saŋa '(to) fork'	>	***n**hak* 'branch off'
***m**asakit 'sick'	>	***n**haxit*

If restrictions were in place on consonants following a linguolabial in clusters, this is something that could have been resolved *after* low-vowel deletion. That is, the same outcome would be observed if there had been syncope of unstressed low vowels and subsequent coronal dissimilation of the linguolabial and following coronal.

This alternative account of dissimilation has the potential to explain all of the data for Tirax. It is necessarily rather complex, but does not invoke irregular change. For one thing, the differential outcome of initial *b in PNCV ***b**arapu 'long' and ***b**ura 'break, smash' requires regressive long-distance dissimilation of the linguolabials operating in the first form only (the conditioning environment would be the intermediate reflex of PNCV *p > *v), since *d̼r and /dr/ would represent licit onsets along the way (c.f. Grassman's law for similar word-level dissimilation). This alternative account for Tirax linguolabials can be schematized as follows:

(1) The plain bilabials *all* become linguolabial, even before *u[12]; then

(2) the unstressed vowels would have been deleted regardless of height; then

(3) linguolabials would dissimilate with other coronals (immediately before a dental stop, or before another linguolabial anywhere in the word); and finally,

(4) remaining linguolabials would undergo mergers with other sounds in modern Tirax (the fricative with the bilabial fricative and all others with the dentals):

[12] Following Dimock (2009) on Nahavaq, I assume that velar/plain bilabials and labiovelars were in complementary distribution even before rounded vowels. Like Nahavaq, these languages likely had front rounded vowel allophones [y] and [ø] of /u/ and /o/, respectively, conditioned by plain bilabials and *y > *i in Tirax. My account does not rely on an irregular change *u > i, unlike Lynch and Brotchie's explanation.

PNCV		Intermediate stages						Modern Tirax
Steps:		1		2		3		4
		Linguolabial		Syncope		Dissimilation		Merger
***b**ata**v**u ‘breadfruit’	>	***d̼**ata**ð̼**	>	***d̼**ta**ð̼**	>	***b**ta**ð̼**	>	**b**ta**v**
***m**ata-gu ‘my eye’	>	***n̼**atag	>	***n̼**tag	>	***m**ta-k	>	**m**ta-k
***m**ala**v**a ‘twins’	>	***n̼**ala**ð̼**	>	***n̼**la**ð̼**	>	***m**la**ð̼**	**>**	**m**la**v**
***b**ara**p**u ‘long’	>	***d̼**ara**ð̼**	>	***d̼**ra**ð̼**	>	***b**ra**ð̼**	**>**	**b**ra**v**
***b**ura ‘break, smash’	**>**	***d̼**ure	>	***d̼**re			>	**d**re
***m**asakit ‘sick’	>	***n̼**ahaxit	>	***n̼**haxit			>	**n**haxit

The pattern of cluster-internal dissimilation would be one whereby an intermediate linguolabial becomes bilabial if:

- it is followed by another linguolabial later in the word,
- the linguolabial is a nasal and immediately followed by a coronal, or
- the linguolabial is immediately followed by a coronal stop.

The first point is particularly noteworthy considering that all of Tirax's labial fricative reflexes are bilabial, because this lends support to the possibility that Tirax previously had linguolabial fricatives. Intermediate linguolabial fricatives are needed to account for the reflex of PNCV ***b**arapu > ***b**rav* 'long' and not otherwise expected ***d**rav*) by word-level dissimilation of linguolabials. Like Nese (Lynch 2019b), the linguolabial fricative would have never yielded a coronal fricative. This is motivated perhaps by economy, since the reflex of *v is non-contrastively voiced, and this would potentially introduce a voicing contrast between existing voiceless /s/ vs. would-be /z/ or /ð/ to a phoneme inventory that otherwise has no contrastive voicing.

More importantly, this alternative account means that the critical rule ordering established for Tirax is not necessary. The not-so-certain fact that syncope blocks linguolabials, and syncope is shared with other languages in the north of Malekula, is the basis for Lynch (2016) dismissing the possibility of shared inheritance of linguolabials. If linguolabials dissimilate to bilabials when followed by other coronals, then Tirax can share both innovations with its neighbors: linguolabials *and* vowel syncope. The processes of dissimilation may be sensitive to sonority and a preference for linguolabials later in the word.

Before moving on to the methods and findings that in fact highlighted this alternative explanation for Tirax, there is some groundwork to be laid first. In the section that follows (§4.2), the synchronic patterns surveyed mean that any account of linguolabials must contend with the fact that assimilation between linguolabials and other coronals and labials operates within morphological paradigms. That section also tempers expectations about the regularity that can ever be found, even using an innovative classificatory statistical tool (§4.3). The final model (§4.4) is not a comprehensive and exceptionless account of Nese. Exceptionless models were only generated for Tirax and V'ënen Taut.

4.2 Evidence in synchrony for complex assimilation and dissimilation

There is enough evidence for assimilation of linguolabials to bilabials that there is reason to suspect that distance assimilation is responsible for at least some cases in some languages in which a plain bilabial consonant would appear to have remained bilabial instead of becoming linguolabial. In Nese, the possessed root for 'father' has two forms: *tamam* 'your (sg.) father', bearing the second-person *-m*; and *tanan* 'their (sg.) father', bearing the third-person *-n* (Takau 2016). Though Nese has some words with linguolabial consonants, most linguolabials have become dental (Crowley 2006d). In this case, it is historically likely that an intermediate form *tan̼a- 'father' yielded a lightly suppletive pattern by asymmetrically assimilating to the second-person form (Table 15). The third-person form could reflect either a regular change to dental or assimilation to the suffixed dental nasal.

	PNCV	Gloss	Nahavaq	Nese
1st person possession	*tatai	'my father'	*tatei*	*tete*
2nd person possession	*ta**m**a-**m**u	'your (sg.) father'	*ta**m**^{w}a-**m**^{w}*	*ta**m**a-**m***
3rd person possession	*ta**m**a-na	'their (sg.) father'	*te**m**e-n*	*ta**n**a-n*

Table 15: Assimilation in Nahavaq and Nese labials contributes to the overall suppletive paradigms for 'father'. Corresponding reflexes of labials are bolded.

It is also possible that bilabial assimilation *predates* the development of linguolabials: Nahavaq, spoken in South West Bay, exhibits such a pattern, with no evidence for or against the language ever developing linguolabial consonants. In that language, plain bilabials in possessed noun roots predictably become labialized when suffixed with the second-person possessor *-m^w*, which contains (nothing but) a labialized bilabial nasal (Dimock 2007)[13]. If this is reconstructible for an older ancestor (perhaps PNCV), the proto plain bilabial consonants would have contrasted with labialized bilabial consonants. In that case, some of the assimilation processes could be older than the linguolabials (and dentals that developed in turn from them).

[13] In Dimock's analysis, plain bilabials are analyzed as *velarized* and labialized bilabials as *labiovelar*.

Since vocative kinship forms are often different from possessed forms, Melanesian languages sometimes have syncretism between the vocative and first-person possessed forms, yielding suppletive patterns. The first-person possessed form for 'father' in Nese *tete* and Nahavaq *tatei* appears to have such an origin when compared with Ninde cognates *tatai* 'father (vocative)' vs. *tama-ŋg* 'my father'. These already suppletive paradigms may have made kinship a semantic domain where divergence is especially tolerated. It is noteworthy that suppletion is also found in many Papuan kinship systems (Baerman 2014); as discussed in §2, contact with Papuan languages is a possible source of many Papuan linguistic features on Malekula. By contrast, other semantic domains may be less susceptible to this kind of divergence in form, and it may be the case that one form or the other provides the basis for a leveled paradigm – this is mere speculation on potential sources of irregularity.

4.3 The limits of sound correspondences

Given the paucity of complete correspondence sets, the data pose a challenge for historical reconstruction using the comparative method. The 32 correspondence sets shown in Table 14 are too many to reconstruct, even considering the conditioning environments described in §4.3.1. Considering factors at play that contribute to genuine irregularities like metathesis and reanalysis (§4.3.2), variable phonetic contexts at morpheme boundaries like the ones just seen in Table 15 (§4.3.3), and the possibility that some of the data may be inaccurate in the first place (§4.3.4), the traditional comparative method would seem to be doomed for northern Malekula.

4.3.1 Conditioned changes

Contributing to the apparent unpredictability of sound change are several patterns that affect labial consonants alternately in internal syllable codas vs. in word-final position. A change from prenasalized stops to nasals can be seen word-finally in Nese non-coronals (compare: Nese *nuam* vs. Tape *nəv^{w}ib* 'fish-poison tree') diachronically, although phonotactically, prenasalized bilabials are otherwise permitted word-finally. This is probably because the word-final prenasalized bilabials were previously linguolabial. Since there may be different outcomes for word-final consonants vs. codas, reduplicated forms like Nese *la^{n}d̼-la^{m}b* 'big'[14] shows that prenasalized linguolabial stops became bilabial word-finally after original prenasalized bilabials became fully nasal in that environment.

As a matter of procedure, where one form was reduplicated, it was paired twice with simplex cognates to reduce the rate of gaps in sound correspondences. Nese *la^{n}d̼-la^{m}b* 'big' really corresponds twice to Tirax *la^{n}d* 'big', but with different environments for Nese. This amplifies the appearance of irregularity presented by increasing correspondence sets if the

[14] This form is from Crowley and Lynch (2006d), but Shimelman et al. (2019) also document *ladlad* and *ladlan* for the two dialects of Nese, which would represent a further change from linguolabial to dental and restricted word-final change of $C_{[+nasal,\ +oral]} > C_{[-oral]}$.

reduplicated root is not identical in reduplicant and base, but it also provides valuable data about the role of phonetic contexts. (On the other hand, these can become complex: Nese *ne*^{m}***b***-*na*^{n}***d̼*** 'swollen glands' appears reduplicated, yet exhibits the opposite sequence of linguolabial and bilabial reflexes as compared to *la*^{n}***d̼***-*la*^{m}***b*** 'big'.)

In other cases, the number of sound correspondences is increased by changes in manner of articulation, even while the place of articulation mirrors other correspondences. In Tape, /p/ and /β/ are in contrastive distribution (e.g., a near minimal pair in *nə**p**el* 'swamp harrier' vs. *nə**v**et* 'rock'), but not word-finally, where [v] ~ [f] ~ [p] are in free variation and are all analyzed as /p/ (Crowley 2006d). Bilabial consonants in reduplicated words like *lə**v**-li**p*** 'mud' show that the underlying status of these sounds may be more complex, since the reduplicand copies a fricative where a stop is phonotactically acceptable.

4.3.2 Metathesis and Reanalysis

There is also some metathesis and assimilation that result in idiosyncratic cases. The potable possessive classifier (meaning something like 'for X to drink') *məna-, appears as V'ënen Taut *n̼an̼a-* and Tirax *nma-*. For these two languages, it is possible that the affinity of the onsets and vowels facilitated a reanalysis of the form as reduplicated in V'ënen Taut on one hand and metathesis in Tirax on the other. Though Nese has no attested potable possessive classifier, it is the only language that has apparently metathesized plain bilabial *m (in bold, perhaps intermediately linguolabial and thus [+coronal]) with *n (underlined) in *n**um*** ‘drink’ and *norru**ma***- ‘chest’ (cf. Tape ***m**ən* ‘drink’ and ***m**ərən* ‘chest’). An even more idiosyncratic change has applied to the *dental* nasal in Nese: *nesi**n***- ‘belly’ was recorded by Crowley and Lynch (2006d), and *nesi**n̼**e*- ‘belly’, recorded by Takau (2016). Such changes are unlikely to represent generalizable patterns of change; some word shapes may be particularly conducive to reanalysis (initial coronal nasals that are not historically the ubiquitous nominal *na-) and metathesis (particularly affecting linguolabial and dental/alveolar nasals).

4.3.3 Morphosyntactic factors

There is a difference between roots that may host suffixes and those which may not, in that the latter may suffer some ambiguity in specifically phonemic representation. Transitive verbs with pronominal objects bear object suffixes. These include an erstwhile transitivizing suffix *-i, reanalyzed as a third-person singular object in Tirax (Brotchie 2005:9) and Nese (Crowley 2006c:71), as well as third-person plural object markers in Tirax *-ɛr* (Brotchie 2005:22), in V’ënen Taut *-r* (Fox 1979:86), in Nese *-er* (Crowley 2006c:71) and in Tape *-ər* (Crowley 2006d:155). Only the latter two have object suffixes for other person-number combinations. If a verb has a root-final bilabial fricative, it may be realized as a stop in unaffixed form (1a) but not before a suffix (1b):

(1) a. i-tep Tape (Crowley 2006d)

3SG.S-push

'they (sg.) pushed it'

b. i-tev-ər

3SG.S-push-3PL.OBJ

'they (sg.) pushed them (pl.)'

In nouns, the same alternation can be found in bare noun roots (2a) vs. nouns with possessive affixes (2b):

(2) a. nipip Tape (Crowley 2006d)

penis.wrapper

'penis wrapper'

b. niviv-ən

penis.wrapper-3SG.POSS

'their (sg.) penis wrapper'

This morphology complicates the data in two ways:

(1) phonetic environments, especially word-final vs. preceding coronal [n] or [r], is dependent on word class and inflected form, and

(2) the paucity of data does not allow for a priori identification of diachronic sources of merged phones where correspondence sets are incomplete.

The bare root that is listed in dictionaries may be abstracted away from even obligatory affixes, further removing reflexes from the conditioning environments contributed by highly common or obligatory affixes.

4.3.4 Inconsistencies in transcription

Finally, there are numerous potential issues with transcription consistency which have entered errors into the record. Fox (1979a) recorded both *ð̼ənmaran* and *vənmaran* ‘old woman’, Crowley and Lynch (2006d) both *neð̼in* ‘arrow’ and *nevin* ‘arrow point’, *tavat* and *teð̼et*[15] ‘woman’. As observed by Lynch (2019b:64-5), there are also linguolabials in words recorded by Takau (2016) that were bilabial in those same words in Crowley and Lynch (2006c). In the data of Shimelman et al. (2019), the Benmara variety of V’ënen Taut has (at the time of the analysis) /ð̼/ transcribed for words that have /t/ or /l/ in the other varieties: *mit*

[15] The difference could be due to ablaut in the vowels of the latter form, which is used only in noun compounds, or compounding could have produced an allomorph by triggering resyllabification of the final /t/ (and thus removed an environment for dissimilation).

‘black’ as *mið̼*; *-lil* ‘big’ as *-lið̼*. However, upon inspection of the audio that is available, it appears the transcriber may have been influenced by brief closure or tapping/flapping involved in the articulation of what is really a word-final [l]. Inconsistency in transcription may not be recoverable for a moribund language like Nese, but does not preclude identification of general patterns.

4.4 Methods

4.4.1 Data

All of the data for this chapter were taken from reference grammars, because these offer phonemic transcriptions informed by phonological analysis: Fox (1979a) for V’ënen Taut, Crowley and Lynch (2006c) and Takau (2016) for Nese, Crowley and Lynch (2006d) for Tape, and Brotchie (2009) for Tirax. The impetus for this data collection was originally to include these lexical items in the reconstruction of Ninde’s immediate ancestor. These four languages serve the goals of this chapter, being the Malekula languages with linguolabials that have the most sizable recorded lexicons to date.

In total, there were 477 tokens of individual linguolabials and their corresponding reflexes in the sister languages. These tokens belonged to 396 unique words arranged in 151 cognate sets. Some words yielded multiple tokens of potential linguolabial reflexes – this is because an individual word may have had multiple plain bilabials in the proto language and/or the root was reduplicated. A mean of 2.6 words were filled in per cognate set, and words with no attested cognates were excluded. The entries were mainly represented as stems with no inflectional morphology – not even the transitive suffix PNCV *-i inherited as a third-person singular object suffix by all four languages.

4.4.2 Coding of Variables

Once collected, the data were coded for place and manner of articulation for target phone (the expected linguolabial), based on the reconstructed word in Proto North Central Vanuatu (Clark 2009) and in the synchronic form. The diachronic and synchronic features were also coded for the consonants immediately preceding and following the target phone. Additionally, the number of expected linguolabials was tallied at the word level (to assess word-level limits on the number of linguolabials permitted). The features of vowel height and roundness were also coded; unlike the consonants, however, these were only coded if the vowels were contiguous with the target phone – this allows for the model to recreate Lynch's model if it is the most powerful. Voicing was excluded because it is not contrastive in any of the languages, and backness was excluded because all back vowels are also round in Clark's reconstructions.

Apart from the phonetic variables, language, word class, and semantic domain were also coded. This is because word class determines what prefixes and affixes are likely to contribute cross-morpheme conditioning environments. Semantic domains were coded with domain labels that may be either associated with trade (like numerals, flora and fauna) and those that may be less relevant (basic and grammatical terms, kinship, social relations, nature, anatomy). These assumptions are not informed by any anthropology of trade on Malekula, but conceptually, the categories are in line with Lynch's implicit characterization that language contact was induced by the salt trade.

The place of articulation for the target phone was treated as a dependent variable and the features of its surrounding (conditioning) environment were treated as dependent variables. The individual tokens and their phonetic variables were identified by computer, using transcribed words as input strings. The semantic and word-class variables were manually coded.

4.4.3 Random Forest Models

The data were then modeled using a random forest model (Ho 1995). A random forest generates a large number of decision trees fitted to random subsets of the data. The optimal decision tree can be read somewhat like a formulation of a conditioned sound change.

Decision trees, which are the basis of a random forest model, divide data into ever-smaller pools of tokens based on levels of the variables. The data are always split into two groups, forming two "branches" of the tree (e.g., data points from Tirax and Nese vs. V'ënen

Taut and Tape). For the purpose of demonstration, the distribution of American English aspirated obstruents in the majority of words can be predicted by pooling tokens into smaller and smaller pools at the intersections of their properties. This successive subdividing is schematized in Figure 8. Voiceless stops and affricates (C[onsonant]Manner = “Stop, Affr”) in simplex onsets must be split into those that follow vowels or approximants (LastC[onsonant] = “Appr, Ø”). Of this smaller pool, the word-initial obstruents (LastV[owel] = “#, Ø”) are all aspirated, whereas word-medial ones (LastV[owel] = “yes”) are only aspirated as onsets of stressed syllables (NextV[owel] = “+stress”). Each of the remaining branches terminate with an unaspirated obstruent.

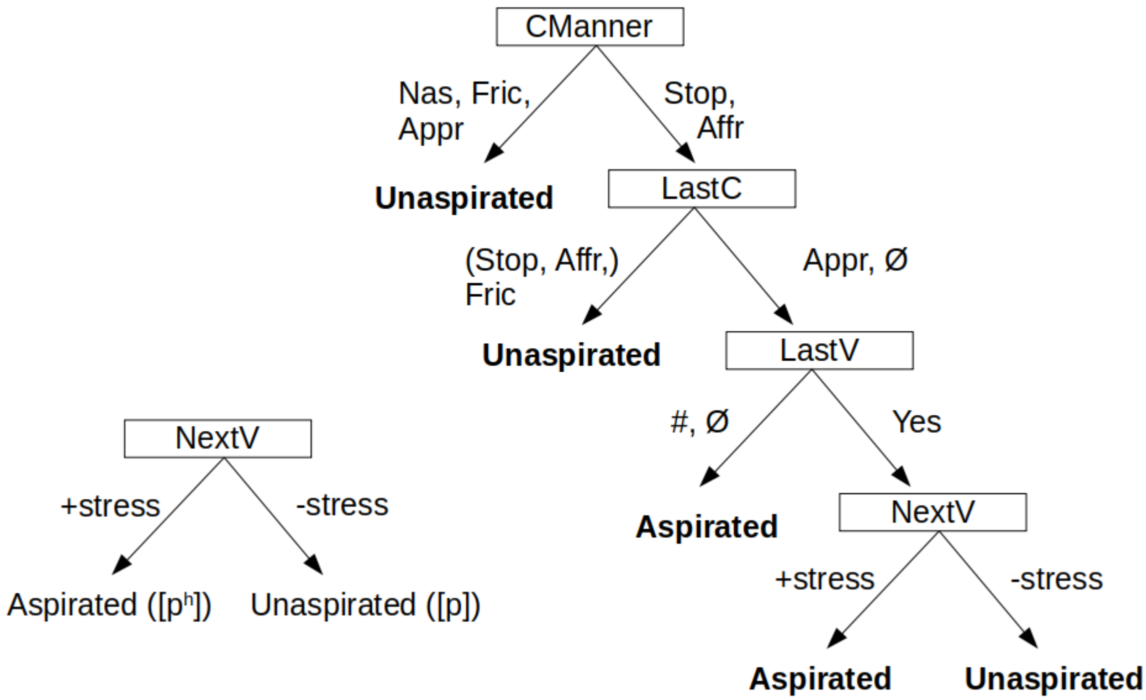

Figure 8: Two optimal decision trees characterizing (on the left) whether American English /p/ is aspirated given the stress status attributed to the following vowel and (on the right) whether any consonant is aspirated based on its manner of articulation (CManner), the manner of the preceding consonant (LastC), the presence of a preceding vowel (LastV), and the stress status of the following vowel (NextV).

For place and manner of articulation in the Northern Malekula data that were actually used, the variable levels were ordered in terms of position relative to linguolabials (glottals are the most distant place feature) and sonority, respectively. This allows for meaningful binary splits at points along these spectra; for example, tokens with a following oral and prenasalized stop may be split from those with a following nasal or fricative.

After an initial split, subsequent splits do not apply to the whole data set, but only to the data satisfying the conditions of the branch (e.g., a second split of consonant tokens that are before low vs. non-low vowels may only apply to the data from Tirax). When the data have been subdivided enough to make a classification of the dependent variable, a branch ends in a *leaf*: a categorical prediction for the dependent variable.

The output of these trees can be read as if they are ordered sound changes. With language as a variable, some of the splitting can also apply to multiple languages. Random forests generate many decision trees and it is possible to quantify the power of the model overall (classification accuracy) or the importance of particular independent variables in differentiating successful trees from unsuccessful ones. The outcome is ultimately a model of sound changes with minimized exceptions. This work might avoid the pitfall that Ladd et al. (2015) observe, that linguists tend to be skeptical of correlational work grounded in new methods if they cannot stand up to traditional standards of scholarship. The random forest model does not, however, provide statistical measures that typically allow us to reject a null hypothesis.

4.4.4 Hypothesis

If Lynch's model of linguolabial innovation is correct, then we expect a number of factors to have high importance as predictors of place of articulation. These include influential contribution of the language as variable (since linguolabials would be irregularly diffused across already diverged languages), semantic domains associated with trade (flora and fauna, social relations, numerals), and the following vowel in the proto language (to explain the blocking effect proposed for Tirax).

By contrast, if linguolabials were innovated rather regularly, then lost to assimilation and dissimilation in complex interaction with other labials and coronals, then the important features may be exactly the opposite. Place features of the environment should play a large role. There should also be an effect of sonority, such that the manner of articulation of target phones and of the consonants in their environment should interact.

The statistical tools used here are not designed to reject a null hypothesis. Since the random forest does not quantify the analysis in terms of probability, what follows in §4.5 is a qualitative interpretation of the variables selected by the model. These are expressed in terms of assimilation and dissimilation environments.

4.5 Findings

Decision trees using semantic domain generally reduced the out-of-bag (OOB) error rate of the model by fewer than eight cases. The OOB error rate of any random forest was high (> 16.8%), but nonetheless there are 39 terminal nodes in the final model with a 100% classification accuracy rate. Of those 39 terminal nodes, 11 represent pathways that can be formulated as assimilation (5) or dissimilation (7) patterns. Note that it is possible for the decision tree to contain pathways for which features of the proto phoneme (input to sound changes) are not a factor, and thus a sound change would represent either assimilation or dissimilation, depending on the input. In other words, an expected linguolabial may uniformly become bilabial whether there is another labial or another coronal in its phonetic environment. There were nine assimilation or dissimilation conditions for Nese, two dissimilation conditions for Tirax, and one assimilation condition for V'ënen Taut (but affecting labialized bilabials). Only two of these conditioning environments represent progressive effects (both dissimilatory): one for Nese and one for Tirax. This would suggest that conditioning environments later in the word tend to affect linguolabials that appear earlier in the word.

4.5.1 Predictor importance

The following paragraphs relate the relative importance of the variables, which were predicted to be phonetic variables in the case of inheritance and related to semantic domain in the case of contact. Generally, variables are named *After* or *Before* if they refer to the phonetic features of the following or preceding (with or without intervening segments) consonant (*C*) or vowel (*V*). Otherwise, consonant features in the variable name refer to the segment (i.e., the labial reflex) for which a place feature is predicted by the model. Features of the environment and the target segment are considered in terms of proto language input if the variable name contains *Proto* or *P*. A variable *PC Manner After*, for example, refers to the manner of articulation of a following consonant. The number of expected linguolabials at the word level (not including the target) is encoded as variable *LLProto*. Non-phonetic variables are self-explanatory: *Language*, *Semantic Domain*, and *Word Class*.

The factors that emerge as important in the model do not suggest effects of contact are necessary to account for most of the data on linguolabials. These are shown in Table 14 in order of decreasing importance as quantified by a Gini score of the predictor's performance across decision trees in the random forest model. Language and manner of articulation emerge as the most important factors, followed by semantic domain (predicted to be associated with contact) and the height of the following vowel.

Model 1		Model 2	
Variable	**Weighted Decreased Gini**	**Variable**	**Weighted Decreased Gini**
Language	63.674609	Language	83.066679
Manner	24.89849	Manner	36.610618
Semantic Domain	14.091069	Proto Place	24.043449
VHeight After	14.059249	Proto Manner	16.815558
Proto Manner	13.200822	PCManner After	13.627837
Proto Place	13.144127	Word Class	12.604382
CManner After	11.245349	CManner After	11.992692
Word Class	10.700606	VHeight After	10.533320
PCManner After	10.36365	PVBackness After	9.919970
VHeight Before	9.49256	PVRoundedness After	9.848501
VBackness After	7.942943	PVHeight After	7.687137
PVHeight After	7.66232	VBackness After	7.244908
CPlace After	6.616386	VHeight Before	6.512732

Model 1		Model 2	
Variable	**Weighted Decreased Gini**	**Variable**	**Weighted Decreased Gini**
PCManner Before	6.248787	PCManner Before	6.354385
PCPlace After	6.06163	CManner Before	5.422313
CManner Before	6.039275	PCPlace After	4.981111
PVBackness After	5.951343	PVHeight Before	4.772588
PVRoundedness After	5.916429	CPlace After	4.465305
CPlace Before	4.866676	VRoundedness After	3.208408
PCPlace Before	4.844059	CPlace Before	3.016776
VRoundedness After	4.816669	VBackness Before	2.914140
PVHeight Before	4.146787	VBackness Before	2.908501
VBackness After	4.109695		
VBackness Before	3.919633		
LLProto Word	3.246032		
PVBackness Before	2.961191		

Model 1		Model 2	
Variable	**Weighted Decreased Gini**	**Variable**	**Weighted Decreased Gini**
PVRoundedness Before	2.806786		
VRoundedness Before	2.55729		

Table 16: Predictor importance in order of decreasing importance to the model as quantified by the weighted decreased Gini score.

At first glance, these appear to be exactly the factors that Lynch and Brotchie (2010) identify for a change in Tirax (language) that regularly affected plain bilabial consonants (manner) after low-vowel deletion (height of the following vowel) and contact-induced change in Nese (semantic domain). The first two predictors will be important in any model, however, and vowel height was specifically selected from the environment in the reflex rather than in the proto language. This means that for Tirax, the model generally did not “consider” whether low vowels were deleted at a separate stage from high vowels. Instead, the reflex vowel quality (including whether a vowel was deleted before the following stop) was used in combination with consonant features as an approximation of factors in syllable and word shape, and there is only one low vowel anyway, which is not rounded.

Additionally, semantic domain did not remain a relevant factor in the branches of the optimal decision tree that reached the lowest rate of exceptions (many of which reached 0%), and were not intuitive semantic groupings when they were used. Its importance in the model may reflect its overlap with other domains, just as reflex vowel height reveals information about roundedness and syllabification. The most common split was between anatomy, flora, kinship, and sometimes including fauna and nature vs. all other categories. These splits grouped basic vocabulary and grammatical morphemes with numerals, social organization, tools, and trade, which may represent words expected to be on opposite ends of borrowability. Words in the kinship and anatomy categories were most likely to be inalienable nouns. This means that they are more likely to bear a nominal *na- prefix and may appear most frequently with reflexes of PNCV suffixes *-mu and *-na for first- and third-person possession (respectively).

A second model was generated excluding semantic domain and the five least important variables. This model performed only marginally worse (17.4% OOB error rate), with 0.6% fewer labial reflexes correctly predicted. Language and reflex manner of articulation were still the most important predictors, but proto place and manner of articulation, as well as word class, took on greater importance. Word class splits most commonly separated inflectional suffixes and independent pronouns from lexical classes. Successive splits in Nese occasionally grouped transitive verbs and inalienable nouns vs. all other classes; this hints at the role of Nese object and possessor suffixes in potentially creating conditioning environments affecting labial reflexes. This confirms that semantic domain was valuable in the initial model as a stand-in for phonetic environments contributed by inflectional morphology; when semantic domain is included, it does *not* select domains that are expected to be less stable and more exposed to effects of contact.

4.5.2 Sound changes

The optimal decision tree generated for the random forest model provides several branches that can be read as conditioned sound changes, since they take into account segments or features of the proto form (input), the place feature predicted for the segment (output), and features of segments in the environment (conditioning factors). There is some information about syllable and word shape, which proves to be relevant, in the form of consonant and vowel variables that were coded as word edges or zeroes. The output of the model does not produce sound change formulations per se, but an R script converted the output into a more familiar form. These were compared against the lexical data to eliminate irrelevant factors, because in combination with other factors, a subset of forms selected by those factor levels was not contrasted with data that did not meet the irrelevant factors. This mainly happened if an early split grouped two or three languages together, but data representing the range of phonetic environments was not available for one of them. The generalizations below contain terms like *all* or *every* if 100% of the relevant data could be predicted with the changes described, and anything below a threshold of 85% was not included.

Since all linguolabial stops in Tirax became dental and all linguolabial fricatives became bilabial, it is only possible to posit coronal dissimilation yielding a bilabial stop reflex, even if other sound changes formally affected other forms. All posited dissimilation patterns are regressive, operating within consonant sequences and syllables. There is only one example of a bilabial reflex of a stop (*b or *m) before a non-coronal: ***m***xo*^{n}*di-* ~ ***n***xo*^{n}*di-* 'grandchild', which could be the result of dissimilation between two linguolabials within a single word (PNCV *maxubi >) ***n̼**xo**d̼**i > ***n***xo*^{n}***di***-*; however, it is unclear why the form would be in free variation, unless word-level dissimilation is not obligatory[16]. The initial linguolabial may have also responded to dissimilation effects from the obligatory possessor suffixes; as a kinship term, certainly it would frequently be used with second-person suffix -*m*, which could also influence the outcome of the linguolabial nasal.

There are a few simple generalizations that go a long way in explaining the changes we see in Tirax. They are presented here, with conditioning environments underlined and target phones bolded:

1. Linguolabial *n̼ > m before *r without exception in Tirax, whether there is an intervening vowel or not:

   PNCV ***m**ar̲aŋ 'dry' > ***n̼**r̲aŋ > ***m**raŋ*

[16] Brotchie (2009) includes both forms but does not characterize their difference; however, she only includes the variant with the bilabial in examples inflected for first-person possession: *mxodi-k* 'my grandchild'. The variation may be explained in terms of allomorphy.

?***m**ar̲ ‘man’ > ***n̼**ar̲ > ***m**ar,* cf. V’ënen Taut ***n̼***ər

2. Linguolabial *n̼ > m before *l if there is an unstressed *a, whether that was deleted or not:

   PNCV ***m**al̲akeja ‘cold’ > ***n̼**l̲axes > ***m**laxes*

   PNCV ***m**al̲e ‘footprint’ > ***n̼**al̲ɛ- > ***m**alɛ-* ‘leg’

3. When linguolabial *n̼ (maybe also *d̼) forms a complex onset with *t after vowel deletion, the reflex is bilabial:

   PNCV ***m**at̲a- ‘eye’ > ***n̼**t̲a- > ***m**ta-*

   PNCV ***b**at̲avu ‘breadfruit’ > ***d̼**t̲að̼ > *ᵐ**b**tav*

   The status of *d̼ in this environment is uncertain, because the second example is the only form attested for Tirax that has PNCV *bVt and it has a linguolabial later in the word that could have conditioned the dissimilation (see generalization 6).

4. In one case, linguolabial nasal *n̼ did not become dental before *n, even though an intervening unstressed *a is retained:

   PNCV ***m**an̲uku ‘bird’ > ***n̼**an̲ix > ***m**anix*

5. Everywhere else, linguolabial nasal *n̼ became dental, even preceding a coda coronal nasal:

***m**in̲ 'drink'	>	***n̼**in̲	>	***n**in*
***m**iD̲ 'ripe'	>	***n̼**i<u>dr</u>	>	***n**i^{n}dr*

6. One exception to the changes above can be further explained as word-level dissimilation between nonidentical linguolabials[17]:

PNCV ***b**arav̲u 'long, tall'	>	***d̼**rað̼̲	>	^{m}***b**rav*
PNCV ***b**atav̲u 'breadfruit'	>	***d̼**tað̼̲	>	^{m}***b**tav*

The plain *b of the first example in generalization 6 is expected to yield a dental reflex in this environment (since *r does not have the same dissimilatory effect on *b as it does on *m). The presence of plain bilabial *v, however, suggests that this was an intermediate linguolabial fricative in pre-Tirax. This lends a crucial piece of support in favor of reconstructing an intermediate linguolabial fricative, making Tirax far more similar to V'ënen Taut and Nese in terms of regular development of linguolabials.

Though the model does not generate exceptionless changes, those that can be inferred explain the distribution of dentals and bilabials in Tirax with fewer exceptions (only two, one

[17] The form *baravu was mistakenly coded as *barav^{w} in the model, based on the rounding of the last consonant in other Malekula languages with word-final vowel deletion, but V'ënen Taut generally has a vocalic reflex of Clark's *u, supporting the reconstruction of *baravu 'long, tall' > *t̼are**i***. The coding error affects the model's error rate and precludes the model from identifying this as word-level dissimilation of linguolabials in Tirax.

of which is in variation with the expected form and the other is in a unique environment). The Generalizations 4 and 5 are not necessarily in competition: the single data point suggesting coronal nasal dissimilation yielding a bilabial nasal involves conditioning across syllables (Generalization 4), whereas the forms that have tautosyllabic intermediate linguolabials and coda coronal nasals have coronal reflexes (Generalization 5). Without knowledge of an earlier state of Tirax that may have had mixed linguolabial reflexes like Nese, or more lexical data, this is only an observation about an exception in Tirax.

In Nese, a different set of dissimilation and assimilation patterns yields a more complex distribution of linguolabial, dental, and bilabial reflexes of plain bilabials. Virtually all of the patterns are also regressive, but there appears to be a reversal of the direction of conditioning in the case of reduplication (the direction of reduplication is not clear). The manner and place of articulation together predict the reflex of plain bilabials in Nese; i.e., the sonority of expected linguolabials and conditioning phonemes interact in Nese:

1. Expected linguolabial fricatives are all bilabial before a liquid *l or *r (coronal dissimilation), but linguolabial before *n, *t, or *s (retention):

***v**ial ‘walk’	>	***ð̼̌**ial̠	>	***v**ial*
***v**aru ‘pig w. circular tusks’	>	***ð̼̌**aɾ̠	>	*na-**v**aɾ*

but

PNCV ***v**ano 'go' > ***ð̼**an > ***ð̼***an*

2. Prenasalized linguolabial stops all become bilabial before heterosyllabic oral coronals including *l (coronal dissimilation), but remain linguolabial before tautosyllabic coda *n or *l[18] (retention):

PNCV ***b**ilaka 'banded rail' > *ni.**d̼**i.l̠ax > *ni-ᵐ**b**ilax*

but

PNCV ***b**ala 'swamp harrier' > *na.**d̼**al > *na-ⁿ**d̼**al*

PNCV *ka**b**ani 'sail' > *na.**d̼**an > *na-ⁿ**d̼**an*

3. Linguolabial fricatives, with one exception (*nararað̼* 'Indian coral tree, Erythrina indica'), became bilabial word-finally:

PNCV *lav-(i) 'take' > *le**ð̼** > *le**v***

PNCV *mawava 'yawn' > *n̼an̼a**ð̼** > *nana**v***

4. All linguolabial nasal stops *n̼ became bilabial before (an optional vowel and) a tautosyllabic rhotic *r or *D in the coda (regressive coronal dissimilation):

***m**er̠tu 'person' > *ne.**n̼**er̠.te > *ne**m**erte*

[18] All of the languages in question have deleted the final vowels in these examples, but this is not reflected in the PNCV reconstructions.

*__m__iD ‘ripe’ > *__n̼__iD > ***m**ire*

5. One example (previously mentioned in §4.2) suggests that linguolabial nasal stops became bilabial preceding a bilabial nasal (regressive assimilation):

   PNCV *ta**m**a-mu ‘2’s father’ > *ta**n̼**a-m > ta**m**am

6. With one exception (the suffix *-n̼i* ‘your’), all remaining linguolabial nasal stops *n̼ (those that did not participate in Change 4) became dental if they were either intervocalic or word-initial and not followed by another linguolabial nasal stop.

   PNCV ***m**as ‘die’ > ***n̼**as > **n**as

   PNCV ***m**alakeja ‘green’ > ***n̼**a.la.xej > ***n**alaxe*^{n}*t͡s*

   PNCV *qa**m**uyu ‘you (pl.)’ > *ka.**n̼**i > *ka**n**i*

   but

   *-**m**uyu ‘your (pl.)’ > *-**n̼**i > *-**n̼**i*

   PNCV *ka**m**ali ‘chief house’ > *na-x.**n̼**al > *na-x**n̼**al*

   PNCV *tano + *maʔeto > *ton-**n̼**ot > *to**n̼**ot* ‘black earth’

7. Linguolabial nasal stops *n̼ became bilabial if they were root-final following *n and an intervening vowel:

PNCV *yum^{w}a ‘house’ > *ne**n̼-** > *nem-* ‘house of (inal.)’

PNCV *inu ‘drink’? > ***n̼**un > *nu**m*** ‘drink’

but

PNCV *yum^{w}a ‘house’ > *nai**n̼**e > *naine* ‘house (al.)’

The PNCV reconstruction for ‘house’ suggests these reflexes should be bilabial, but compare Tirax inalienable *nen-* ‘house of’ and alienable *nain* ‘house’, which suggest a linguolabial can be reconstructed for Tirax and Nese (but not V’ënen Taut and Tape).

8. Reduplicated forms with a prenasalized linguolabial stop yield one linguolabial and one bilabial prenasalized stop, but the order is not predictable. This occurred after *b^w > m word-finally, because it counterfeeds that change:

PNCV *laba ‘big’ > *la**d̼-**la**d̼** > *la**d̼**la**b***

? ‘swollen glands’ > *ne**d̼-**na**d̼** > *ne**b**na**d̼***

9. Multiple expected linguolabial sonorants similarly show no unified pattern, except that they yield the same output if they are identical:

PNCV ***m**ea ‘tongue’ > *ne**n̼**e**n̼** > *ne**n̼**e**n̼***

? > ***ð̼**e**ð̼** > ***ð̼****e****ð̼****-naul* ‘read’

PNCV *ga**m**a**m**i ‘we (excl.)’ > *ka**n̼**a**n̼** > *ka**n**a**n***

PNCV ***m**awava[19] ‘yawn’ > ***n̼**an̼av > ***n**anav*

The form *nanav* ‘yawn’ is predictable, because the second intermediate linguolabial appears intervocalically and does not precede another linguolabial, so Sound change 6 applies here, affecting both linguolabials. The other form *kanan* ‘we (exclusive)’ is not the form that is predicted.

10. Dentals and expected linguolabials could be particularly susceptible to metathesis with coronal nasal *n in Nese:

PNCV *inu ‘drink’? > ***n̼**un > *nu**m*** ‘drink’

? c.f. Tape ***m**ərə**n*** ‘chest’[20] ***n̼**oru**n**a- > ***n**oru**m**a-* ‘chest’

Based on the Tape form, the final /a/ in Nese ‘chest’ is not expected, so the metathesis could have put intermediate linguolabials in word-final position chronologically between Sound changes 6 and 7.

[19] But compare Tape *mamau* ‘yawn’, V’ënen Taut *n̼an̼ao*, and Tirax *mamav* (the latter are two from Shimelman et al. 2019).

[20] Most languages of Malekula have a form resembling *mVn* ‘drink’ (including Tape *mən*), and the form for

In one case, a sound change registered in the model as assimilation (Sound change 6) can be stated more broadly as a change conditioned by word-inital and intervocalic environments. All but four of the 16 words in this environment are additionally followed by reflexes of *t, *s, *l, *r, or *n (corresponding to the natural class selected by the model). When explained as implied by the model, there are four exceptions, but as explained here, there is only one exception.

The following Nese items are exceptions and in many cases are the only representatives of the particular environment in which they are found:

- PNCV ***b**aravu ‘long, tall’ > **d̼**aɾa**ð̼** > ^{n}***d**aɾav*

  This could be dissimilation between two intermediate linguolabials, but yielding a uniquely dental reflex word-initially. Unlike in Tirax, the PNCV minimal pair is not a minimal pair in Nese, since the noun PNCV *batavu ‘breadfruit’ has taken on a nominal prefix in Nese: *na: *na**d̼**atav.*

- PNCV *mata ‘eye’, ***b**are ‘blind’ > ***d̼**aɾ > *nenet*^{m}***b**aɾ* ‘blind person’

  This is the only other reflex of intermediate *d̼ that is word-initial.

- PNCV *rara**v**i ‘Indian coral tree’ > *narara**ð̼** > *narara**ð***

‘chest’ is not attested for other regional languages, but could be reflexes of *m^{w}adu-n ‘their (sg.) back’ – in that case, there would be no expected linguolabial in this form at all, but the absence of other reflexes leaves this irreconcilable.

This is the only intermediate linguolabial fricative that did not become bilabial word-finally. The expected form is *nararav*.

- PNCV *qa**m**a**m**i ‘we (incl.)’ > *ka**n̼**a**n̼** > *kanan*

  This contains an intervocalic linguolabial nasal, but it is followed by another linguolabial, so Sound change 6 does not apply. The expected form is *kan̼an̼*.

- *-**m**a**m** ‘ours (incl.)’ > -***n̼**a**n̼** > -*nan*

  Similarly, this form has both an expected linguolabial nasal that may appear following a vowel or consonant and also preceding another linguolabial nasal, so Sound change 6 does not apply here either. The expected form is -*n̼an̼*.

- *ta**v**at ‘woman, girl’ > *ta**ð̼**at > *tavat*

  There is no apparent conditioning environment for the intermediate linguolabial fricative to become bilabial; however, this is a doublet with *teð̼et* ‘woman’, which retains the linguolabial and otherwise resembles noun-modifying variants with characteristic ablauting of /a/ → [e] described by Crowley and Lynch (2006d:51).

- PNCV *sa**v**a ‘what?’ > *-s**ð̼**e > *tasve* ‘when’

There is no discernible conditioning environment for the bilabial reflex of the intermediate linguolabial fricative, but this is the only reflex immediately after /s/.

- Finally, one sound represented by PNCV *b has an inexplicable set of reflexes:

  PNCV ***b**ea ‘where’ > *xa**d̼**e > xa**d**e

  but

  PNCV *bani~*ka**b**a-u > *naxa**d̼**e > naxa**d̼**e

  This reflex corresponds to a prenasalized dental stop in Tirax *a**d**ɛ*, a bilabial fricative in Tape *e**v**i*, and a voiceless linguolabial stop in V’ënen Taut *it̼i*. Tape is noteworthy, since a fricative reflex is not expected for PNCV *b, which generally merged with /p/. Perhaps this represents a proto sound that Clark (2009) does not reconstruct for PNCV – like *p – that in Nese and Tirax yielded an early prenasalized dental (before any intermediate linguolabial *d̼ became dental). This could mean that some irreconcilable reflexes in Nese represent an inherited contrasting phoneme that tended to merge with *b.

- PNCV ***b**aiqa ‘green snail’ > *na**d̼**aike> *na**d**aike*

The expected form is *na***d̼***aike*, absent any changes yielding a dental reflex of an intermediate prenasalized linguolabial stop.

The model also selects apparently erroneous patterns reflecting limitations of the data. For example, dissimilation in V'ënen Taut if the rhotic (trill or prenasalized rhotic) is in a tautosyllabic coda (i.e., no syllables of the shape /**n̼**Vr̠/) – recall that this is one of the environments for linguolabials to dissimilate to /m/ in Tirax (Sound change 1). This appears to be merely a coincidence, since the data do not include any V'ënen Taut forms that would have inherited a syllable *m^wVr. On the other hand, PNCV *m^wera ~ *m^wara 'person of' is the closest semantic and phonological match for ***n̼****rtu* 'person' (Tirax ***m****orti*, Nese *ne***m***erte*) and ***n̼****ər* 'man' (Tirax ***m****ar*), two sets which were modeled without a proto form. Additionally, PNCV *mawirne 'left hand' frequently yields a labialized bilabial as product of coalescence in the *m and *w, as in Avava *m^wi:r* 'left hand' (Crowley 2006a). Tape ***m****orne-* 'left hand' permits the possibility of a labialized intermediate *m^w (because *m^w and *w alike could account for vowel rounding), but V'ënen Taut ***n̼****irne* 'left hand' suggests *m (and the form is unattested in Nese and Tirax). The model could instead be accurate, representing an unexpected change in V'ënen Taut: bilabial nasals (inherited from labialized bilabial nasal *m^w) partially merged with linguolabials wherever it was the onset of a syllable followed by an alveolar trill in the coda.

4.5.3 Shared changes

If Nese and Tirax both inherited linguolabials, and both have lost them, then shared changes would suggest characteristics of a shared proto language. Shared changes would be any coronal dissimilation yielding a bilabial reflex for stops in both languages or coronal assimilation yielding a dental in Nese. Additionally, bilabial reflexes of linguolabial fricatives as found in Nese would be possible to reconstruct to a node ancestral to Nese and Tirax, since it cannot be determined if Tirax represents a one-time merger of linguolabial fricatives with bilabial fricatives. A mere two shared changes are suggested in Table 17.

Linguolabial	Environment	Tirax	Nese	Shared
Fricative	Before *l or *r	Bilabial	Bilabial	Yes
	Before *n, *t, or *s	Bilabial	Linguolabial	No
Nasal	Before *(V)l	Bilabial	Linguolabial	No
	Before tautosyllabic *n or *D	Dental	Bilabial	No
	Before *t	Bilabial	Dental	No
	Intervocalic and word-initial	Dental	(Dental – but most data also before coronals)	Maybe

Linguolabial	Environment	Tirax	Nese	Shared
Prenasalized stop	Before tautosyllabic *n	Dental	Linguolabial	Retention
	Before tautosyllabic *l	(Dental – but limited data)	Linguolabial	Retention
	Before heterosyllabic *l	Dental	Bilabial	No
	With another non-identical linguolabial	Bilabial	Dental	No
	In reduplicated forms	(No data)	Bilabial	

Table 17: Generalizations about intermediate linguolabial reflexes in Nese and Tirax with respect to environments in PNCV.

A shared proto language ancestral to Nese and Tirax could have already lost some inherited linguolabials. These would be: *ð̼ before a liquid and *n̼ intervocalically and word-initially. Although no attested cognate pairs are reduplicated in both Nese and Tirax, the word-level dissimilation in Nese only resembles that of Tirax in reduplicated words. Otherwise, there are also shared lexical developments in and intermediate alienable *nain̼e 'house' vs. alienable *nen- 'house of'. Most intermediate linguolabials would have been retained by an immediate ancestor to Tirax and Nese, then rather straightforwardly lost in Tirax.

Comparatively little linguolabial loss seems to be shared with V'ënen Taut or Tape. The only matter of overlap is that Nese, Tirax, and V'ënen Taut seem to have a merger of earlier *m and *n̼ before liquids, but in different directions. V'ënen Taut has no bilabial nasals before *(V)r, even when there is evidence that these might be reflexes of labialized bilabials. Nese and Tirax, by contrast, have restrictions on linguolabial or dental nasals in that environment, even in cognates. A conditioned neutralization could be convergent, or it could be reconstructed for an ancestor common to all four languages considered here.

4.6 Conclusions

The analysis of patterns uncovered with the help of random forest modeling suggests that some of the linkages in Lynch (2016) rely on superficial shared changes and dismiss other important shared changes. The data from Tirax can be explained without the blocking effects of vowel deletion, meaning that linguolabials could easily be inherited in Tirax. By comparison, Nese has more exceptions, but patterns can be observed that would similarly explain the distribution of dental, bilabial, and linguolabial reflexes by appealing to assimilation and dissimilation processes. Some of the exceptions are phonotactically unique cases and may not be explainable unless there is more lexical data available, especially from Nese and Tirax. Relatively orderly acquisition of linguolabials is significant for phylogeny in this region, since Lynch's Peripheral Western Malekula linkage includes V'ënen Taut, Tape, and Tirax as a group unrelated to the other languages that have linguolabials. This chapter does not demonstrate that linguolabials are a diagnostic feature of genealogical relationships, but it suggests that they could have been inherited even in languages where they appear irregular in their distribution.

The use of random forest models proved to be fruitful for a task of this nature. This approach began as a way to investigate suspected assimilation and dissimilation between [+coronal, +labial] linguolabials and other coronals and labials. The model has limits in explaining the data, but brought to the fore previously unconsidered interactions between sonority and syllable shape.

In Lynch's Western Malekula Linkage, Ninde shares with V'ënen Taut the post-vocalic deletion of PNCV *p and a conditioned neutralization of PNCV *j and *s when followed by *i – but these are not clearly changes diagnostic of shared innovation. In most languages, *p became a bilabial fricative in contexts where it is lost in Ninde. There is evidence that Ninde has lost *p more specifically in intervocalic position – this followed final vowel deletion that did not occur in V'ënen Taut: PNCV *bata**p**u 'breadfruit' > V'ënen Taut *t̼atei*, Ninde *nəmbəte**p***. The neutralization of *j and *s occurs in contexts of palatalization where many Malekula languages underwent palatalization of *t, often resulting in mergers with *j or *s. Of the three languages in its group, Lynch indicates Tape shares a lexical idiosyncrasy with a neighboring language grouped with the Western Malekula Linkage: Larëvat. This idiosyncrasy is metathesis in PNCV *katabola 'Dracontomelon species' yielding *kataloba > Tape *nitlip*, Larëvat *nəxatelab*. This shared metathesis, together with shared natural and widespread sound changes, is the basis for grouping three Northern Malekula languages with languages that more resemble Larëvat. On the whole, the reasons *for* including languages of northern Malekula with languages distributed across the entire western coast are slim.

This chapter begins the important work of identifying conditioning factors in the realization of linguolabials. This is important for the subgrouping of Malekula languages, but it also sheds light on how linguolabials might change over time. The data suggest that they are in some ways unstable and prone to loss. Prior work suggested that linguolabials pattern phonologically as labial (Olson et al. 2009), yet the results of this investigation suggest they are susceptible to dissimilation and assimilation with coronals in diachrony.

In many ways, the changes brought about by the innovation of linguolabials render the cognate sets fully unrecognizable *a priori* to a comparative linguist who might be doing initial cognate detection. This could explain in part the limits in time-scale of the comparative method. At great time-depths, the likelihood of such an innovation that obscures patterns of inheritance probably increases. If linguolabial consonants existed in any prehistoric language, the correspondence sets between its daughters – much changed with time – between labial and coronal reflexes could be unrecoverable. This is especially so if those sounds interact in complex ways with each other.

5 Ninde as a Mixed Language

Ninde is an Oceanic language of the Mewun people in South West Bay, the coastal southwestern region of Malekula island. The oral record describes two Ninde languages that together came to form a single modern language. In Chapter 3, an effort to find bundled isoglosses representing a stable layer of innovations shared by inheritance identified a potential subgroup of languages spoken in Northern Central Malekula (NCM): Avava, Neverver, and with less certainty Neve'ei and Naman. Large parts of the lexicon, phonology, and grammar are shared, however, with other languages spoken in the South West Bay region, which also form a potential subgroup (SWB): Nahavaq, Naati, and possibly Na'ahai/Naha'ai. In this chapter, I will demonstrate that Ninde is a mixed language with contributions at various levels of language structure from Northern Central Malekula and South West Bay languages when core parts of the lexicon and grammar are considered at scale.

In this case of language mixing, the source languages were relatively closely related languages with many cognates that would have been virtually interlinguistic homophones. It is difficult to address parentage for Ninde, not solely because of the magnitude of contributions from two sources, but for the large number of words and grammatical elements that could have been from either, or both, sources. Whatever the type of language mixing that occurred, it is clear that neither of the source Ninde languages continues to exist. Disentangling lexical form, morphosyntactic structure, and lexical semantics from one another is key to understanding the oral-historical accounts of the inextricable mix that is modern Ninde.

This chapter considers Ninde in terms of the first two: shared innovations in lexical forms, which have long been the basis for classifying languages genealogically; and morphosyntactic structures, which are often the basis for discerning whether apparent cognates are the result of contact or shared inheritance (Ersheidat & Tahir, 2020). This discussion begins with coverage of prior work on the subjects of mixed languages generally, and specifically relating to Mewun people's mixed linguistic and cultural heritage (§5.1). Some lexical contributions from various sources are discussed in §5.2. The clearest cases of mixed lexical origin, contextualized with Ninde's phonology, in turn inform §5.3, which provides an account of Ninde's diachronic phonology for each of its lineages where sound change generated divergent reflexes. In light of the disentangled lexical and phonological history, I address morphosyntactic structures from each source language in §5.4. Each of these levels of linguistic structure support a conclusion of Ninde as a mixed language (§5.5). The following chapter (§6) reconstructs the lexicons more specifically for the two language subgroups (NCM and SWB); the discussion is rounded out in §7 with special consideration of lexical systems in Ninde, some of which are blends of forms and semantics from disparate sources.

5.1 Prior work

5.1.1. Defining mixed languages

There is presently no consensus on how to delineate mixed languages from other contact phenomena, but Ninde fits the profile of many languages that have been described as such. Oft-cited criteria seek to determine mixed status based on the nature of their genealogy or the sociolinguistic circumstances of their creation or maintenance. Most importantly, a mixed language has lexical and/or structural elements from at least two languages, to an extent that it is not possible or defensible to trace a genealogy through one of these languages alone.

If a diachronic framework allows for primary linguistic descent through only one language, then mixed languages must be excluded from classification altogether or one of the source languages must be identified as contributing an uninterrupted core of the language (Meakins 2013:184). While this is certainly the basis for some of the objections to the very possibility of mixed languages, Versteegh (2017:220, 233) denies that they represent a unitary phenomenon, given the typological diversity and absence of a single mechanism to account for their origins.

Ambiguity in genealogic classification has been underscored as a prominent feature of mixed languages (Thomason & Kaufman 1988) or what defines them (Bakker 2015). This is generally understood to mean that, even when the origin of words and structures in a language are known, there is approximately as much reason to classify the language with one heritage as there is with the other. Barranquenho, however, is a contact language with transparent contributions from Extremeño Spanish and Portuguese, where these are divergent, but a good deal of ambiguity in the form of overlap (Clements et al. 2011:20). It is a candidate for reclassification as part of a Portuguese-Spanish dialect continuum (Meakins 2013). This classification would mean that some dialect continua are carved out by innovations diffused across once-distinct languages, in addition to those that come about via incomplete diffusion across an originally more homogeneous language.

Compared to pidgins and creoles, mixed languages tend to be abrupt developments affecting societies with at least two common languages available for communicative purposes (Matras 2003:152-4). A mixed language could be immediately adopted as a community language instead of as a system of communication across ethnolinguistic lines. The mixed languages of Mbugu/Ma'á (Bantu and Cushitic) appears to be a case of slow mixing as speakers of Ma'á faced persistent pressure from Bantu languages (Thomason 1997:6), or alternatively, Mbugu speakers who identified as culturally and phenotypically separate from the socioculturally dominant Pare (Bantu) speakers developed the language as demographic shifts made Ma'á speakers and their linguistic features accessible (Mous 2003:83). In any case, a mixed language emerges in communities with speakers connected to at least two languages; either they identify with both languages (or neither fully), or they shift from a dominant language, one that is antithetical to their identity, toward one that is in line with their sociocultural distinctness. While the motivation to develop a pidgin may be communicative need, the motivation for a mixed language may be expressive (Golovko 2003:191, Muysken 1997:375).

Compared to phenomena like code switching, mixed languages are autonomous; they do not require continual input from (or even continued familiarity with) the source languages (Matras & Bakker 2003). The nature of the mix may pair a lexicon from one source with a grammar from another, or there may be dual inheritance within each of these domains (Meakins 2013). Hybrid lexicons may maintain grammatical properties of some word classes (e.g., Bantu gender on Cushitic words in Ma'á/Mbugu) and syntactic phrases may determine the source language (e.g., French noun phrases and Cree verb phrases in Mitchif, which also classifies words with both French and Cree gender categories) (Rhodes 1977:6). Like code-switching, the sociolinguistic circumstances of language mixing favor, or at least allow for, greater complexity in form than early pidgins. Thomason (2009) supports a possible pathway to a mixed language status via ad hoc code-switching. Unlike code-switching, the distribution of source languages in a mixed language is highly conventionalized.

Compared to adstrate influences arising in situations of prolonged contact, truly mixed languages defy genealogical classification because the indispensable core of a language, however determined, is not simply inherited from one language. A natural approach to quantitatively diagnosing mixed language status might establish some minimal unit (e.g., words or clauses) and determine whether the rate of tokens from one origin surpass an arbitrary threshold – or, at the extreme, whether it is at all possible to form units from material of a single origin. Some have asserted that there is a bimodal distribution of borrowing rates, with few instances of languages that have between 40% and 90% of the lexicon borrowed from another language, and this may inform a threshold differentiating intense contact from lexicon mixing (Bakker 2003). Lexical ratios alone are not enough to distinguish language shift with retention of substrate forms, which would comfortably allow for genealogy to be traced through the target of the shift alone (Versteegh 2017:225-228). Basic and naturalistic English clauses, for example, can be formulated without any French loans.

Phonological systems of mixed languages are as variable as their lexicons and grammars, but no language is a mixed one on the basis of their phonology alone. One input language may rephonologize the contributions of the other or two distinct phonological subsystems may coexist; subsequent phonological innovations can influence the mixed language and its parent languages independently (Meakins 2013). Phonology, then, does not serve to differentiate loanword phenomena from lexical mixing.

Linguists' willingness to accept the status of a contact language could depend on the typological profile of the language and the genealogical relationship of the source languages. Can a language be structurally mixed if core properties of its grammar or lexicon are already shared between the languages in contact? Perhaps there are combinations of languages that can never qualify as mixed, or for which the burden of proof is greater.

Mixed languages with closely related source languages could comprise a distinct subtype of language mixing phenomena, as structural similarity could allow for different kinds of mixed grammatical structures. Law (2017) characterizes Tojol-ab'al as a mixed language formed from closely related Chuj and Tseltal (Mayan, Mexico), highlighting the methodological difficulty in demonstrating this status. In this case, basic structures of case, number, and negation share affinities with both sources, and he suggests that the nature or degree of language mixing is distinct for related languages. Gómez Cruz (2017) demonstrates mixed morphology in classifiers, positional verbs, and verbs of motion (in use as main verbs and as verbs of motion). Like Ninde, Tojol-ab'al would represent language mixing of languages from closely related subgroups of the same language family, but unlike Ninde, the languages involved are morphologically complex.

Ultimately, the crucial factor in language mixing may be the recognizability of the product as a new language and as one that is mixed. The act of mixing may be motivated by secrecy, ritual, or play (Auer & Hakimov 2021:154), or reaffirming ethnocultural distinctness, sameness, or continuity. The creative product of language mixing need only contain salient elements of two languages that are recognizable to non-specialists. Where languages are particularly complex, a successful mix may exploit bilinguals' competence in manipulating those structures.

Within the context of Austronesian languages, the community on the island of Rapa Iti in French Polynesia had shifted to Tahitian (before an ongoing shift to French), but with knowledge of their Old Rapa language remembered by community elders, they have engaged in a process of language mixing with Tahitian to create Reo Rapa (Walworth 2017:204-238). This is a language adopted for use as a community language in some contexts and is associated with a new Rapa cultural identity. This mixing is partially represented in the lexicon, though elicited words tend to contain more Old Rapa than the words used in casual speech, which tend to contain more Tahitian. Nonetheless, lexical and grammatical choices are consistent across generations in sentence elicitation. There is no systematic distribution of source material from Tahitian vs. Old Rapa and a large number of forms are identical between the two sources. Nonetheless, the community strongly identifies their language as mixed.

Ninde's mixed genealogy will be addressed in the following sections in terms of lexicon, morphosyntactic structure, and phonology, all of which show potential for mixing. Within each domain, much of the remainder cannot be ascribed to a single language, or innovation has differentiated Ninde from both of its ancestors. Tracing Ninde's linguistic heritage through two sources has the effect that there is both regional continuity of Ninde as a local language and distinctness from the surrounding ethnic groups.

5.1.2 Historical identity of the Ninde language(s)

In oral history about Ninde, there was a *Ninde Tlepe* 'big Ninde' and *Ninde Tloulou* 'small Ninde', which were once distinct languages that are now inextricably fused as modern Ninde (Letpen 2018a,b). Some older Ninde speakers remember lexical and grammatical differences between the two languages, but neither language is spoken today. Ninde takes on another form in songs: both dance music accompanied by beats on a slit-gong (Kaiseng 2018) and text attributed to characters in traditional tales. This language is markedly different from spoken Ninde, and speakers generally do not offer translations of individual words during the transcription process. Speculation about the origins of sung forms is common practice, with some believing it to be archaic Ninde, and others believing it to be another, unidentifiable Malekula language.

Spoken Ninde shares with its sung counterpart many unique sound changes and lexical forms; furthermore, spoken and sung Ninde have core vocabulary and basic grammar uniquely shared with both NCM and SWB languages, but some of those innovations are attested only for one language (or for only sung or only spoken Ninde). The sound changes responsible for Ninde's phonology and lexicon can be ordered on the basis of isogloss bundling, feeding and blocking relationships, and comparison with historical documents containing older forms. I consider these documents to include sung forms in the oral-historical record and missionary texts in the written historical record. Based on the chronology of sound change, there are innovations variously inherited from the ancestors to NCM and SWB, but all of them further undergo changes that are unique to Ninde.

Letpen (2018a) reports that Big Ninde was spoken in the interior highland villages of *Mela'ai, Hlambutep, Nambüye, Nimbülyes, Venelu, Venemaxamb, Timinimb, Hlayemb, Luwanowoi, Hloxmare, Hloxtou, Opmomba*, and *M(ew)un*. Small Ninde was spoken along the coast in *Timbüse, Vivou, Minduwo, Milvou, Hloxominduwo,* and *Hloxotaxat*. Of note are *Minduwo* and *Hloxominduwo,* which would appear partially cognate (*hloxo* [l̪ˤoʁo] is a designated place for ceremonies) with one of the Neverver speakers' autonyms: *Mindu*. Barbour (2012:2-7) notes that a map drawn by the anthropologist Deacon places a *Mindu* tribe (perhaps in error) in Malekula's southwest, rather than near the Mindu River in the northern highlands. She supports the oral record over Deacon's map, in part on the basis that Neverver lacks much of the "rich coral reef vocabulary" of other Malekula languages. Small Ninde is a likely candidate for the source language for many marine terms in Ninde, including *nimiyal̪ˤe* 'fish' and *naʔau* 'reef', which have cognates in SWB, but no attested cognates in Neverver or Avava. The map in Figure 9 represents the location of languages referred to in this chapter, but the specific location of relevant locations are not identified.

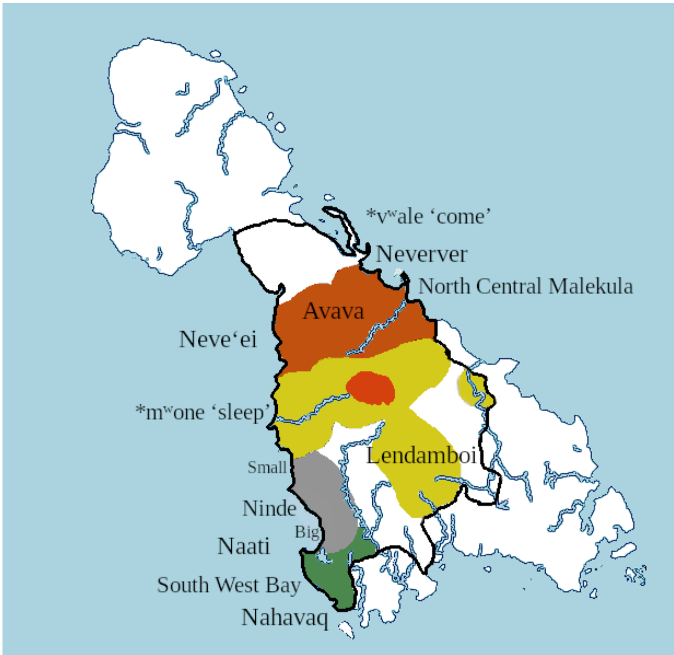

*Figure 9: Map of approximate traditional lands of language communities most referenced in this chapter. Colors are **red**: Northern Central Malekula languages, **green**: South West Bay languages, **yellow**: Remaining languages that have *vʷale 'come', **grey**: Ninde.*

5.1.3 Ethnocultural Identity

While Ninde is spoken as a first or second language by many newcomers and their children, the traditional custodians of the language are ethnically Mewun. Their traditional home is in the Blacksands region of South West Bay and the adjacent highlands, but many individuals live throughout Vanuatu and the greater Pacific region. Nonetheless, there is a one-to-one relationship between linguistic and ethnocultural identity, at least in popular conception: Ninde belongs to Mewun people, and Mewun identity is traced patrilineally.

Ethnographic details about Mewun people (the ethnic group associated with the Ninde language) suggest cultural ties to areas north of South West Bay. Mewun women traditionally wore rolled banana-fiber skirts (Mangke Kaisul 2018), whereas the women of other groups in South West Bay wore mat-like skirts woven from pandanus leaves and tied around the waist (Deacon & Wedgwood 1934b). On the other hand, Deacon also notes that a unique style of sand drawing existed in South West Bay, and he includes geometrical shapes from Mewun people in his survey of this distinctive regional style (Deacon & Wedgwood 1934a). These observations support popular wisdom that Mewun people are most like people of Dixon Reef (north of South West Bay).

People are traditionally associated directly, or by marriage, to a chiefly house called *nemel* (or *nakamal* in Bislama), while *nasara* were public meeting places for people of various nakamals for purposes of community decision-making and ritual. While there are generations-old nasara along the coast, many coastal-dwelling Mewun people identify with their postcolonial resettlement from the edge of Malekula's vast interior region. While the language is identified as mixed, the modern-day ethnicity of Mewun people is not mixed, but not monolithic. Within the village of Lawa', there are four so-called *stesin* (from English *station*) representing clans that play an important role in ritual life and determining in-law taboos: *Raki*, *Layax*, *Laindua*, and *Lalpupa*. These stations were designed to keep together communities that were displaced from the highlands. There is comprehensive documentation of the complex lineages now living in South West Bay, motivated by a sense that contemporary people are no longer aware of their pre-Christian genealogy (Kastow 2013).

5.2 Lexical Contributions

This section considers lexicon that is not generally shared across NCM and SWB languages. These are ordered in three ways: First, the Ninde words that appear cognate with NCM forms; then words that are cognates with SWB forms; and finally, synonym pairs made up of cognates with each of the two groups. The forms that have been chosen are generally ones that have attested, non-cognate forms in each of the language groups. This section presents basic words that are semantically equivalent with etymological differences, regardless of whether cognates with narrowed senses exist. This sets up the following section (§5.3) on irreconcilable sound changes, the consideration in the following chapters of words equally inherited by both language groups (§6), and lexical systems (§7), which may have lexical form and semantics from different sources.

5.2.1 Words with NCM cognates

Ninde shares a number of unique word forms with Avava, Neverver, and marginally also Neve'ei and Naman. Starting with the most inclusive isoglosses and moving onto the most restricted, Table 18 samples the many ways that Ninde's lexicon can pattern with that of other NCM languages (and not SWB) on the basis of both form and meaning. In these cases, the forms I referred to as *shared* are most likely *cognate*. Where possible, I refer to Proto North Central Vanuatu (PNCV) for the most recent ancestor of Malekula languages that has been reconstructed (Clark 2009).

NCM			SWB	
Neverver	**Avava**	**Ninde**	**Naati**	**Nahavaq**
lavlav 'harvest' (*ɣavax* 'plant (yams)')	*lap* 'plant (crops)'	*l̪ˤap* 'plant (anything)'	*imbʷiʔ* 'plant (crops)'	*(ʔi)ʔambʷiʔ* 'plant (crops)'
nuɣutn 'trunk'	*wutn* 'base, trunk'	*nuŋgute* 'trunk'	*mpati-* 'head, trunk'	–
ni-mda-li 'door'	*mata-li* 'door'	*ni-mit-lu* 'door'	*mboʔon jum, mbʷysys* 'door'	*mbusus* 'door'
nimɣut 'person, man'	*mu:t* 'person'	*numyŋgut* 'person, man'	*nemurut* 'man, husband'	*mʷor* 'man'
(*nilβu-* 'tooth')	*boroh* 'tooth'	*nə-mboʁi* 'tooth'	*ne-liβo-* 'tooth'	*ni-liβo-* 'tooth'

Table 18: Rare lexical forms are shared between Ninde and both NCM and SWB languages.

In each of these sets, the NCM and SWB languages – excluding Ninde – can be differentiated as non-cognate. Ninde patterns in each case with other NCM languages:

1. Forms for ‘door’ are derived either from PNCV compounds *mata + *liu[21] ‘eye of the door’ (NCM + Neve’ei and Naman) or *bʷaŋo + *yumʷa ‘mouth of the house’ (Naati). Note that there is expected homophony with reflexes of PNCV *liua ‘arrow’ after final vowel loss. Naati and Nahavaq both have a form *nemb(w)ysys* ‘door’, which resembles Ninde *ambus*- ‘behind’ and *ambuse* ‘at the back of the house, opposite the door’. Perhaps Ninde shares this form with SWB and its function as a part of the house, but the meaning could be influenced by *nembyse*- ‘tail, end of’.

2. Forms for ‘trunk, base’ are either a putative *xut-n (NCM + Neve’ei, Naman) or *bʷati-n < POc *buRit (Naati and Nahavaq). Clark (2009) does not reconstruct a form for this set for PNCV, but POc *buRit accounts for V’ënen Taut *pla~ple* and Tirax *bot*. A connection is often drawn in Ninde between near-homophones *nuŋgute* ‘trunk’ and *nuŋguwute* ‘octopus, squid’ because of the tentacle-like appearance of roots; the NCM forms could be doublets. Naati has merged its form with the word for ‘head’ < PNCV *bʷatu.

3. Verb forms for ‘plant (v.tr.)’ may be from PNCV *lavo ‘plant’ or another form reflecting putative *ka(vʷ,bʷ)ik. Neverver reserves *lav* for planting seed crops and *xavax~xavux* for planting yams, but Avava, Ninde, and Neve’ei use *lavo reflexes for planting seeds and tubers (Ninde has a different verb for propagation from

[21] My own reconstruction based on comparison with forms V’ënen Taut *liu* and Tirax *nali* ‘door’.

cuttings). Naman, Naati, and Nahavaq have generalized the forms resembling *ka(v^w,b^w)ik for all kinds of garden planting. Nominal forms of *ka(v^w,b^w)ik refer to the earth mounds formed over a planted yam, so an ultimate origin in PMP *qi(m)pun 'heap' > POc *opun ~ *upun 'heap' is likely.

4. A form for 'man' reflecting a putative intermediate proto *mukut or *muxut (Avava, Neverver, and Ninde) vs. one that is apparently a narrowing of PNCV *m^{w}ara ~ *m^{w}era 'person. Reflexes of *r from both NCM and SWB sources are the uvular fricative/trill in Ninde, but they do not surface as /ŋg/, whereas *k occasionally does after a nasal stop. The NCM form superficially resembles PNCV *ata-m^{w}aʔane 'man, male'.

5. Only Ninde and Avava have a putative *borVx- 'tooth' instead of more widespread *liv^{w}o- 'tooth'. The more widespread form may be retained in a more restricted sense as Ninde *nelwoʁawoʁa-* 'molar'.

This runs the gamut from SWB innovations that exclude Ninde (example 1) to innovations affecting NCM (2-5). Within the greater NCM group, we find innovations that exclude Naman (3), Neve'ei (4), and Neverver (5).

5.2.2 Words with SWB Cognates

Other forms are uniquely shared between Ninde and the other SWB languages (Naati and Nahavaq). Since documentation of Na'ahai/Naha'ai is comparatively less advanced, it is not considered here. In many cases, there is no attested cognate; this is due to the sparse documentation of Naati, but many words for coastal flora and fauna cannot be found in the NCM languages. The forms are presented in Table 19.

NCM		SWB		
Neverver	**Avava**	**Ninde**	**Naati**	**Nahavaq**
(*nibutuan* ‘hill’)	(*butuan* ‘hill’)	*noʁku* ‘hill’	*neruʔuh* ‘hill’	*niruʔuh* ‘hill’
nibet ‘breadfruit’	*ebet* ‘breadfruit’	*nimbytep* ‘breadfruit’	*mbatap* ‘breadfruit’	*mbetep* ‘breadfruit’
nelag ‘wind, sky’	*alaŋ [nan auh]* ‘cloud’	*nimiləŋge* ‘cloud, cloud cover, sky’	*meliŋg* ‘sky’	*meliŋg* ‘cloud’
norgo- ‘finger’		*nəmbaʁaʔ* ‘finger’	*mbaruʔ* ‘finger’	*aŋgar* ‘finger’
(*le-sbo* ‘middle finger’)	(suʙu- ‘finger’)			
–	(*toʙur* ‘hibiscus’)	*ne<saŋga>saŋga wol* ‘hibiscus’	*nesiŋge͡ul* ‘hibiscus’	–

Table 19: Cognate sets including Ninde, Naati, and Nahavaq.

1. Forms for ‘hill’ reflect either some form *rukuh (SWB – cf. POc *(k)oro ‘hill’, which does not account for the final syllable) or PNCV *vitu ‘hill, mountain’ > [bVtuan] (Avava, Neverver, Neve’ei, and Naman).

2. Forms for ‘breadfruit’ are all from PNCV *batavu, but SWB languages have retained more of the form (intermediate *batap) as compared to NCM languages (intermediate *bet).

3. Forms for ‘cloud’ are either from PNCV *maligo ‘cloud, dark’ (SWB) or *laŋi ‘wind’ (Avava, Neverver, and Naman). This form is sometimes used for ‘sky’ or ‘heaven’ in Ninde and the Naati and Nahavaq forms exist alongside a word (also) meaning ‘cloud’.

4. Forms for ‘finger’ are divergent from PNCV *bisu in all languages except Avava. Curiously, Nahavaq, Naati, Ninde, and Neverver have typical reflexes of *r and *g, but only Nahavaq has a rhotic following the velar. Ninde and Naati alone share an initial [mba] and a glottal instead of a velar, which is common for word-final velar stops.

5. A form for ‘hibiscus’ reflecting PNCV *b^{w}akala and *varu is apparently not found (the Avava form could reasonably be derived from the latter), but the Naati and Ninde forms are rather similar. There is no form attested for Nahavaq.

5.2.3 Words with blended NCM and SWB senses

Illustrating the problems with assessing parentage, a number of Ninde words seem to be influenced in their semantics by both sources. These are particularly compelling when only one sense is listed, but only a speaker of the language can confirm the limited semantic range of the source items. Particularly strong examples are represented in Table 20.

NCM			SWB	
Neverver	**Avava**	**Ninde**	**Naati**	**Nahavaq**
ra- 'on'	*ra-* 'on'	*(a)ʁa-* 'on, with (instrumental'	*ra-* 'with (instrumental), to, under, beneath'	–
–	*ndam* 'shout, yell, (of dog) bark'	*ndam*^{w}*e* 'respond, allow, call (of *nesaŋge* bird)'	–	*ndam*^{w} 'agree, accept'
–	*(kil)kila* 'look, open the eyes'	*(ŋgo)ŋgolou* 'look, open the eyes, look for'	*(ŋgi)ŋgile͡u* 'look for'	*(ŋgi)ŋgilew* 'look for'

Table 20: Ninde forms that have blended senses of NCM and SWB cognate sets.

1. Ninde *ʙa-* is both a locative ‘on’ like its NCM cognates and an instrumental like its Naati cognate. Avava oblique-marking *i-*NOUN and Neverver applicative VERB-*ix* add instruments to clauses (and may represent differing parses/analyses of the same historical structure), and Naati has a general locative *le-* that also serves as a supraessive. In fact, the locative sense of Naati makes it a partial antonym of Ninde’s locative.

2. Ninde *ndamʷe* means ‘to respond’ (usually in the affirmative), but it is also means ‘call (of *nesaŋge* bird)’. In Avava, the near-homophone *ndam* is only provided with the sense ‘bark (of dogs)’.

3. Ninde *ŋgolou* can mean ‘open the eyes, flash (of lightning), signal, look’ like its Avava cognate, but (especially in combination with *woʙ* ‘about’), it means ‘look for’ like in the other SWB languages. Avava *(la) wahi* and Neverver *llaŋ* ‘look for, seek’ are the forms otherwise used in the SWB-like sense, and no forms in SWB seem to mean ‘open the eyes’.

This suggests another way that a lexicon can be mixed: using form and meaning from two sources together.

5.2.4 Synonym pairs with mixed cognacy

Ninde has some sets of synonyms and doublets that are shared with both NCM and SWB languages. Section §5.1 focused on lexical items exhibiting competing sound changes, but this section presents lexicon that strongly points to language mixing that draws from two lexifiers. Some words are doublets, ultimately inherited from the same PNCV and POc sources. Others are synonym pairs that share the unique lexical innovations of both groups. Finally, some of Ninde's lexicon exhibits unique changes associated with one or the other group, including metathesis and morphological changes. In this section, the tables will represent only one NCM and one SWB cognate for comparison per doublet.

Some words are near-synonyms, but the origin of each synonym in a pair can be linked to NCM cognates and the other to SWB cognates (Table 21):

Other NCM	Ninde (NCM)	Ninde (SWB)	Other SWB
sxen 'not be so' (Neverver)	*səke* 'not exist'	*eheʔ* 'no'	*eheʔ* 'no' (Nahavaq)
levlep 'big' (Avava), (*tə)leb* 'big' (Naman)	*tlepe* 'big'	*l̪ˤapo* ~ *l̪ˤapʷe* 'many, big'	*ilambʷ~ilamb* 'many' (both)
as-n 'jaw' (Avava)	*nes(i)-ne* 'cheek'	*nene-ne* 'jaw bone'	*nehe*- 'chin, jaw' (both)
tabatn 'begin' (Neverver)	*tupatne*	*tumbute*	*tumbʷatin* (Nahavaq)
anam 'mosquito' (Avava)	*nanŋgamʷe* ~ *nanŋgamo*	*nudukas*	*te͡uʔas* (Naati)

Table 21: Synonym pairs in Ninde with non-cognate sources from the NCM and the SWB sources.

1. While *səke* 'not exist' is used in negative existential constructions and counterfactual clauses, *eheʔ* 'no' is uninflected and cannot be combined with other forms. The interjection is one of very few words containing /h/ in Ninde – words of SWB origin

have deleted glottal fricatives. While the debuccalized form is virtually identical with that of Nahavaq, it is also present in Neve'ei and identified as a loan (Musgrave 2007: 9).

2. While *tlepe* 'big' is a singular modifier of size, *l̪ˤapʷe~l̪ˤapo* generally means 'many' (i.e., it functions as a quantifier), as in the other SWB languages; in some frozen lexical constructions, the latter can also mean 'big', as in *mbaʁaʔ l̪ˤapʷe* 'thumb (lit.: big finger)'. Apart from the semantic matches, Ninde shares in *tlepe* the vowel fronting with Avava and Naman and a fused relativizer *t-* is in one variant in Naman. Like the SWB forms, Ninde formerly had the fused third-person singular subject prefix *i-* (which conditioned the pharyngealized reflex of inherited *l). In Avava and Naman, this form is fused to all monosyllabic verbs (but not if they are reduplicated or relativized as they are here). In SWB, they are only used with quantifiers. In Ninde, both of these words can be used predicatively or within the noun-phrase without relativization – quantifiers are simply a subclass of stative verbs. The erstwhile relativizing *t-* of *tlepe* is never removed, and the form can be relativized as *ti-tlepe*.

3. The two forms for 'jaw' have slightly different meanings: *nes(i)-ne* is used for the cheek of a living person (with flesh), but *nene-ne* is used for the jawbone. SWB *h was virtually always lost in Ninde, but the expected form exists as another word:

nene ‘of it’, so *nenene* may have a reduplicated root as a form of homophony avoidance.

4. In one case, the distribution of forms appears to be a matter of interpersonal variation: *tupatne* vs. *tumbute* ‘begin’. Both of these are straightforward compounds of *tu* ‘put (on)’ and *nu-mbute-(ne)* ‘(its) head’. The only evidence in favor of attributing *tupatne* to an NCM source is that /n/ (like /m/) is permitted as a syllable nucleus, whereas the epenthetic vowel of SWB might favor removal of the final *n.

5. The form *nudukas* ‘mosquito’ is the term most people use; however, clan-based taboos dictate that some people may not use this form. People affected by the taboo instead say *nanŋgam^{w}e* < PNCV *namu-ki. Since the word in Neverver is *nituɣaz*, both forms could have been inherited from NCM alone.

These represent some of the ways that doublets of mixed origin can be differentiated in Ninde. Some are differentiated by a complete change in word class (1) or a more subtle one (2); a narrowing of a broad sense available in only one of the languages (3); or simple synonyms used by different people at will (4) or for indexically prescribed purposes (5).

5.3 Phonology

Ninde shares sound changes with both NCM and SWB languages, but generally these affect complementary parts of the lexicon. This has been a problem for assessing phylogenetic subgrouping, since different parts of the core vocabulary can be used to assert linguistic closeness to either language.

This section addresses inherited phonological traits that describe how the phoneme inventory is organized. Diachronic phonology will be addressed in depth in §6, where sound changes will be modeled as taking input from two source languages (one identifiable as an NCM language and the other as a SWB language). In this section, I first discuss in §5.3.1 some of the source-independent parts of the diachronic phonology that will contextualize later claims about morpho-syntactic inheritance. In §5.3.2, I consider whether Ninde shares phonology with NCM or SWB, specifically the status of prenasalized stops in SWB and the possibility that proto NCM had geminate consonants.

5.3.1 Diachronic phonology

There are some changes that allow for a small amount of internal reconstruction within Ninde: the distribution of Ninde's two lateral liquids and palatalization of root-initial fricatives. Both of these diachronic changes reveal: (1) that Ninde has once had a third-person singular subject prefix *i- fused to monosyllabic verb stems, (2) that this initial vowel was devoiced, (3) and that sibilants were intermediately debuccalized, conditioning devoicing before they were ultimately deleted.

Development of the Lateral Contrast

Evidence for pre-Ninde *h lies almost exclusively in the distribution of one phoneme in particular, a lateral phoneme in Ninde transcribed here as /l̪ˤ/ that Mewun people often call *heavy /l/*. This diagnostic phoneme has resisted phonetic description, but is produced as some type of lateral coronal continuant. It has been described as a voiced lateral fricative, a lateral interdental fricative, and as a lateral retroflex approximant (Lynch 2016). Analysis by Caroline Crouch (p.c.) show that it is not a true fricative (based on periodicity), but that it has greater variation in phonation, tending toward breathiness for some speakers (based on spectral tilt). It is more accurately described as a pharyngealized dental lateral approximant, transcribed [l̪ˤ], and in usual articulation is laminal and interdental (at least in hyperarticulated speech), with a secondary constriction at the upper lip and often lip rounding. At least one speaker reported the intuition that the primary point of constriction is linguolabial.

Ninde's contrastive heavy /l/ represents a conditioned sound change whereby Pre-Ninde *l in some environments changed in quality somewhere along the way; this much is undisputed, since the alternative is that Ninde is the sole language to retain an older lateral contrast reconstructible to Proto Austronesian. Lynch (2008) has analyzed the diachrony of heavy /l/ in terms of its vocalic context: instances of *l became pharyngealized unless adjacent to a front or high vowel; i.e., if the vowels on either side is either *a or *o at a syllable boundary, or if the vowels on both sides are *a or *o (Lynch 2012). This longstanding account would appear to explain the back quality of heavy /l/ and suggests that pharyngealization and the other articulatory changes are incidental phenomena associated with tongue root retraction. Indeed, our lexicography work has only identified seven unique word roots with plain lateral approximants between low vowels and 12 word-initial or in complex onsets, 16 word-final following a low vowel, and two with vocalic contexts that included a vowel invariably recorded as schwa (which can be an allophone of high or low vowels).

In many cases, the pharyngealized lateral corresponds to sequences of what was likely to be pre-Ninde /h/ and /l/, with or without an intervening vowel. This /h/ was the reflex of Proto Oceanic *s, as shown in Table 22:

Protos							
POc	PNCV	Ninde	Neverver	Avava	Lendamboi	Nahavaq	Naati
PMP *suluq 'torch'	*sulu 'torch, shine onto'	*je/joḻ̪ˤo*	*su/sul* 'shine, glow'	*sil* 'fish by torch-light'	–	*hul/hul* 'kindle fire'	–
PMP *salaq 'wrong, miss'[22]	–	*-pi**j**aḻ̪ˤe*	*pu**s**eḻ*	*biḻi**h***	–	*piḻe**h***	–
(PMP *hikan 'fish')	(*ika)	*ni/mi**j**aḻ̪ˤe*	*(netas)*	*(iki)*	*(na)me**s**eḻe*	*ma**h**aḻ*	*ma**h**aḻ*
*salan 'road'	***s**aḻa	*na/ḻ̪ˤe*	*nesal*	*a/**s**aḻ*	*(ne)**s**eḻe*	*nahal*	*(na)hal*
*las̱oR 'testicle'	*ḻa**s**o	*no/ḻ̪ˤo-*	–	*ḻe**s**e-~liso-*	–	*no/ḻo**h**o-*	*ni/ḻo**h**o*
*qalo 'sun' (+ PMP *sinaR 'sunbeam')	*yaḻo + ***s**ina	*na-ḻ̪ˤa-**s**ne*[23]	*niaḻ*	*a/naḻ*	*(ne)jeḻe*	*ni/naḻ*	*(ni)naḻ*

*Table 22: Cognate sets with *l in the environment of a voiceless fricative (bolded), which conditions pharyngealization (underlined) in Ninde.*

[22] But cf. PPh *pus<el>it 'slip away'.

[23] Ninde is either a NOUN-VERB compound with *sne* 'shine', or a rebracketing of *nahla snesne* > *nahlasne sne*

Note those contexts where *l is cannot simply have been pharyngealized in the context of *a or *o: *suluq 'torch, shine on' (but Clark identifies *wul* 'light a burning thing' and derivatives as the sole descendant form in Ninde) and *pijal̪ˤe* 'err, miss', which has viable cognates (affected by metathesis), but none with low vowels. At least in some cases, Ninde developed pharyngealized laterals in the context near a (usually deleted) *s.

Often, where there was no preceding *s (except later reflexes of *z), the lateral *l was not pharyngealized word-medially (Table 23):

'the sun is shining'. Either way, the word-final /a/ does not undergo word-final raising and the form *nahlasne* 'sun' is provided in isolation.

Protos						
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
*keza ‘intense blue-green’	*mal̲a-keza ‘blue, green’	*mel̲(u)k(i)se*	*xaz*	*melih*	*ʔasen*	*(mal)ʔasan*
PWMP *haluh ‘large lizard sp.’	*ŋgala ‘lizard’?	*nel̲ei*	*le/ɣul̲i/aŋ*	*jal̲a*	–	*nil̲͡ei*
PMP *k(u,i)lat ‘open eyes’	*kila-la	*ŋgol̲ou*	–	*kil̲a*	*ŋgil̲ew*	*ŋgil̲eu*
*kanon, PWMP *butil ‘kernel’	*kutu/*biri ‘seed’	*nuwul̲e* ‘seed, kernel’	*(ni/ʙulun)* ‘seed’	*(ʙul)* ‘(breadfruit) seed’	*nuʔul̲i-* ‘kernel’	*noʔul̲i-* ‘seed’
–	–	*l̲aŋgas* ‘hold with tongs’	–	*kal/kal̲at* ‘tongs’	*ŋgal̲as* ‘hold with tongs’	–

*Table 23: Reflexes of *l (underlined) that are not in environments for pharyngealization in Ninde.*

Crucially, the form *laŋgas* ‘hold with tongs’ is in the environment before /a/ in any language that has an apparent cognate, but Ninde’s form would be affected by metathesis. The reflexes of PNCV *ŋgala ‘lizard’ or *malakeza ‘blue, green’ raise questions about the time depth of such a change, since this is the shallowest proto language reconstruction and *l is in an environment Lynch postulates for pharyngealization.

Ultimately, the conditioning relationship could just as well have been the reverse: heavy /l/ provided the context for vowel lowering. Synchronic patterns of noun stems appearing with the *nV- nominal prefix vs. a locative **l̪ˤ*V- prefix reveal the lowering effect that heavy /l/ has on adjacent vowels (Table 24):

Nominal form	**Locative**	**Gloss**
nu̲-ŋgute	*l̪ˤa̲-ŋgo̲te*	‘tree trunk, base’
ni̲-mi̲ləŋge	*l̪ˤa̲-me̲ləŋge*	‘cloud, sky’
ne̲-mel	*l̪ˤa̲-mel*	‘chief’s house’
no̲-mbo(ŋo)-ne	*l̪ˤa̲-po(ŋo)-ne*	‘their (sg.) mouth
nə̲-mbepete	*l̪ˤa̲-pepete*	‘yam platform’
ne̲-tene	*l̪ˤa̲-ta̲ne*	‘ground’

Table 24: Vowel alternations (underlined) between simple nominal forms and their corresponding locative, which contains a (generally) pharyngealized lateral.

It is just as likely that heavy /l/, with its lowering effect on non-final vowels, blocked or reversed diachronically lowered vowels. Such alternations could be explained away as a difference in the vowel of the prefix (harmonizing *nV̲*- vs. invariant *l̪ˤa̲*-), but note the distribution of /l̪ˤ/: it only appears as a simple onset of syllables with low-vowels /a/, /o/, or word-final [e] – /a/ is synchronically raised to [e] word-finally. For this reason, it patterns phonologically with uvular /ʁ/, which also has the effect of lowering /e/ to [a] and /u/ to [o] or [a]. The uvular, however, also lowers vowels preceding it and word-final /a/ is not raised after it.

The paucity of plain /l/ in the context of low vowels can be compared to the distribution of /t/: only six instances of /t/ are found between low vowels in unique word roots (plus three including invariable schwa), seven instances preceding one syllable-initially (plus four including schwa), and four word-finally (plus one including schwa). These forms include two redundant forms that are reduplicated *tVtV* sequences and five that are influenced by lowering effects of /l̪ˤ/ or /ʁ/. On the whole (though not corrected for morphological complexity), there are 688 instances of [t], compared to 601 of [l] and 317 [l̪ˤ] in the lexical database of 1,986 native forms. A comparison of these rates would suggest that *t was similarly lost in the context of low vowels, but the more likely story is that the reflex of *a is generally [e], except where there is a post-velar consonant. Remaining plain /l/ still found in the environments of low vowels tend to become heavy /l/, but these are often the cases in variation like *nalaŋ~nal̪ˤaŋ* 'wind; taro', *nomboloŋgo~nombol̪ˤoŋgo* 'bone; cottonwood'.

Ultimately, Lynch's account of low-vowel conditioning as the sole factor in lateral pharyngealization, as the reconstructed vowel qualities themselves, is not falsifiable without an accompanying shallow linguistic reconstruction. It relies on reconstructions at a time scale too deep (PNCV and POc) with an understanding that Ninde shares some of the innovations of its neighbors. This difficulty to reconcile ancient forms with recent innovation is also at play in his later treatment of final vowels. Lynch (2012:6-11) argues that Ninde, along with just four languages of Malekula's north (Nese, Vovo, Botovro, and Vao), is among the only Malekula languages to have retained most final vowels. Since he considers Ninde to be related to languages that share the final vowel loss of the majority of Malekula languages, he postulates that the same pattern was innovated seven times after divergence. Finding no conditioning environments in Ninde's unique patterns of retention, he concludes that final-vowel deletion is in progress, yielding only tendential loss of especially *i but also *u, and around half of the instances of *o and *e. As an alternative to this account, Ninde's final vowels will instead be explained as recent innovations specific to that language.

It seems, then, that the distribution of /l̪ˤ/, at least in the context of high vowels and at word edges, reveals the existence of some former conditioning environment with a pharyngealizing effect on *l. In at least one case, the pharyngealizing (or otherwise altering) effect of the *s/*z appears to work at a distance, or more likely, metathesis is suggested. Table 25 shows correspondences and reconstructions for Ninde *namal̪ˤaʔ* 'cold' (including Lynch's reconstruction for Proto Malekula) that suggest that *k and *s were at some point glottal in pre-Ninde:

Protos						
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
–	PMal *malakaso	*namal̪ˤaʔ*	(*ne/tgar*)	*melekit*	*malʔah*	*malʔah*
*pisiko 'flesh, lean meat'	*visiko	*no/woʔ* 'body'	*ni/visxo-n*	*i/pso-*	–	*ne/vyʔoh*

Table 25: Ninde forms which suggest glottal metathesis involving a stop and (now deleted) fricative.

Assuming an original order of *k followed by *s is reflected by Avava ([k] and [t]) and SWB, then metathesis is needed to explain the final glottal stop of Ninde. Such a process would mirror that seen in Naati's reflex for 'flesh', where a debuccalized *s and *k are prone to metathesis. Pre-Ninde likely retained segmental /h/ as a reflex of *s, and it is this that conditioned pharyngealization of laterals.

Word-initial Insertion (Prothesis) of *h before Verbal Roots

A prefixed *i-, if realized as *hi-,* would have provided the context for laterals to become pharyngealized, but only in monosyllabic verbs. Word-initially, the conditioning environment for *l pharyngealization appears to be irregular. The great majority of words beginning with /l̪ˤ/ are verbs or adjectives (some of which are shown in Table 26), in line with the fact that virtually all nouns in Ninde begin with a fused *na- nominal prefix (not an environment for pharyngealization):

Protos						
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
PAN *ala ‘take, marry’	*laki	*l̪ˤaʔ* ‘marry, married’	*lav* ‘marry, take, get’	*dak/daka* ‘married’	*leʔ* ‘marry’	*leʔ* ‘marry’
–	*lavo	*l̪ˤap*	*lav/lav* ‘harvest’	*lap*	–	–
(PPh *lak(e)táw ‘jump over’)	*lakau ‘pass over’	*l̪ˤou* ‘go over’	*lɣa*	*jala*	*leʔ͡eu*	–
*lakas ‘take a step’	*lako	l̪ˤo(ɲe) ‘go’	(*vavu* ‘walk’)	(*sol* ‘walk’)	*loŋ* ‘go, walk’	*loŋ* ‘go, walk’

Table 26: Ninde forms affected by root-initial lateral pharyngealization in verbs, compared to cognates.

In each of these examples, there is no evidence to support the development of a pre-Ninde /h/ preceding the /l/ if one is to examine only the verb roots.

A possible explanation for this additional distribution is in patterns of vowel devoicing in nearby Lendamboi. In more recent years, David Kaiar of South West Bay has recorded word lists and connected speech in a variety he calls Nevitangiene. Approximately

180 minutes of recordings he recorded with Crouch and me in 2017 and approximately five minutes that he recorded with Leslie Su in 2019 show clear patterns of vowel devoicing. Crucially, one context for devoiding involves the reflex of *i- 3SG.SUBJ that has been retained in this form by NCM and SWB languages, except for Ninde. A spectrogram in Figure 10 shows the initial devoicing of this morpheme at a phrase boundary:

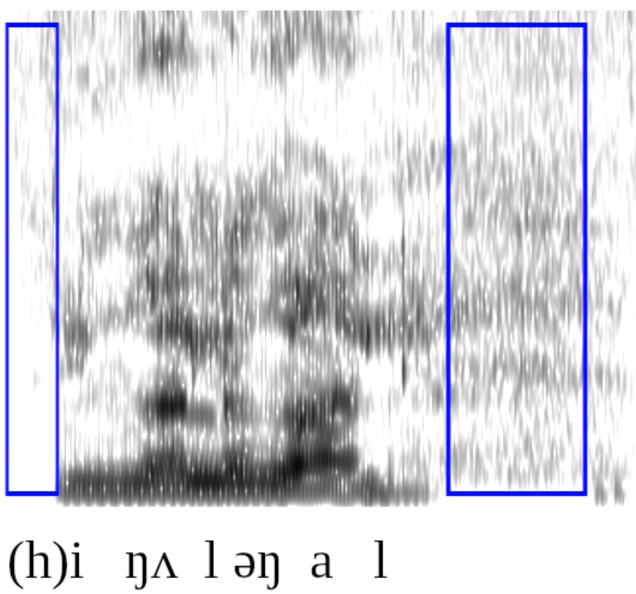

Figure 10: Spectrogram of [iŋʌləŋal] 'they (sg.) are happy' with the devoiced portion of the prefix i- in the first blue box. The final [l] is partially devoiced and followed by a nasal outbreath (in the second blue box).

Singular third-person subjects are unmarked in modern Ninde (though oral and written history also represent use of the widespread *i-), but there is strong comparative evidence that it was once fused to some verb roots. Fusion of the third-person singular realis subject prefix with phonologically short (usually monosyllabic or smaller) verb roots is attested in some languages of Malekula. In Avava (Crowley 2006a) and Naman (Crowley 2006b), inflectional prefixes have been reanalyzed as part of monosyllabic verb roots, at least when there are no suffixes or reduplication. In all NCM and SWB languages, this is true for numerals generally (in many ways a subclass of verbs), but it pertains to all monosyllabic verbs in Naman and only some of the verbs in Avava (Table 27).

Protos		**NCM**		**NCM+**	
POc	**PNCV**	**Ninde**	**Avava**	**Naman**	**Neve'ei**
*pano 'go'	*va	*ve*	*ip*	*iv*	–
*tolu 'three'	*tolu	*tl*	*itl*	*itl*	*itl*
*pati 'four'	*vati	*ves*	*ivat*	*ives*	*ivah*

Table 27: Monosyllabic word forms in Avava, Naman, and Neve'ei (compared to Ninde cognates) affected by fusion of the third-person singular subject prefix i-*, which serves to create a minimal disyllabic word.*

Avava (Crowley 2006a:67-75) and Neverver (Barbour 2012:166-173) additionally both have short and long forms of inflectional prefixes, and the latter involve an extra /i/ between the prefix and root. While the variants in Neverver can be explained phonologically by limits of syllabification, Avava has inflectional classes (Table 28) that could reflect a much broader domain (like that described for Naman).

Short Prefixes		(continued)		Long Prefixes	
tur	'stand'	*riit*	'punch'	*per*	'work'
va	'do, make, say'	*lik*	'tie'	*ka*	'call'
jor	'chase, grate'	*jet*	'feed'	*sil*	'go torch-fishing'
val	'fight'	*dah*	'descend'	*kin*	'peel (tuber, banana)'
v^{w}el	'come'	*sap*	'dance'	*soŋ*	'cook in bamboo'
jan	'eat (tr.)'	*sak*	'ascend'	*kan*	'eat (intr.)'
la	'see'	*jeŋ*	'remove (pudding)'	*tem*	'step'
luk	'stay, live'	*tn*	'roast'	*pih*	'tie'
tok	'be located, exist'			*pom*	'go ahead'

Table 28: Forms taking short prefixes vs. long prefixes with an additional /i/. The first inflectional class is split across two columns for space. Reproduced from Crowley and Lynch (2006a:72).

In this case, Avava *tur* 'stand' is inflected with a short prefix (*at-tur* 'they stood') and *per* 'work' is inflected with a long prefix (*ati-per* 'they worked'). Some forms with "short" prefixes still show possible effects of a fused *i-: *jan* 'eat' (PNCV *kani) and *jor* 'grate' (PNCV *kori) have unexpected /j/ reflexes of *k; *ka* 'call' has a rare verb-initial /k/, but the Neverver cognate *k:e* 'call' has an initial geminate that suggests a proto NCM *kka < pre-proto *kaka (reduced reduplication). None of the long prefix verbs are attested in texts or examples provided by Crowley in reduplicated form; seven of the short prefix words appear in reduplicated form; and *v^{w}el* 'come' is used in texts with both short and long prefixes. While the allomorphy is not strictly selected by synchronic phonology in Avava, the diachronic explanation is likely rooted in the phonology of its ancestor.

Ninde and Nasarian have an unusual and irregular change *s > i (like Avava's *k > j) in select words (Lynch 2012), but all of the affected words are verbs with the reflex in word-initial position. It is likely that the /i/ (usually now /j/ in Ninde) is either the palatalized reflex of *s > *h or a reduced remnant[24] of the (now lost) third-person inflected form. Alternatively, Pre-Ninde root-initial *h (< PNCV *s) and *x (< PNCV *k) in verbs were palatalized as /j/. The palatal approximant also surfaces as a reflex of PNCV *k in Avava. Table 29 shows the patterns of palatalization in Avava and Ninde, including Ninde irrealis forms, as compared to other NCM and SWB languages:

[24] Either through a change of *i > j / _V, or an excresent /j/ was retained after deletion of /i-/.

Protos		NCM				SWB	
POc	**PNCV**	**Ninde (Real.)**	**Ninde (Irr.)**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
*kani ‘eat’	*kani	*jen*	*pan*	*ɣan*	*jan/jan*	*ʔan*	*ʔan*
*ka(Ra)ti ‘bite’	*kaRa-ti	*jes*	*pas*	*ɣas*	*jat*	*ʔas*	*ʔas*
*korit ‘scrape, grate’	*ko(r,R)i	*joʁ* ‘scrape out, scratch’ *ŋgoʁ* ‘scratch’	*poʁ*	*ɣur* ‘itch, scrape’	*jor/jor*	*ŋoro/ŋor* ‘scrape with teeth’ *ŋgar* ‘scrape straight’	–
*sake ‘ascend’	*saka	*jaʔ*	*paʔ*	*sax*	*sak*	*haʔ*	*haʔ*
(PWMP *sayaw ‘dance’)	*savʷa	*jawo*	*pawo*	*sav*	*sap*	*hapʷ*	*hap*
*jaRu ‘blend voices in song’	–	*jal̪ˤe* ‘sing’	*pjal̪ˤe*	*jer* ‘sing’	*jal/jal* ‘sing’	*jal/jal* ‘sing’	*al* ‘sing’
–	*sale ‘hover’	*jele* ‘fly’	*pijele*	*jal*	*jal/jal*	–	–
–	–	*jaʁ* ‘make, do’	*pjaʁ*	–	–	*jar* ‘finish’	–

*Table 29: Forms affected by palatalization of *s and *k in Ninde and Avava, with forms compared across mood (in Ninde) and languages.*

Layered origins are reflected today in Ninde irrealis forms, which take on a prefix *p(e)-*; this replaces the initial /j/ if it is a reflex of *x or *s, but not if it corresponds to /j/ in the other languages (those which do not have corresponding /j/ as a reflex of any fricative). This strongly suggests that Ninde *jele* ‘hover’ is not a reflex of PNCV *sale (and an exception to the account of conditioning by vowel quality, not to conditioning by voice quality).

It seems very likely that monosyllabic verb roots, when not reduplicated or suffixed, were propped up by a prefix *i-* or *hi-*, (morphologically) like Avava and Naman and (phonologically) like Lendamboi. Whether the third-person inflectional marking would have been incorporated into verb roots or simply provided the phonetic context for sound change in the root, the process like the one in Ninde and a number of its Malekula relatives is formulated elsewhere as Watkins’ law, whereby a form inflected for third-person serves as the basis for the whole paradigm. The changes discussed up to this point are rather well established diachronic changes.

Unpharyngealized *l in Root-initial Devoicing Contexts

In the cases where initial *l is unaffected, reduplication may play a significant role. The number of /l/-initial verb roots is much fewer than those with initial /l̪ˤ/, but seem to be present in adjectives and verbs that have full or partial reduplication in Ninde, or in every attested cognate to the exclusion of Ninde (Table 30):

Protos						
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
(*aŋo-aŋo-ana 'yellow')	(*aŋo-aŋo)	*lit*	(*jaŋjaŋ*)	*lit*	*lutlut*	*lytlyt*
*lua-ki	*lua	*liluwo*	–	*lualu*	*(luwe)lu*	*l*
(*mamis 'sweet')	–	*lum* 'sweet; salty'	*l:um* 'be tasty'	*lum* 'sweet'	*(kaskas)*	*(kaskas)*

Table 30: Exceptions that could be backformed from reduplicated words.

Note that 'yellow' is syntactically unlike other verbs – in Neve'ei, Musgrave (2007) identifies *lulut* 'yellow' as an "adjectival verb". In Ninde, *lit* 'yellow' behaves just as she describes this class[25] in Neve'ei: the word can serve as the head of a predicate and can host verbal affixes, but a relativizer is not required for attributive use in the noun phrase. Reduplication is used in Malekula languages to derive intransitive verbs from transitive ones and/or adjectives from verbs. Reduplication and syncope also yields Neverver initial geminates – it is not clear if Avava or Ninde formerly had geminates because of the dearth of correspondences. These forms may have resisted lateral pharyngealization if, as adjectival

[25]Ninde does not have this distinction between adjectives and adjectival verbs. Unlike in Neve'ei, all underived adjectives can be used predicatively.

verbs or reduplicated forms, they did not undergo fusion with the third-person singular subject.

Other Word-Edge Contexts for Lateral Pharyngealization

The context for vocalic devoicing in Nevitangiene extends beyond just the initial vowel of *i*-, as Figure 11 shows in the partial devoicing of the final /l/. Lendamboi languages may have been rife with allophonic vowel devoicing in the past like Kaiar's Nevitangiene today. He identifies as the last full speaker of this language, and comparison with other language varieties traditionally spoken in Lendamboi suggest that he may uniquely retain what is described as post-tonic "destressing" in the form of vocalic devoicing: Charpentier (1982) and Tryon (1976) record word forms from what is likely a Lendamboi dialect continuum. The names of four mutually intelligible varieties are recorded as Ayiauleián^{a}, Mbotkót^{e}, Natanggan (which was moribund at the time), and Nioleien (Repanbitipmbangir and Niolenien/Repanbitip in Tryon 1979 and 1996, respectively). Charpentier describes the superscript vowels as post-tonic vowels that are "so destressed as to become unrecognizable (1982:63)." This was likely to involve changes in vowel and voice quality based on modern languages.

Archival data (Shimelman et al. 2019) has the following varieties of Lendamboi represented: Natingatlang, Aingelemolesa, Newotenyene, Nevatanyene. None of these varieties are transcribed with voiceless vowels or prothetic [h] (although other languages in the database do have a prothetic [h] at least in transcription); the audio does not suggest either of these exist in the speech of those speakers; and none of the final vowels seem meaningfully “destressed”. This may mean that the phenomenon attested in historical sources is today only present in Kaiar’s variety.

Charpentier attempted to discriminate vowel quality of these vowels nonetheless, representing those post-tonic vowels as superscript. However, the exact phonetic realization and phonemic status of the process he described, as well as the domain at which it operates (his data took the form of words elicited in isolation), are unknown. If Pre-Ninde had more vowel devoicing like Lendamboi, this could additionally account for apparent word-edge effects.

There appear to be various such processes in Kaiar’s speech: in addition to the prothetic [h] in vowel-initial words, word- or perhaps prosodic phrase-final vowels are optionally partially voiceless as in Figure 11 if they are not entirely deleted (compared to voiced and phrase-medial Figure 12).

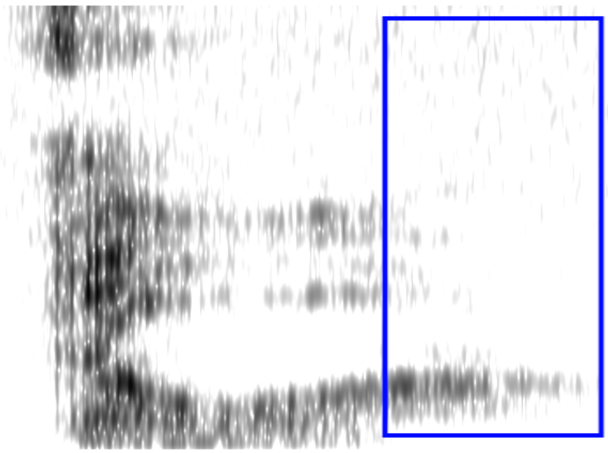

Figure 11: Spectrogram of /towo/ 'rooster, cock' with the devoiced and reduced vowel in the blue box

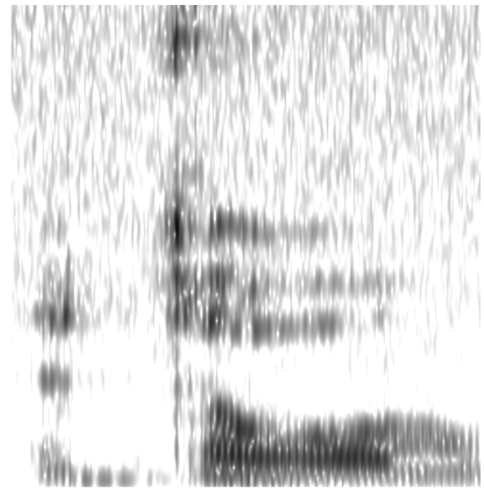

Figure 12: Spectrogram of /netowo/ 'rooster, cock' in phrase-initial position followed by [mboŋ] 'common'

In phrase-final syllables, even vowels following sonorant onsets (Figure 9) may be voiceless for the entire duration of the vowel:

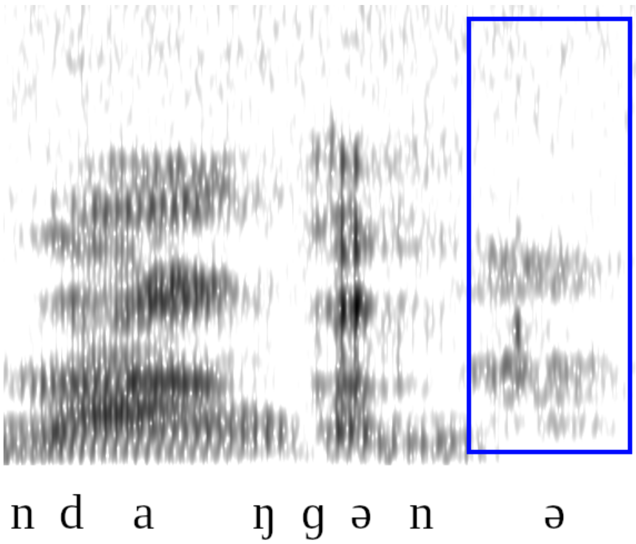

Figure 13: Spectrogram of [ndaŋgənə] 'first' with the devoiced and reduced vowel in the blue box

A final generalization is that liquids can be devoiced in a phrase-final coda. This includes /l/ (as seen previously in Figure 10) and /r/ of Figure 14:

If word-initial devoicing could have conditioned pharyngealized laterals, then it is less unlikely that phrase-final contexts also conditioned the change word-finally.

Ninde has relatively few word-final consonants, having either retained or innovated vowels where its sisters have not. Devoicing synchronically affects the few consonants that are word-final: the underlyingly voiced /β/, /ʁ/, and prenasalized stops are allophonically devoiced at the ends of words. Though no longer active, allophonic devoicing of /l/ would account for the exceptions that remain to generalizations about the realization of *l in Ninde. For a number of these exceptions, the reflex of *l is word- and/or syllable-final in other languages, whereas the unaffected *l was followed by additional syllables at an earlier time (represented by PNCV in Table 31).

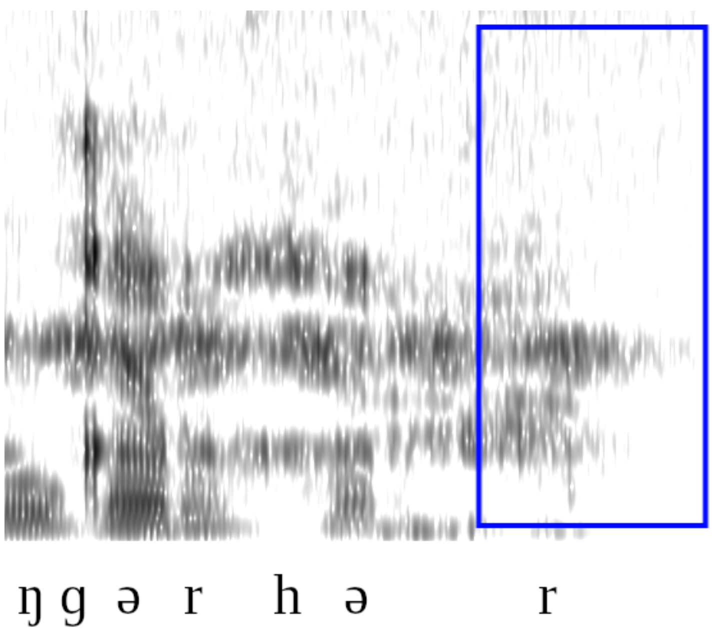

Figure 14: Spectrogram of [(armun)ŋɡərhər] 'they (dual) are just friends' with the devoiced and reduced consonant in the blue box

Protos						
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
(*palala ‘bald’)	*mʷala ‘naked’	*mal̪ˤe*	*malmal*	*malamal*	–	*malamal*
PMP *ma-rara ‘red’	*miala ‘red’	*mijal̪ˤe*	*mial*	*mial*	*(ma)mal*	*(mø)mal*
*loki ‘bend, crooked’	*ma-luqi ‘bent’	*mele*	–	–	*miliw(lew)*	–
*k(a,e)li ‘dig up’	*keli ~ kili	*kil*	*ɣil*	*il*	*ʔil*	*ʔil*

*Table 31: Forms with complex reflexes of *l.*

In the first pair, Ninde does not share the final vowels of the other languages and either uniquely retained or innovated them. Perhaps cues to a reduced final vowel – including the vowels linking reduplicated roots and with final devoicing of a coda consonant in a final stressed syllable – conspired to reverse vowel deletion. In *mele* ‘bent’, the final vowel of Ninde is really penultimate in origin – it yielded a diphthong in Nahavaq. Ninde *kil* ‘dig’ is harder to explain. The final liquid is in a context for devoicing in Nevitangiene, but the lateral did not become pharyngealized in Ninde and the root did not gain an additional vowel. As a transitive verb, however, this never appears in *phrase*-final position or perhaps even in word final position (if Ninde shared what Avava retains as a transitivizing suffix *-i*[26]), since the word is always followed by a lexical object. An intransitive form exists, but is formed by incorporating the root of *nowol* 'hole': *kil-wol* 'dig (intr.)'.

Final vowels could be restored from phrase-final voicelessness cues. Devoiced vowels may be perceptually difficult to detect and discriminate, as Charpentier (1982) suggested. Devoiced vowels involve only optional lingual gestures in Tokyo Japanese (Shaw & Kawahara 2018). As much as vowel devoicing may favor ultimate deletion of affected vowels (Blevins 2018), it may also be the case that they are subject to revoicing. In the process of revoicing voiceless or partially devoiced vowels, the addition of vowels in environments where they were previously absent may be perceptually less marked, or a matter of hypercorrection.

Given that voicing is absent as a distinctive feature in Ninde and around Malekula, it is likely that an allophonically devoiced sonorant would give rise to a contrast employing secondary articulation rather than a voicing contrast. Some phonemes in Ninde's inventory are, however, specified for non-final voicing values. The lateral fricative is always voiced in Ninde and never word-final. If the presence of a devoiced reflex of *l gave rise to final vowels, it may have also set up the environment for the phoneme to become voiced.

[26] Ninde *vaʁi* ~ *ʁevi* 'pull' idiosyncratically retains the transitive suffix; its cognates do not have a final /i/, suggesting that Ninde could have inherited and then lost that suffix.

If word-edge or prosodic boundary phenomena are relevant, then word class is of prime importance. Lynch (2012) finds that transitive verbs were more resistant to final vowel deletion in Ninde, and attributes this to the POc transitivizing suffix *-i. Considering the word-final environment, historical and present-day morphology of alienability and transitivity, for example, can differentiate between sets of roots that frequently occur in the same phonetic environments. Word class also has distributional implications within prosodic phrases. If final-vowel devoicing is an effect found in the prosody of a language, and not at the word level, then it may impact words differently that are likely to be at the end of a phrase: for SVO Malekula languages, intransitive and stative verbs (and adjectives) and nouns that tend to complement verbs.

5.3.2 Origin of the phonological system

Phonological processes in Ninde are generally rather specific to the phonemes Ninde has acquired: the uvular trill/fricative and the pharyngealized lateral, both of which have the effect of lowering adjacent vowels. Morphophonological changes attributable to NCM languages may have created the conditions for heavy /l/ to proliferate, but how much of Ninde phonology is directly inherited from each language? The evidence is considered for two areas of segmental phonology: geminate consonants in NCM and the status of homorganic nasal-oral consonant sequences.

Geminate Stops

One of the unique characteristics of Neverver is its geminate stops /l: s: n: m: p: t: k:/. These are generally from partial reduplication and syncope (Barbour 2012:176-177); a previously discussed exception (§3.3.3) *nevennimzo* 'star' was most likely from place metathesis of the nasal stop and prenasalization in some earlier *nevenmi^{n}zo. Given its uniqueness in Neverver, gemination may have been innovated there, but there are a few correspondences to stops that resist lenition in Avava. Some, but not all, geminate consonants could be reconstructible to Proto NCM.

Clark reconstructs an asymmetrical labial series with no *p for PNCV and no fricative *x alongside *k, but all of the NCM languages have /p/ contrasting with the bilabial fricative /β/ everywhere but word-finally. The majority of instances are at the word edge (with allophony) or in complex onsets, but Avava /p/ often corresponds to Neverver /p:/ and /k/ to /k:/. Additionally, there are only two Neverver forms with geminate lateral approximant /l:/ that have Avava cognates, but they both have a corresponding /d/ in Avava. These correspondence patterns are shown in Table 32.

Protos		NCM			SWB	
POc	PNCV	Ninde	Neverver	Avava	Nahavaq	Naati
PMP *kunut 'pinch'?	*kisi 'poke'	*ŋgi-lim* 'peel (with knife)'	*k:is* 'peel'	*kih* 'peel fruit'	–	–
*kani 'eat'	*kan(i) 'eat'	*kaʔan* 'eat (intr.)'	*k:an* 'eat (intr.)	*kan* 'eat (intr.) '	*ʔaʔan* 'eat (intr.)'	*ʔaʔan* 'eat (intr.)'
–	(*talawa 'web')	*naŋgu-* 'spider web'	*nek:a* 'spider web'	–	–	–
(*kara(ŋ,n)i 'near')	(*rivita 'near')	(*sa[ŋe]/ sa[ŋe]*) 'near'	*sup:ax ~ sup:ak* 'near'	*supak* 'near'	(*saŋ/saŋ* 'near')	–
(PMP *berber 'flutter')	(*ʔalo-vi 'wave'	*pul/pul* 'wave, fan'	*p:ul* 'wave'	*bul/bal* 'fan'	–	–
*papaq 'under'	(*vava 'under')	(*ve-ne* 'under')	*lap:an* 'under'	*lapan* 'under'	(*evu-* 'under')	(*ra-* 'under')
–	*livuka 'middle'	*livete* 'night'	*livɣat* 'night'	*lupat* 'night'	*levʷaʔhat* 'morning'	*levʷaʔat* 'night'

Protos		NCM			SWB	
POc	PNCV	Ninde	Neverver	Avava	Nahavaq	Naati
*sokol ‘choke, strangle’	*liko-ti ‘hang, strangle’	*like* ‘hang (to kill)’	*l:eŋ* ‘hang’	*daka* ‘hang’ *lik* ‘tie up’	*liʔ* ‘tie’	–
*ta-liŋi ‘spilt’	*ligi ‘pour’	(*ŋgol* ‘pour’)	*l:ivix* ‘pour’	*deveh* ‘pour’	–	(*susu* ‘pour’)

Table 32: Scant correspondence sets suggesting that some geminate consonants could be reconstructed for NCM languages.

There are few complete cognate sets, but those that are available suggest no systematic correspondence between Neverver geminate consonants and Ninde or Avava. Even where reduplication is the apparent origin of an initial geminate, there are some forms in Neverver that appear to be reduplicated again (13 entries that begin with the pattern C_1VC_1: could be explained as repeat reduplication). This means that what appears to be intact reduplication in Avava or Ninde could be reduplicated after syncope.

Nevertheless, there are hints of inherited gemination in Avava and Ninde. Ninde has prenasalized allophones of /l/ ([^{n}d]) and /p/ ([mb]), each conditioned by like consonants. Plain (and voiceless) PNCV *k generally became a fricative in Neverver and was deleted or palatalized in Avava. The form for ‘night’, with its fricative sequence of /vɣ/ in Neverver, has a simple /p/ in Avava, probably as a regular fortition of coda bilabial fricatives before the velar fricative was deleted. Considering this, Avava has occational /p/ and /k/ where fricatives are expected, but they only surface as stops where Neverver has geminate /pː/ and /kː/, and /^{n}d/ corresponds to Neverver’s /lː/ in at least its reflex of ‘pour’. If Ninde inherited some geminates from Proto NCM, then it may have prenasalized reflexes represented by the first syllable of *ŋgilim* ‘peel with a knife’ from possible *kːis ‘peel’. Generally, realis forms in Ninde rarely begin with /p/ (even if they correspond to /p/ in sister languages) and when they do, they are homophonous with their irrealis counterparts – this makes verbs like *pulpul* ‘wave, fan’ also potential reflexes of former geminates. Ultimately, there may never be enough data to identify whether Ninde ever had geminates.

Homorganic Nasal-Oral Consonant Sequences

Ninde’s phonology does not treat all homorganic nasal-oral stop sequences as unitary segments – this may be an areal feature in South West Bay languages. Most work does not address native speakers’ intuitions and the phonological reasons for treating these sequence as unitary segments or sequences, but Barbour (2011) has considered this for Neverver and we have addressed the issue in Ninde.[27] Terry Crowley represented prenasalized stops in

[27] Following my colleague, Caroline Crouch, I will analyze nasal-oral sequences as phoneme combinations synchronically.

Avava with single letters (2006a), but with sequences in Naati (1998). This can be taken as tacit evidence that prenasalized stops in South West Bay are generally perceived as sequences.

Contrary to the native speaker perceptions reported for Neverver (Barbour 2011), Ninde speakers overwhelmingly reject single-letter representations of these sequences (except for /^{n}d/) as "too light" to represent the sounds. Whereas Neverver speakers have strong intuitions to represent /^{m}b/ as ⟨b⟩ within morphemes (but ⟨mb⟩ when the irrealis *m-* is present), Ninde speakers entertained analyses of *levendes* 'in the ocean (beyond the reef)' both as a reduced /liβe-ne-tes/ 'middle of the sea' or as /liβe-tes/ with allophonic prenasalization. By contrast, several Ninde speakers commented that the set of labialized bilabials ⟨mbw mw⟩ was represented inaccurately with an overt ⟨w⟩, expressing a dispreference for a digraph to represent a single segment. While anecdotal, these experiences have led to community choices in orthography to represent nasal-oral stop sequences with overt nasal graphemes.

There is some evidence in the orthographic representations developed for Malekula languages, since linguists have generally worked with communities to develop practical orthographies. For the NCM languages Avava (Crowley 2006a), Naman (Crowley 2006b), Neve'ei (Musgrave 2007), and Neverver (Barbour 2012), the representation of the prenasalized consonants /^{m}b ^{m}ʙ ^{n}d ^{n}d^{r} ^{n}z ^{ŋ}g/ is, respectively, ⟨b bb d dr j g⟩ (though not all of the languages have all of these phonemes), except that /^{n}z/ is represented with ⟨ns⟩ in

Neve'ei. In the SWB language of Naati, Crowley (1998) chose to represent /$^{\text{m}}$b $^{\text{n}}$d $^{\text{ɳ}}$ɖ$^{\text{ɽ}}$ $^{\text{ŋ}}$g/ as ⟨mp nt ntr ngk⟩ respectively, and for Nahavaq, Dimock (2009), in tweaking an existing missionary orthography with the community, used ⟨mb nd nd gc⟩, with no orthographic representation of the obsolescent nasal rhotic. Ninde speakers by public vote unanimously chose to represent /mb nd ŋg/ as ⟨mb nd ngk⟩, though regrettably, the options were presented for the whole set and not differentiated by place of articulation. On the whole, these orthographic choices may reflect native speaker intuitions to some extent; especially compelling is that Crowley made divergent choices for Avava and Naman vs. Naati[28].

There is little evidence for prenasalization in Ninde phonology. As discussed in §6, there are diachronic reasons to consider the nasal-oral stop sequences to be reflexes of unitary segments, but not for the entirety of Ninde's prehistory. While the prenasalized rhotic is analyzed as a single segment in languages that have it, Ninde recently (since missionary texts were written) lost it to /r/. This optionally affected /nr/ sequences across morpheme boundaries, so that a type of bird called *nemen-roʔoi* is optionally *nemeroʔoi*. The synchronic phonology does not favor an analysis of these sequences as unitary segments. In intervocalic positions, /p k/ may be realized as voiced [b g] and only /t/ surfaces for some speakers as [$^{\text{n}}$d]. Unlike the NCM languages, the nasal portion of these sequences is not deleted after

[28] Dimock (2009) reports via personal communication that Crowley later adopted her phonological analysis of Nahavaq vowel fronting for Naati, which he had analyzed in his 1998 sketch grammar as having phonemic front vowels /y ø/ instead of allophonic velarized and labialized bilabial stops. This suggests his analysis of Naati phonology could be seen as underdeveloped, but I consider here that the orthography may have been developed

another nasal; words like [nəmndaŋge] 'puzzle tree *(Kleinhovia hospita)'* have a sequence of stops. This is highly perceptible, and by my L2 perception, the nasal portion is generally syllabic following another oral or nasal stop. The definite enclitic *=nge* /ŋe/ surfaces with a labialized velar nasal after bilabials and /u/, but the velars of the proximal enclitic *=ngke* /ŋge/ do not become labialized. When a predicate ends with a /p/, a (probably syllabic) [m] is inserted before the perfective enclitic *=paʔ* /paʔ/, voicing the oral stop:

/ʀop/ 'run' + /=**p**aʔ/ PFV → [ʀəp**m̩b**aʔ] 'ran'

The inserted nasal could alternatively be considered a process of dissimilation, whereby the oral stop becomes a prenasalized one after an identical stop, but voicing also occurs across morpheme boundaries where there is no insertion:

/sim/ 'cut with a large blade' + /**p**oʁ/ 'break (s.t. round)' → [sɪm**b**əχ] 'cut (s.t. round)'

These processes can be understood without categorizing the sequences as unitary segments in Ninde.

Much of the phonological patterning that supports an analysis of unitary segments in the phonology of NCM languages does not apply in Ninde phonology. Small Ninde most likely retained many nasal-prenasalized sequences inherited from NCM (*neme**nŋg**aʀai* 'flying fox'), whereas Big Ninde lost the nasals in these sequences to regressive assimilation (***sumb**u* 'forget' has historically incorporated the once-obligatory subject *sin 'guts').

in consultation with Aiar Rantes, the native speaker who provided those data.

Prenasalized stops of both sources subsequently became sequences of nasal stops followed by oral stops, at least in perception – this change likely also affected Naati and Nahavaq.

5.4 Mixed Grammar of Ninde

All of the NCM and SWB languages fit a similar syntactic typological profile, making detectible influences rather restricted to clitics and affixes. All of the languages have SVO word order with various strategies for combining verbs into complex lexical constructions, sometimes with different valency frames. Most of the words can be loosely classified into nouns and verbs – the latter contains subclasses including stative/adjectival verbs and quantifiers. When combined, verbs that appear later in the verb complex (often designated V2, whether they appear as the second verb or later) may have specialized usage related to, but not predictable from, their meanings when used as main verbs. These are generally in an irrealis form of the verb (Dimock 2009:157; Barbour 2012:325). Some frozen prepositions have transparently verbal origin, where the use as a main verb has apparently been lost. Nouns virtually always begin with some reflex of *na-, but bare roots are used in compound nouns and incorporated objects. Locative nouns instead have reflexes of *lV-, but they bear a prefix /t(i)-/ when serving as noun modifiers – a homophonous /t(i)-/ is also prefixed to relativized verbs.

In terms of its morphological structure and form, Ninde has core elements in common both with NCM and SWB languages, with a great deal of innovation not shared with either group. This section will first consider two domains that bear approximately equal influence from NCM and SWB:

1. Experiencer 'gut' constructions

2. Irregular irrealis forms

Following that, the remainder of this section will consider innovations shared with each of these groups – first NCM (§5.4.1) and then SWB (§5.4.2) – then those that could best be described as creative blends (§5.4.3). There are comparatively few changes shared with Lendamboi languages; these are shallower influences that do not warrant claims of language mixing (§5.4.4). Each of these sections will begin with low-level structures and move toward more complex ones, from shared material form to shared patterns and paradigms, and from unproductive to more productive.

1. Experiencer 'gut' constructions

Ninde has in common with both NCM and SWB sources several verb forms that include a reflex of 'gut' (or PNCV *tinaʔe 'intestine'). In NCM languages, this is a stand-alone subject inflected for the experiencer. In SWB languages, the reflex of 'gut' is incorporated into the verb, which is inflected for subjects in the experiencer role. Ninde has both types, represented by the constructions *nese*-POSSESSOR *pele* 'for POSSESSOR to be angry' and *sumbu* 'forget, not think about' (< PSWB *si-n-bwuŋ 'their (sg.) gut does not think about'), which is inflected like monomorphemic verbs.

At least two verbs in Ninde are used in the 'gut' subject construction. These verbs relate internal states and can have the possessed subject 'gut', which in Avava is *itnen* (1) and in Ninde is *nese-* ~ *nesa-* (Table 33).

p. 145

(1)	itnen	na	i-ʙuleŋi	isan	oŋ	Avava
	gut	1SG.POSS	3SG.SUBJ.REAL-forget	name	2SG.POSS	

'I forgot your name'

Person	Singular	Dual
1	nesa-ŋ pele ‘I am angry’	–
2	nesa-m pele [nesambele] ‘you (sg.) are angry’	–
3	nese-ne pele ‘they (sg.) are angry’	nese-ne pele-pele rax ‘they (du.) are angry with each other’

Table 33: Ninde pele *‘be angry’ with ‘guts’ as its obligatory subject.*

The paradigm as recorded is incomplete, but this is likely because third-person singular forms with *-ne* combined with independent pronouns (in any combination of person and number), like lexical noun phrases, have virtually supplanted non-singular possessive constructions. To date, the only other attested lexeme that is inflected like this is *joʔjoʔ* ‘be nauseous’ (cf. Avava *jok* ‘vomit’ and perhaps somehow Neverver *joŋ* ‘throw out (food)’). By comparison, Naati and Nahavaq have no such usage of ‘gut’ attested in current documentation.

What SWB languages have, by contrast, is verbs with incorporated *si-n ‘their (sg.) gut’. Nahavaq has *sinmbʷuŋ* ‘forget’ (partially cognate with Avava *ʙuleŋi* and fully cognate with Ninde *sumbu*,[29] Naati *simbuŋ* ‘forget, not think about’) and *sinkon* ‘be angry’ (PNCV *kona ‘bitter, sour’). In each SWB language, the erstwhile incorporated subject is not inflected for possessor; instead, the entire verb is inflected for experiencer with subject prefixes, as in Ninde ***na**-sumbu-paʔ* ‘**I** forgot’ and ***na**-snawaʁ* ‘**I** don’t want’. Ninde *s(e)nawaʁ* ‘not want’ appears to contain the incorporated reflex of ‘gut’ and a cognate of Nahavaq *vʷer* ‘want’ and/or Neverver *ver* ‘want, say’. The encoding of the negative sense as well as the source of the internal /a/ are unclear.

2. Irregular negative verbs

Each of the (SWB and NCM) languages discussed here has a set of verbs that are semantically negative, but do not appear in any productive negation constructions. At least in Ninde, they are also grammatically affirmative (taking *=pei* ‘still’ instead of the negative *(s)-…=ŋgevei* ‘not yet’). In Neverver, negative verb stems have proliferated through selective retention of a simple negative prefix *si-* (Barbour 2012:326-327), in contrast to the productive independent negative particle *si,* which follows the verb. A sample of these forms

[29] The loss of verb-final /ŋ/ is typical for Ninde and sometimes reversed in formal speech and the gloss reflects unmarked usage of *sumbu* ‘not think about’ and usage with the perfective enclitic *sumbu=paʔ* ‘forget’.

is provided in Table 34 – representing all of the negative verbs described for Naati and Nahavaq, but only some of the forms for Avava and Neverver.

Gloss	Neverver	Avava	Ninde	Naati	Nahavaq
'have'	*tʙux*	*tok-i*	*vijan*	–	*veʔen*
'not have'	*(tʙux si)*	*tik-i*	*səke*	–	*si(ʔ)siʔ*
'exist, stay, be'	*tox*	*tok*	*toʙ/koʙ*	*toʔ/koh*	*tur/koh*
'not exist'	*(tox si)*	*titik*	*s(ə)kes(ə)ke*	*sasa*	*si(ʔ)siʔ*
'think about'	*setːa*	*sieda*	*ru-woʙ*	*ŋɽimtrim*	*ŋɽimŋɽim*
'forget, not think about'	*setvun*	*(itnen i)-ʙuleŋ-i*	*sumbu*	*simpuŋ*	*sinmbʷuŋ*
'want'	*ver*	*wah-i*	*ʙo(ŋo)-nda(ŋe)*	–	*vʷer*
'not want'	*ros-ix*	*lip*	*s(e)nawaʙ*	*wuswus*	*ŋot*
'know, be able'	*roŋ-il*	*rokut*	*laŋgəre ~ raŋgəle*	*roŋ-hur*	*roŋ-hur*
'not know'	*si-balbal, melmel-ix*	*deleŋi*	*lame/lambaʔ*	–	*mbʷit*

Table 34: Irregular negative verbs in NCM and SWB languages. Though Ninde shares cognates with both language groups, it has important structural similarities with NCM negative verbs.

Even though the majority of these Ninde forms – affirmative and negative – have cognates with similar meanings in Nahavaq in particular, there are important characteristics shared with NCM languages. Ninde shares with the other NCM languages a distinction between 'not have' and 'not exist' and with Avava a partial cognate in 'not know'. Avava *deleŋi* 'not know' most likely has a transitive suffix *-i*, while Ninde has lost word-final *ŋV* in most verbs, possibly also in *la-* 'not know, be unable'. Ninde's negative *la-* never occurs without a suffix or enclitic, and specifically prenasalized allomorphs (stative *-me*, perfective *=mbaʔ*, or *-mbei* 'yet') – it seems to be the only verb that requires these – and probably converged with the affirmative counterpart *laŋgəre ~ raŋgəle,* influenced by analogy. The affirmative verb *laŋgəre* 'know, be able' is clearly a metathesized from of the synonymous *raŋgəle*, which closely resembles Neverver *roŋ-il* 'know, be able'. If they are cognate, then all the NCM forms contain a clause subordinator: Neverver *il* is attested in this usage,[30] Ninde has *tele* 'in order to' (which could contain *t-*, a relativizing prefix), and Avava *rokut* is most likely a compound of *roŋ* 'perceive' and PNCM *kut 'where' (cf. Avava *o-ut* 'place', but in *rokut* the *k survives as /k/ in the context after a consonant). By contrast, the *hur* of Naati and Nahavaq *roŋ-hur* is equivalent to 'about'. In summary, Ninde's affirmative-negative pairs generally resemble SWB forms, and even when they do not, they appear to use subject-incorporation structures that are unique to SWB.

[30] Barbour (2012: 378) considers this derivation, but finds *roŋ* 'want, sense' + *lel* 'be wise' to be the more likely source.

5.4.1 Grammatical features of NCM languages

Ninde appears to share exclusively with NCM languages the following:

1. Form of the irrealis
2. Post-verbal modifiers
3. Determiners
4. Counterfactuals

1. Form of the irrealis

Irrealis is marked in varied ways across NCM languages with respect to person and number of the obligatory subject prefixes and negation, but every language has a form that could be a reflex of a proto form *b^we- in all irrealis inflections, with the exception of second-person irrealis in Avava (abstractly represented in Table 35). The unabstracted pronouns can be found in §7, where the focus is on reconstructing more than the mood prefixes. This form can be analyzed as *b^wi-* ~ *b^we-* in Avava and *b(i)-* in Neverver, which has generally lost labialization in the bilabial consonants. In Ninde, the form is *p-* and has a voiced (but not prenasalized) allophone, but missionary writings represented it with a vowel harmonizing with the vowels of verb roots as *pV-* and rarely as *b^we-*, the latter especially in hymns. In each of these languages, the form suspiciously resembles the word 'go', which in each language has a distinct irrealis stem (rather than mood marking in the prefix alone): Avava *-ap* (realis -

ip) Neverver -*ʙu* (realis -*uv* ~ -*vu*), and Ninde *pe* (realis *ve*). One possibility is that a common ancestor retained only one irrealis form from an older system of general irrealis marking by initial consonant mutation. Verb compounding of the irrealis motion verb and a realis verb stem would have supplanted virtually all the stem alternations.

	Avava		**Neverver**		**Ninde**	
Person	**REAL**	**IRR**	**REAL**	**IRR**	**REAL**	**IRR**
1SG (INCL)	*na-sa-*	*na-sa-bʷV-*	*ni-*	*nib(i)-*	*na-s-*	*na-s-pe*
1NSG (INCL)			*ni-*	*nib(i)-*	*te-s-*	*te-s-p-*
1NSG (EXCL)			*na-*	*na-bi-*	*e-s-*	*e-s-p-*
2SG	*o(ŋ)-sa-*	*kV-, o(ŋ)sakV-*	*ku-*	*ku-b(i)-*	*ku-s-*	*ku-s-p-*
2NSG			*ka-*	*ka-b(i)-*	*e-s-*	*e-s-p-*
3SG	*i-/e-, sa-*	*(e)bʷV*	*i-*	*i-b(i)-*	*Ø-s-*	*Ø-s-p-*
3NSG			*a-*	*a-b(i)-*	*re-s-*	*re-s-p-*
DUAL	*ar(i)-*	*(a)bʷir(i)-*	*ri-*	*bi-ri-*	*ʁ-s-*	*ʁ-s-p-*
DUAL EXCLUSIVE			*nari-*	*nab(i)r(i)-*	*ʁa-s-*	*ʁa-s-p-*
PLURAL	*s-at(i)-*	*s-(a)bʷit(i)-*	*ti-*	*b(i)t(i)-*	*re-s-*	*re-s-p*
PLURAL EXCLUSIVE				*nambit-*		
IMPERSONAL	*(a)ra-*	*(a)bʷir(i)-*			*e-s-*	*e-s-p-*

Table 35: Abstract comparative paradigms for the NCM languages. The following generalizations and exceptions apply: where there are nonsingular (NSG) forms, they are followed by specific dual or plural number prefixes. Avava prefixes are complete as shown, but all are represented negated. The negative dual forms are not attested for realis or irrealis. For Neverver inflections ending in b(i)C(i)-*, only one /i/ surfaces depending on the*

initial segment of the root (whether it is a consonant or vowel) and any /b/ can become a full nasal [m] at the end of a syllable, such that any /bt/ resulting from vowel deletion is realized as [mt]. For Ninde and Avava, the corresponding positive simply omits the morpheme containing [s].

While Naati *p-* and Nahavaq *vi-* inchoatives look like candidate cognates, they have important differences from Ninde irrealis. They are prefixed exlusively to noun roots (with no nominal *nV-*). This makes them related to the Ninde copula *vi*, which does not have an incorporated predicate nominative in Ninde. A copular usage with a noun phrase complement is also found in Nahavaq.

While there are similiarities in form and patterns of irregularity across NCM, there are differences in the morphological template of stacked inflectional prefixes. In Avava, there is syncretism in the nonsingular inflections, which level person and clusivity distinctions (though these are marked by the independent pronouns). The template for nonsingulars has a structure analyzable as having polarity (Ø- for positive, *sa-* for negative) followed by mood (realis Ø- or irrealis *(a)b^{w}i-*) and finally by number (*ri-* dual or *ti-* plural) – an additional element [a] could be attributed to either of the zero forms, but is obligatory in realis and optional in irrealis forms. By contrast, singular first- and third-person stacked prefixes have an initial fusional prefix marking singular number and person (*na-* first person, *i-* or *e-* third person) – which is also the only indication of number – followed by polarity and mood. The second-person singular alone has a fusional prefix *kV-* marking the second person, singular, and irrealis. Neverver inflection shares the order of Avava singular forms, but this order excludes a negative prefix entirely. This loss appears to be specific to Neverver, but was retained in a handful of words. Neverver nonsingular forms consist of a unique prefix, each of which resembles a singular prefix matched for person, mood (Ø- for realis, *b(i)-* for irrealis).

Ninde differs substantially from what could be reconstructed for proto NCM. Subject inflectional prefixes do not include both dual and plural prefix sets – all dual forms in Ninde contain a cognate with common NCM *ri-* in Ninde *ʁ-*, but this combines with unmarked plural person prefixes and varies in its placement (following the person prefix except in first-person exclusive or second-person dual ʁa-, which could be perhaps be analyzed as having the plural *e-* preceded by the dual with regular vowel lowering). The NCM final plural prefix *ti-* could be connected with Ninde's rarely used paucal suffix *-t* (on independent pronouns), and would represent a set of innovations that Ninde did not participate in (*tl 'three' > *-t(i) PAUCAL > *ti- PLURAL). Very little material of Ninde's inflectional paradigm resembles that of the other NCM languages.

In summary, it appears as though an irrealis prefix was innovated from the motion verb 'go' in all NCM languages, but subsequently each daughter has reduced the complexity of the paradigm in some way. Avava most likely lost person distinctions in nonsingular inflections, Neverver lost negative prefixes, and Ninde largely regularized the order of prefixes in line with singular NCM inflection. Further analysis will show that the properties that make Ninde stand out from other NCM languages are shared with the SWB languages.

2. Post-verbal modifiers

A number of verbal modifiers are unique to NCM with respect to SWB languages. These range from verbs that are compounded non-initially with other verbs but share subject, object, tense, and aspect with the compound root (V2), enclitics that appear at the end of complex verbs, and postverbal verbs, which can add oblique arguments not shared by the main verb but may repeat the subject prefixes. There is also one preposition. These are shown in Table 36, but this section focuses on the completive/perfect and a form meaning 'err' which serves as a frustrative – both of these are shared between Ninde and Neverver and exhibit similar semantics and morphosyntactic and discourse distribution.

Ninde	Avava	Neverver	Position
paʁ ‘break (something round)’	*bʷir* ‘break’	*pir* ~ *bir* ‘break’	V2
pal̪ˤe ‘break (something long)’	*wel* ‘split’	–	V2
tata ‘(grasp) tightly’	*tata* ‘hard, tight’	*tata*	V2
luwo ‘out, finish’	(*lu* ‘out, quickly’)	*lu* PERFECT	V2
nditip ‘try, taste’	*titipm* ‘try’	(*glek* ‘to taste’)	V2
pijal̪ˤe ‘err’	–	*pusel* ‘err’	
=*nde* ‘some’	*da* ‘a little’	*da* PARTITIVE	Encliticd
=*me* ?	*meh* ‘only’	*me* ‘only’	Enclitic
spaʁ ‘until, up to’	*siber* ‘reach’	*sber* ‘all the way up to’	Post-verbal
tele BENEFACTIVE, CAUSE	(*wal* ‘for, because of’)	*il* BENEFACTIVE, CAUSE	Preposition
ŋgawes ‘cross’	(*kawat* ‘cross’)	*gwas* ‘cross, over’	Post-verbal

Table 36: Verbal modifiers only attested for Avava, Neverver, and/or Ninde.

A development from *lua ‘out’ to an aspectual morpheme appears to be restricted to Ninde and Neverver. In both languages, this is homophonous or polysemous with a word meaning ‘shoot (with an arrow)’ and in Ninde also used with expressions of extraction in V2 position (cf. *sevi-luwo* ‘pull out’). The usage in (2) is representative of a potential bridging context, where the extractive usage is compatible with urinating.

Shemson Thompson [ninde035]

(2)	“Val̪ˤat	misaʁa↘	nije	mʷismʷis=luwo	ka↗	ʁa-pul-lis↘”	Ninde
	go.ahead	outside	3SG	urinate-COMPL	TOP	2DU-come.IRR-back	

‘Go on outside. When he has finished peeing (alt.: after he pees) you two come back.’

What is unique to NCM is the general usage of this word as a marker of event completeness beyond contexts where something is being emptied. In Ninde, this adds to two other verbal enclitics that signify different types of event completion (Table 37 demonstrates its usage with a transitive accomplishment verb *jen* ‘eat’ and an intransitive stative verb *mes* ‘be dead’): *=paʔ* ~ *=jaʔ* PERFECTIVE (most likely from the homophonous irrealis form of *jaʔ* ‘ascend’) and for transitive verbs, an object quantifier *=jaʁa* ‘all’, used when the entirety of the group or mass specified by the object has been affected.

Aspect enclitic	***jen* ‘eats (tr.)’**	***mes* ‘is dead’**
=paʔ ~ *=jaʔ* PFV	*jen=paʔ* ‘eat/ate’	*mes=paʔ* ‘die(d)’
=luwo COMPL	*jen=luwo* ‘finish eating, after they (sg.) eat’	*mes=luwo* ‘after they (sg.) die(d)’
=jaʁa ‘all’	*jen=jaʁa* ‘eat all of’	–
=luwo=paʔ ~ *=luwaʔ* ‘already’	*jen=luwo=paʔ* ‘has/had already eaten’	*mes=luwo=paʔ* ‘has/had already died’

Table 37: Aspectual clitics in Ninde for completed events.

The relevant forms in Ninde are *=luwo* COMPLETIVE and *=jaʁa* ‘all’, because they appear to come from different sources (NCM vs. SWB). Specifically, *=luwo* marks an event that has come to a conclusion, especially in relation to a sequence of events – regardless of how much of the object has been affected and sometimes in imperfect contexts. Since it does necessarily mark events that come to their natural end, it could be called a terminative in contrast to a true completive (following Poletto 2008). In both Neverver (3) and Ninde (4), this is common in tail-head linkage when reporting sequences of events.

Lerakhbel NVDL12.9-10

(3)	Ale	ku-jas↘	ku-jas	lu↗	ku-sxav↗	Neverver
	then	2SG.R-stone.cover	2SG.R-cover	PERF	2SG.R-cover	
	ku-tvin-ix	nibit:an↗				
	2SG.R-bury-APPL	soil				

‘Then you cover it with stones. Having covered it with stones, you cover it with leaves and bury it with soil.’

Letin 146

(4)	alei	nije	tu	nende-ŋe	a	nije	l̪ˤatane,	tes↗	nijou, ↗	Ninde
	so	3SG	put	basket-DEF	POSS	3SG	on.ground	unwrap	laplap	

jen=ŋe.↘	jen=**luwo**=ŋe, ↗	alei	kombaʁane,↗	ŋgom=lis.↘	tu
eat=3DEF.OBJ	eat=**COMPL**=3DEF.OBJ	so	and.then	get.up=back	put
nijou	le	nende	lis.↘		
laplap	in	basket	back		

‘So she put her basket on the ground, unwrapped the laplap, and ate it. She finished eating it and then (alt.: **after** she finished eating from it,) she got back up and put the laplap back in the basket.’

The Neverver example in (3) is from a procedural text for baking laplap (a dish based on grated tubers) in an earth oven. Similarly, in the Ninde example (4) from a traditional story, a woman is running from her husband, all the while carrying laplap that is cooking in a basket by the heat of rocks extracted from a fire. She stops multiple times, here for the first time upon realizing her hunger, to eat small amounts at a time while on the run. In both contexts, the cognate reflexes of *lua are unambiguously signifying the chronology of sequential events, rather than the completion of an action to its natural limit. The usage of the completive together with the perfective clitic can be read as a (plu)perfect construction in non-sequential contexts, as in (5) (the first line is segmented into larger constructions and included only for context):

Shemson Thompson ninde035

(5)

[raʁjel l̪ˤo]	te	raʁpe-li	nesum-a-pusuwoʁ	[te noʁpo tuwa sei]	Ninde
3DU.bring	to	2DU.go-see	grandmother-POSS-boy	for some reason	

e-mbiti	wut	nesum	a	pusuwoʁ	ka	mes=**luwo**=paʔ
3NSPEC-say	CPZR	grandmother	POSS	boy	TOP	die-**COMPL**-PFV

‘they bring it (a basket of food) so they can go see the boy’s grandmother for some reason, but it is said that the boy’s grandmother **had already** died.’

Ninde and Neverver also share a frustrative in V2 position meaning ‘err, do badly’ with other languages more broadly, but this pair of languages uniquely shares metathesized consonants and a semantic narrowing. Though there are no examples for the syntactic structure in Neverver, the lexical constructions include *xab-pusel* ‘throw (a long object at a target) and miss’[31] vs. *tuv-pusel* ‘throw (a round object at a target) and miss’ (Barbour 2012:327). Assuming from the omission that Neverver retains a transitive argument structure, Ninde (6) is more aligned syntactically with Neverver with its transitive usage:

Group elicitation

(6) ʁkin=pijal̪ˤe nemen — Ninde

pelt=FRUST bird

‘They (sg.) threw (something) at the bird and missed.’

Neverver generally makes use of the applicative *-ix,* a verbal suffix, whereas Ninde and Nahavaq allow for oblique arguments to be introduced by a preposition (*ʁa-* and *ra-*, respectively) – Ninde does not use this oblique construction for the frustrative, but Nahavaq does (7):

[31] In all other examples for Neverver, *xab* means ‘be full’, but *r:ik* is another word meaning ‘throw’ and the latter is partially cognate with Ninde *ʁkin* ‘pelt’.

(7)	i-vini	nemen,	i-ven	pileh	ra-n	nemen,	nemen	i-top^{w}	Nahavaq
	3SG.R-shoot.TR	bird	3SG.R-shoot	miss	at-3	bird	bird	3SG.R-jump	

‘They (sg.) shot the bird, but they missed and the bird escaped.’

The origin of Ninde’s frustrative is additionally important for phonological reasons. Like Neverver *pus̱eḻ* and Naman (where it is also a V2 but resembles a nominalized form) *ne/vs̱iḻ/ian*, the order of the second and third consonants of Ninde *pij̱aḻ̪ˤe* reflects *s and then *l, whereas the opposite order is found in Nahavaq *piḻeẖ*, and Neve’ei *biḻiẖ* ‘do badly/carelessly, do all over’. (Avava has a similar V2 form *bivil* ‘do all over’, but it does not have any clear cognates – Ninde *pele* ‘do badly’ is only a phonological fit.) The specificity of this change lends support to shared inheritance.

Other shared enclitics and compounding verbs (Table 20) lack the data to make a compelling case, because this would rely on negative evidence from still underdocumented languages. The absence of an attested cognate could additionally be the result of grammatical innovations replacing once-widespread forms. With no identifiable cognate, peculiarities of the diachronic phonology cannot be identified either. Nonetheless, verb modifiers localized to NCM include V2 forms ‘break’, ‘tightly’, ‘until (a time), up to (a goal location)’, and a partitive enclitic meaning ‘some of OBJ’. Those shared between Ninde and Avava include ‘split’ and ‘try, taste’, and those shared between Ninde and Neverver include a V2 usage of ‘cross’ and an oblique preposition for benefactives and causes.

3. Determiners

Definite morphology is somewhat unique in Ninde, but comparison mainly with NCM languages points to origins of determiners and deictics in grammaticalized verbs. Generally, NCM languages mark definite noun phrases with antecedents and optionally mark indefinites, which are not included in this discussion. SWB languages, excluding Ninde, mark indefinite noun phrases and definite noun phrases are generally unmarked. These are presented in Table 38, with forms from other classes that I will claim are diachronically associated.

Position	Neve'ei	Neverver	Avava	Ninde	Naati	Nahavaq
NP	*ŋe* DEFINITE	*-aŋ* ANAPHORIC	*ki ~ ti* DEFINITE	*-ŋe* DEF	–	*nin* DEF
VP	*i* 3SG	*ei* 3SG	-*i* 3SG.OBJ	*=ŋe* 3SG.DEF, *nije* 3SG	*amataŋ* 3SG.OBJ	*ʔin* 3SG.OBJ
NP	*ie* 'this; here it is'	*-ax* 'this'	(*ki ~ ti* DEFINITE)	*-ŋge* 'this'	–	*asiŋ* 3SG.PROX
NP	–	*tzax* 'this here'	(*itnan* 'here')	*tsəŋge* 'this'/ *təŋge* 'this'		(ti)jiŋ(g) 'this'
NP	*ziŋ* 'be there, sit'	*tziŋ* 'that (visible)'	(*itnani* 'there')			((*w)utin* 'there')
NP	*iaŋ* 'that thing'	*taŋ* 'that (non-visible)'	*teki* 'over there'	*təŋe* 'that',	*etaŋ* 'there'	*etaŋ* 'there'
(Adv)	–	–	–	*jenda* 'here'	*iʔenda* 'here'	–
VP	*zax* 'be here'	*jax* 'be here'	–	*jaŋge* 'here'	–	*ejaŋ* 'there'

Table 38: Deictics in NCM and SWB languages with diachronically associated verbs.

The proximal determiners (those corresponding to *=ŋge* in Ninde*)* appear to be the only set (partially) cognate across all three NCM languages, with an unreconstructable velar in each language and variable vowels. Ninde and Neverver further share a general anaphoric, which forms a minimal pair with the proximal affix: Ninde has proximal *=ŋge* vs. general anaphoric definite *=ŋe*, while Neverver has proximal *-ax* vs. anaphoric *-aŋ.* Avava *ki~ti* is glossed as a demonstrative 'the', DEM, or occasionally more specifically as 'this' (proximal). Neve'ei, which has an indeterminate place either in or closely related to the NCM languages, shares with Ninde the form of its determiner *ŋe*. The Neverver forms are suffixed to nouns as in Ninde, but Ninde additionally has two sets of homophonous forms that can serve as anaphoric object pronouns when cliticized to the verb phrase (but nouns in any case role can bear the suffix). In both nouns and verbs, they suppress raising of word final archiphoneme /A/. Given that Ninde appears to have innovated final /A/ in many words, one possibility is that the determiners initially had /a/ (as in Neverver), but these were lost (and gained by nouns and verbs) through rebracketing of determiners. Furthermore, Ninde /ŋg/ corresponds to Neverver /x/ in two of these forms ('here' and 'this'), even though this is not a general correspondence across these languages (the corresponding Neverver reflexes are /ɡ/ and /ŋ/ word-finally). These facts support a suffix pair tentatively reconstructible as *-aŋ(V) DEF and *-aɡ(V) PROX, but instead of the expected Neverver homophones *-aŋ*, the proximal (like related deictics) was realized with /x/ as the velar reflex.

Despite the homophony of nominal affixes and object enclitics in Ninde, evidence does not unequivocally support a common origin. In addition to suppressing final vowel raising, both the verbal and nominal forms are uniquely affected by rounding in the environment after a final /p/ or /u/. Despite these similarities, the verbal clitic *=ŋe* and nominal suffix *-ŋe* are represented differently in missionary texts: the verbal clitic as a separate word ⟨gin⟩ [ŋin], which Letpen (p.c.) identifies as a verb meaning 'take', and the nominal suffix is represented attached to its host (without word spacing) as ⟨ga⟩ [ŋa] – many instances of [e] are represented as ⟨a⟩ in those texts in both word-final and internal contexts, suggesting that these vowels were raised in the last 150 years. Adding to the mystery, Neve'ei *ŋe* optionally marks the relativized object of transitive verbs in relative clauses (Musgrave 2007:69-70). A determiner following a relative clause with transitive verbs has high potential to be ambiguously part of the relative clause or the larger noun phrase, and this is also a viable pathway to a general object pronoun through reanalysis. The available data from written texts and comparative work simply do not tell a consistent story in this case.

A comparison between nominal affixes with presentational predicates supports an origin in reduced verbs in NCM. Neverver *jax* 'be here' can be compared to the proximal definite *-ax*, and Neve'ei presentational verb *iaŋ* 'there it is, it's that thing' to Neverver anaphoric definite suffix *-aŋ*. Neve'ei also has *zax* 'be here' and *ziŋ* 'be there, sit', which can be compared to Neverver *tzax* 'this here' and *tziŋ* 'that (visible)'; *t(i)-* may have been a relativizer, since it is a relativizer in Ninde and forms possessive pronouns in Avava and Neverver, and specifically a relativizer for definite nouns in Nahavaq. Ninde has no verb resembling the general definite suffix, but it does have a verb *ŋge* 'be this one' (which we analyze as a distinct word because vowels of the preceding noun undergo word-final raising). Reduced forms of this pair of words – or forms unprefixed with an invariable third-person singular subject *i- – may have become noun suffixes from uninflected statives modifying the nominal head.

Now that Ninde's presentational verb and nominal suffixes have a potential explanation, the mystery remains as to the origin of the homophonous object enclitics. The best likely individual cognate is the Nahavaq third-person object pronoun *ʔin*, which is also an object pronoun, but the expected corresponding form in Ninde should be *kin* in the absence of a preceding nasal context. Rarely, Ninde /ŋg/ corresponds to Nahavaq /ʔ/ word-initially (Ninde *ŋgənis* and Nahavaq *ʔinis* 'pinch'), and commonly after a nasal – see §5.4.5 (part 2 on Form of the irrealis on consonant mutation in verbs) – but this would yield *ŋgin*, not contemporary *=ŋe* or archaic *ŋin*. Letpen's characterization of this word as a verb is difficult to substantiate comparatively, and the closest matched forms are found in Neverver *gen* and Avava *ka* 'be like' – both of which are again more consistent with expected reflex *ŋgin*. While the origin of the verbal object enclitics remain obscure, they were almost certainly modified on analogy both with each other and with the nominal suffixes.

4 Counterfactuals

Like Neverver, counterfactuals in Ninde are formed using a grammaticalized verb: *səke* 'not have' in Ninde, and *besi* in Neverver, which bears an irrealis prefix *be-* typical of verbs, even though *si* is not attested as a verb. Nevertheless, the position and function of the construction in Neverver (8) matches that of Ninde (9).

Emlina, Limei Simo, Miriam Simo, or Helen-Rose Peniyas								NVCV06
(8)	besi	im-dak	ei	lon	naxsan	...		Neverver
	if	3SG.SUBJ.IRR-fall	3SG.REFL	LOC	base			
	‘If it had fallen on the base...’							

Letpen							p.c.
(9)	səke	nije	s-l̪ˤo=vei	ʁa	plen	...	Ninde
	if	3SG	NEG-go=NEG	INST	plane		
	‘If they (sg.) hadn’t gone on the plane...’						

By contrast, Nahavaq counterfactuals are marked with *konoʔ* ‘if’ and always take irrealis verbs (9).

(9)	**konoʔ**	**ke**-tip	ra-n	niruʔuh	ʔet	i-s-pwaras	veʔ	Nahavaq
	if	3SG.SUBJ.**IRR**-grow	on-3SG	hill	PART	3SG-NEG-strong	NEG	

‘If it grows on a hilltop, it’s not strong.’

No usage of ‘if’ appears in the glosses of any examples in the sketch grammar of Naati (Crowley 1998), and a search of the lexicon and texts of Avava reveals *ba(na)* ‘if’ in the lexicon and a general subordinator *det*, which is in counterfactual clauses with ‘if’ in the free translation of texts. That is in line with elicitation data: multiple speakers provide Ninde *wut* for both Bislama *sapos* ‘if’ and *wea* ‘where’ (compare *nuwute* ‘place’), but counterfactual usage with *wut* has not been attested in the corpus. None of these attested counterfactual strategies bear any formal resemblance to Ninde, with the exception of Neverver.

5.4.2 Grammatical features of SWB languages

Ninde and its sung form appear to share exclusively with SWB languages the following:

1. The locative/existential verb paradigm
2. Preverbal modifiers, only in sung Ninde

3. Negation

4. Noun case, but only in sung Ninde

5. Irrealis prefixes and residual consonant mutation

6. Several post-verbal modifiers

7. Complex nominalized verbs

1. Locative/existential verb paradigm

Ninde shares with Nahavaq – and as far as description of Malekula language allows, only Nahavaq – a paradigm of root changes in the existential and locative verb *tur. Ninde's reflex *toʁ* bears superficial similarity to Avava *tok* and Neverver *tox*, which is an existential verb for animate subjects in both languages, but evidence points to Nahavaq *tur* (cf. Avava and Neverver *tur* 'stand') as the cognate. Ninde has *taʁse* 'stand', cognate with Naati *ta:ris* and Nahavaq *taris* 'stand'. Sung Ninde has the locative form *tor*. Additionally, the paradigm is like that of Nahavaq.

Nahavaq has two roots, one for plural or collective subjects: *tur* is strictly singular, but *koh* is used for plural subjects (Dimock 2009:132-133). Naati also has existential verbs *toʔ* and *koh* (Crowley 1998), but the documentation does not reveal whether these differ in subject number,[32] and the former is at any rate cognate with the NCM existential verbs and not from 'stand'. With singular inflection, Nahavaq *koh* can also be used for mass nouns. Ninde has a plural form *koʁ* – clearly influenced in form by analogy with singular *toʁ,* with its irregular correspondence of /ʁ/ to Nahavaq /h/ instead of /r/. While this form in Ninde has not been recorded inflected for a singular mass noun as its subject, it can appear uninflected in serial verb constructions. This usage is restricted to constructions with the verb *tu* 'put', which has an object complement, where *toʁ/koʁ* has a locative complement and agrees with the number of the object.

Incidentally, Ninde may have other plural-marking suppletion patterns in common with Nahavaq. The verb *tu* 'put' and *lip* 'take, give, get' have forms used with non-singular (dual or plural) objects: *jaʁa* and *ŋgaʁa*. Nahavaq has two verb forms glossed with 'take', but they are *lip* and *her* (which could be congate with *jaʁa*, even if these glosses do not match). Of the examples used in grammatically unambiguous contexts in Dimock (2009), 39 unique tokens of *lip* have unambiguously singular objects (discernible from the English glosses), eight have (English) mass nouns, two have dual objects (both cases of coordinated nouns), and only two have plural objects. The plural objects are glossed as 'a bow and arrows' and 'things'. On the other hand, *her* has three tokens of plural nouns and one mass noun ('fish',

[32] Similarly, Crowley and Lynch (2006a) make no mention of subject number in locative verbs, but because of the comparative breadth of examples and included texts, it is observable that all Avava examples of *tok* have singular subjects and *luk* (another locative verb) has exclusively dual and plural subjects.

which is also used with *lip*). These patterns would suggest that Nahavaq too could have a number-marking suppletive paradigm for *lip*/*her*.

The other NCM languages, by contrast, do not have any such alternation. Neverver has cognates *lav* 'take', with reduplicated *lavlav* used with plural objects, and *tox* (which has no number – the corresponding reduplicated form means 'wait' (Barbour 2012:240). Avava has a form *rep* 'take' and a cognate *tok* 'be (located)', also in use with singular and plural objects and subjects, respectively, throughout examples and texts.

2. Preverbal modifiers

In spoken Ninde, verbal prefixes are rather limited to bound person and number, negation, and mood, but sung Ninde appears to preserve one additional prefix. The usage in (10) is clearly preverbal, even if the meaning of this element cannot be recovered from context:

Shemson Thompson ninde035

(10) a. nawu mun ndam^{w}e Sung Ninde

b. nu *mun ndam^{w}e spoken gloss

breast respond

'the breast replies' / 'breast, reply!' (?)

From context, it is clear that the sung forms correspond to the spoken forms provided in the Ninde interlinear gloss (10b); in this story, the malevolent spirit of a grandmother who has just breastfed her grandson is chasing him, calling out to her breast to reveal their location, since it is in the boy's stomach. This repeated line of song is followed by *wou!*, presumably the response of the breast. This usage of *mun* as a preverbal modifier can only be reconstructed by comparison with the SWB languages, exemplified in continuous usage in Naati *min-* (11) and its recent past relative to a reference point in Nahavaq *min-* (12):

Aiar Rantes p. 121

(11)	i-min-ʔoi	nimeːtu	Naati
	3SG-CONT-grate	coconut	
	‘They (sg.) are still grating coconut.’		

Massosoh Robert MR01

(12)	Ru-min-koh	len	nehew	ʔet	nimbʷunoŋ	tijaŋ	ar	ra-tal	Nahavaq
	3DU.REAL-recent-be	in	garden	then	child	DEF.DIST	PL	3PL-return	
	‘They (two) were in the garden when the boys returned (to a different location)’								

Note that spoken Ninde has a resumptive post-verbal marker *mun* ‘continue to, for the time being’, but it only ever *follows* verbs. Given the usage of other SWB languages, the sung Ninde may mean ‘the breast is replying *wou!*’, with an archaic (or foreign) progressive *mun*. Given all the possibilities, sung *mun* cannot be related to any contemporary morpheme, but all the evidence points to an aspectual function; if it is an aspectual construction, the position before *ndamʷe* ‘respond’ (supported by spoken Ninde homophony and contextual appropriateness) resembles the SWB usage.

3. Negation

Negation in Ninde is strikingly similar to that of Naati and Nahavaq. All negatives involve a prefix *sV-* which precedes the verb root immediately after the person/number and aspect prefixes (but in Nahavaq, they precede preverbal prefixes). A second element (Table 39) is placed after all core verbs and enclitics but before the object and other verbs (Dimock 2009:141-142; Crowley 1998:122). Simple negation takes the form *=vei* in Ninde, making it homophonous with the stative verb *vei* '(be) strong' (13):

Letpen ninde010

(13)	kəne	ka,	na-p-s-mitoʁ=vei	nduwan	nuŋk	Ninde
	1SG	FOC	1SG-IRR-NEG-sleep=NEG	join	2SG	

'Me, I won't be sleeping with you (sg.).'

This is rather similar to Naati's *-ve* in (14):

Aiar Randes p. 122

(14) wa-sa-metur-ve Naati

2SG.IRR-NEG-sleep-NEG

'You (sg.) will not sleep'

Nahavaq too has the form *veʔ*, as in (15):

Lesley Salei LS01

(15) ku-s-mataʔ veʔ Nahavaq

2SG.IRR-NEG-fear NEG

'Don't (sg.) be scared'

These structures can be contrasted with those of NCM languages. Avava negatives take the initial form *sa-*, but they follow only person prefixes (only differentiated for singular subjects in the verbal morphology) and precede the irrealis prefixes *kV-* 2SG.IRR and *bʷV-* IRR (for all other person/number combinations) and the plural subject prefix *ti-*; the discontinuous verb-final suffix is *-mu* (Crowley 2006a:82-83). In Neverver, the plain negative takes a simple form *si* which typically follows the verb and is rarely used with other post-verbal modifiers (Barbour 2012:279-281). In terms of form and order of morphemes, Ninde closely mirrors the negatives of SWB languages and is rather unlike Avava and Neverver.

The remaining negators suggest that negation was inherited wholesale from an SWB language. They are compared and contrasted in Table 39.

	SWB			NCM	
Gloss	**Ninde**	**Naati**	**Nahavaq**	**Avava**	**Neverver**
'not'	*s(V)-…=vei*	*sa-…-ve*	*sa-…veʔ*	*sa-…-mu*	*si*
'not yet'	*s(V)-…=ŋgevei*	*sa-…-ŋgeive*	*sa-…malas veʔ*	*sa-…-va*	*vas(i)*
'not anymore'	*s(v)-…=səkəvei*	–	*sa-…veʔ lis*	–	*simo*

Table 39: Negation across SWB and NCM languages. Ninde forms are in line with SWB forms.

Where there is data, Naati negation seems most similar to what is seen in Ninde. Avava and Neverver have in common an element *va-* in 'not yet' (which is not attested outside of this construction) which resembles the post-verbal negative of SWB languages, but only in form.

4. Noun case, in sung Ninde

While Ninde lacks the large case paradigms of Naati (Crowley 1998:126-128) and Nahavaq (Dimock 2008:20), Ninde nonetheless has an expanded set of forms resembling prepositions and one form in sung Ninde bears a stronger resemblance to its Naati counterpart. These include relational nouns like *tuko* 'behind' (compare *nituko-* 'back', Naati *tuʔo-* 'behind') and *opmo* 'in front of' (compare *nopmo-* 'face'), but the remaining forms appear to be verbal in nature or constructed from the locative *l̪ˤV-* prefix on noun roots with possessive morphology. The irrealis forms of *ve* 'go' and *wul* 'come', *pe* and *pul*, can be used with their nominal complements to specify goal and source for motion and spatial verbs. The form *pe* in particular can be used with a locative prefix *lele* 'in, at', which was perhaps reduplicated in Ninde or restructured as though reduplicated from an intermediate *le-ne* (cf. cognates in Naati, Nahavaq, and Avava *le-n* 'in/at it' and Neverver *lon*). This combined usage is exemplified in (16).

Raobong Serau ninde018

(16)	*te*	*kaiwut*	*toʁ*	*pe*	*lele*	*nesoʁo*	Ninde
	so	old.man	exist.SG	go.IRR	in	coconut.petiole	

'… so that the old man there could go into the coconut petiole'

This usage closely resembles the combination of direction *mpi* and locative *len* in Naati (17) – note that third-person possessive *-n* is often accreted in prepositions:

Aiar Randes p. 128

(17) *i-loŋ* ***mpi*** ***len*** *lempinwen* Naati

3SG.REAL-go **ALL** **LOC** beach

'He went to the beach'

An allative prefix *mbile-* appears to replace the nominal *nV-* of *netes* 'sea' in the following line of Ninde song in (18), resembling Naati usage of *mpi len* 'to' and Nahavaq *mbʷeleŋ* 'toward':

Edwel Kaiseng [ninde121]

(18) a. *U-ve-ndo* *mbile-tes* *mijale* Sung Ninde

2SG-go-PFV ALL-sea red/fish

Possibly: 'you went to the red sea/sea of fish'

Since this song is not part of a story, but a traditional song for dances (*levaʔale*), there are no contextual cues and the glosses are all arrived at through comparative reconstruction alone. Two forms are not present in spoken Ninde at all: a 2SG subject marking and probably realis *u-* (spoken Ninde has *ku-* for realis and irrealis) and a perfective-marking *-ndo* is interpreted as cognate to Nahavaq *ndoh* (spoken Ninde has *=paʔ ~ =mbaʔ).* This Ninde example, if correctly interpreted, makes this a strong case for inheritance from an SWB source, even though this form is not part of spoken Ninde.

5. Irrealis marking and residual consonant mutation

Irrealis marking in Ninde is similar to that of Naati and Nahavaq in two main ways: like Naati (and to some extent also Nahavaq), the irrealis prefix is independent from person and number marking. This position close to the verb root allows for uninflected irrealis forms to serve as clausal modifiers (V2); these generally follow the first verb but come before the object if shared with the independent verb, and after the object if they serve as resultatives or have their own complements. In Ninde, verbs in their V2 function sometimes take part in consonant mutation.

SWB languages have complex patterns of consonant mutation that distinguish realis roots from irrealis roots. This is an active pattern in Naati (Crowley 1998:124-125), but only unproductively distinguishes verb roots from post-verbal modifiers in Nahavaq (Dimock 2009:46) and Ninde. In all three languages (as Crowley and Dimock point out), the alternation is unusual in that the irrealis or dependent verbal forms tend to be prenasalized, whereas other Vanuatu languages exhibit the opposite pattern for mood-based consonant mutation (Crowley 1991:180-183). Avava similarly has consonant mutation for verbs in V2 position if they begin with /k/ or /t/, in both cases surfacing as prenasalized [g] (2006a:91). The patterns showing these related phenomena are represented in Table 40:

Naati		Ninde		Nahavaq	
Realis	Irrealis	Realis	Irrealis	Realis	Irrealis
v	mb	w	p	v	mb
t	nd	v	p	v^w	mb^w
r	ndr	w	mb	t	nd
w	mb^w	p	mb		
ʔ	ŋg	l	d		
k	ŋg	t	d		

Table 40: Patterns of consonant mutation in the SWB languages: Naati and Nahavaq patterns are reproduced from Crowley (1998) and Dimock (2009), respectively.

In Ninde, active consonant mutation patterns primarily affect initial labials like *wul* 'come' (irrealis *pul*) or *ve* 'go to' (irrealis *pe*). In rare cases, a complex onset has been elicited for irrealis forms, like the copula *vi* (irrealis *p-vi*). New patterns have been introduced with deleted initial *s and *k > *ɣ, which have reflexes of /j/ that are generally replaced by the /p/ of the irrealis (presumably because the consonant removed these deleted segments in monosyllabic verb roots from the context of an obligatory prefix *i-*). These active, newer patterns in Ninde may have replaced an older consonant mutation pattern.

The main evidence for this is that sound change exceptions are often initial consonants in Ninde verbs and statives. An irrealis may have been fashioned from uninflected initial nasals: Ninde *lumb* 'heavy' corresponds to Naati *ndip* and Nahavaq *ndip*^{w}*;* Ninde *waʔwaʔ* 'short' to Naati *mbøʔambø* and Nahavaq *mbuʔmbuʔ* better than to Avava and Neverver *mut*; Ninde *ndaŋgal* 'hang' resembles Nahavaq *taŋgar* 'hang' (but also Nahavaq *ŋgal* 'hang'); and within Ninde, *loulou* 'small' resembles verbal diminutive suffixes *-lou* and *-ndou* ~ *-ndu*. Each of these are viable cognates if Ninde previously had a realis and irrealis pair with the same consonant alternations that are found in SWB languages. In each pair, only one would survive, and usually it would have been the realis form.

Ninde irrealis forms are marked by the addition of *p(e)-,* similar to irrealis verbs in NCM (Barbour 2012:165, Crowley 2006a:74-5) as discussed in §5.4.1, but the placement in the morphological template (after person and number) is more like that of SWB languages (Table 41). The order of prefixes in SWB languages can be schematized as follows:

Ninde:	Person + (Dual)	+ (Negative)	+ (Irrealis)	+ Verb root
Nahavaq:	Person + (Dual)	+ (Irrealis)	+ (Negative)	+ Verb root
Naati:	Person + (Dual/Plural)	+ (Irrealis)	+ (Negative)	+ Verb Root

	Ninde		**Nahavaq**		**Naati**	
	Realis	**Irrealis**	**Realis**	**Irrealis**	**Realis**	**Irrealis**
1SG	*na-*	*na-p-*	*ne-*	*(ni)ŋge-*	*ni-*	*na-*
2SG	*ku-*	*ku-p-*	*u-*	*ku-*	*u-*	*wa-*
3SG	*Ø-*	*Ø-p-*	*i-*	*ke-*	*i-*	*ʔa-*
NSG	NSG-	NSG-*p*-	NSG-	NSG-*ʔ*-	NSG-	NSG-*a*-

Table 41: Realis and irrealis subject prefixes in SWB verbs, suggesting a common fixed order and formerly transparent realis morphemes. Where a prefix is identified as NSG*, this specifies a fusional person-number prefix, elaborated in the text that follows.*

In Nahavaq, an irrealis morpheme takes the forms *ke-*, *ŋge-*, or *ʔ-* (Dimock 2009:137). in singular-subject irrealis verb forms, it surfaces in first-person *ni-ŋge-* in religious registers (ŋge- colloquially) and the third-person irrealis prefix is *ke-*, with no corresponding third-person *i-*. In all nonsingular forms, the irrealis is marked by a ʔ- following a nonsingular subject prefix. The irrealis morpheme could reflect an older *ke-, which is affected by vowel deletion yielding coda [ʔ]. In Naati, the irrealis prefixes are even more transparent, since they reliably surface simply as *a-*. (Crowley 1998:120). The singular third-person forms, however, resemble the Nahavaq patterns in that realis *i-* is replaced by a fusional irrealis prefix. The Naati fusional form *ʔa-* suggests a shared origin with pre-Nahavaq general irrealis *ke-. (These might both be compared to Ninde *ka-,* which inflects complement clauses that share a subject with the matrix clause.)

Another similarity is that, unlike the structure previously shown for NCM languages, all three SWB languages derive a dual from a separate, non-singular stem (Table 42). These dual forms have an additional prefix *r-*, though the vowel changes in Nahavaq (and the similarity to PNCV*rua 'two') suggest a shared *ru-. Naati is unique in that plural forms are also marked. In each of the three languages, however, a distinct irrealis prefix reliably follows the dual marker in the morphological template. Recall that Avava bound pronouns do not distinguish nonsingular subjects, and Neverver dual *ri-* follows the irrealis *bi-* and is added to singular prefixes.

The only differences in order and function between Ninde person inflection and the other two SWB languages are in the form of the irrealis prefix and its order relative to the negative prefix: Ninde *p-* follows negative *s-*, whereas Nahavaq and Naati show reflexes of *ke preceding negative *sa-. This makes the placement of the negative prefix in Ninde more like that of Avava; unlike Avava, however, the order of person and mood prefixes is fixed in all three SWB languages.

	Ninde		**Nahavaq**		**Naati**	
	Plural	**Dual**	**Plural**	**Dual**	**Plural**	**Dual**
1INCL	*te-*	*taʁ-*	*nde-*	*ndu(r)-*	*ntal-*	*ntar-*
1EXCL	*e-*	*ʁa-*	*mi-*	*mu(r)-*	*ŋgal-*	*ŋgar-*
2	*e-*	*ʁa-*	*a-*	*wa(r)-*	*m^{w}al-*	*m^{w}ar-*
3	*re-*	*raʁ-*	*re-*	*ru- ~ ur-*	*al-*	*ar-*

Table 42: Abstracted prefix templates for SWB verbs inflected for nonsingular subjects, not including mood or polarity.

In summary, all SWB languages have some vestige of an old and unusual consonant mutation system for mood marking, but have dedicated irrealis prefixes used for virtually all persons in a fixed position. Unlike the non-Ninde NCM languages, dual and plural forms are distinguished with at least a dual prefix *r- that precedes the irrealis prefix. Furthermore, the dual is formed from the plural in SWB languages, whereas it is formed from the singular in NCM languages (excluding Ninde).

6. Several post-verbal modifiers

Several of the post-verbal modifiers in serial verb constructions are shared with Naati and especially Nahavaq. It is possible that Naati shares more forms that simply have not been documented in the sketch grammar. The commonalities between Ninde and mainly Nahavaq are presented in Table 43.

Ninde	Naati	Nahavaq	Position in Nahavaq
maʁsu ‘all over’	–	*morsu* ‘unrestricted’	PVM1
ma/ʔas ‘well’	–	*kos* ‘directly/correctly’	PVM1
taʁtaʁ ‘always’	*tartar* ‘always’	*tartar* ‘always’	PVM1
mun ‘first, while’	–	*mʷin* ‘first’	PVM1
lis ‘back, again’	–	*lis* ‘again’	PVM2
=saʁa ‘somewhere’	–	*sar* ‘remain’	Verb/PVM1
=jaʁa ‘all’	–	*jar* ‘finish’	Verb/PVM1

Table 43: Verbs in Nahavaq that correspond to Ninde verb-phrase enclitics. PVM stands for post-verbal modifier, and PVM1 is a position for non-initial serialized verbs that precedes the negative particle and PVM2 is a position following the negative particle.

Of these, all are exceptionless cognate forms with similar semantics, but only the root of Ninde *maʔas* ‘(do) well, (do) correctly’ corresponds to Nahavaq *kos* ‘(do) directly, (do) correctly’, but many verbs in Nahavaq appear with and without the reflex of a now unproductive stative prefix POc/PNCV *ma-. In terms of grammar, Nahavaq *lis* ‘again’ is a PMV2 and follows the second negative particle, but all others precede it. All of the forms in Ninde precede the object enclitic *=ŋe*, but negation data is lacking. Ninde has apparently lost main verb usage of these forms, but Nahavaq retains two of them as main verbs. These facts of shared form, usage, and grammar make this set compelling.

Although a cognate of *lis* '(do) again, (do) back' is not reported in Avava or Neverver, but Naman *lis* and Neve'ei *lieh* are clearly cognate. Nahavaq *sar* is not reported as a productive modifier, but it is present in *tusar* 'leave in a place' and *toʔsar* 'be left behind'. In Ninde, the cognate *-saʁa* is used as a valency-reducing suffix with a range of verbs that otherwise require a location as a complement when the source, goal, or location is unspecified. It may also be compared to Nahavaq *sara-* 'place for'. Finally, *jaʁa* in Ninde is used as an object quantifier when the event effects the object in its entirety (e.g., *jen-jaʁa* 'eat all of'). This makes it similar to the completive *luwo* that was discussed in §5.4.1 (Post-verbal modifiers), which instead centers the completeness of an action on the agentive subject (e.g., *jen-luwo* 'finish eating'). In terms of function, Nahavaq *i-jar en* 'after that' shows the same sequential function of Ninde *-luwo*.

This area is a challenge for comparison within SWB since no texts were used for the description of Naati, and it seems that very few serialized verbs were used in elicited material. Nevertheless, each of these represent relatively unpredictable serialized usage, given their semantics as main verbs. These semantic functions are fulfilled by other verbs in Avava and Neverver.

7. Complex nominalized verbs

The NCM and SWB languages also differ in terms of the complexity of verbs permitted in nominalized constructions formed by adding the general nominal prefix *na- and a nominalizing suffix *-ian to a verb root. Avava appears to have the most restricted nominalization patterns, whereas Nahavaq allows the most complex verb phrases to be nominalized; however, the comparative paucity of Avava and Naati data may account for some of this. Avava and Neverver allow reduplicated verbs and the generic subject *na-ut 'place' to be nominalized with this pattern; Neverver additionally allows at least one modifying verb *gor* 'prevent' to be part of the nominalized verb complex. SWB languages allow nominalization of a wider range of verbal compounds.

Tables 44-50 present lexical forms as they are attested for each of the NCM and SWB languages with increasingly restrictive distribution (and increasingly freer combinations of elements). At the extreme, only Nahavaq allows copulas with incorporated copular complements to be nominalized; a corresponding example is found for only one word in Ninde, and only in the translation of the gospels: *nevivetlaien(a)* 'Lord', which contains *vetla* 'wealthy man' (an archaic form of *petla*). This word in modern Ninde is unusual in that it has a word-final /a/ which does not raise to [e], but there is no evidence for a blocking final historical /ŋe/. This nominalized copular construction does not appear to be in use in Ninde today.

With respect to its permitted verbal root types, the most varied nominalization is found in Nahavaq, where the verb *roŋndow* 'believe' can bear both the discontinous nominalizing morphemes *nV-* + *-jen* and the discontinous negative construction *s-* + post-verbal particle *veʔ*. In this case the post-verbal negator appears after the final nominalizing suffix. In sum, SWB languages are much freer in terms of what can be nominalized, and some words in Ninde (like those in Table 46 and Table 47) perfectly parallel constructions in Nahavaq and Naati, while the other NCM languages do not appear to support such nominalization at all.

Reduplicated verbs

Avava	Ø-	lum-	lumuh	-ian
Neverver	ni-	tos-	tos	-ian
Ninde	no-	mbone-	mbone	-jene
Naati	ne-	ʔap-	ʔap	-ian
Nahavaq	na-	hap^{w}-	hap^{w}	-jen
Gloss	N-	REDUP-	wash/write/heap/circumcise/dance	-NMLZ
	'washing', 'writing', 'group', 'circumcision', 'dance'			

*Table 44: Nouns formed from reduplicated verbs with reflexes of discontinuous nominalizing affixes *nV- + *-ian*

Incorporated subject *na-ut 'place'

Avava	o-	ut-	ran	-ian
Neverver	na-	ut-	ran	-ian
Ninde	nu-	wuta-	ʁan	-ijene
Naati	ne-	wut-	reŋ	-ian
Nahavaq	ne-	wut-	reŋ	-jen
Gloss	N-	place-	light	-NMLZ
	'daylight, dawn'			

*Table 45: Nouns formed from verbs with incorporated subject *ut 'place' and reflexes of discontinuous nominalizing affixes *nV- + *-ian*

With *-gor 'prevent'

Neverver	ne-	tata	-gor	-ian
Ninde	ni-	ti	-ŋgoʁo	-ijene
Nahavaq	ne-	vʷer	-ŋgor	-jen
Gloss	N-	promise/say/speak	-prevent	-NMLZ
	'engagement', 'law', 'law, rule'			

*Table 46: Nouns formed from verbal compounds with reflexes of *gor 'block, prevent' and discontinuous nominalizing affixes *nV- + *-ian*

With *-sur 'follow'

Ninde	ni-	ti	-woʁ	-ijene
Naati	ne-	tiŋ	-hur	-ian
Nahavaq	ni-	ndiŋ	-hur	-jen
Gloss	N-	say	-about	-NMLZ
Translation	'story'			

*Table 47: Nouns formed from compound verbs with reflexes of *sur 'follow' and discontinuous nominalizing affixes *nV- + *-ian*

With incorporated object

Ninde	nu-	mul	-mijaʔ	-ijene
Gloss	N-	shed	-sick	-NMLZ
Translation	'healing'			
Nahavaq	ni-	veneven	-men	-jen
Gloss	N-	hunt	-bird	-NMLZ
Translation	'bird hunting'			

*Table 48: Nouns formed from verbs with incorporated objects and reflexes of discontinuous nominalizing affixes *nV- + *-ian*

With copula

Ninde	ni-	vi	vetla		-ien(a)
Gloss	N-	COP	-rich.man		-NMLZ
Translation	'Lord'				
Nahavaq	ni-	vi	-hala-	-n	-ien
Gloss	N-	COP	sibling	-3SG.POSS	-NMLZ
Translation	'sister'				

*Table 49: Nouns formed from a copula *vi with incorporated objects and reflexes of discontinuous nominalizing affixes *nV- + *-ian*

With negation

Nahavaq	ni-s-roŋndew-jen	veʔ
Gloss	N-NEG-believe-NMLZ	NEG
Translation	'non-believers'	

Table 50: Noun in Nahavaq formed from a negated verb with reflexes of discontinuous nominalizing affixes *nV- + *-ian. Negation is also discontinuous in Nahavaq, and the negative particle which follows verbs follows the whole nominal construction.

5.5 Conclusions

The most important factor in considering Ninde a mixed languages is that Mewun people identify Ninde to be a mix of two languages, even if the source languages are not identified with language varieties that would be recognized today. Ninde shares basic vocabulary and highly productive grammatical elements with both NCM and SWB languages, including some word forms and grammatical structures that could be attributed to either subgroup. The remaining parts of the language attributable to a single origin rely on negative evidence. Despite these methodological limitations and the limited morphological complexity of Malekula languages (in contrast to Mitchif, Mednyj Aleut, Media Lengua, etc.), the blended semantics, form, and usage of Ninde's lexicon and grammar are undeniable when compared to what is known about Neverver, Avava, Nahavaq, and Naati.

Comparative evidence shows that it is nearly impossible to formulate a basic clause, let alone a morphologically complex word, without elements from two distinct linguistic sources. Ninde's lexicon is shaped by two distinct diachronic trajectories of sound changes that have converged and then, as a single language, further diverged in the form of an autonomous lexicon. In this lexicon are doublets, synonyms, and words with blended form and meaning. In terms of grammar, much of the verbal inflectional system uses NCM parts to calque SWB morphological templates, yet negation is virtually taken wholesale from a SWB source. Common structures like existential, presentational, and deictic predicates and definite affixes also have mixed origins. This is a contact phenomenon that is deeper than typical large-scale borrowing.

Ultimately, the creation of Ninde may have been in part a creative endeavour. Based on anecdotes from the community, the language is perceived as distinct among Malekula languages, yet it seems to share some of its rarer properties with many of its neighbors. These changes are often masked by subsequent developments that Ninde seems to share with no other living relative, including the insertion of word-final /A/, a pharyngealized lateral approximant, and a uvular rhotic. These phonological changes have allowed for rebracketing of lexical word forms and analogical change that further make Ninde distinct. While speakers of other languages may have initially recognized Ninde as a mixture, a unified Mewun identity is no longer a novel one, and this recognizability may have given way to more patterned change. It is Ninde's status as a mixed language coupled with accumulated phonological opacity that account for Ninde's elusive genealogical classification.

6 The Models: Computational Modeling of Sound Change for Large Lexical Samples

6.1 Introduction

Now that prior work has established Ninde's likely dual parentage, this chapter represents an attempt to deploy a diachronic model over most of the lexicon. Two proto languages were reconstructed: North Central Malekula (NCM), including mainly (Small) Ninde, Avava, and Neverver; and South West Bay (SWB), including (Big) Ninde, Naati, and Nahavaq. This reconstruction was performed on virtually all the attested lexicon of each group using computational modeling of specific phonetic changes in the history of each language. Since Ninde is considered a member of both clades, the lexicon was divided by this method into NCM words, SWB words, words compatible with either origin, and words whose origin is as yet unidentifiable as a reflex from these sources.

Proto forms can be reconstructed from a single word using such a model; predicted cognates can be compared against the lexicons of related languages, identifying a potential cognate that may have been missed in the process of wordlist assembly. With comparative data, proto forms reconstructed on the basis of individual words can either converge, validating the diachronic models and supplying a singe proto form, or they fail to converge. In the case of failure, the diachronic model and/or cognacy relationship is reevaluated.

For this work, the consideration of Ninde as a mixed language means that words from two separate phylogenetic groups are in some sense cognate with forms from two different subgroups and not simply borrowed. The source language from the NCM group is geographically removed from the communities speaking related languages. For that reason, contact with other NCM languages has not been strong and all the forms shared in common with them are assumed to be cognates. In the case of the SWB source, which likely supplied much of the regional toponymy held in common with Naati and Nahavaq, there are both cognates (words that were involved in the original language mixing process) and more recent borrowings (some of which are perceived by speakers as such). The unconventional use of the word *cognate* in this sense differentiates words incorporated from the SWB and NCM sources at the time of mixing as cognates, and later adoptions from any Malekula language as borrowings, even if the borrowing reflects a cognate at a deeper time depth. This usage reflects a strong stance that Ninde's lineage can be traced to both sources, and it allows for a meaningful distinction between mixing and borrowing, both of which are at play.

Overall, this computational approach was successful in reducing the number of cognate sets, determining which Ninde forms were from each source involved in language mixing, and in generating viable proto forms that account for modern-day reflexes. This type of modeling en masse can help identify exceptions, provide valuable analysis of words from a range of usage frequency, and lessens biases imposed by cultural outsiders about what counts as a lexicon's core.

The sections that follow report the procedure and findings. First, I expand on the specific methods used in this task (§6.2), then I present the sound change models (§6.3), the results of the computational implementation and metrics for success (§6.4), and then conclusions (§6.5).

6.2 Methods

The comparative data described in Chapter 2 were used for comparative reconstruction, and the sound changes identified by this process were implemented computationally as diachronic models.

Sound changes were assessed and modeled in the following steps: cognates were assembled and analyzed, the comparative method was applied to find sound correspondences, and a model of sound changes was applied computationally to the lexical data. The first set of data focused on words that were only cognate with either the forms found in one or both of the other languages of South West Bay (SWB) – Nahavaq and Naati – or one or more languages spoken generally in northern central Malekula (NCM) – Avava, Neverver, and marginally Neve'ei and Naman. For Neverver, Nahavaq, and Ninde, comparison with the historical record (song forms and missionary texts) also aided in ordering some of the sound changes in diachrony.

The flowchart in Figure 15 demonstrates how cognate sets are reconstructed to a proto form and validated. By applying sound change formulations both forwards in time, from a proto word to a daughter, and backwards, from a daughter to all potential protos yielding the same form, cognate sets can also be collapsed. This is accomplished by searching the data for viable cognate forms, then prompting user action.

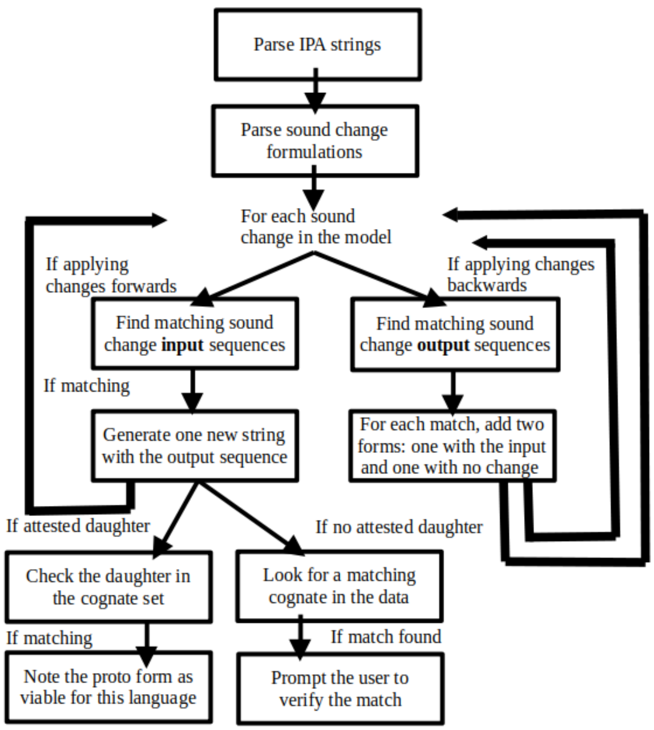

Figure 15: A flowchart schematizing both forward and backward application of sound change formulations. Both directions of reconstruction are used in tandem to assess cognate sets and identify potential missing forms from a larger database.

Since Ninde is considered in this dissertation as a mixed language with input from two source languages, one in the NCM group and one in the SWB group, the first step was to assess sound changes from doublets and one-sided cognates. This ultimately informed the reconstruction of two proto languages, provisionally named Proto North Central Malekula (PNCM) and Proto South West Bay (PSWB). In order to identify a Ninde form as either NCM or SWB in origin, possible expected forms were generated for both of these small language families, and then matched to a Ninde form if possible.

Models of sound change were applied to large amounts of lexical data en masse using software that I developed for the purpose. This software parses character strings in the International Phonetic Alphabet and applies sound changes that have minimally an input speech sound or natural class of sounds and an output, as well as a conditioning environment. Sound changes can be applied forward (generating a reflex from a proto form) or backward (generating sets of any proto form that can account for the reflex, given the model). Forward reconstruction, applied to proto sets generated from incomplete reflex sets, allowed cognates that were previously missed to be identified from among the singletons in the data.

Backward reconstruction and free variation pose limits on computational efficiency, since any sound change that involves deletion, merger, or large natural classes dramatically increases the number of potential proto forms at each step, and a change in free variation creates doublets. To mitigate the limits on computational power, several steps were taken: where possible, sets of proto forms generated from cognates were restricted to the intersection of each set; deletion environments were overspecified considering phonotactic constraints (e.g., word-final vowel deletion would apply after a consonant, but not after a final vowel); and natural classes were limited to those purported to exist in the proto language. Some changes (applied forward) were allowed to generate two expected forms with documented free variation. In this case, the word form that served as input to the change was also passed as output alongside a form in which the sound change applies. These free-variation changes were restricted (strictly on practical grounds) to the most shallow end of the sound change model sequences.

Though much of this processing could have been mitigated by beginning with sound correspondences, large numbers of entries were initially singletons. The lexical database was itself restructured by scanning the data for any possible reflexes of the proto form for any language missing from a set. The software displays the gloss for any form matches, allowing the user to block or allow sets and singletons to be combined with the stroke of a key. The inclusion of more than one language in a set reduced the intersection of possible proto forms.

The ordering and inclusion of sound changes was manipulated by the user to maximize the rate of successful reconstructions. The model produced a successful reconstruction if one or more proto forms were generated for a cognate set that accounted for every reflex; when a singleton was reconstructed, it was a success if a cognate was identified by scanning the expected reflexes of every potential proto form. Explicit errors, where words in a set were divergent from the expected reflex, were compiled in an error report that was used to reassess sound change models.

After hundreds of iterations, the comparative database was reduced in number of cognate sets (and thereby also the number of singletons) through the merging of complementary partial sets, and a success rate was determined for the final models (one for SWB and one for NCM). It is additionally possible to assess for Ninde whether each form is the expected reflex of the reconstructed PNCM or PSWB word form.

6.3 The sound changes

In this section, the sound changes ascertained from comparative reconstruction are addressed. These changes will be addressed in various subsections: first, the most recent changes, which can be identified by comparison with forms from the historical record (§6.3.1); then, Ninde sound changes are sorted in terms of whether they are found in cognates with NCM forms or cognates with SWB forms (§6.3.2); finally, the remainder of the changes, representing prehistorical developments (§6.3.3). The full model can be reviewed in the appendix.

6.3.1 Sound changes suggested by historical records

Only Ninde and Neverver have been compared with historical texts written in earlier forms of those languages. In the case of Neverver, Barbour (2012) examines data from a language identified as Nessan (incidentally *nessan* is the Neverver word for 'thing') recorded in phonetic notation by Deacon in 1927. Sound changes in Ninde were all identified first by comparison with other languages, but when missionary writings and a body of sung speech later became available, they were identified across media. This work assumes some phonetic values used for (Latin) letters in the missionary orthography, and suggests that sung Ninde preserves archaic word forms. Since the sound changes are also found in cross-language comparative work, the risk of misidentification is low.

Neverver

Barbour (2012:8) highlights the following changes since the time that Nessan data were recorded in 1926-7:

- Geminate consonants were already part of the phonology: <mmap> 'heavy' corresponds to *m:av* [m:aɸ], and <llag> 'seek' to *l:aŋ*
- A change from [^{n}ʤ] > [^{n}s] "was underway"; <nituğans> 'mosquito' corresponds to *nituxaz* [nituɣa^{n}s].

- Labiovelar consonants have merged with plain bilabial consonants since then: <nağambwir> ‘dog’ corresponds to *naxabir*, <pwis> ‘smart’ to *p:is* ‘hurt’, <nambwer> ‘mushroom’ to *neber* ~ *nebe*ᴅ

- An excrescent [d] has emerged in the prenasalized rhotic: <nenre> ‘blood’ corresponds to *ne*ᴅ*e* [ne^{n}dre]

- Front rounded vowels have since been unrounded or deleted: <nivüs> ‘bow’ corresponds to *nivis*, <tölas> ‘undo’ corresponds to *tlas* ‘untangle’

Changes in Ninde since Missionary Texts:

The forms that can be ascertained from written Ninde likely represent speech in the late 1800’s, when the first missionaries began their activities in South West Bay. Contrary to usage elsewhere in this dissertation, the voiced stops here represent allophonic voicing in intervocalic and post-nasal positions. This is still an allophonic pattern today, and the phonetic value of voiced and voiceless consonant alternations can additionally be verified by comparison with English loans like <neprofet> ‘prophet’ and <kupbaptais> ‘you will baptize’ (which includes a 2nd person singular subject *ku-* and an irrealis prefix *p-*), and comparing those with native forms like <latembi> ‘wilderness’ (corresponding to *l̪ˤatembi*). These changes are:

- Word-final /a/ was since raised to [e], or [o] if preceded by a syllable with a round vowel or rounded onset; the /a/ retains its status as phonemically distinct from /e/ if the word takes any suffixes or enclitics: <liveta> ‘night’ corresponds to *livete* and <iloga lua> [iloŋa#lua] ‘he went out’ to *l̪ˤo(ŋe) luwo*.

- The vowel was not raised if it preceded by uvular /ʁ/: <nevaha> [nevaʁa] ‘hand, arm’ corresponds to *nevaʁa.*

- Word-medial /a/ is raised if the following vowel is [e] (except in the context of a labiovelar, uvular, or pharyngealized lateral): <tlaba> ‘big’ corresponds to *tlepe*.

- Monomorphemic words ending with /ŋa/ lost this final syllable in all but formal and religious registers, creating a new invariant final /a/ (in contrast to the archiphoneme /A/ which is affected by word-final raising): <ndaga> [ndaŋa] ‘good’ corresponds to *nda* in informal usage.

- Some affricate represented by <j> and restricted to word-final position fully merged with /s/: <netemij> [netemits] ‘devil’ corresponds to *netemes.*

- The prenasalized rhotic represented by <nr> became oral /r/: <enrei> ‘up’ corresponds to *erei*, <nanre> ‘blood’ to *nare.*

- Front vowels were rounded before a following round vowel: <nemucut> [nemugut] ‘man’ corresponds to *nymŋgut.*

- Several words underwent syncope (deletion of medial vowels): <nohobo [noʁobo] ‘thing’ corresponds to *noʁpo*, <motina> ‘right’ to *motne* [modne], and <isu̲podau> [isupodau] ‘rested (lit.: sat small)’ to *s_pondou*.

- An excrescent nasal is inserted where syncope resulted in an adjacent nasal and oral stop with different places of articulation: <ne̲mucut> [ne̲mugut] ‘man’ corresponds to *nymŋ̲gut.*

- The diphthong represented by <au> is now generally [ou]: <nicenau> ‘crowd’ corresponds to *nikinou.*

- Many vowels became [ə] in the environment of a dorsal consonant or unrounded bilabial. In many cases the underlying vowel quality cannot be established synchronically: <kina> 1SG.PRONOUN corresponds to *kəne*, <sike> ‘not, if’ to *səke.*

Changes in Ninde since sung Ninde

Compared to sung forms, the following differences can be observed:

- A form pronounced /ŋg/ sometimes corresponds to /j/: *ŋganjara* ‘ate all of’ corresponds to *jenjaʁa.* It could be a preserved voiced velar fricative /ɣ/, subsequently changed in Ninde song to the existing voiced velar as a reanalysis of its phonotactically illicit initial voicing.

- Sung [r] corresponds to modern /ʁ/: *ranmbaŋgarou* ‘on Nembangkaxou (an outcrop near Lawa‘)’ corresponds to *ʁa nəmbaŋgaʁou*.

- Modern Ninde allows /a/ and /o/ before the (now uvular) rhotic, but song allows high vowels. These were most likely lowered in the context of the uvular, phonemecizing allophonic [w] onsets of initial /u/: *raneureure* ‘on the island’ corresponds to *ʁa nawoʁawoʁa*.

6.3.2 Irreconcilable sound change as a diagnostic for parentage

Several classes of phones have been notoriously difficult to account for when comparing high-level reconstructed proto languages to their modern descendants, since contrasting phonological features may have swapped values across the lexicon within languages. These features include the nasality of prenasalized consonants and the voicing of velar stops across the Austronesian languages (Blust 1996), the place and/or manner contrasts between *r and *R in Northern Oceanic (François 2011), and post-velar consonants generally in Austronesian (Lynch 2009). Blust (1996) rejects a neogrammarian orientation that assumes a satisfying explanation has not yet been worked out, but suggests that irregularity could be inherent to intersections of natural classes (like voicing status in velar consonants). Others have narrowed in on the sociolinguistics of language diversity as an explanation for such distributions.

Though I elsewhere espouse the comparative method as a methodology, the following section considers irregularity in Ninde's inheritence of particularly post-velar consonants and sibilants. Within the low-level groups previously identified (NCM and SWB), the exceptional reflexes number very few. The complexity of Ninde's reflexes can be accounted for if the lexicon is considered mixed.

Clark (2009) reconstructs Proto North Central Vanuatu with a comparatively limited number of proto forms in these areas of reflex complexity. He reconstructs plain velar *k and prenasalized *ŋg (represented as <q> in that work) alongside a glottal stop *ʔ. He only reconstructs two sibilants: an oral *s and a prenasalized sibilant represented by *z (which can be realized within and between languages as an alveolar or postalveolar fricative or affricate). Among Ninde's reflexes of *k, he lists [k], [ʔ], and Ø, and of *s, [s] and Ø, while *g is straightforwardly retained and *z realised as [s]. While phonetic context plays a role in determining the outcome, the origin (NCM/SWB) of a lexical form is also a key factor.

Velars

Ninde has multiple possible reflexes of PNCV *k, including /k, j, w, ʔ, ŋg/ and deletion. Different reflexes will be shown to appear in similar sets of environments. Instead of complexifying the proto languages with extra phones, these reflexes will be explained as a result of language mixing. Ninde clearly has overwhelmingly deleted *k from NCM sources, while it retains most reflexes from SWB as glottal stops. Nonetheless, in some environments, Ninde has debuccalized *k from PNCM and, conversely, has a velar /k/ as a reflex of PSWB*ʔ. This makes for complex correspondences with both NCM and SWB phones, but these will be broken down into various contexts.

The PNCV*k can be reconstructed as a velar fricative (*x, but possibly with a voiced allophone *ɣ) for PNCM between low vowels and with a corresponding glottal stop in SWB languages with velar stop allophones (Table 51).

Protos		**NCM**			**SWB**	
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
*kamaliR	*kamali	*ne_mel*	*naɣamal*	*_amal*	–	*naʔamel*
‘men’s house’		‘chiefs’ house’	‘house’	‘house’		‘meeting house’
*papian ‘firewood’	*kabu	*ne_mb*	*naɣaʙ*	*a:_ʙ*	*naʔamp*	*naʔamp*
		‘firewood’	‘fire, firewood’	‘fire’	‘fire’	‘fire, firewood’
*kape	*kave	*ne_ve*	–	–	–	*naʔap* ‘rock crab’
‘t.o. crab’		‘hermit crab’				
*kaso ‘coconut-leaf basket’	*kato ‘basket’	*ne_nde*	*naɣat/xat*	*_at/_at*	*noʔon(d)*	*noʔontr*

*Table 51: Regular reflexes of PNCV*k in Ninde when inherited from the NCM source.*

PNCM*x is deleted in Ninde between low vowels *a or *o, but generally those vowels coalesce into a front vowel *e; in Neverver, the fricative is retained as a voiced velar fricative in word-medial contexts after voiced segments; in Avava, *x is deleted intervocalically, yielding a long vowel via coalescence of the preceding and following vowels. If sequences of PNCV*aka (> NCM *axa) yielded a fronted vowel *e in Ninde, the form can be traced through NCM. This allows us to tentatively identify an NCM source for Ninde *neve* 'hermit crab', even when no cognates are attested there.

Where there are unexpected reflexes in the other NCM languages as in Table 52, Ninde often remains consistent – this allows us to ascertain a critical ordering of the sound changes affecting PNCM *x in Avava and Neverver.

Protos		**NCM**			**SWB**	
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
*k̲apik{a,u} '*Syzygium malaccense*'	*k̲avika '*Syzygium sp.*'	*ne_vyke*	*ni_vɣo* 't.o. tree'	*_avik*	*naʔ̲aviʔ*	*naʔ̲aviʔ*
*k̲ayu 'tree, wood'	*k̲ayu	*ne_i*	*naɣ̲a*	*ag̲a(i)*	*naʔ̲ai*	*naʔ̲ai*
–	–	*no_ʁo* 'men's ceremony'	cf. *nox̲aro* (Neve'ei)		*neʔ̲erew*	–

*Table 52: Reflexes of *k in Ninde that can be attributed to NCM (whether a cognate is attested or not), even when there are SWB cognates.*

A form can be reconstructed as PNCM *naviku ‘Syzygium sp.’, but comparison with POc and PNCV suggests *naxaviku (the final round vowel is necessary to explain Ninde and Neverver’s round vowels). Both forms would predict the Ninde form, but unusually, Neverver and Avava have also both lost the *x to deletion. In *naxa(i) ‘tree, wood’, Avava has a prenasalized /g/ as a reflex, which could represent a different reflex in post-tonic position, though Ninde and Neverver both have the expected reflexes of *x. Finally, though no cognate is attested for Avava or Neverver, PNCM *nVxVro ‘men’s ceremony’ can be reconstructed with any combination of two low vowels; even an intermediate realization of *nero from *naxaro would yield a harmonized vowel /o/ seen in Ninde’s *noʁo*. Reconstructions from Ninde will likely fail to predict forms like these in Avava and Neverver, and it may have something to do with underdescribed stress placement.

Modern Ninde /e/ can surface as [o] if followed by another rounded vowel and as [a] in the non-final context before /l̪ˤ/ or anywhere immediately before or after /ʁ/, as in the forms in Table 53.

Protos		NCM			SWB	
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
–	–	*l̪ˤa_mb* ‘on the fire’	*na-xaʙ* ‘fire’	*aː_ʙ*	*na-ʔ̲amp* ‘fire’	(*na-ʔ̲amp* ‘fire’)
*paRu ‘*Hibiscus tiliaceus*’	*vaRu	*na_ʁi* ‘beach hibiscus’	*naxari* ‘beach hibiscus’	(*balaka*)	–	(*mpalaŋkeu*)
–	(*vera *‘Abelmoschus sp.’*)	*no_moi*	–	(*oweŋ*)	*na-ʔ̲avʷus ne-weŋ*	*naʔ̲avʷus*

*Table 53: Reflexes of *aka that would be /e/ in Ninde, but are lowered in the environment of a dorsal consonant or rounded in the environment of a round vowel.*

These changes likely affected a phonemic /e/ and are active processes in the synchronic morphophonology of Ninde. This suggests that vowel lowering in the context of uvular and pharyngealized consonants was a comparatively late development. Furthermore, Ninde's *nomoi* ‘*Abelmoschus sp.*’ resembles at least the head of corresponding SWB compound forms, but has the coalesced vowel and deleted final *s are expected from a corrsponding PNCM*naxavʷis. This means that nasalization of *vʷ in the context following a nasal happened at distance with an intervening *x, or after deletion of *x.

Ninde's reflexes of PNCV *k in this context that were inherited via SWB languages are realized as glottal stops. This is also the case for Naati and Nahavaq. Nonetheless, the data (Table 54) require brief explanation.

Protos		NCM			SWB	
POc	**PNCV**	**Ninde**	**Neverver**	**Avava**	**Nahavaq**	**Naati**
PMP*sakaRu 'reef'	*sakaRu	*naʔau* 'reef'	–	–	*naʔhaw*	*nahaʔau*
(*qatop 'roof, thatch')	(*ato 'thatch, Sago palm')	*naʔaivər* 'thatch'	(*niat*)	(*iat* 'Sago')	–	*naʔaivør*
–	*makobu 'gecko'	*nemaʔamb*	*numuxum*	*omo:b* 'skink'	–	*meʔemb*
–	–	*-maʔas* '(do) well, correctly'	–	–	*kos* 'correctly (v2)' (Nahavaq)	–

*Table 54: Intervocalic reflexes of *k for Ninde forms that were inherited from an SWB source and not NCM.*

Included in this group are: *naʔaivər* 'thatch', identified as a loan into Ninde (Letpen, p.c.); *nemaʔamb* 'skink', which has a semantic change found in Avava but the expected SWB form; and *maʔas* 'well', which is not attested with the stative prefix *ma-* in Nahavaq's cognate. In some cases, these changes mirror language-internal variation (e.g., see Dimock 2009:276-9 on verbs with vestigial stative *ma- and their bare-stem counterparts in Neverver). These are considered reflexes of PSWB *ʔ.

PNCV *k yields a Ninde glottal stop reflex syllable-finally after a low vowel *a or *o whether it was inherited via NCM or SWB (Table 39).

POc	PNCV	Ninde	NCM (Neverver)	SWB (Naati)
(*maridriŋ ‘be cold’)	*mala(ka)so ‘cold’	*namal̪ˁaʔ* ‘cold’	**malah~melekit* (Avava)	*malʔah*
*sake ‘ascend’	*sake ‘go up’	*jaʔ* ‘climb up, ascend’	*sax* ‘go up’	*haʔ* ‘climb up’
*qasi ‘no, not’	*tika-i ‘not’	*eheʔ* ‘no’	*i-sxen* ‘it is not’	*sasa*, *siʔsiʔ* (Nahavaq)
(*droŋan ‘with’)	–	*maʔ* ‘and, join, be with’	*ma_* ‘with, and’ (Avava)	–
*matakut ‘fear, be afraid of’	*mataku	*metaʔ*	*matak* (Avava)	*mataʔ*
PAn*ala ‘take, get, marry’	*laki ‘marry’	*l̪ˁaʔ* ‘marry’	(*lav* ‘marry, get’) *dakdaka* ‘married’ (Avava)	*leʔ* ‘marry’
PPh*lukluk ‘hide’	–	*l̪ˁaʔl̪ˁaʔ* ‘hide’	(*solix*), *lalak-i* ‘secretly’ (Avava)	*luʔluʔ*
(*poki ‘clear the ground for a garden site’)	–	*ʁaʔ* ‘work, clear ground’	*rak* (Avava) *rax* (Neverver)	*raʔ*
*pi{s,j}iko ‘meat, flesh’	*visiko ‘meat’	*nowoʔ* ‘body’	*ipso_* (Avava), *nivisxon* (Neverver)	*nevyʔoh*

*Table 55: Word-final reflexes of *k in Ninde and NCM/SWB. Data in each cell are from the representative language indicated in the header unless followed by a language name in parentheses.*

When the word ends in *h, there was a tendency for the glottal stop and fricative to metathesize in the SWB source before *h was deleted (e.g., ‘cold’ in Ninde and ‘meat’ in Naati). In many cases, the metathesized form exists alongside the original form. Naati and Nahavaq retain the regular glottal stop and undeleted *h, providing a clear basis for comparison. This could be reconstructed as *k for PNCM and *ʔ for PSWB – the latter debuccalized nearly all *k. Avava word-final /k/ as in *pinok* ‘steal’ is sometimes deleted non-finally in *vi-pino_-ian* ‘theft’, and this likely reflects an older distribution of intervocalic *x allophones of *k. This retained final *k is seen in intransitive *rak* ‘clear a garden site’ but not in inalienable *ipso-* ‘flesh, meat’ and *ma* ‘with, and’ which could reflect the domain at which this sound change occurred. While these changes can be explained by phonetic context within syntactic constructions congruent with word class, they appear to be irregular to a reconstruction algorithm operating on the domain of single notional words, adding to the overall error rate of the approach in the results reported in §6.4.

If the previous consonant is also a *k, then intervocalic *k is realized as a glottal stop after *a or *o (Table 56) from PNCM – these can be reconstructed as geminate *k:. The second Ninde form, *koʔone* ‘bitter’, is reduplicated in a manner consistent with NCM (partial reduplication).

POc	PNCV	Ninde	NCM (Neverver)	SWB (Naati)
*kani 'eat'	*k̲ani 'eat'	*k̲aʔ̲an* 'eat (intr.)'	*k̲:an* 'eat (intr.)' *k̲an* (Avava)	*ʔ̲aʔ̲an* 'eat (intr.)'
–	*k̲ona 'bitter'	*k̲oʔ̲one* 'bitter'	*k̲:on* 'taste bad'	*k̲onk̲on* 'bitter, sour'

*Table 56: Reflexes of repeated *k, generally in reduplicated forms.*

These sequences yielded geminate /k:/ in Neverver and a short /k/ in Avava. The contrast in NCM languages between transitive *xan 'eat' and intransitive *k:an 'eat' can be reconstructed as geminate on the basis of Avava's inflectional classes. Long prefixes in Avava incorporate the fused third-person *i- and affect words that were monosyllabic – words like *kan* and *ka* 'call' that take long prefixes have corresponding geminates in Neverver (*kek̲:e* 'call'). In Ninde, these can also be contrasted with the word *nukuk* 'nest', which has no cognates but did not undergo the change. The absence of an inserted nasal suggests an origin in the SWB source. Ninde could have inherited a reduplicated form that had not yet yielded the geminate and undergone assimilation, or it could have simply realized *k: as /kVʔ/ with a vowel quality copied from the following vowel.

Before and after high vowels, a reflex of PNCV*k is generally always /k/ in Ninde, whether inherited from PNCM or PSWB (Table 34).

POc	PNCV	Ninde	NCM (Neverver)	SWB (Naati)
*k{a,e}li 'dig	*keli 'dig'	#***k**i̱l* 'dig'	***x**il*	***ʔ**il*
*kapik{a,u} '*Syzygium malaccense*'	*kavika '*Syzygium sp.*'	*nevy**k**e*	*niv**x**o* 't.o. tree'	*na**ʔ**aviʔ*
*qumwaŋ 'hermet crab'	*rakumwa 't.o. crab'	*noʁ**k**um* 'crab'	*nuru**x**um* (Neve'ei)	*neru**ʔ**um* 'land crab'
*keza 'intense blue or green'	*malakeza 'blue, green'	*meli**k**ise* 'blue, green'	*melih* (Avava)	*mal**ʔ**asan*
*(ŋ)isa 'name', *ŋacan 'name'	(*asa 'name')	*ni**k**i* 'name'	*ni**x**izan* (Neverver)	*ne**ʔ**ia-*
?*loki 'bend, crooked'	?*luku-ni 'bend (esp. limbs)'	*lu**k**(lu**k**)* 'cover (oneself)'	*lu**k**u* (Avava) 'cover'	?*lu**ʔ**lu**ʔ*** 'hide'
*liko 'commit suicide by hanging'	*liko-ti 'strangle, hang'	*li**k**e* 'hang (oneself)'	?*la**x*** 'hang' (Neve'ei)	–
(*bulos 'turn around')	(*vilo-si 'turn')	*vi**k**is* 'turn, change'	(*vilih* 'turn' or *vuk* 'change' [Avava])	*vi**ʔ**is*
(*ibe 'mat')	*doŋi 'mat'	*nimwe/ri**k*** 'mattress'	*eᴅe**ŋ*** 'woven mat' (Avava)	*nentri**ʔ*** 'woven mat'
(*ñamuk 'mosquito')	(*namu-ki 'mosquito')	*nandu**k**as* 'mosquito (taboo avoidant)'	*nitu**x**az*	*teu**ʔ**as*

Table 57: In the context of high vowels, inherited velars and glottal stops are realized as Ninde /k/ from both NCM and SWB sources.

From these cognate sets, it would appear that PNCM*x underwent fortition in this environment, or these must be reconstructed as PNCM*k, which merged with the fricative in Neverver and (like the fricative) was deleted in Avava, when it did not geminate in partially reduplicated stems. Likewise, PSWB can be reconstructed with *k, with Naati and Nahavaq later merging this with the glottal stop of other environments, or *ʔ, with a Ninde reflex of /k/.

After a syllable with a nasal onset, Ninde has /ŋɡ/ as the reflex of PNCM*k (Table 32).

POc	PNCV	Ninde	NCM (Neverver)	SWB (Naati)
(*puna)	–	*nu**ŋg**ute* 'trunk'	*nu**x**utn* 'trunk'	*(mpati-)*
*kuRita	*kuRita 'octopus'	*nu**ŋg**uwute* 'octopus'	*no**x**owit*	*noʔoijit*
*maRuqane	–	*nymy**ŋg**ut* 'man'	*nim**x**ut* 'man'	*(nemurut)*

Table 58: Ninde /ŋɡ/, with an inserted nasal after a syllable with a nasal onset, are generally from the NCM source.

The distribution of these inserted nasals before high vowels lends further support to the reconstruction of *k for PNCM in that environment. The inserted nasal can be dated by comparison to a stage after divergence from PNCM, and after deletion of *k before low vowels (which is shared with the patterns already discussed with Avava and Neverver). Incidentally, Neve'ei has glottal stops corresponding to this environment, and a fricative velar elsewhere.

Ninde /k/ corresponds to Nahavaq and Naati /ʔ/ after a syllable with a nasal onset (Table 29). The expected reflex, if these words are from the NCM source, is a velar stop with an inserted nasal /ŋɡ/.

POc	PNCV	Ninde	NCM (Avava)	SWB (Nahavaq)
–	(*walu ‘group’)	*nik̲inou* ‘crowd’	–	*niʔ̲inew*
PMP*kamuniŋ ‘*Murraya sp.*’	–	*nik̲amou* ‘t.o. tree’	–	*nuʔ̲umow*
PMP*hikan ‘fish’	*ika + *miala ‘red’	*nik̲imijal̺ˤe* ‘redfish’	*iki melih*	–
*kumwi ‘beard’	(*ase ‘chin, jaw’)	*nuk̲um* ‘chin’	*k̲imwi* ~ *g̲imwi* ‘chin’	–
*(ŋ)isa ‘name’, *ŋacan ‘name’	(*asa ‘name’)	*nik̲i (a)* ‘name’	*ni**x**izan* (Neverver)	*neʔ̲eh* ~ *neheʔ̲*

Table 59: Ninde /k/ with no inserted nasal are likely from the SWB source.

A reconstruction of PSWB *ʔ is viable, because there are no glottal stops in Ninde attested in the environment before a high vowel (regardless of source). This means that synchronic alternations, now unproductive (but observable in *me-taʔ* ‘be afraid’ vs. *tək-əs* ‘be startled’) could provide the basis for a change of *ʔ to /k/. The alternative is that Ninde simply had prenasalized *k in nasal contexts in the NCM source, but not in the SWB source, and these differences persisted after mixing.

Word-initially, PNCV*k yields Ninde /k/, presumably unchanged from both PNCM and PSWB (Table 27):

POc	PNCV	Ninde	NCM (Neverver)	SWB (Naati)
*kani ‘eat’	*kani	#***k****aʔan* ‘eat’	***k****:an*	***ʔ****aʔan*
(*matuqa ‘older person’)	–	#***k****aiwut* ‘old man, husband’	***x****avut* ‘husband, old man’	***ʔ****aivwut* ‘husband’ (Nahavaq)
*kadrut ‘scratch an itch’	*kaRa-ti ‘itchy’	#**k**aʁas ‘scratch’	**x**orxor ‘itch, scratch’	***k****aras* ‘itch, scratch’
*k{a,e}li ‘dig	*keli ‘dig’	#***k****i̱l* ‘dig’	***x****il*	***ʔ****il*

*Table 60: Word-initial retained *k in Ninde.*

Crucially, these are either found in polysyllabic verb roots, which would not have a fused prefix *i-, and in a rare set of nouns that begin with *ka-*. Nahavaq and Naati have a mix of /k/ and /ʔ/ in the broader set of reflexes, but the contexts are not clear from PNCV.

Verb-initial *k became /j/ in monosyllabic morphemes from PNCM, probably regularized in the context of the 3rd person singular subject prefix *i- (which Ninde has uniquely lost) (Table 26):

POc	PNCV	Ninde	NCM (Neverver)	SWB (Naati)
*kani 'eat'	*kani	***j****en* 'eat (something)'	*(i-)****x****an*	*(i-)****ʔ****an*
*ka(Ra)ti 'bite'	*kaRati 'bite'	***j****es* 'bite'	*(i-)****x****as*	*(i-)****ʔ****as*

*Table 61: Palatal approximant reflexes of *k in Ninde.*

These are reconstructed as PNCM *x, allowing for the generalization that PNCV*k > PNCM *x before *a. It also allows for the generalization that fricatives PNCM*x and PSWB*h in the environment after *i- yield the palatal approximant /j/, with one exception, *iʔaras* 'gasp for air', which is the only verb documented in Ninde with an initial /i/.

The patterns here suggest that Ninde acquired much of the vocabulary containing glottal stops from the NCM source, even though Avava and Neverver have no glottal stops. Neve'ei, which shares a large amount of vocabulary with these two languages, also developed glottal stops in similar contexts, and it is possible that *k was debuccalized in Ninde before contact with languages of South West Bay. Unlike Ninde, the patterns in Neve'ei allow for glottal stops in the environment before high vowels (*nibiʔiŋ* 'giant turban shell'), and between two low vowels (*naʔai* 'tree, wood'); where Ninde reliably has /k/, Neve'ei has /x/, like before high round vowels (*nuxut-* 'base, tree [of a named type]') and in partially reduplicated stems (*xoxon* 'bitter'). If Ninde and Neve'ei share any innovation here, it can be exclusively debuccalization of word-final *k, and lenition of PNCV *k that is also shared by Avava and Neverver.

Sibilants

The reconstructed sibilant contrast of PNCV must have been maintained in PNCM, yet dramatically changed: alveolar fricative *s and a prenasalized coronal fricative *z (Table 62).

Reconstructions		**NCM**				**SWB**	
POc	**PNCV**	**Naman**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
***s**ale**s**ale 'float'	***s**ale	***s**əl*	*salsal*	***s**al**s**al*	(***j**ele**j**ele* 'hover')	–	***h**al*
–	***s**avwa 'dance'	***s**av*	*sav*	***s**ap*	***j**awo*	***h**ap*	***h**apw*
***ñ**amu 'chew'	***z**amwa	*[xasxas]*	*z̧e(m,mw)*	***s**em*	***s**a(mwe,mo)*	--	***s**emw*
*qu**s**an 'rain'	*u**s**a	*no͡u**s***	*nau**s***	*au**h***	*nuwo*	*nuwu**h***	*newu**h***
*boRok 'pig'	*buka**si**	*bue**s***	*niʙwa**s***	*aʙua**h***	*nəmbuwa**s***	*n(i,ø)mbua**s***	*nimbwuwe**s***

*Table 62: Reflexes of PNCV *s and *z, with relevant palatalizing environment distinguishing Ninde cognacy with SWB to the exclusion of cognacy with NCM sets.*

Note that Ninde *jelejele* 'hover' corresponds to Avava *jaljal*, and has been falsely included in the set of cognates from POC *sale *sale 'float' in prior work. These data show that PNCM *s and *z correspond to PSWB *h and *s, but in SWB languages, older *s in the environment before *i were not debuccalized. This difference allows for a distinction between Ninde *nuwo* 'rain', which on the basis of deletion could represent either the NCM or SWB source, and *nəmbuwas* 'pig', which is only explained as coming from the SWB source. In Neverver, the sibilant contrast was retained, but the reflex of *z is merged with *s in some people's speech (Barbour 2012:36). In Avava, all these sounds have reflexes of [s], but do not represent full mergers: word-finally, *s yields [h] in Avava and *z yields [s]. This is much like the general pattern in SWB languages, but with no distinct reflexes for palatalized *s in any NCM reflexes.

Both NCM and SWB languages palatalized coronal stops before *i, but the patterns can be distinguished from each other (Table 25).

Reconstructions		**NCM**				**SWB**	
POc	**PNCV**	**Naman**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
*pati 'four'	*va**ti**	*i-ve**s***	*i-va**s***	*i-va**t***	*ve**s***	*i-vø**s***	*i-ve**s***
(*taŋop 'touch')	***ti**bwa-ri	***č**əber*	*[se]**s**ber*	***s**iber*	***s**paʁ*	*[vurvur]*	*[soʔ]*
***p**udi 'banana'	*vu**di**	*nevə**z***	*navu**(s,z)***	*ap**m***	*ny**s***	*[ne:ŋgit]*	*[n(i,u)ŋgut]*

*Table 63: Contexts for debuccalization of *s in SWB languages.*

In Neverver, *t and *d in the environment before *i have merged with *s and *z, respectively. In Avava, in the context CV_#, where *z followed an unstressed high vowel word-finally, it yielded a homorganic syllabic nasal (e.g., PNCV *navuz 'banana' > *apm*); and *t yields [t] word-finally, after deletion of a following high vowel. A proto language for NCM could be reconstructed with both *s and *z, but the status of *t is unclear – like the Southeastern Malekula languages, there could have been an intermediate phoneme *č for palatalized *t, which only in Avava would be realized as [t] in word-final position. As noted in §6.3.1, Ninde clearly had a distinct reflex represented by <j> and crucially never <nj> (vs. nasal-oral sequences <mb nd gc>) in missionary texts, and this was restricted to word-final codas (except in English loans like <wurjip> 'worship') and corresponds to /s/ today. This older Ninde form supports a reconstruction of PNCM *č for PNCV *ti, while PNCV *di would have merged with *z.

In rare cases, *s has not been debuccalized in Naati and Nahavaq where predicted by the generalizations stated, and these are in word-final contexts preceded by a front rounded vowel (Table 64).

Reconstructions		NCM				SWB	
POc	**PNCV**	**Naman**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
*pus{a,o} 'heart'	*v**u**so 'heart'	*[nuvidamda am]*	*[nol:o-n]*	*[mu:t nan]*	*nawa/nev**y**s*	*niv^{w}a/nav**y**s*	*nav^{w}a/nev**us***
PAn*pijax 'how many'	*v**i**sa 'how many'	*iv**i**s*	*iv**i**s*	*ivih*	*v**i**je*	–	*v**i**h*
*susu 'female breast'	*s**u**su	*nisə-n*	*nas**us***	*as**i**h*	*n**u**_*	*nis**y**s**y**- ~ neh**u**h**u**-*	*as**u**s ~ nuh**u**h**u**-*
*maqati 'ebb, dry reef'	*ma-masa 'dry'	*mes~iməs*	*m:as*	*mah*	*[mem]*	*mah*	*meh*
*qase 'jaw'	*ase 'jaw'	*na:se-*	–	*as-*	*nes**i**-* 'jaw',		
					ne_ne- 'cheek'	*neh**e**-*	*neh**e**-*
*pusuR 'hunting bow'	*vusu 'bow'	*nivəs(ax)*	*niv**y**s*	*ih[bwari:n]*	*na[ʁ]ve_* 'arrow'	*nev**ø**h*	*ne[re]v**u**h* 'arrow'

*Table 64: SWB has debuccalized (deleted in Ninde) reflexes of *s (underlined) escept in contexts of following high vowels. Avava debuccalization, by contrast, is conditioned by the word-final environment.*

The following morphological context could be important. Alienable and intransitive forms *vih 'how many', *meh 'dry, (of reef) low tide' have final *h, whereas inalienable *nav^{w}anevus 'heart' does not. Regardless of the origins, Ninde is largely consistent in cognate sets with SWB languages. As exemplified in §5, the doublet in Ninde of *nesi-* 'jaw' and *nene-* 'cheek' represents forms from each of Ninde's sources – the deleted sibilant corresponds to PSWB*h, whereas the retained one corresponds to PNCM*s, even though these in turn are inherited by the proto languages from the same older sibilants.

Final vowels

It is difficult to predict whether a Ninde form will have a final vowel or not, but the following generalizations go a long way:

- Final vowels are inserted in places where a final continuant or entire syllable would be expected to have undergone devoicing;
- some final consonants are affected by preceding high vowels, which may have palatalized or rounded the final consonant; and
- if a form is reduplicated, an inserted vowel between the reduplicated roots will be copied to the word edge.

Generally, these may represent word-final articulatory gestures with a high potential for perception as a reduced final vowel: voiceless continuants and secondary articulation.

First, some observations about the phonotactics of Ninde will establish some expectations about vowel insertion. Nearly all of the word-final vowels in Ninde are [a] or [o]. Before a suffix or enclitic, this vowel becomes [a] and is represented as /A/, an archiphoneme distinct from /e/ and /o/. Nevertheless, some final vowels are /e/ and /o/, even when they are inserted. Final [u] and [i] are generally deleted, and where a vowel is "retained" in this position, it is typically [e]. Since [e] is not licit after /ʁ/ in Ninde phonotactics, it is lowered in that environment to [a]. From this distribution of final vowel qualities, and the unpredictable patterns of final vowel retention (Lynch 2014), it is likely that the vowels represent innovations after these vowels were lost in Ninde's ancestors.

Consonants condition allophony that might offer insight into the patterns found in Ninde. Like the SWB languages (and as shown in decades-old written records of Neverver), the back rounded vowel /u/ has a front allophone [y] (but not so /o/, as suggested for these other languages). Unlike the pattern described for Nahavaq, this is restricted to vowels before a coda /s/ or /t/, not involved in productive morphophonemic alternations, and /y/ has become phonemicized in other environments. Most vowels become [ə] before or after velar stops, but not all vowels are affected the same in everyone's speech. After /t/, /l/, /s/, or /r/, no [o] is inserted, even if the preceding vowel is rounded, but after /l̪ˤ/, /ʁ/, bilabial consonants, and rarely velar ones, an inserted [o] is possible. In fact, after /t/ and /s/, final vowel /e/ often corresponds to /o/ in cognates. This means that the final consonant may determine the quality of an inserted vowel.

Generally, a completely voiced final syllable does not have an inserted final vowel in Ninde, regardless of the source (Table 65).

Reconstructions		NCM			SWB	
POc	**PNCV**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
(*matanV 'sharp')	(*makani)	*m*^{ŋ}*gal*	*ma*^{ŋ}*gal*	*maŋgal*	(*makan*)	(*makan*)
PWMP *lagu 'way, manner'	–	–	–	*niliŋ*	*neliŋ*	*neliŋ* 'behavior'
(*unus 'pull out')	(*tuku 'release)	^{n}*deŋ* 'remove, pull out'	^{n}*daŋ* 'remove, take out'	*ndiŋ* 'release, empty'	–	–
–	*maʔuri 'alive'	*maur*	*maur*	*mowoʁ* 'be alive'	*mewur*	*mewur*

Table 65: Final continuant codas after a voiced onset generally remain final codas in Ninde.

The semantics of Ninde *ndiŋ* 'empty (a pool of liquid)' could be influenced by an unattested cognate of Neverver ^{n}*daŋ* 'breach (a dam)'. Generally, no final vowel is inserted if voicing persists throughout the entire syllable.

If the final syllable has a voiceless onset and a continuant coda, a vowel is usually added to the end. This vowel is usually /A/ (which is generally raised word-finally) or /e/ (which remains [e] before a suffix or enclitic) (Table 66).

Reconstructions		NCM			SWB	
POc	**PNCV**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
(*ŋaŋa ‘bitter, poisonous’	*kona ‘bitter’	*kːon* ‘taste bad’	*kon*	*ko**ʔon**e̲*	*kon**kon*** ‘bitter, sour’	*kon**kon***
*zalan ‘path, road, way’	*sala	*asal*	*nesal*	*na**l̪ˤ**A̲*	*na**hal***	*na**hal***
(*kau ‘catch, as with a hook’)	(*gau ‘catch with hook’)	*sil* ‘fish at night by torchlight’	–	***sl**e̲* ‘hook’	–	***sel*** ‘hook’
(PEMP*ikan ‘fish’)	*ika ‘fish’	–	–	*nimi**jal̪ˤ**A̲* ‘fish’	*ma**hal***	*ma**hal***

Table 66: Final vowel insertion (underlined) after a continuant coda preceded by a voiceless onset within the same syllable (bolded).

This would suggest that a formerly devoiced final liquid or nasal was reinterpreted as the onset of a reduced final syllable and restored, whereas a fully voiced final syllable was clearly perceived as final.

When the vowel of the final syllable is a high front vowel with a continuant and/or coronal coda, /e/ and specifically not /A/ is inserted (Table 67).

Reconstructions		NCM				SWB
POc	PNCV	Neverver	Avava	Ninde	Naati	Nahavaq
*kuRita 'octopus'	*kuRita	*noɣow**it***	*ko**it***	*nuŋguw**ute***	*noʔo͡ij**it***	*noʔoj**it***
*sipi, pwirip 'parrot sp.', *siri 'parakeet'	*siviri 'lorikeet'	–	*siβir*	*nesvaʁa*	*nesiβir*	*nisiβ**ir***
*turu, *tuqur 'stand',	*tuʔu-ru	(*tur*)	(*tur*)	*taʁ**se***	*ta:r**is***	*(ti)tar**is***
(*qumun- 'roast in an earth oven')	*bulu 'oven'	–	*iβleh*	*nal**se***	*neβul**ys***	*niβul**us***

Table 67: Vowel insertion after a final coronal preceded by a high front vowel.

Though Ninde *nuŋguwute* 'octopus' now has a rounded vowel in the syllable /wu/, comparison with cognates reveals that it was previously *i (and the inserted nasal points to an NCM origin). PSWB*taris 'stand' > *taʁse* and *nevulys 'earth oven' > *nalse* also undergo syncope of the conditioning vowel, whereas exceptions like PSWB*viʔis 'turn, change' > *vəkəs* and *teʔis 'yank' > *təkəs* do not undergo either change. In these cases, the vowel insertion could follow syncope of high vowels in the unstressed *-is suffix (with harmonizing allomorphy), blocked in some words by a phonotactic constraint on complex codas with a glottal stop. Alternatively, the velar reflex in Ninde, which conditions [ə], could apply first in bleeding order, removing the front-vowel context for /e/ insertion. The first explanation cannot simply be framed in terms of natural classes (coronal consonants following a high front vowel), and the latter explanation would leave important exceptions, like first-person inclusive plural pronoun PNCM*git > *kəte*. As a resolution, NCM-origin words are subject to the high-front-vowel context and SWB-origin words are subject to a change repairing complex final codas.

After a high back rounded vowel *u followed by a continuant and additionally after word-final labialized bilabials, [o] is inserted (Table 68):

Reconstructions		NCM			SWB	
POc	**PNCV**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
(*bituŋ 'type of bamboo')	*bue 'bamboo'	*ni/βinᵐʙu*	*βʷiniᵐʙuh*	*na/ nambuwo*	*n/e:mbu*	*na/mbʷu*
*qusan 'rain'	*ʔusa	*naus*	*auh*	*nuwo*	*nuwuh*	*newuh*
*pwiRa 'elephantiasis'	*bura	*ᵐʙur* 'swell'	*ᵐʙur* 'swell'	*mboʁo* 'swell'	(*timb* 'swell')	(*timb* 'swell')
*silak 'shine'	*sulu 'shine'	*susul* 'shine, glow'	–	*jejol̪ˤo* ~ *jol̪ˤo/jol̪ˤo*	–	*sulsul*
(*qusan 'to rain')	(*ʔusa)	*βoβ*	*βop*	*wuwo* 'rain'	*wop*	*βʷupʷ~βopʷ*
(PWMP *sayaw 'dance')	*savwa	*saβ*	*sap*	*jawo*	*hapʷ*	*hap*
*ñamu 'chew'	*zamwa	*ze(m,mw)*	*sem*	*samo* ~ *samwe*	–	*semw*

*Table 68: Final vowel insertion after labialized bilabials (phonemically rounded) and continuants following *u (probably allophonically rounded).*

This results in a constraint on final rounded consonants, whether they are allophonically or phonemically rounded. The final /h/ of Avava *βʷiniᵐʙuh* 'bamboo' is unexplained diachronically, but it corresponds to a context for final vowel insertion in Ninde. Ninde *jejol̪ˤo* and *jol̪ˤojol̪ˤo* 'shine, glow', with its pharyngealized lateral approximants, may alternatively be explained in terms of final voiceless continuants (considering /l̪ˤ/ to have developed from a voiceless allophone of *l) as above. In fact, the vowel insertion affecting continuants is generally consistent with vowel insertion after final devoicing.

A final *u before word-final *t or *ŋ may have been allophonically fronted, conditioning the insertion of a final /e/, but exceptions abound in this context (Table 69).

Reconstructions		**NCM**				**SWB**
POc	**PNCV**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
*kalapo 'rat, mouse'	(*karivi, *kasuve 'rat')	(*ni^{m}bis^{m}box*)	(^{m}*b*^{w}*isi*^{m}*bok*)	*noḻˁombute*	*nelembut*	*nalamb*^{w}*ut*
(*puna 'base of a tree')	*boto 'bottom'	*nuɣut-n* 'trunk'	*wut-n* 'trunk'	*nuŋgute-ne* 'trunk'	(*mbwati-n*)	–
*qumata, *mwata 'snake'	*mwata	*nemat*	*am*^{w}*at*	*namate*	*nøm*^{w}*at*	*nam*^{w}*at*
(*maRuqane 'male')	(*ata-mwaʔane 'man')	*xaβut* 'husband'	*it* 'husband'	*kaiwut* 'old man, husband'	–	*ʔajβ*^{w}*ut* 'husband'
(*maRuqane 'male')	(*ata-mwaʔane 'man', *ʔata 'person')	*nimɣut*	*mu:t*	*nym(y)ŋgut*	(*nemurut*)	(*m*^{w}*or*)
*ŋutu 'louse'	*kutu	*noɣut*	*aut*	*nuŋgut*	*noʔut*	*naʔut*
– 'mackerel'	–	cf. Neve'ei	*na*^{m}*buŋ*	*nəmbəŋe*	*nimbuŋ*	*nemb*^{w}*uŋ*
(PWMP*liqu 'forget')	(*lolo)	*set/βun*	*(itnen)* ^{m}*ʙu/leŋi*	*su/mbu(ŋe)*	*si/mbuŋ*	*sin/mb*^{w}*uŋ*

*Table 69: Cognate sets with indeterminate final vowel insertion patterns after final syllables with labiovelar onsets and *t or *ŋ in the coda.*

Final vowels after *t and *s are never rounded in Ninde, even when they are inserted in an environment after a *t coda of a rounded vowel. In a sense, an inserted vowel, and specifically unchanging /e/, could resolve the conflicting environments for vowel allophones – preceding labiovelars condition a back vowel, but /t/ in the coda conditions a front vowel. The addition of the vowel /e/ possibly allowed for the newly penultimate vowel to be clearly perceived as rounded and unaffected by the fronting of the following /t/ or /s/. It is probably the case that high rounded vowels had back-vowel allophones if adjacent to a labialized bilabial (labiovelar) consonant and were fronted elsewhere (including environments with plain bilabial consonants). Forms with closer matches in SWB cognates have both final /t/ and inserted vowels: *kaiwut* 'old man, husband' and *nol̪ˤombute* 'rat' (but cf. Neve'ei *nela^{m}but*); the same applies to forms matching NCM cognates: *nuŋgut* 'louse' and *nuŋgute-* 'trunk'. The models reflect an inserted /e/ after *t for both origins, after a rounded vowel, but a SWB-specific insertion of /e/ after *uŋ.

Finally, vowels are inserted if the final consonant is *(ŋ)g with few exceptions (Table 70).

Reconstructions		NCM			SWB	
POc	**PNCV**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
–	*baiga ‘green snail’	–	*i^mbiŋ*	*nəmbəŋge*	*mbe:ŋg*	–
*bulaka ‘swamp taro’	*buaga	*ni^mʙuaŋ*	*^mʙuaŋ*	*nembijaŋgA*	*nembʷiaŋg*	*(ni-)mbujaʔ*
(*pau ‘*Kleinhovia hospita*’)	(*matala)	*nemiⁿdaŋ*	*miⁿdaŋ*	*nəmndaŋgA*	*nemindaŋg*	–
*waga ‘outrigger canoe’	*waga	*nuwaŋ*	*awaŋ*	*nowoŋgo*	*nuwaŋg*	*nawaŋ(g)*

*Table 70: Reflexes of word-final *g.*

These inserted vowels are /e/ if the final vowel was *i or *u, /o/ if it was *o, and /A/ if it was anything else. Comparison of cognate sets shows that they are frequently lost in other languages, except in words suffixed with first-person singular possessive *-g. In some cases, these changes could represent homophony avoidance with possessed forms. The remaining word-final /ŋg/ forms in Ninde cannot be linked to any cognates.

Many of the remaining exceptions have reduplicated forms and vowel insertion can be predicted on the basis of whether a vowel is inserted between the reduplicant and base (Table 71).

Reconstructions		NCM			SWB	
POc	**PNCV**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
–	–	*βaⁿɽβaⁿɽ* ‘rub noisily’	–	*waʁa* ‘wipe dry’	***mbwar-(a-*** *mbwar)* ‘rub’	*mbwar* ‘wipe’
–	(*vono-ti ‘block’)	*kor* ~ *ⁿgor*	*kor(or)*	***ŋgoʁo***- *(ŋgoʁo)*	–	*ŋgor*
–	(*ago-ago ‘yellow’)	(*jaŋjaŋ* ‘yellow’)	*lit*	*lit_* ‘yellow’	***lyt****lyt*	***lut****lut*
(*kalu ‘stir’), *kiri(k) ‘tickle’	*gali ‘stir, rub, wipe’	?***xer****xer* ‘stir’	–	***kit****kit_* ‘tickle, rinse’	–	***kit****kit* ‘stir’

Table 71: Exceptions to change modeled on a reduplicated form.

If it is the case that some basic roots are reformed on the model of the reduplicant, then this suggests the insertion of a final vowel depends on whether there are phonotactic limitations on a root’s last consonant followed by its first. These exceptions include both instances of expected vowel insertion (*lite* is expected for ‘yellow’, but it is *lit*) and unexpected vowel insertion (*waʁa* ‘wipe, towel dry’ is fully voiced and expected to be *waʁ*). A pair of simple and reduplicated variants is not always available, and at any rate, cannot be implemented by the software as exceptionless change. To generalize, a number of these forms begin with a bilabial consonant and end with a coronal consonant. Such insertion in reduplicated forms are part of the computationally implemented model.

Final *u inherited from both sources underwent deletion (and syncope), but only after a preceding front vowel was rounded (Table 72).

Reconstructions		NCM			SWB	
POc	**PNCV**	**Neverver**	**Avava**	**Ninde**	**Naati**	**Nahavaq**
(PMP*niuR ‘coconut’), PCEMP*matu ‘dry coconut’	*matu-i, (*niu) ‘coconut’	(*nani*)	(*[n]ani*)	*nimyt*	*me:tu*	*metu*
*pudi ‘banana’	*vudi	*naβu*^{n}*s*	*apm*	*nys, nesu-*	(*neːŋgit*)	*n(i,u)ŋgut*
–	*madou ‘thirsty’	^{ɳ}*ɽom*	^{ɳ}*ɽim*^{ɳ}*ɽim*	*mar*	*maɳɽu*	–
(*maRuqane ‘male’)	(*ata-mwaʔane ‘man’, *ʔata ‘person’)	*nimɣut*	*mu:t*	*nym(y)ŋgut*	*(nemurut)*	*(m*^{w}*or)*

*Table 72: Fronting of pre-Ninde *e, including before a deleted final /u/.*

Of this set, the vowel deletion is clearest in PSWB*metu ‘coconut’ > *ni-myt*, with the same rounding effect of intermediate *e observable in PNCM*navuz ‘banana’ > *nevus > *nys*, *nemukut ‘man’ > *nym(y)ŋgut*, which was attested at the turn of the 19th century as [nemugut]. In light of this vowel deletion and the fact that some cases of final *u “become” /A/ in Ninde, these patterns do not support the analysis (Lynch 2014) that Ninde uniquely retained final vowels, participating selectively in areally diffused sound changes.

Remaining changes

The remaining sound changes were generally restricted to very few cognate sets, sometimes as few as one. The sound changes gleaned from comparison were ordered and reordered to maximize the success rate of the code. The final models are included with examples in the Appendix.

6.4 Results

The success of sound change models was determined using two metrics: reduction of partial cognate sets into larger, combined sets; and the convergence of proto forms from within NCM and SWB groups. As an added result, the lexicon of Ninde can be identified in terms of parentage for individual forms, expressed as a ratio.

Out of a sample of 300 random forms, 147 had enough data to attempt reconstruction from more than one language, and an additional 99 were initially singletons. Of those, only eight sets failed to converge, meaning that the sets of possible proto forms did not have intersecting forms. The number of proto forms reconstructed per set was high (m = 17.97, stdev = 103.86), but the median was low (2). These values are reduced when all three NCM languages are attested (mean = 6.0, median = 2, stdev = 8.55). In many of these cases, if Ninde is deemed a fit, then it constrains the possible proto forms of the set.

The high values are due to mainly one outlier: one set that generated 1048 possible proto forms (Avava *prura*, Neve'ei *viviri*, Ninde *ʙoi* ‘spit’) These particular sets have particularly high possible proto counts because the forms in particular generate many possible proto forms, and the forms are only partially cognate. With multiple syllables, there are many locations for deleted segments and possible vowel harmony could obliterate once distinct vowel qualities.

For SWB languages, there were 127 successful protos reconstructed and 28 failures to converge. The mean number of reconstructed forms was similarly high (m = 18.70, stdev = 62.22) and the median number was 2. For sets with data from all three languages, the data were much more likely to reach a small number of possible proto word forms (m = 1.51, stdev = 0.77, median = 1). This may reflect a shallower time depth in the SWB languages.

Additionally, 99 forms were reconstructed from singletons from NCM languages and 56 from the SWB languages. Many of these are idiomatic constructions, undetermined species of flora and fauna, or simply rarer words not available in the source data.

6.5 Conclusions

When Ninde’s lexicon is split into two subsets, it can contribute to the reconstruction of two separate proto languages representing each source in the mixing process: NCM and SWB. Moreover, the implementation of a computational approach deploying a chronologically ordered series of sound changes has some success in generating possible proto forms, identifying expected reflexes, and merging partial cognate sets that were not identified in the manual data entry. This process has highlighted complex correspondences involving in particular coronal, velar, and glottal consonants, as well as vowels in word-final contexts. What appear to be exceptions, in many cases, are consequences of limited representation of affix boundaries, which represent word-internal contexts that are dependent on the overall morphology of the word. Additionally, the potential influence of word class on phonetic context for bound roots, as well as analogy afforded by simple and reduplicated alternants, are not implemented in such a model. Nevertheless, a cycle of computational processing and human model tweaking has significantly assisted the recognition of these factors.

Even if the language families identified (SWB and NCM) do not stand the test of time, this work may precipitate that discovery. The models of ordered sound changes ultimately generate proto forms with accurate reflexes. Inaccuracies in the phylogenetic model that was used can easily be falsified by either identifying alternative sound changes and diachronic ordering that result in a more accurate model, or by demonstrating that other languages participate in a sequence of those changes.

The failures to reconstruct proto forms do not uniformly represent inaccuracies of the model. There are many reasons to expect exceptions in the correspondence sets: variation not documented, borrowing (as distinct from mixing), analogy, and metathesis. All of these have been exemplified in this dissertation at some point, and many of those examples were identified through the limits of the model.

Nevertheless, the data are too limited in their present form to guarantee accurate reconstruction. With no representation of stress, any sound changes that interacted with stress are missing. Although there was information about word class and inflectional classes, it was not always used by the model. The forward-reconstructed forms may offer a means to probe missing or inconsistent phonetic transcription, and so enhance lexicography in a region that has been underserved.

7 Lexical Systems in Diachrony

Beyond the questions addressed in §6 about whether the comparative method is a reliable tool for assessing the relatedness and clade structure of groups of languages, there is robust evidence of genealogical relationships in the form of lexical systems. Lexical contributions from multiple diachronic sources are briefly addressed in §5.3, but lexical systems are subsets of the lexicon that form taxonomies of some semantic domain. If they are ordered sets, they may be susceptible to list effect – preferential analogy with forms that are nearby in recited sequences; there may be multiple dimensions of synonymy and antonymy, such that forms may become polysemous, supplanting adjacent "cells" by leveling semantic contrasts. Across lines of symmetry, parallel developments often affect lexical forms that do similar work. In the context of this dissertation, Nince's lexical systems lend support to a creative and conscious mixing process, whereby word forms and taxonomies appear impossibly blended when compared to other Malekula languages.

Faced with high rates of cross-language diffusion, great intra-language variation, and scant documentation, adequate data may not exist for comparison of words on phonetic grounds alone; cognates may simply be lacking for comparison. In this chapter, I discuss five lexical systems in Ninde with a comparative perspective and evaluate potential sources of innovation.

The diachronic picture is different for each domain. I argue that the kinship system resembles those of North-Central Malekula (NCM) as much as other South West Bay (SWB) languages, and that lateral gender distinctions (i.e., for siblings and cousins) show recent innovations in Ninde largely shared with innovations in Neve'ei. Cardinal directions are largely stable, but are localized to the topography of each location – nonetheless, there are systematic changes affecting this domain. The calendrical terms are much more limited, but Ninde shows strong similarity to the patterns in North-Central Malekula, but there is no calendrical information to compare with other SWB languages. The number system shows signs of more recent restructuring with highly varied influences. This aligns with scholarship that has characterized the shift to quinary number systems as a recent diffusion across Malekula, replacing most of the decimal systems that previously existed. Finally, the pronoun system of Ninde is in many ways an innovation-rich mix of NCM and SWB sources.

In reconstructing the lexicon of Proto NCM, I consider that proto words had not only proto forms, but also proto meanings. Those proto words were parts of syntactic constructions and had properties like word class, or they belonged to specific constructions. The design of the lexical systems would have influenced how words and meanings were borrowed. For example, the oppositions between months in indigenous calendars were similar to the ones in the Gregorian calendar, so it is easy to borrow December as a conceptual unit with an indigenous name. The relationships words have to each other within a lexical system can also influence how words are altered on analogy with each other, and also how the system can extend to new parts of the semantic space.

The lexicon on the whole can be thought to function as a system of Saussurean oppositions, with each linguistic signifier having the semantic value that it has by virtue of having opposing words that allow it to carve up the conceptual world. I consider these lexical systems to be domains that together carve up restricted semantic domains by creating some taxonomy. Taxonomic sisters in such systems might have parallel paradigms (or not), and they may be more susceptible to reanalysis that generalizes properties of the morphology. In domains that involve items that can be associated with one another in a linear fashion, like numbers or months of a week, it may be common practice to list the members of a system, even if only in child-directed usage. There may be list effects that cause morphophonological change to propagate between adjacent members. Dyen and Aberle (1974) reconstruct a Proto Athabaskan kinship system using the categories that exist in each of the daughter languages. In some cases, it is possible to identify kin terms that existed in the proto language, but the actual forms are not always cognate across languages. Moreover, social factors appear to account for the statistical patterns in changes to kin terms: polygyny, double descent, cross-cousin marriage, and non-neolocal marriage.

As a diagnostic for genealogical relationships, lexical systems appear to reaffirm Ninde's status as a mixed language. In some systems (kinship, calendrical terms, and to some extent, pronouns), semantic contrasts mirror the patterns of one source family but resemble forms of the other source family. There is a case for a core NCM semantic structure with SWB calques, but in some cases the semantics are from SWB and the forms from NCM.

Perhaps more important than any theoretical importance, these lexical systems are often important technological innovations for a culture. They are often the parts of a lexicon that are explicitly taught to children. An understanding of the parts of such a system can facilitate, and often reflect, cultural norms and practices relating to the physical and spiritual worlds. For example, kinship might track intergenerational relationships, or important intra-generational age and gender differences. It is useful for learnability and for future applications of tradition-based lexical systems to analyze how the system has been organized diachronically, and this is what has really motivated most of the contents of this chapter.

7.1 Kinship Systems

Ninde has a rather complex kinship lexicon in several regards. Words in this domain are the most morphologically complex nouns in Ninde. This is owing to the fact that first-person possession may require a different root from second- and third-person possessed forms; the two roots may be different yet from vocative forms. These complex forms classify some kin by generation, their place in a birth order relative to ego or one's parents, and their gender with respect to either the gender of ego or the parent through which the relative is related. Several affixes are unique to this domain.

With up to three persons for possessed forms and a vocative, the already complex classification is made still more complex in its grammar. Vocative forms are usually of the same form as the first-person possessed form, but often suffixed with a first-person possessive *-(a)ŋg* or (counter-intuitively) with a second-person possessive *-(a)m*. Possessed forms for the other persons are usually suffixed with a unique (to Ninde and to kinship) *-woʁ*, which will be discussed at length in this section. However, this complexity strengthens the case for mixing, as it is dual inheritance that has allowed for the multitude of word forms.

Ninde's kinship system, like that of Melanesian languages generally (Hocart 1937), is by and large classificatory. The basic forms corresponding to 'mother', 'father', and 'child' are broad in usage, and do not identify specific individuals in a parent-child relationship, but classify entire generations of parents and their siblings. Gender is generally distinguished in the root forms in other Malekula languages (but not as much in Ninde). Like other Malekula languages, the root also encodes relative birth order of siblings. Most Malekula languages differentiate same-sex from cross-sex siblings, and occasionally have an additional word that does not encode a specific gender of ego and/or kin. Such general terms for sibling are very commonly also extended to 'friend' or 'relative'.

Diachronically, the kinship system of Ninde reflects mostly NCM forms and categories for vertical kinship relationships (related through direct descent) and SWB forms combined on NCM models for aunts and uncles. This could reflect the fact that the same terms are used for aunts and uncles by marriage. While this pattern alone would suggest that a core lineage is fundamentally NCM in origin, with heavy contact in the home with SWB languages, the generalization breaks down when considering vocative kinship terms. Ninde has taken most of the vocative forms, even for vertical relationships, from the SWB source. In many ways, this makes the kinship system emblematic of the mixed nature of Ninde.

7.1.1 Siblings

Ninde forms for siblings have been generalized for birth order across genders, just like they have been in Neve'ei (Table 73). These basic roots are *tuwaʔ* 'older sibling' and *tisu-* 'younger sibling' (only the second is a bound root). These correspond in cognacy and semantics to Neve'ei *tuʔa-* and *tas-*, but cognates are present in both Nahavaq (*tuʔa-* and *tesu-*) and Neverver (*tas(t)n*), where they are more specifically male. Avava would seem to have undergone generalization along the other axis, generalizing gendered sibling terms across birth order, with a special term for same-sex siblings (or brothers specifically, but usage for same-sex sisters is not available for analysis). These are *vivini-* 'sister', *manu-* 'brother (of a female)', and *sala-* 'brother (of a male)', which also have cognates in both Nahavaq and Neverver. While it is possible that Neverver represents the more conservative classification for NCM languages and Neve'ei a similar model for the system in Ninde, *tisu-* is the expected corresponding form only to Nahavaq *tesu-*, even if both the SWB and NCM sources would have yielded very similar forms in Ninde.

Language			Ninde		Neve'ei	Neverver	Avava	Naman	Nahavaq	Naati
Kin	Age	Ego	**Voc.**	**Ref.**						
Sister	Older	of male	*tuwaʔ*	*tuwaʔ-lepe-woʁ* ~*nele-lepe-woʁ*	*tuʔa-*	*vivin*	*vivin(i-)*	*vəvən, mox amu* (first born)	*vene-* ~*ni-(vi)-venenjen*	*vene-*
		of female						*tuxan, mox amu* (first born)	*hala-*	
	Younger	of male	*tisu-ŋg*	*tisu-lepe-woʁ*	*tas-*	*vivin, nebat* (last born)	*vivin(i-)*	*vəvən, mox atox* (last born)	*vene-*	
		of female				*vivin, nebat* (last born)		*tuxan, naarwon, mox atox* (last born)	*hala-, tesu-*	

<table>
<tr><th colspan="3">Language</th><th colspan="2">Ninde</th><th rowspan="2">Neve’ei</th><th rowspan="2">Neverver</th><th rowspan="2">Avava</th><th rowspan="2">Naman</th><th rowspan="2">Nahavaq</th><th rowspan="2">Naati</th></tr>
<tr><th>Kin</th><th>Age</th><th>Ego</th><th>Voc.</th><th>Ref.</th></tr>
<tr><td rowspan="3">Brother</td><td rowspan="2">Older</td><td>of male</td><td rowspan="2">tuwaʔ</td><td rowspan="2">nel-na-woʁ,
~tuwaʔa-ne-woʁ</td><td rowspan="2">tuʔa-</td><td>neman, tatan</td><td>sala-n</td><td>xavən, mox amu (first born)</td><td>tuʔa-, wowow (first born)</td><td rowspan="3">hala-</td></tr>
<tr><td>of female</td><td>neman, tatan</td><td>manu-</td><td>man, mox amu (first born)</td><td>tuʔa-, wowow (first born), m^{w}ene-</td></tr>
<tr><td>Younger</td><td>of male</td><td>tisu-ŋg</td><td>tisu-woʁ</td><td>tas-</td><td>neman, tas(t)n, nebat (last born)</td><td>sala-n</td><td>xavən, mox atox (last born)</td><td>hala-, tesu-</td></tr>
</table>

Language			Ninde		Neve'ei	Neverver	Avava	Naman	Nahavaq	Naati
Kin	Age	Ego	**Voc.**	**Ref.**						
		of female				*neman, tas(t)n, nebat* (last born)	*man(u-)*	*man, mox atox* (last born)	*mʷene-, tesu-*	
		Generic		*kala-ne* ('sibling, friend)	*xava-* ('sibling /friend')	*niar* 'sibling, (fence)'		*niaŋgan* ('same-parent sibling')	*hala-* 'same-sex sibling'	

Table 73: Sibling terms across NCM and SWB languages.

While kinship has been explicitly analyzed for Neverver, Ninde, and partially for Avava and Nahavaq (mainly in terms of grammatical properties), some properties of the kinship system presented here are inferred from glossing. The fact that Neve'ei appears to have a collapsed form for older siblings and younger siblings, regardless of the genders of ego and kin, is not explicit in the grammar by Musgrave (2007), but strongly suggested by the glosses. Additionally, examples of usage (Musgrave 2007:36) also show first-person possessed forms used vocatively, just as a few of the Ninde vocatives are formed. On this basis, Ninde and Neve'ei seem to be the most similar, but perhaps only superficially so.

7.1.2 Vertical relationships

Kinship terms for people related through direct descent show a split between vocative and possessed forms (Table 74). The Ninde term *nesu-* 'grandmother' is innovated and has no cognates, unless it is from NCM*nasik 'kingfisher' (*nambolmije* 'kingfisher' is often an epithet for chiefs, as well as the name for one of the support beams in a house). Otherwise, each form for two generations older and younger than ego have a third-person possessive form cognate with a form in the NCM+ languages, and more specifically Neve'ei. Both have *teme-* 'father' and share Ninde *sa-na-* 'mother' and Neve'ei *səne-* 'mother', which are clearly more distantly cognate with Nahavaq *(ni)hine-* 'mother'. Otherwise, Ninde vocatives have cognates in SWB languages: *tatai* 'father' and *pap* 'father, uncle' show regular correspondence with Nahavaq equivalents. These vocative-possessed paradigms form the basis for terms for aunts and uncles, greatly inflating the presence of vocative SWB forms throughout the kinship system.

	Neve'ei **[Naman]**	**Avava**	**Neverver**	**Ninde**	**Naati**	**Nahavaq**
Mother	*naŋ* (voc.) *sǝne-* (others) [–]	*galu* (voc.) *galu te* (inal.)	*nida*	*naijei* (voc.) *sa-na-* (others) *mama* (neologism)	*amoʔ*	*amoʔ* *(ni)hine-*
Father	*mam^{w}e* (voc.) *teme-* (others) [*tate* (voc.) tavǝm (2nd person) teme- (others)]	*mom* (voc.) *mom de* (inal.)	*tata* (voc.) *mama* (others)	*tatai/pap* (voc.) *tam-* (voc. and 2nd person) *teme-* (others)	*tatai*	*tataj* (voc.) *papap,* *apap,* *pap*

	Neve'ei [Naman]	**Avava**	**Neverver**	**Ninde**	**Naati**	**Nahavaq**
Child	*nat-* [*netən, netite*]	*atda*-n *vanat-n*	*niterix*	*net-*	*nembunoŋ*	*mbʷunoŋ*
Grandmother	*bʷetah* [*bubu, zəbə tevet*]	*ʙuah*	*niʙua*	*nesum* (voc.) *nesu-* (others)	*avu (kakap)*	*avʷu-kakapʷ*
Grandfather	*tabu-* [*bubu, zəbə-*]	*ʙum*	*ʙuʙu*	*mbumbum/ təmbək* (voc.) *timbu-* (others)	*avu (tot)*	*avʷu-tot*
Grandchild	*naʔaibu-* [*na:bə-*]	*eʙu-*	*naxaʙun*	*neimbu-*	–	*mʷaʔaimbu-*

Table 74: Vertical kinship terms (involving descent relationships) in NCM and SWB languages. Coloring reflects cells with cognate or partially cognate forms.

Ninde generally has a greater proportion of SWB forms in terms for 'aunt' and 'uncle', but they are combined to form a classificatory scheme most like that of Neverver (Table 75). Specifically, aunts and uncles are identified by their gender, birth order relative to one's parent, and the gender of the related parent for all aunts and uncles in this group. At one extreme is Neve'ei, which classifies uncles by age only for one's father's brothers. Avava is intermediate, classifying at least the relative ages of both parents' same-sex siblings, but there is no coverage of paternal aunts.

Most of the Ninde forms here could be called SWB calques on the model of NCM forms, using semantically equivalent modifiers 'small' and 'big' to derive kinship terms for aunts and uncles from *nai(jei)* 'mother (voc.)', *sa-na-* 'mother', *pap* 'father (voc.)', *tam* 'your father', and *teme-ne-* 'father'. In Nahavaq (there is no vocabulary available for Naati), there is no differentiation by relative age of one's parents' siblings; all aunts are apparently simply *av^{w}ejv^{w}ej* (most likely partially cognate with the *-woi* and *-poi* found in terms for Ninde paternal aunts), but a distinction is made between paternal and maternal uncles. Nonetheless, every Nahavaq form has some partial cognate in Ninde, whereas many of the NCM forms do not have known cognates in Ninde: the common *bibi* 'uncle', Avava *dat* 'uncle', or any of the base forms meaning 'mother, aunt' (*naŋ*, *galu*, or *nida*).

Kin			Neve’ei/ [Naman]	Avava	Neverver	Ninde	Nahavaq
Mother’s	**older**	**sister**	*naŋ*	*gal(u)-lam*	*nida* ***toxtox***	*nai-****l̪ˤapʷe*** (voc.) *sa-m-/nai-*+***l̪ˤapʷe*** (2nd pers.) *sa-na-* + ***tlepe***	*avʷejvʷej*
		brother	*bibi* [*bibi* (voc.)] *[bəne-]*	*dat*	*bibi* ***toxtox***	*mituwa-*	*papap/* *apap/pap* *mituwa-*
	younger	**sister**	*naŋ*	*gal(u)-****lele***	*nida* ***lele***	*nai-****lou*** (voc.) *sa-*POSS*-/ nai-* + ***tloulou***	*avʷejvʷej*
		brother	*bibi* [*bibi* (voc.)] [*bəne-*]	*dat*	*bibi* ***lele***	*mituwa-*	*papap/* *apap/pap* *mituwa-*

Kin			Neve’ei/ [Naman]	Avava	Neverver	Ninde	Nahavaq
Father’s	older	brother	*mamʷi-**lam*** [*matər **varəx***]	*mom-**lam***	*mama **toxtox***	*teme-* (3rd pers.) *tam-* + ***l̪ˁapʷe*** (others)	*taj-**vʷariʔ**,* *p/a/pap*
		sister	– [*nane **varəx***]	–	*nida **toxtox***	*sa-poi* + ***tlepe*** (voc. and 2nd person) *sa-* + *-woi* + ***tlepe***	*avʷejvʷej*
	younger	brother	*mamʷi-**ŋaʔ*** [*matər **varəx***]	*mom-**lele***	*mama **lele***	*pap-**lou*** (voc.) *tam-* + ***tloulou*** (2nd pers.) *teme-* + ***tloulou*** (all others)	*taj-**vʷariʔ**,* *p/a/pap*
		sister	– [*nane **varəx***]	–	*nida **lele***	*sa-poi* + ***tloulou*** (voc./2nd pers.) *sa-* + *-woi* + ***tloulou***	*avʷejvʷej*

Table 75: Terms for aunts and uncles in NCM and SWB languages, framed in terms of age relative to a parent. Bolded forms mean ‘small’ or ‘big’ in each language and usually encode whether the relative is younger or older than the parents of ego, but varəx *and* vʷariʔ ‘small’ *in Naman and Nahavaq show no such pattern. None of these forms are attested for Naati.*

These terms for aunt and uncle are expressed in terms of aunts and uncles related by descent, but the same terms are used for aunts and uncles by marriage. When a relation is tracked through the maternal side, but related through marriage to an aunt or uncle, the term used is the same as what would be used if the relation were a sibling of the father, and vice versa (paternal in-laws have the same terms as maternal aunts and uncles). This creates a tempting parallel between linguistic descent and familial descent – if the linguistic influence of SWB languages is greatest for kinship terms more affected by marriage, then could an identity with the cultures associated with NCM languages have been historically more important? Considering traditions of patrilineal descent, are SWB cognates like vocative *tatai* and *pap* 'father' more representative of a core than NCM-aligned referential forms like *tamaŋg* 'my father'?

In summary, Ninde shares numerous innovations with Neve'ei, which has proven difficult to include in the NCM group proper on the basis of shared innovations. Both of these marginal NCM languages have lost gender distinctions in lateral kin terms (siblings), at least in the morphological roots. This leveling of the paradigm eliminated same-sex and cross-sex distinction, but unlike Naati, it did not result in gendered roots. (It is also likely that the terms of Naati were only recorded as would be used by a man, since Crowley worked with one male speaker, Aiar Rantes.) Neve'ei also forms the kinship terms for parents' siblings with similar modifiers *lam* 'big' and *-ŋaʔ* 'small' (Musgrave 2007:39). The form for 'small' is not cognate with Ninde's *-lou*, but it is noteworthy that they share a unique property: they are always reduplicated except in these kinship terms. If anything, kinship points to the possibility of retention of a system that predates NCM, with the system of NCM likely obfuscated by systematic leveling of kinship distinctions.

Ninde innovations from relativized verbs

There is ample evidence that Ninde expanded its inventory of kinship terms by extending morphology from verbal sources. The properties that make Ninde's kinship terms resemble verbs are two suffixes: *-woʁ*, which is homophonous with the verb *woʁ* 'follow along', which as a serialized final verb (V2) has the sense of 'along (something long), about (a topic)'; and *-woi* or *-poi*, which pertains to the relationship between a paternal aunt and fraternal niece/nephew and could be realis and irrealis (respectively) V2 forms of a verb *poipoi* 'raise (a child)'. At least in kinship terms, the unreduplicated form is clearly cognate with Nahavaq *av^{w}ejv^{w}ej* 'aunt', which refers to both maternal and paternal aunts. A series of reanalyses of verb forms could explain Ninde's singularly complex kinship morphology.

Though the function and origin of *-woʁ* are not obvious, there appears to be no source in noun morphology. No source on a Malekula language describes any reasonably cognate suffixes. The suffix also has a peculiar position in the noun: it occurs after bound possessive markers, which otherwise (in Ninde and Malekula at large) occupy the last morpheme slot in the noun.

There are no apparent sources for *woʁ* outside of the verbal senses, but it is further extended from kinship nouns to alienable nouns. The words for ‘wife’ and ‘child (offspring)’ can polysemously be used inalienably. The first, *watawoʁ* ‘woman, wife’, shares a frozen bound root with *watevilaʔ* ‘female in-law’, and *netnowoʁ* ‘child’ can also be used without a possessor, despite sharing the *-woʁ* that is unique to kinship terms. We also recorded usage of a variant *Atuwoʁ* ‘God’, which is a documented borrowing via European missionaries of the Samoan *Atua* ‘God’. In Ninde, the more common form is *Atuwo*, with predictable rounding of the final vowel. The reanalysis of this final /wo/ demonstrates the pervasiveness of this suffix in the domain of kinship and adjacent domains.

Possible routes from verbal PNCM *sur or PSWB *hur ‘follow along’ to nominal kinship affix exist via (1) noun-verb compounds and/or (2) nouns of verbal origin. One possibility is that *-woʁ* was initially a verbal element in *pusuwoʁ* ‘son, boy’, which would originally be morphologically complex and clausal in nature:

pu-su-woʁ

IRR-come.back-follow

‘(one who) will come back to’

(A less likely alternative is that the word is from *wusuwoʁ* ‘ask’, which has a homophonous irrealis form *pusuwoʁ*.) This could have existed as a headless relative (but it is troubling that this form lacks a highly productive relativizing *t(i)-*), or as a truncated noun phrase (‘the child that will come back'). Like Neve'ei *natiti* 'child', the missing head of a relative clause could have been reanalyzed from an unpossessed *nati* 'child' with relativized modifiers relativized with *ti-*. Neve'ei also has compounds *natiti nemʷen* 'boy (lit.: male child)' and *natiti ŋaŋaʔ* 'baby (lit.: small child)'. Crucially, this hypothesis relies on the assumptions that the nominal object of transitive *woʁ* 'follow', which adds a locative complement to verbs of motion, was reanalyzed as a possessor, in order for the form to be reanalyzed as a viable noun in its headless form.

Alternatively, the *-woʁ* suffix may have origins in a repetitive suffix. Neve'ei *sur* 'follow along' and V'ënen Taut *-ur* 'follow along', both of which have a V2 sense 'repetitively, over and over', have significant overlap with the semantic range of Ninde cognate *woʁ* 'follow along', which does not have any such documented function. Like Ninde, these forms in V2 position can function like 'about, on the topic of' for verbs of speaking; 'alongside' for verbs of motion; and with verbs of seeing, are part of applicative constructions meaning 'look after, care for'. It is possible that Ninde once had a repetitive sense of *woʁ*, and the form *pusuwoʁ* would originally have meant something like '(who) will keep coming back'. There is comparatively less evidence for this possibility.

The first step in reanalysis could well have been an extension to in-law forms, which generally have the suffix(es) *-l̪ˤowoʁ* following the possessor suffixes (*-ŋg*, *-m,* and *-ne*). The in-law suffix resembles the end of *vel̪ˤowo* 'holy, taboo'[1], (cf. V'ënen Taut *lalau* 'taboo'). Since most social restrictions (taboos) affect the interactions between people and their spouses' families, it is reasonable that this might have formerly been a word meaning 'taboo'. (There is also a parallel in Bislama and Tok Pisin usage: *tambu blong mi* 'my in-laws', or literally 'my taboos', and Nahavaq has borrowed *tambu* as 'father-in-law'.) Parents-in-law could then be 'taboo parents' – the taboo-avoidant register and practices are used with them. If this were the case, then a putative *temene-l̪ˤowo* could have been reanalyzed as *temene-l̪ˤo-woʁ*, allowing for the proliferation of this *-woʁ* suffix throughout the kinship system (except for *nelmen* 'nephew' and *nelmen wylepe* 'niece'). Though unusual, a comparable phenomenon can be observed in the Windua form previously described, where the relatively recently borrowed Samoan A*tua* 'God' has become A*tuwoʁ* 'God' for some speakers (and *Atuwo* for others).

Another suffix that appears to be verbal is *-woi* and its variant *-poi*. The variation does not make much sense with an interpretation of mood, but is parallel to verbs that begin with /w/ that is replaced by a /p/ in irrealis usage. This suffix is part of the forms listed relating to the relationship between paternal aunts and fraternal nephews, narrowly glossed as 'paternal aunt':

sa-**poi**

mother-**paternal.aunt**

'my aunt (vocative)'

sa-na-**woi**-woʁ

mother-3SG-**paternal.aunt**-?

‘their (sg.) paternal aunt’

nitu-**woi**-woʁ

child-**paternal.aunt**-?

‘your (fraternal) nephew’

net-na-**woi**-(lepe)-woʁ

child-3SG-**paternal.aunt**-female-?

‘their (sg.) fraternal nephew/(niece)’

sa-**poi**-woʁ

mother-**paternal.aunt**-?

‘your paternal aunt’

nitu-**woi**-(lepe)-woʁ

child-**paternal.aunt**-(F)-?

'your fraternal (niece)/nephew (of a woman)'

By contrast, the root for a niece or nephew via a sister, or of a man via any sibling, is *nelmen*. This is the only word in the Ninde kinship system that is sensitive to the gender of the ego. A partially homophonous *poipoi* 'raise, bring up (a child)' (Bislama 'leftemap pikinini') was provided in a literacy workshop in 2018 in Labo without a corresponding realis form (or irrealis if this is the realis form); we did not record or elicit the mood paradigm and it is unclear from the context – the task was to identify words with the /p/ phoneme – whether this is a citation form of the word. Nonetheless, this could suggest that there was some association between paternal aunts and child-rearing, and it may be related to the fact that women uniquely call their brothers' children *nitu-* 'children (offspring)'.

It is unusual in today's context that paternal aunts and their fraternal nieces and nephews should contain an element of 'raise', since women are highly likely to move away from home after marriage. Kinship terms for relations between maternal aunts and their sororal nieces and nephews are modeled on roots *sa-* 'mother' and *nitu-* 'child', respectively, and this is reflected in care-taking practices today. By contrast, female blood relatives very often live in other villages as a result of patrilocal marital practices, making an appellation like ‘raising mother’ somewhat unexpected. A highly speculative explanation could be that terms for aunt and son came about at a time when women were raised by fathers’ sisters away from home, only for the sons to return to claim their inheritance upon reaching adulthood or engage in ceremonies like circumcision. Men did not traditionally live with their wives and had no role in the upbringing of their children until these manhood initiation rites (Deacon & Wedgwood 1934a). Given these characteristics of traditional society, an older sister might raise her younger brothers’ children, ensuring some cultural and familial continuity through the fathers’ relations; this would require boys (but not girls) to relocate to their fathers’ homes.

(Ninde-affiliated) Mewun society, like all Malekula societies, is patrilocal, with sons inheriting land from their fathers (under most circumstances, only men directly own land). Although there is no evidence that sons were expected to leave their communities and subsequently return, it is considerably less likely that a daughter would return to her parents’ community, especially after an exogamous marriage.

Other suffixes remain which have no transparent cognates in any of the Malekula languages surveyed:

- *-ko* ~ *-kou* used exclusively in forms for fathers’ sororal cousins *ta-ko* and *tama-kou* ‘great cousin (vocative)’. Cf. Bislama *tawi* with the same meaning, from some northern Vanuatu (and cognate) source;

- *-lepe* 'female', used with ungendered terms that are generally interpreted as male;
- the alternation between *nai-* 'mother' found in compounds and the vocative *naijei* 'mother'; and
- *-ŋgəlei* with *sa-* 'mother', the in-law vocative for 'mother-in-law'. The rest of the paradigm is undocumented and unknown to most, if not all speakers of Ninde. There is a tempting resemblance in Avava *galu* 'mother', which appears to have no cognates. If proto NCM had a form *sa- 'mother', the third-person possessed form *san* would risk homophony with the highly common Avava *san*, Neverver *nes:an*, and Ninde *nsa* 'thing'; Avava could have adapted an older term for 'mother-in-law'

7.2 Cardinal Directions

All of the NCM and SWB languages for which cardinal directions have been described have at least one axis of cardinal directions that runs from the sea to the interior highlands. Like many Austronesian languages (Fox 1997), two of these directions depend on the landscape, with one term for points and directions that are relatively higher in elevation than the origin of movement or static reference point, and another term for points that are relatively lower in elevation (and closer to the sea). A second axis in Ninde and Neverver involves two opposite directions that are undifferentiated (i.e., designated by the same term). This axis is roughly perpendicular to the first, and points on this axis are only identified as existing at the same elevation.

	Avava	**Neverver**	**Ninde**	**Naati**	**Nahavaq**
‘uphill’	*l/aka* ‘in/to the bush’ *ɒaka* ‘in the bush’	*axus* (absolute) *axsuŋ* (deictic)	*ʔa-*	–	*eraʔai*
‘downhill, toward the sea’	–	*avev* (absolute) *aviviŋ* (deictic)	*suwa-*	–	*ejten*
‘same elevation’	–	*atl*	*ʁa-*	–	–
‘up (vertically)’	*ɒaka*	*arxa*	*erei*	–	*liβʷaʔanmehe p*
‘down (vertically)’	*kutn*	*bis(t)n*	*l̺ˤata-(ne)*	–	*leten*
‘on’	–	*ar/an-*	*ʁa-*	*ra-*	*ra-*
‘in the middle of’	*lupan*	–	*live(ne)*	–	*liβʷaʔan*
‘under’	*lapa-*	*lap:an*	*ve-ne*	*ra-*	*evun*
‘body, trunk’	*lupan*	*ni/livxan*	–	–	–
‘base, trunk, bottom’	*wuti-*	*nuxutn*	*nuŋgute*	–	–
‘tree’	*aga(i)*	*na/xa*	*n/ei*	*naʔai*	*naʔaj*
‘ground, earth’	*(a)tan*	*l/ot:an* ‘on the ground’	*nete-ne*	*nitan*	*neten*
‘base of, trunk’	*beh/nan* ‘base of a coconut tree’	–	*ne/mbyse* ‘end, tip, tail’	*mbati* ‘trunk’	*mbusu* ‘tail’
‘ashore (from the sea)’	*out*	*aut*	*ewute*	–	*mbi-βerʔeh* ~ *mberʔeh*
‘on the shore’	*l/awal*	–	*l/owute*	–	*l/aw*

	Avava	Neverver	Ninde	Naati	Nahavaq
'inland person'	*nat/aut*	*nid/aut*	–	*mal/ewus*	–

Table 76: Cardinal directions, vertical directions, and various forms that are proposed to be related in some way.

Due to the location of each language community, corresponding forms do not refer to the same axes of the sun cycle. While Neverver *axus* 'uphill' corresponds generally to the South, Ninde *ʔa-* 'uphill' corresponds to the Northwest in frozen usage (including names of regions, and the people, flora, and fauna found in them). Like Neverver (Barbour 2012:95) and distantly related languages of other languages like Ambae (Hyslop 2002:52), these forms are used to identify locations and objects at a very local scale (preferred over 'left' and 'right') in Ninde, even though they are rooted in an absolute frame of reference that is calibrated to the topography one is in.

Every language for which such a deictic system is described has at least one form that is derived from another source, often extended from words related to 'tree, forest', 'ground', and 'base' – the NCM languages, including Ninde, reserve these more transparent forms for the vertical axis. Locative forms of 'tree', like Avava *laka*, often refer to the wooded interior of Malekula – since the interior is also higher in altitude, it can mean both 'in(to) the woods' as well as 'in(to) the island interior'. The word for 'tree' can form the basis for both vertical and diagonal senses of 'up'. Barbour (2012:95) suggests that the Neverver form *arxa* 'up' is historically from a fusion of ***aran***-*ar* 'on (locative)' and *na*-***xa*** 'tree' with a combined sense of 'up a tree'. This is supported by the converse vertical direction, which in Neverver *bis(t)n* 'down' corresponds to elements in other languages for 'base'. This reflects a generally widespread polysemy of a word meaning 'base' and 'trunk (of a tree)', or rather, a general prototypical meaning of 'trunk' associated with a more general word for 'base'. The NCM languages, however, all have reflexes of a shared *nokut 'trunk of' – correspondingly, the Avava source of *kut* 'down (vertically)'. In the vertical sense, Ninde *l̪ˤata(ne)* 'down' is in line with Nahavaq *leten* 'down'; both words are locative forms of *netene* and *niten* 'ground, earth', respectively. Neverver *avev* and *aviviŋ* the directional and locative forms corresponding to locations closer to the sea, could be related to PNCV*vava 'under, beneath' and to Ninde prepositional *vene*. Crucially, Ninde *erei* 'up' corresponds in form and sense to NCM cognates.

Nahavaq alone seems to have swapped vertical directions with horizontal/diagonal axes to some extent. This direction of change is suggested by the composition of these forms, which are are transparently *eraʔai* 'up in a tree' > 'uphill' and *ejten* 'on the ground' > 'downhill'. (There is unfortunately no data from Naati.) Oddly, the form meaning 'up' on the vertical axis, identified as an opaque compound (Dimock 2009:202), contains *livʷaʔan* 'in the middle of' and an unidentifiable *mehep*, which could well be cognate with Ninde *nemep* 'elevated flat land'. This would represent the reverse – a shift in sense from a topographical axis to a more local vertical one. While Ninde shares a form for the vertical sense of 'on the ground' > 'down' with Nahavaq *leten*, Ninde lacks the form derived from the less productive locative prefix *e-* found in Nahavaq. Insofar as Nahavaq represents the SWB group, Ninde directional terms are overall in line with NCM, containing only a single SWB form.

The most robust systems described are those of Neverver and Ninde, yet the forms do not immediately appear to be cognate. Ninde *suwa-* 'downhill' could be related to Nahavaq *sip*^{w} 'go down' (PNCV*sivo 'down, go down'). This would be a doublet with a corresponding Ninde verb *sip* '(of bird) alight, (of wind) go downhill', which has comparatively restricted usage. A source NCM form would likely have been *siv^{w}, of which it would be a regular reflex. Neverver *axus* 'uphill' and Ninde *ʔa-* could be partially cognate, since postvocalic *s would ultimately be deleted in Ninde and /ʔ/ is a reflex of older NCM velar consonants. This may be related to PNCV*sake 'up, go up' (a doublet in both languages with Neverver *sax* 'go up' and Ninde *jaʔ*); both Neverver and Ninde would have had to undergo the same metathesis of *s and *k exclusively in the adverbial usage. The rounded vowel of the Neverver form is not explained by this etymology.

A few forms are left with no explanation at all. These include Ninde *ʁa-* and Neverver *atl*, both corresponding to either direction on the same axis (neither uphill nor downhill). Nonetheless, the shared polysemy is curious and no other discussion of cardinal direction in the other languages includes this axis. Furthermore, unique constructions exist for the three Ninde adverbs identifying cardinal directions on perpendicular horizontal axes. Like verbs, adjectives, and locatives formed with *l̪ˤV-* prefixes, these adverbs can uniquely take a relativizer *t-* (which for *ʔa-* results in the only known token of a consonant cluster involving a glottal /tʔ/); this reinforces the proposal of a verbal origin. There are two affixes found nowhere else in Ninde: *-t*, which functions like Neverver *-iŋ~-uŋ* in nominal forms and derives a static locative, and *-indeve* (c.f. *jaʔndep* 'reach, arrive at', which could contain *jaʔ* 'go up' and a related verb), which indicates movement in the direction. They can also be combined with each other with rather unpredictable meaning and morphology: *suwa-suwo-t* 'north', *suwa-ta-ʁa-t* 'east', and *suwo-t-ʔa-t* 'south'. Higher order reconstructions do not offer any insight: POc*laur 'toward the sea' and PNCV*ʔuta 'inland' are both likely found in reflexes meaning 'ashore (from the water)' or 'onshore' throughout NCM and SWB, and no candidates are found among many words reconstructed for 'up'. These forms may never be explained, unless more data is made known from other Malekula languages.

Once again, the cardinal directions represent a blended system. While the horizontal plane mirrors the taxonomy of NCM languages (and the vertical axis is taxonomically the same in SWB), most of the forms are SWB in origin. As seen in kinship terms, Ninde innovates its own forms from verbs, exploiting the affordances of serialized verbs for reanalysis and extension to adverbial switch contexts.

7.3 Calendrical terms

Unlike kinship and cardinal directions, the purpose of calendars requires a lexical system that is fundamentally sequential rather than structured on oppositions and reciprocal relationships; this means that it is closely linked to number systems. I argue in this section that the weekdays employ what I call an intrinsic temporal frame of reference, following Levinson's (1996) taxonomy of spatial frames of reference. The use of an intrinsic frame of reference for weekdays makes for a more complex lexical system than the weekdays of the Gregorian calendar; weekdays are points in time relative to a mutable reference point, which allows for more than just a linear relationship between fixed sequences of days. Traditional months, on the other hand, were probably absolute – but not without a high degree of innovation and localization. On Malekula, the same roots used for days of the week were combined with lunar months and moon phases. These are fixed points in time, so the relative weekdays can also be used to identify fixed points in time far from the present day.

For Mewun people, the traditional calendar constrained resource consumption, and was used to mediate people's relationship to the physical environment. Given that calendrical data for NCM and SWB languages primarily exists for Avava, Neverver, and Ninde, it is important to reconstruct the proto language calendar systems to understand the contemporary record.

There are several challenges to identifying cognates across the three calendrical systems considered here.

- Apparent cognates can be at virtual opposite regions of the system: Avava *<u>ra</u>da* ‘January’, *<u>ra</u>wel lam* ‘February’, and *<u>ra</u>bʷisisi* ‘March’ share PNCM *rak > Avava *rak* ‘make a forest clearing’[2] with the *ʁaʔ* ‘work, make a forest clearing’ of Ninde *ʁaʔpape* which roughly corresponds to July. These refer to the preparatory phases of yam cultivation, and so it should be expected that they would be winter months (June, July, August).
- Forms that are clearly cognate across languages can have difficult-to-explain differences: Avava *uwah*, Neverver *was* ‘second day’, and Ninde *wa* ‘third day’, which represents the second or third day of the week depending on language.
- There are also secondary forms that have cognates, but are rarely used: Ninde’s *nese<u>lm</u>enene* ‘fifth (day) of something’ exists alongside a form with no cognates, *wa pal̪ˤat* ‘fifth day from’. Compare Avava *t<u>lim</u>* and Neverver *ti<u>lim</u>*, both ‘fifth day of’.
- The forms may be semantic equivalents, but not cognate: Avava *<u>mʷeta</u>gadoli* ‘April’ and Ninde *<u>netal̪ˤai</u> papal̪ˤe* both start with non-cognate words for ‘axe’

These differences require a holistic approach to shared inheritance, where cognacy and shared semantic composition consider calendrical units as both cyclically ordered sequences with mutable starts and ends as well as descriptive names for changing weather patterns in a climatically diverse topography.

In this subsection, I discuss lunisolar calendars broadly and what has been said of Melanesian calendars. I then assess the state of the data available for NCM and SWB calendars. The days of the week and the notion of an intrinsic temporal frame of reference are exemplified, then compared to other systems with non-cognate forms in related languages. Finally, I carry out a reconstruction by contrast for the months of the year, which have minimal cognates but follow the growing cycle of yams.

Vanuatu is predominantly Melanesian in its modern identity and history, belonging to a region that relies on the yam as a dietary staple. Much of the year is shaped by strict social taboos imposed during sensitive periods of yam growth. These taboos traditionally limit sexual activity, violence, and the movement of women for parts of the growing cycle and human reproductive cycles.

In more recent times, changes in waste management and the influx of tourists, plastics, and new toxins all put a strain on local ecosystems. Overfishing has also removed the natural predators of the Crown-of-thorns starfish, which preys on coral and destroys reefs. Reefs no longer protect against the threat of erosion, which is made worse with rising sea levels. Today, Vanuatu's population is at considerable risk with high levels of exposure to natural disasters (in 2022, the World Risk Report lists Vanuatu as the 48th most exposed country, down from 1st place in its 2016 edition) (World Risk Report 2016, 2022).

7.3.1 Intrinsic temporal frame of reference

The forms in Table 77 are days of the week in Ninde, which are by default relative to the present day. However, if the context provides another day as a reference point, typically the first day of a sequence of days, these terms are relative to that day. The sequence also extends into the past, relative to 'today', or another reference point. The forms for the past bear a resemblance to the future days, except that the days of the past tend to have an initial *na-*.

Back in time		**Forward in time**		**Ordinal value**
Ninde	**Gloss**	**Ninde**	**Gloss**	
kəsaʁ	'today'	*kəsaʁ*	'today'	'the first day'
l̪ˤombun	'yesterday'	*maʁan*	'tomorrow'	'the next day'
wije motne	'-2 days'	*wije*	'day after tomorrow'	'the third day'
nawa	'-3 days'	*wa*	'in three days'	'the fourth day'
nawa tuwa	'-4 days'	*wa tuwa*	'in four days'	'the fifth day'
nawa pal̪ˤat	'-5 days'	*wa pal̪ˤat*	'in five days'	'the sixth day'
nawa naki	'-6 days'	*waŋgi*	'in six days'	'the seventh day'
nawa su	'-7 days'	*wa su*	'in seven days'	'the eighth day'

Table 77: Days of the week in Ninde.

The long week is likely motivated by the use of lunar phases to keep track of the days. In Ninde, there are named moon phases providing reference points at least every seven days of the month, meaning that any day in a month can be identified within a 15-day week that is centered on a distinct lunar phase. These phases are: *neve tivou* ‘new moon’, *nemelmbəle* ‘first quarter’, *nambaŋg* ‘half moon’, *namal̪ˤo* ‘full moon’, and *vilivilu* ‘waning moon’ – as the moon wanes, stages of waning are identified with ordinal numerals. In Avava and Neverver, there are only five weekdays recorded in either direction, but these may have also sufficed.

The traditional months are more poorly documented. There are between 11 and 14 month names per language, generally recorded in relation to 12 months of the Gregorian calendar. The numbers of months do not match because there are repeated names in the sequence and some months have several alternatives. It is not generally known what time periods the months initially corresponded to, or how many months there originally were. However, the system of weeks and months appears similar in many ways, and some information about a proto calendar can be assumed from these similarities.

All lunisolar calendars are by definition both lunar and solar, meaning that moon and sun cycles together provide the basis for the units of calendar-scale time. Lunisolar calendars have to coordinate synodic months (the 29.53 days from one new moon to the next) with the solar year, which is approximately 365.24 days long.

In order to have 12.37 months in the year, calendars have been developed with various strategies. Intercalation, or the insertion of additional units into a sequence, is one strategy. Calendars in Melanesia tend to have intercalary months (rather than days or weeks units). There may also be a mismatch, meaning that the sequence of named months does not cover an entire solar year, or may be longer than a lunar year. Finally, unit lengths may have been varied, with some months longer than others. This is not likely for any NCM calendar, since it has not been described for any Melanesian languages. In terms of function, such a month type would mean that the lunar phases cannot be used to identify far-off dates using a five- or seven-day week. The existence of an intercalary month is most likely, given the cultural significance of the calendar.

Natural phenomena trigger recalibration by the addition of an intercalary month in different ways for traditional Vanuatu calendars. The aquatic Palolo worm, for example, releases its light-sensitive reproductive organs into coastal waters in October and November. It is sensitive to the time of year and the moon phase, which makes it particularly useful for coordinating the two cosmic cycles. There may be astronomically important alignments of the sun with features of the landscape, or the appearance of constellations like the Pleiades in the night sky.

An alternative to documented practices is that no mechanism existed to keep months aligned with the years. When the descriptive name of a month no longer matched the environmental reality, it could have been removed or shifted in the sequence. Since months are used to schedule ceremonies related to yams, it is not entirely unexpected that there is so much variation in the months. These include prescribed times of the year for warfare and violence, and times to display yams to gain status. Yams are highly sensitive to local microclimates, and Malekula's steep mountains make climate patterns highly varied. Month names describing yam-based work might be hyperlocalized, and in that case we expect linguistic variation.

Reconstructing the week

There are 11 weekdays recorded for Avava and 13 month names – although Lynch objects there are only 11 named months, the missing month of September is named in the lexicon of Crowley's grammar (2006a). Barbour (2012) provides only 9 weekdays for Neverver, but 14 month names. Letpen recorded 15 weekdays for Ninde, and Edwel Kaiseng recorded 11 unique month names covering 12 months. Lynch (in a footnote in his edit of posthumously published Crowley 2006a) notes that there are various means of adapting indigenous calendars to Gregorian months, assuming that the traditional calendars should have had 13 named months. These adaptations have obscured the nature of the original systems.

The days of the week in each of the three languages – Avava, Neverver, and Ninde – employ an intrinsic temporal frame of reference. In spatial frames of reference, a relative frame of reference, like the one we use in English, uses both the point of view of an observer and a reference point to cast a coordinate system over the wider landscape. In Figure 16, the star would be located to the 'left' of the house in a relative frame of reference.

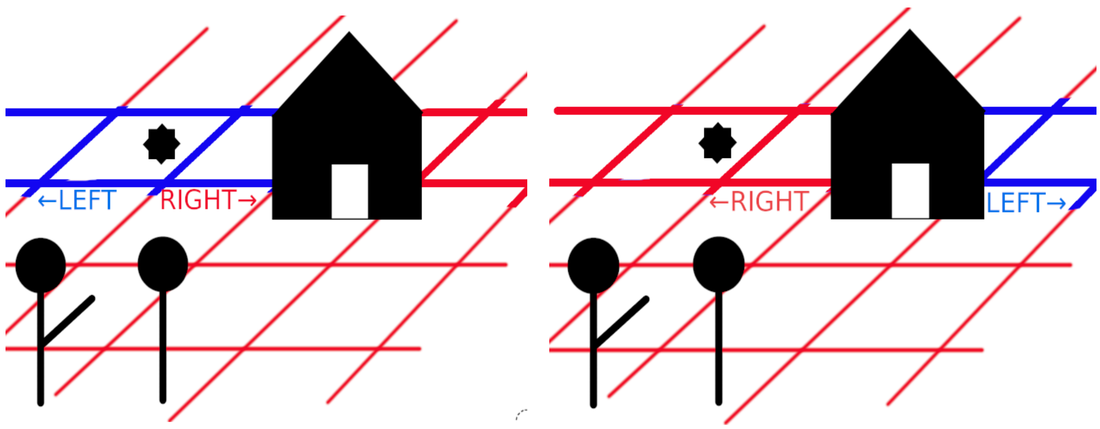

Figure 16: Schematization of the relative (left) and intrinsic (right) frames of reference from the perspective of the viewer. In the relative frame on the left the star is to the left of the house, but in an intrinsic frame, the star is on the house's right.

In an intrinsic system, objects have their own *left*, *right*, *front*, and *back*. The coordinate system in an intrinsic system emanates from the reference point. In Figure 16, the front of the house would likely be the side with the door. In an intrinsic frame of reference, the star would be located to the right of the house.

Expressions like *the day after tomorrow* have been described as employing a relative temporal frame of reference (Radden 2003). If observers and events are thought of as moving through time, or facing the future, for example, then it is possible for events to have an intrinsic front and back that is at odds with present day observers. An expression like *move* *the event two days forward* can be interpreted using a relative frame of reference – the event will be postponed. If the event is thought of as moving in through time *toward* the present, then *move forward* means that it will be rescheduled for a sooner, earlier time. The intrinsic frame of reference was found to be consistently used in this scenario in some Germanic languages; it was also found to be inconsistently used by English speakers in the US (Rothe-Wulf et al. 2015). Zinken (2015) predicts that absolute frames of reference are more likely used for large timescales, but relative ones for immediate events. Although he makes no prediction about intrinsic frames of reference, we see that they are more flexible for time than relative frames of reference.

In a relative temporal frame of reference, a timeline always emanates from the present day. In such a system, the star will always be located 'yesterday', no matter what reference point might be established by context. In an intrinsic frame of reference, however, the timeline emanates from the reference point. If the reference point is in the past, then 'tomorrow' can refer to a day that has already passed. In Figure 17, the star would have happened 'tomorrow of the house', where the house stands for an event.

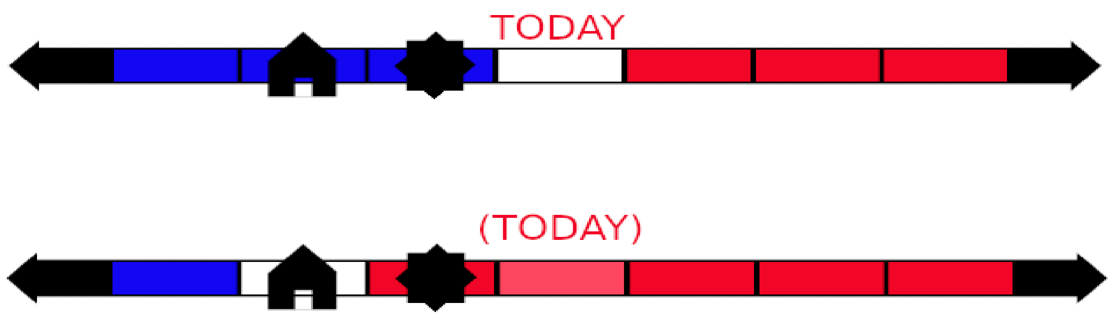

Figure 17: Relative (top) vs. intrinsic (bottom) frames of reference schematizing the day before the present ('yesterday' in English) on a timeline. In an absolute frame of reference, an event occuring on the prior day is indicated in terms of 'today' as a fixed reference point, but in an intrinsic system, if a narrative begins with an event (the house) that took place two days ago, then the equivalent of 'tomorrow' refers to one day prior to today, the day after the event began.

Unlike English *tomorrow,* PNCM *maran can refer to days in the past like this. This has the same logic as relative tenses (Comrie 1985), which also have an explicit or discourse-contextual reference point that can be temporally moved. The relative tense is relative to that mutable reference point, and in a way could also be said to represent an intrinsic frame of reference. Again unlike relative terms like *the day after tomorrow*, a word that represents either 'the day after tomorrow' or 'the third day' can be used with fixed points in an absolute frame of reference. In Ninde, the time frame can also be on the scale of years, since these forms can be combined with the word *netou* 'year' with a relavizing *t-*: *netou t-maʁan* 'next year'. In stories, Ninde speakers often recite the sequence of weekdays intervening between the reference point and the target day, instead of naming only the target. The reference point can also be shifted partway through a story.

Comparing the weekdays across languages (Table 73), Ninde stands out as having largely unrelated forms from the other NCM+ languages. For one thing, the system is more extensive in Ninde, so many forms have no attested cognates. The NCM+ languages generally have (instead of a *na-* prefix as in Ninde) a root *ʙuŋ followed by an ordinal number *(n)itl 'third', *(n)ivats 'fourth', or *(n)ilim 'fifth'. Future days are generally formed from the ordinal form, but with a relativizing or genitive *ti- instead.

Day (after RP)	PNCM	Avava	Neverver	Neve'ei	Ninde	Naati	Nahavaq
-7					*nawa su*		
-6					*nawa naki*		
-5		*ʙunlim*			*nawa pal̺ˤat*		
-4		***ʙu**nivat*			*nawa tuwa*		
-3	*mbuŋ nitl 'third day past'	***ʙu**nitl*	*na**ʙuŋ** itl*	***buŋ**nitl*	*nawa*		
-2	*(ta)-na-was	*tanawah*	*nanas*	*tenoah*	*wije motne* 'true future day'	*nembuŋ-ru*	
-1	(*ʙuŋ)	*tenep*	*a**ʙuŋ***	*tenev*	*l̺ˤ**ombun***	*lambum*	*lambʷum*
0		*livani*	*barnax ~ barnix*	*marnaŋ*	*k(i)saʁ*	*ŋaːr*	*ʔor*
1	*ma-ran 'morning'	*maran*	*maran*	*meran*	*maʁan*	*lavøʔ*	*lavʷuʔ*
2	*(i)-was	*uwah*	*was*	*uah*	*wije*	*mbu-ti-ru*	

Day (after RP)	PNCM	Avava	Neverver	Neve'ei	Ninde	Naati	Nahavaq
	'future day'				'future day'		
3	*ti-tl 'the third'	*titl*	*titl*	*madl*	*wa* 'future day'		
4	*ti-vats 'the fourth'	*tivat*	*tivas*		*wa tuwa* 'other day'		
5	*ti-lim 'the fifth'	*tlim*	*tilim*		*wa pal̪ˤat* 'day ahead'		
6					*waŋgi*		
7					*wa su* 'day returns'		

Table 78: Days of the week in the NCM languages, with a proposed literal gloss where possible in Ninde. Days are numerated relative to a mutable reference point (RP).

Clark reconstructs PNCV*boŋi 'night, day (unit of time)', in turn from POc*boŋi 'night'. These form the base for past days, and as such retain an association with the night, though no language has retained this exact reconstructed meaning (Table 79). Adverbial usage of its reflexes can mean '(at) night', as in Tirax, or 'yesterday' in Neverver and Ninde. Most likely, this sense is extended from a reference to the night immediately prior to the entire preceding day. Ninde is unique in extending a word for 'daylight', the nominalized form of *ran* 'be daytime' to units of measure as well. A primary meaning of 'yesterday' also found in Ninde, could easily have been extended in every language but Ninde.

	Avava	**Neverver**	**Ninde**	**Tirax**	**V’ënen Taut**
‘night’	*aʙuŋ* ‘day’	*naʙuŋ* ‘day’ *aʙuŋ* ‘yesterday’	*l̪ˤombun* ‘yesterday’	*(ne)buŋ* ‘past day’ *labuŋ* ‘night’	*napən* ‘day’
‘daytime’	*outranian*	*nautranian*	*nutaʁanijene* ‘day’	*nɔtrɛn* ‘day’	–

Table 79: Two words, ‘night’ and ‘day’, but as a unit of time, the word for ‘night’ may be translated as ‘day’. Tirax and V’ënen Taut are included to illustrate widespread association with ‘night’ and days in the past.

There are unexpected reflexes of the nasals in these past-day forms. The idiosyncratic coronal nasal in Ninde *l̪ˤombun* makes it superficically like V’ënen Taut *napən* ‘day’, but that language has an exceptionless change of velar nasals to coronal ones that Ninde does not share. This idiosyncratic Ninde form could be a hyper-correction from the SWB source, where coronal nasals at the end of a word become bilabial nasals after round vowels and velar nasals elsewhere (cf. Nahavaq *lambʷum*, Naati *lambum* ‘night’ and in both languages *reŋ* ‘be daytime’). Alternatively, Avava and Neverver (but not Neve’ei) seem to have lost different nasals in the sequence formed at the morpheme boundary in an inherited *ʙuŋ-ni-tl ‘three days ago’ (where *ʙuŋ ‘past tense and the ordinal prefix *ni- meet). Ninde, like Avava, could have formerly lost only the velar nasal, then extended this as the base form to ‘yesterday’, before eventually replacing all the numbered weekdays.

The weekdays generally have important semantic extension relationships with times of day (Table 80). Incidentally, times of day in Ninde are generally from the NCM source, but two (noon and afternoon) are from the SWB source.

	Naman	**Avava**	**Neverver**	**Ninde**	**Naati**	**Nahavaq**
'morning'	***metebəx***	***mataʁux***	***mitaʁux***	***mindumbuko***	***levʷaʔhat***	***levahat***
'midday'	–	*lupanal*	*lonial*	*venelis*	–	*livʷaʔaljes*
'afternoon'	*revrev* 'evening'	*kinkiɒap*	*livrav*	*l̪ˤaʁap*	*revrev*	*larap*
'(at) night'	***delva:t***	***lupat***	***livxat***	***livete***	***levʷaʔhat***	***livʷaʔat***
'darkness, dusk'	–	*out-met-ian*	*met*	*nuwute-mite* 'dusk'	–	*newut-ponopon-jen*
'daylight, dawn'	***iren***	***out-ran-ian***	***naut-ran-ian***	***nuwute ʁan* '(it is) dawn'**	***newut-reŋ-ian***	***newut-reŋ-jen***

Table 80: Times of day in NCM and SWB languages. Colors reflect cognacy; red cells reflect the inherited form in Ninde. Bolding reflects forms that appear to replace older ones, or that are innovated

Most of these forms are from *livʷa- ‘in the middle of’ or *mata ‘eye’ (the same reconstructions apply for NCM and SWB). A compound of PNCV*mata- ‘eye of’ and PNCV*vuko ‘tomorrow’ is reconstructed as the etymon for widespread words for ‘morning’ that are reflected in NCM languages (Clark 2009:229); the term for ‘night’ that apparently replaces PNCV*boŋi is *livʷa- and an unidentifiable PNCM*xat (and probably PSWB*ʔat). In Ninde, the word for ‘day’ that serves as a unit of measure is *nuwutaʁanijene* (which generally surfaces as *nu[wu]tʁaine*), the nominalization of ‘be daylight, dawn’ that in all the other NCM and SWB languages is used for daylight.

A major challenge lies in the fact that the whole system is transparently derived from ordinal numbers in NCM languages and the minimal attestation of SWB, but not for Ninde. In some contexts, Ninde also allows for ordinal numbers to identify days in similar contexts. Once a day has been identified as “the third day” using a noun phrase headed by *nuwutaʁanijene* ‘day’ (with too many casual-speech variants to list here), it is possible to omit the word ‘day’, using headless ordinals, as: *the third day, the fourth, the fifth*. Ninde effectively has two sequences in regular use for the days of the week, one of which is based on numerals (but is only recorded for sequences of days). This potentially represents older usage that predates the contemporary Ninde forms, given the present-day usage of numeral-based forms in other NCM languages.

Words for ‘today’ come from various sources. In Avava, the word *livani* is clearly from ‘the middle’ (compare also ‘midday’ and ‘night’); Neverver *barnax* and Neve’ei *marnaŋ* are polysemous and also mean 'now', but importantly contain the proximal suffix -*ax* and -*aŋ*, so a cognate would best match *barVn. The Ninde form has no clear cognates anywhere on Malekula, but Naati *ŋaːr* and Nahavaq *ʔor* 'today' come close. Naati long vowels are sometimes from vowel coalescence after the loss of an intervocalic *h (most likely when the first syllable is stressed), and Nahavaq initial /ʔ/ may alternate in a residual consonant mutation pattern with /ŋg/, but a reconstructible form *ŋgV(s,h)ar 'today' leaves much to be explained phonologically.

Similar phenomena seem to have provided the basis for the future days as they do for past days. A PNCM form *was (cf. POC *mapua, cf. PPh *buás 'tomorrow') can be reconstructed from Avava and Neverver's 'day after tomorrow' and Ninde's 'third day'. The reconstructed form resembles the Proto Philippines form more closely than Proto Oceanic; it is possible that the data informing the reconstructions of time-keeping vocabulary are generally scarce. Nevertheless, Ninde *wije* 'the day after tomorrow' shows some metathesis of the erstwhile third-person subject prefix *i- (which could also explain the initial /u/ in Avava *uwah*) with the initial reflex /w/, so the stative POC *ma- was lost, if it was ever obligatory to begin with. In Ninde, this *wa* has uniquely proliferated through not only all future days, but also the past days.

Ninde has a doublet, and it is possible that a PNCM*was has provided both the forms for the 'second day', the 'third day', and the root for all the future days. Since the future days in Avava and Neverver are headless, this is also a strong candidate for a missing head (if there ever was one). One Ninde form *wa* is unusual phonologically, because it is resists word-final raising of the /a/ → [e], and it does not undergo the metathesis of the fused erstwhile subject *i- with the initial consonant. Monosyllabic verbs in Avava (and probably also Pre-Ninde) do not have a fused *i- when the verb is compounded with other verbs, so the alternation between *wije* and *wa* could have initially been allomorphy between free and compound forms. The word final *h of a Pre-Ninde *iwih would be consistent with final vowel insertion after other *ih sequences (cf. PNCM*i-vis 'how much/many' > Ninde *vije*). The Ninde *wa* 'in three days' could have been back-formed from the compound set, and placed in the position of the first (closest to the reference day) *wa* day, after *wije*.

The word 'tomorrow', like 'yesterday', transparently comes from a verb form PNCM and/or PSWB *ran 'be early, become dawn or day' with a stative prefix *ma-. In context, the morning can refer to the next day. The future sequence in Ninde is formed from *wa* and various words that appear to have verbal origins. In other contexts, *pal̪ˤat* means 'go first, go ahead' and makes sense here as a relatively early form in the sequence of compound forms. From the final weekday, the form *su* is also a verb and means to 'return, come back'. Since this is the seventh day, it could be that this is the day that the week resets, and in a way conceptually "comes back" to the start. These interpretations leave only two weekdays in each direction with no recognizable internal structure: *wa̱ŋg̱i* (which unexpectedly has only an oral velar in the corresponding past day *nawa nak̠i*) and *(na)wa tuwa* (*tuwane* is an indefinite marker of nouns and has been recorded in back-formation as *tuwa*, but this does not account for the form).

As for the days in the past, a form reconstructible as PNCM *na-was for the 'day before yesterday' looks simply like a noun form of *was, the second day in the future. This could be the result of analogy. Since days of the past are generally numerals preceded by the noun *na-buŋ meaning 'a past day', the word class could have been generalized for the sequence. It could alternatively be that the nominalized forms of the verbs in some way lent perfective aspect, which is generally compatible with events completed prior to a reference point. The verb forms for the future words, however, are not correspondingly irrealis (which is typically marked for clauses denoting future events). If this is the origin of the word, it is a very old origin. V'ënen Taut, which is more distantly related to Avava, Neverver, and Ninde, also has a noun-like form (i.e., one beginning with *na-*) for the second day of the past: *naua* 'two days ago' corresponding to *paua* 'in two days'. The Neverver *nanas* appears to be doubly nominalized, with an additional *na-* prefixed to an already nominal form.

In fact, days of the past show signs of multiple word class changes. A verbal origin works for the future sequence in Ninde, since verb-verb compounds are common, but less so for the past, since noun-verb compounds are more limited. At any rate, the usage of all weekdays in PNCM languages can be verbal (i.e., they can appear in the predicate without a copula). Avava, however, has what looks like relativizing morphology on the ostensibly nominal roots in *<u>te</u>nep* 'yesterday' and *<u>ta</u>nawah* 'the day before yesterday'. By comparison, NCM forms *ran 'be daylight' are part of nominalized constructions *naut-ran-ian 'dawn' (nominalized with a subject-incorporated *naut 'place'). As with perfective aspectual markers, nominalization has the effect here of interacting with the stative verb's lexical aspect by making it inchoative. The residual irrealis form *Dan (with the prenasalized rhotic associated with unproductive consonant mutation) serves as a V2 meaning 'to _____ through the night until morning'. Both constructions may occur post-verbally, and it is the noun form of the words that serves as an expression of a time past.

In contrast to kinship and cardinal directions, the weekday in Ninde is not meaningfully mixed. There appear to be two schemas, one closely resembling the numerated days found in NCM languages, and the other is wholly innovated or borrowed from an as of yet unidentified source. There is no evidence of such a week in SWB languages, but it is possible to speculate that it was not a familiar practice for speakers of SWB languages. With nothing to mix, Ninde speakers may have asserted a distinct, mixed identity by innovating an entire set of forms to replace the conceptual units inherited from NCM.

Months of the year

There is much to say about the weekdays, but comparatively little to make of the months. Nonetheless, a comparison of forms can reveal differences in adapting the month names to the Gregorian months. Avava's calendar starts in January, but the Neverver calendar starts in

May. Of cognate semantically equivalent forms (Tables 81-83), it can be seen that there is a consistent offset of five months. It seems that Avava speakers mapped the first month of their calendar to January, whereas Neverver mapped their months to the Gregorian months that are temporally closest, or most overlapped. Observe that there are numerous family resemblences with no core shared by all NCM.

	Avava	Gregorian		Neverver	Gregorian
1	*rada*	Jan	1	*ni**blongmur*** **(12a)**	May
2	*rawel lam*	Feb	2	*nekkanbor*	Jun
3	*rabwisisi*	Mar	3	*nekkanbrokhari*	Jul
4	*mwetagadoli*	Apr	4	*navulbrongnaj*	Aug
5	***lapda*** **(6)**	May	5	*nisuda*	Sep
6	***lav vwelam*** **(7a)**	Jun	6	*nelavlav*	Oct
7	***lap mwisisi*** **(8)**	Jul	6	*ni**lavda*** **(5)**	Oct
8	*reptamal*	Aug	7a	*ne**lavlavlab*** **(6)**	Nov
9	*reptulu*	Sep	7b	*nelavlavran*	Nov
10	***ivlerongo*** **(11)**	Oct	8	*ni**lav**la**varikh*** **(7)**	Dec
11	*ivlemial*	Nov	9	*nibongvkhal*	Dec/Jan
12a	***bolongmur*** **(1)**	Dec	10	*nibilkhenbet*	Jan/Feb
12b	*abasi mial*	Dec	11	***navulbrang*** **(10)**	Feb/Mar
			12	*nepraskha*	Apr

Table 81: Pairwise comparison of Avava and Neverver months, with cognates or semantic equivalents in bold with a number represented in the month in the corresponding calendar.

	Avava	Gregorian		Ninde	Gregorian
1	***ra**da* **(1)**	Jan	1	***ʁaʔ**pape* **(1-3)**	Jul
2	***ra**wel lam* **(1)**	Feb	2	***netal̪ˤai*** *papal̪ˤe* **(4)**	Aug
3	***ra**bʷisisi* **(1)**	Mar	3	*katis*	Sep
4	***mʷeta**gadoli* **(2)**	Apr	4	*nespin*	Oct
5	*lapda*	May	5	***l̪ˤapl̪ˤapo*** **(6)**	Nov
6	***lav*** *vʷe**lam*** **(5)**	Jun	6	***ʁiluk**ijan tlepe* **(9)**	Dec
7	*lap mʷisisi*	Jul	7	***ʁiluk**ijan tloulou* **(9)**	Jan
8	*reptamal*	Aug	8	*ʁilukijan tlepe*	Feb
9	***rep**tu**lu*** **(6-8)**	Sep	9	*noʁtaʁu*	Mar
10	*ivleroŋo*	Oct	10	*nelinawi*	Apr
11	*ivlemial*	Nov	11	*nesuwaloujene*	May
12a	*boloŋmur*	Dec	12	*nanmaʁi kake*	Jun
12b	*abasi mial*	Dec			

Table 82: Pairwise comparison of Avava and Ninde months, with cognates or semantic equivalents in bold with a number represented in the month in the corresponding calendar.

	Ninde	Gregorian		Neverver	Gregorian
1	*xaʔpape*	Jul	**1**	*niblongmur*	May
2	*netahlai papahle*	Aug	**2**	*nekkanbor*	Jun
3	*katis*	Sep	**3**	*nekkanbrokhari*	Jul
4	*nespin*	Oct	**4**	*navulbrongnaj*	Aug
5	***hlaphlapo* (6a/7a)**	Nov	**5**	*ni**suda*** **(11)**	Sep
6	*xilukiyan tlepe*	Dec	**6a**	*ne**lavlav*** **(5)**	Oct
		Jan	**6b**	*nilavda*	Oct
7	*xilukiyan tloulou*	Feb	**7a**	*ne**lavlavlab*** **(5)**	Nov
		Mar	**7b**	*nelavlavran*	Nov
8	*xilukiyan tlepe*	Apr	**8**	*nilavlavarikh*	Dec
9	*noxtaxu* **(10)**	May	**9**	*nibongvkhal*	Dec/Jan
10	*nelinawi*		**10**	*nibilkhenbet* **(9)**	Jan/Feb
11	*ne**su**wa**lou**yene* **(5)**		**11**	*navulbrang*	Feb/Mar
12	*nanmaxi kake*	Jun	**12**	*nepraskha*	Apr

Table 83: Pairwise comparison of Avava and Ninde months, with cognates or semantic equivalents in bold an conceptual links underlined with a number represented in the month in the corresponding calendar. Alignment reflects the shared start of the year in winter.

From the perspective of Ninde, there are only four months that cannot be related to a form in either Avava or Neverver: *katis, nespin, nelinawi, and nanmaxi kake*; Avava has two: *ivlemial* and *abasi mial*; and Neverver has five: *nekkanbor*, *nekkanbrokhari*, *navulbrongnaj*, *nibongvkhal*, and *nepraskha*. Several of these, however, have conceptual links to one another.

It is also clear that the start of either Ninde's or Neverver's year was probably shifted by a single month. Avava *boloŋmur* (11th month) and Neverver *nibloŋmur* (1st month) both mean 'leaves fall' and describe the death of the yam vine. The loss of leaves signals the end of the yam growing cycle and the transition to the next year. This transition could be interpreted as the end of one growing cycle, or the month before the new growing cycle begins.

Comparing the Avava calendar with the Ninde one, we can assume that the first months of the Avava calendar happened in the (southern hemisphere) winter. The first three months in Avava have the form *rak*, cognate with Ninde's *ʁaʔ*; both mean 'make a clearing in the forest'. This is most likely the first month or months of the year, and suggests that Neverver's first month *nebloŋmur* was probably an innovation.

Comparing the three languages, we can make some assumptions about the proto calendar's months. The year likely began with one or more months of garden clearing (*rak). This was potentially followed by a time to chop larger saplings and bamboo, and these months begin with a form meaning ‘axe’ (perhaps *ne-mweta ‘blade, knife, axe’, if Ninde’s *netal̪ˤai* ‘axe, clamshell’ was a later semantic extension). The period of planting has forms from *lap ‘plant (yam)’ in all three languages – this was the most important month, and was preserved by all the daughters. There then was likely a period for collecting other things besides yams. Yams are dug out (*kil ‘dig (yam)’), and not pulled out (*rep ‘pull out plant’/*ri ‘extract’). These months have in common (with the exception of the month associated with splitting by ‘axe’) that they may have a month of light activity (Avava and Neverver partitive/‘some of’ *-da*, Ninde *lou[lou]* ‘small’, Neverver *varix* ‘small, insufficient’ and *ran* ‘early’ if it does not mean ‘through the night’), a month of peak activity (Avava *lam*, Neverver *lab*, Ninde *l̪ˤapo~l̪ˤapʷe* ‘big’), and another month of diminishing intensity.

Other months do not seem to be climactic sequences, but only appear once when they do at all. There was likely a single month for the construction of yam platforms, which were built to store and display the yam harvest: Ninde’s *noʙtaʙu* has *noʙo* ‘yam platform’ (and possibly *taʙuwo* ‘second’); Neverver *nibilxebet* (cf. *nebelxa* ‘yam platform’, and perhaps troublingly *ni-bet* ‘breadfruit’). Each of the three languages has months explicitly named ‘month’ or ‘moon’ – this is something peculiar to these languages. These include a ‘useless, ordinary month’ in Avava *ivle roŋo* and Neverver *navulbroŋnaz*, and a ‘red month’ in Avava *ivle mial*. The ‘useless’ months are candidates for an intercalary month, if only for the name, or they could simply be months during which little work is available. Lynch (in Crowley 2006a) speculates that the ‘red month’ was a month of fighting, which would make it similar to Neverver's *niboŋvxal* 'fight month'. In this case, it would also be practical for fighting to be limited to the time after harvest. Finally, the year ended with or shortly after the death of the yam vine, in a month that was likely called *boloŋmur 'leaves fall'.

So far, I have suggested that semantic contrasts in lexical systems reveal unique relationships between forms. Any calendar can borrow from the number system, since they share similar sequential logic (except that the calendrical cycle periodically resets). Relative and intrinsic frames of reference allow for added symmetry, however, since there are the same time points in two temporal directions.

Using an absolute frame of reference – the months of the year and lunar phases – motivates a seven-day week in either direction. The symmetrical week in the intrinsic frame of reference has weekdays associated with numbers and fingers. There are also patterns concerning future days vs. past days. Between the two sequences, there are also pairs of days sharing the same absolute value from the reference point.

Prior to missionization, it is likely that calendars were adapted to local microclimates, as Lynch suggests in Crowley and Lynch (2006a). A balanced ecosystem was critical to human survival before globalization, and it is once again critical Ni-Vanuatu people restore local ecosystems. Month names have the clearest connection to resource use, since they are derived from horticultural activities. Unfortunately, months are more difficult to reconstruct, and therefore also to recover for modern languages. Since months are not as complexly interrelated, all of the intuitions about month names in diachrony come from the yam growing cycle. This helps reconstruct historical change, since it constrains the search space for etyma of unidentifiable morphemes.

There is no available data on calendrical systems of South West Bay (or the other SWB languages). To the extent that the absence of coverage, even in Dimock's sizeable grammar (2007), can be taken as negative evidence, it seems viable that Mewun people brought with them into the mixed language of Ninde a calendar that shared general elements of NCM languages.

7.4 Numeral system

Ninde's numerals show signs of layered change, starting with the restructuring of a PNCV decimal system into one that is quinary for values smaller than ten (Table 84). Ninde is unique among both NCM and SWB languages in having a base *ndumo* for 'six' and *nduma-ne* for 'seven', 'eight', and 'nine', whereas other languages have reflexes of PNCM*zau/PSWB*seu. The form *selme* 'five' (and very rarely *selime*), however, does appear to have a reflex of *zau.

PNCV	Neve'ei	Avava	Neverver	Ninde	Nahavaq	Nevitangiene
*zikai '1'	*sevax*	*-sapm*	*i-sxam*	*sei*	*i-siʔ*	*esuaʔ*
*rua '2'	*rua*	*iru(i)*	*i-ru*	*ʁuwo*	*i-ru*	*erʊʌʔ*
*tolu '3'	*tl*	*itl*	*i-tl*	*tl*	*i-tul*	*etl*
*vati '4'	*ivah*	*ivat*	*i-vas*	*ves*	*i-ves*	*evas*
*lima '5'	*ilim*	*ilim*	*i-lim*	*sel(i)me*	*i-lim*	*elɪmɛ*
*ono '6'	*zou-h*	*sou-t*	*i-zo-s*	*ndumosei*	*(i-)sow-siʔ*	*ɛlʌv-suaʔ*
*bitu '7'	*zu-ru*	*sou-ru*	*i-zo-ru*	*ndumane ʁuwo*	*(i-)sow-ru*	*ɛlavə-rʊʌʔ*
*walu '8'	*zu-tl*	*se-tl*	*i-zo-tl*	*ndumane tl*	*(i-)sow-tul*	*ɛlʌv-tl*
*sivwa '9'	*za-vah*	*sa-vat*	*i-zo-vas*	*ndumane ves*	*(i-)sow-vej*	*elav-vɛs*
*saŋavulu '10'	*naŋavil*	*laŋal*	*naŋavul*	*l̪ˤaŋal*	*(i-)laŋavul*	*ìlʌŋavl*

Table 84: Basic numerals 1-10 in NCM+, SWB, and Nevitangiene (fine phonetic transcriptions).

The identity of Ninde's base for numerals 6-9 is the same form that separates tens from ones and cognate with the separator in at least NCM and SWB (Table 85). However, the fact that Ninde *selme* 'five' seems to contain a reflex of *zau suggests that this was extended by some anticipatory list effect to 'five' before the tens and ones separator *ndumane* was extended to serve as the quinary base. The initial extention of the quinary base to 'five' is shared with Nesarian (only the variety recorded for Wileven) and Angavae (Shimelman et al. 2019) – though they are not well-documented enough to determine how more complex numerals are structured. On the whole, Ninde's numerals have quinary and decimal bases at different orders of magnitude, like the numerals of Avava, Neverver, and Neve'ei. On the other hand, Ninde does not have any vigesimal base for numerals greater than 20, but languages like Nahavaq and Nevitangiene do.

Function	**Neve'ei**	**Avava**	**Neverver**	**Ninde**	**Nahavaq**	**Nevitangiene**
5+	*zV-*	*sV-*	*zo-*	*nduma-(ne)*	*sow-*	*elav-*
10+	*neɒemʷen*	*ɒuman*	*niɒuman*	*ndumane*	*(ni)ndumʷen*	–
10×	*naŋavil*	*laŋal*	*naŋavul*	*nal̪ˤaŋal*	*(i-)laŋavul*	*ìlʌŋavl*
20×	*NA*	*NA*	*NA*	*NA*	*mʷorlalaʔ*	*(nə)magəte*
20+	*NA*	*NA*	*NA*	*NA*	*noʔorond*	–
100×	*naŋat*	*aŋat*	*naŋat*	*nal̪ˤaŋal l̪ˤaŋal ŋoʙo ve-*	(*mʷorlalaʔ ilim* 'five twenties' = 100)	(*magəte lɪme* = 100)
1000×	*netar*	*atar*	*netar*	–	–	–

Function	Neve’ei	Avava	Neverver	Ninde	Nahavaq	Nevitangiene
million×	*namul*	*amul*	*namul*	–	–	–

Table 85: Numeral bases, modifiers, and place separators in NCM+, SWB, and Nevitangiene. Cells with NA are incompatible with either decimal or vigesimal systems used in those languages. Numerical values followed by a plus sign (+) represent forms that are bases for complex numerals, or place separators if there are multiples that do not exceed the order of magnitude. Values followed by a multiplication sign (×) are explicit bases combined with ones quantifiers and are used alongside separators if there are lower orders of magnitude.

Though Ninde has a rare list-effect extension of the quinary base in common with nearby languages, this fact does little to support a genealogical relationship with them. The number system common to NCM is decimal-quinary, where a base of five is merely a historical source of numerals 6-9, but larger values are divided into place values that are exponents of a decimal base. If SWB and Nevitangiene are representative of numerical systems areally, then a mixed vigesimal-decimal-quinary system was widespread in southern Malekula. This is a system where place units are not exponents of a single base, but divided into twenties, then of the remainder, tens, and then fives and ones. Ninde’s initial modification of ‘five’ reveals that it was formerly much like nearly all Malekula languages in having a quinary root *sV-*, but it was not vigesimal like languages of the region, as hundreds were expressed as *nal̪ˤaŋal l̪ˤaŋal ŋoʙo ve-*X ‘ten tens X times’. More documentation work would be needed to assess with a comparative approach whether Nesarian and Angavae share a vigesimal system with their traditional neighbors or have any other similarities with Ninde.

In summary, numerals in Ninde represent a domain where larger areal factors may obscure its relationship with both NCM and SWB subgroups. Like most languages of Malekula, all the NCM and SWB languages have innovated a quinary system over an older decimal one. Unlike SWB languages and Nevitangiene, there is no trace of any vigesimal order of magnitude. Unlike NCM languages, there is no knowledge of simplex words for values greater than one thousand. While this domain is areally rooted in cross-linguistically widespread changes, it is not one that appears to blend NCM and SWB sources.

7.5 Personal Pronouns

Crowley and Lynch (2006a:81) suggested that a compelling diagnostic of language relatedness may be the presence of a generic inflectional verb form that backgrounds the subject. In keeping with the assumption that pronouns represent a lexical core resistant to change, I argue here that Ninde's mixed and innovative pronoun system reflects its status as the product of mixing rather than prolonged contact. Crowley identified Avava, Naman, and Neve'ei as members of a single clade since all have the generic/nonspecific subject, which is passive-like in demoting the subject. Lynch (2008:297) has pointed to a shift from verb-root consonant mutation patterns encoding mood (represented in central Malekula languages) to analyzable mood prefixes as a source of irregularity in sound change. Affinities in the bound and free pronouns support the idea that lexical systems in Ninde tend to have two different origins for lexical forms, semantic oppositions, and patterns of combination.

Both NCM and SWB languages share the locus of mood marking on the prefixes rather than the verb roots, although NCM and Ninde both have irregular verbs that have apparently been influenced by the conditioning environment of the prefixes. The irrealis forms are generally marked with the addition of some partially productive prefix that intervenes between the person-number inflection and the verb stem; this is in contrast to the interior languages, for which Lynch (2008:297, citing personal communication with David Healey) analyzes unmarked irrealis and marked realis forms.

In this section, the inflectional paradigms are presented by the conventions described as follows. The bound pronouns are considered in §7.5.1 and the independent pronouns in §7.5.2. Within each of these two subsections, the predominant forms are represented across three tables: singular, dual (as applicable), and plural. The trial/paucal pronouns are not discussed, nor are other described independent pronoun types: indefinite, negative, reflexive, and relative pronouns. Where the forms are presented, typically those chosen as the main representation by the grammar writers have been chosen. Capitalized letters represent vowels (archiphonemes) susceptible to harmony, generally with the first vowel of the verb stem; most often this is /e/ that can become rounded [o].

The nature of Ninde's inherited pronouns is complex. Syncretization of realis and irrealis forms has worked in both directions, but has affected all persons in Ninde. The forms of both bound and independent prefixes seem to blend together morphemes of NCM origin with those of SWB origin.

7.5.1 Bound Pronouns

For singular bound pronouns (shown in Table 56), Ninde is most similar to NCM languages in the ways discussed in §5.4.1, but not without SWB influence (§5.4.2). The translated gospels (discussed in Chapter 1) inflect verb forms for second-person singular subjects with *i-*, but this is no longer present in Ninde. Naman, like Ninde, has lost this morpheme, but this appears to be a parallel development. Both languages have word-initial sound change best explained as palatalization after the conditioning environment of the /i/ of the erstwhile prefixed form. For instance, PNCM *s̲ak 'climb, ascend' yielded Ninde *ja̰ʔ* and Naman *s̲iəx* 'go up'. Naman has fused the third-person prefix to monosyllabic verb roots across the inflectional paradigm (Crowley 2006:53-54); it is likely that Ninde and NCM did the same (see §5.3.1 for evidence that "long prefixes" in Avava and Neverver incorporate 3rd-person singular *i-).

The pronouns in Ninde are generally complex in their variation, and especially so for the second person. There is some disagreement between Ninde speakers about second-person singular inflection of verbs when they are represented in writing. Some speakers produce the verb with no bound form, but with an obligatory form identical to the independent pronoun *nuŋg* – this resembles the grammaticalization of the independent pronoun *oŋ* into the realis prefix *o(ŋ)-* in Avava. Otherwise, the realis form of Ninde resembles what is generally the realis form of other languages, but Naman, Neverver, and Ninde suggest an initial contrasting velar was once present (most likely *x that underwent fortition). In Ninde song, the second-person singular *u-* is attested in *<u>u</u>vendor mene* 'where did you go', suggesting that it was the irrealis prefix that was generalized in Ninde. This part of the inflectional paradigm may be subject to layered stages of renewal – perhaps because of the rarety of inherited word-initial vowels – and Ninde appears to maintain every stage of renewal.

Ninde is unique in having a subjunctive prefix *ka-* used for sequential events and verbs of complement clauses that share a subject with the matrix clause. This verb form appears to be a repurposed cognate of the third-person singular irrealis subject prefix of Naati *ʔa-* and Nahavaq *kE-*, which is unique to SWB languages. The function is inconsistent with this source, since the realis *ka-* has an irrealis counterpart *ka-p-* in Ninde. Verb forms with *ka(p)-* are not serial verb constructions, which by contrast always have an inflectional *p-* (homophonous with the irrealis prefix) and no person marking in non-initial verbs. The use of irrealis for non-initial serialized verbs is a grammatical feature of all NCM and SWB languages; perhaps the SWB *ka- survived in a subset of non-finite verb forms of Ninde before expanding in usage[33].

[33] Some speakers also use finite *ka-* prefixation for first-person singular inflection, but this is underrepresented in our documentation and we do not understand the extent of its usage.

All irrealis verb forms in Ninde take an analyzable prefix of the form *p-* after the subject prefix. This form is *pe-* if *p-* would produce an illicit consonant cluster, *po-* if the first vowel of the verb root is round, or *m-* before a nasal. Many verbs that begin with *i, j, w,* or *v* undergo replacement of the initial segment in the irrealis form; this includes the verb *ve* 'go' (irrealis *pe*) which can form compounds in the realis. Compounding with the irrealis verb *pe* is a likely source for the person-general irrealis prefix in Ninde and for the third-person irrealis forms of the other NCM languages.

Grammatically, the bound pronouns are patterned vis-à-vis other pronouns more like the other SWB languages. Naati and Nahavaq, like Ninde, have irrealis prefixes, *a-* and *ʔ-* respectively in non-singular forms. These follow the subject pronouns and unmarked verbs are realis. The singular paradigm in Naati can be partially explained as an extension of an *a-* prefix via fusion to the singular subject pronouns. In this way, Ninde behaves most like Naati in terms of the order and function of inflectional prefixes, and for non-singular forms, it behaves like Nahavaq.

To generalize, NCM and SWB languages all shifted the locus of mood marking from the verb stem to the person-marking prefixes. The more widespread, and therefore presumably older, phenomenon in Northern Vanuatu languages is one whereby an unmarked verb is irrealis, and initial consonants undergo predictable mutation to inflect verbs for realis (Lynch 2008). Both NCM and SWB languages developed separate paradigms for realis vs. irrealis singular pronouns, but SWB developed irrealis prefixes separate from the verb stem and plural bound pronouns. Naati and Ninde seem to have extended the irrealis prefix to singular inflection; Ninde and the other NCM languages appear to have generalized compounding forms of 'go' to build up parts of their irrealis paradigm.

Person	Naman	Neverver	Avava	Neve'ei	Ninde	Naati	Nahavaq
1st	*nə-* *bə- ~ ba-*	*ni-* *nim-*	*na-* *nabʷ(e/i)-*	*nV-* *nebʷe- ~ nabʷi-*	*na- ~ ka-*	*ni-* *na-*	*ne-* *(ni)ŋgE-*
2nd	*xə-* *kə-*	*ku-* *kum-*	*o(ŋ)-* *ki-~ke-*	*u-* *ku-*	*ku- ~ nuŋg* *Ø-*	*u-* *wa-*	*u-* *ku-*
3rd	*Ø-* *bə-*	*i-* *im-*	*i-~e-* *(e)bʷV-*	*(i-)* *bʷe/i-*	*Ø-*	*i-* *ʔa-*	*i-* *kE-*
Subjunctive					*ka-*		

Table 86: Singular Bound Pronouns in NCM+ (left) and SWB (right) languages. The top line within each cell represents the realis form; the bottom represents the irrealis.

Across the language families, bound dual subject pronouns (Table 87) and plural subject pronouns (Table 88) are related to each other in a principled way. In NCM, dual forms end in /r/ (and plural forms in /t/), with epenthetic vowels in some cases according to phonotactic constraints. In Naati, dual forms also end in /r/, but the plural forms end instead in /l/. Nahavaq and Ninde have basic plural forms; in Nahavaq, the vowel of the plural prefix is replaced with /u/, and in Ninde the /ʀ/ ⟨x⟩ is added before or after the prefix and the vowel is correspondingly lowered to [a]. Since these forms are generally derived from the plural prefixes, the bases are discussed in the following praagraph. The addition of *r as well as the vowel replacement by *u could both reflect an older addition of a reduced form of PNCM/PSWB *rua 'two', which surfaces as *-ru in the independent pronouns (see §7.5.2).

	Naman	**Neverver**	**Avava**	**Neve'ei**	**Ninde**	**Naati**	**Nahavaq**
1st exclusive	mər- ᵐbər-	nar(i)- namr(i)-	ar(i)- (a)ᵐbʷir(i)-	er- ᵐbʷur-	xa-	ŋgar-	mu-
1st inclusive	tər(a)- tər-	nir(i)- nimr(i)-	ar(i)- (a)ᵐbʷir(i)-	er- ᵐbʷur-	tax-	ndar-	ndu-
2nd	xər- kər-	kar(i)- kamr(i)-	ar(i)- (a)ᵐbʷir(i)-	ar- aᵐbʷ(e~i)r-	xa-	mwar-	wa-
3rd	rə- ᵐbər-	ar(i)- amr(i)-	ar(i)- (a)ᵐbʷir(i)-	ar- aᵐbʷ(e~i)r-	rax-	ar-	ru-

Table 87: Dual Bound Pronouns. The NCM languages have realis forms (top of cell) and irrealis forms (bottom).

The plural bound pronoun paradigm (shown in Table 88), and thus the base for the dual forms, suggests greater influence instead from SWB languages on Ninde and especially the otherwise more distant Nahavaq. On the whole, a number of innovations are grammaticalized independent pronouns – these are discussed in greater detail in following §7.5.2 – or they may be reflexes of PNCM *mʷer 'person' or PSWB *mʷor 'man' (which have no known reflexes in Ninde). From this source, Naman may have *mər-* 1st person dual exclusive, and Naati has *mʷar-* 2nd person dual, with plural forms formed on analogy.

Only NCM languages maintain a distinction between the generic subject pronoun and the third-person plural subject form. Like Nahavaq, Ninde seems to have a third-person plural subject prefix that is cognate with the generic subject of NCM. Dimock (2009:259-260) explicitly identifies this as a polysemous form in Nahavaq, also serving the function of nonspecific subject. The third-person plural *re-* is idiosyncratic in Ninde, where the plain alveolar trill is expected to correspond to a *uvular* trill or fricative. This prefix could be a late borrowing, or it could be the result of rhotic dissimilation after the independent pronoun *naʁa*[34]. Ninde has innovated a new generic subject form *e-* by extending the already syncretic form for exclusive 1st person and 2nd person plural subjects. This unique innovation is used by some for citation forms of verbs and (parallel to the generic subject in NCM) passive-like constructions. Notice in the usage exemplified in (1), the interpretation of a second-person plural or first-person exclusive plural is incompatible with the meaning:

[34] This is a regular and predictable process that occurs across word boundaries, but it is not known to happen when there is an intervening vowel.

Example from unrecorded interaction

(1) e-veu kəne ... Ninde

NSP-give.birth.to 1SG ...

'I was born...'

	Naman	**Neverver**	**Avava**	**Neve'ei**	**Ninde**	**Naati**	**Nahavaq**
1st exclusive	mət- ^{m}bət-	nat(i)- namt(i)-	(kom)at(i)- (a)^{m}b^{w}it(i)-	it- ^{m}b^{w}it-	e-	ŋgal-	?
1st inclusive	tət- tət-	nit(i)- nimt(i)-	at(i)- (a)^{m}b^{w}it(i)-	it- ^{m}b^{w}it-	te-	ndal-	ndE-
2nd	xət- kət-	kat(i)- kamt(i)-	at(i)- (a)^{m}b^{w}it(i)-	at- (a)^{m}b^{w}it-	e-	mwal-	a-
3rd	at- ^{m}bət-	at(i)- amt(i)-	at(i)- (a)^{m}b^{w}it(i)-	at- a^{m}b^{w}it-	re-	al-	re-
Generic	rə- rə^{m}bə-	ar(i)- a^{m}(bi)-	ra-	re- rV^{m}b^{w}V-	e-	?	re-

Table 88: Plural Bound Pronouns. The NCM languages have realis forms (top of cell) and irrealis forms (bottom). The first-person plural exclusive form is not described for Nahavaq, and the generic subject is not described for Naati.

To summarize, the bound pronouns in Ninde represent a blend of paradigms from SWB and from NCM. Like SWB, however, pronominal prefixes are separate from mood prefixes across the board and the dual pronouns are made up of plural counterparts followed by a dual prefix. Like many isolated pronouns in NCM languages, however, the realis and irrealis are syncretistic – this has been extended to the entire paradigm.

7.5.2 Independent pronouns

On the whole, independent pronouns (Table 89) in Ninde are intermediate in complexity between NCM and SWB pronouns; that is, they appear to "add" a morpheme to stems similar to those found in NCM, but "remove" one that is found in SWB. Like the other SWB languages, Ninde has an initial *n-*, perhaps the same nominal *nV-* prefix found on virtually all lexical nouns, on the second-person singular pronoun. Ninde has uniquely extended this prefix to the third-person singular pronoun *nije*. Unlike SWB, however, Ninde does not have an added *-ŋg* in the first-person singular pronoun – this could be extended from the second person, but it is also one of the possible first-person singular possessor suffixes in all three SWB languages. The independent pronouns generally suggest the Ninde paradigm has a closer relationship to SWB.

Person	Naman	Neverver	Avava	Neve’ei	Ninde	Naati	Nahavaq
1st	*kine*	*(i)na*	*na*	*no*	*kəne*	*kinaŋg*	*(ki)naŋ(g)*
2nd	*a(xu)ⁿg*	*(i)ox*	*oŋ*	*ⁿgu*	*nuŋg*	*inuŋg*	*(i)nuŋ(g)*
3rd	*ai~i*	*e͡i* *maŋ* ‘he’ *vinaŋ* ‘she’	*e~ei*	*i*	*nije*	*ʔe͡i*	*ai~i*

Table 89: Independent Singular Pronouns. Only Neverver has innovated gendered third-person pronouns, not differentiated by grammatical case.

Of the dual forms (Table 90), Ninde is most similar to Naman and Neve’ei, but also shares key similarities with Naati but not Nahavaq. Neverver stands out as having noncontrastive nonplurals (i.e., the duals are syncretic with the plural forms). Most NCM languages suggest -*(a)ru* or -*ⁿɽu* ‘two’ (probably reflexes of a frozen realis and irrealis pair from PNCV*rua ‘two’ with a regular consonant mutation) is suffixed to the plural form to yield an independent dual pronoun. Ninde and Naati support innovation on an older -*ⁿɽu*, despite the absence of a prenasalized portion; the missing final nasals of the Naati roots could be explained as regular coalescence of a nasal stop followed by a prenasalized stop, with regressive place assimilation. Ninde shares with Naman and Naati the absence or loss of the prenasalization in the dual suffix: only *r yields Ninde ⟨x⟩ /ʁ~ʀ/. The third-person dual pronoun of Ninde is unique in the ordering of prenasalized and plain rhotic reflexes. Neve’ei has both rhotics, but in reverse order, and Naati and Nahavaq have only plain rhotics. The

Ninde form could simply be extended from the bound subject prefix, which unlike all other pronouns is homophonous with the independent form.

	Naman	**Neverver**	**Avava**	**Neve'ei**	**Ninde**	**Naati**	**Nahavaq**
1st exclusive	*kam(em)(a)ru*	*(i)nam~(ᵑgu)mam*	*kopⁿɽu*	*ᵑgememru*	*kamaʁ*	*ŋgaru*	*(nuʔu)mʷem*
1st inclusive	*(i)ᵑg(e~ə)taru*	*(i)ᵑgit*	*ᵑgitⁿɽu*	*ᵑgetⁿɽu*	*kətaʁ*	*ndaru*	*nuʔun(d)*
2nd	*(i)ᵑg(e~ə)m(a)ru*	*(i)ᵑgam*	*kamⁿɽu*	*ᵑgemru*	*kamoʁ*	*amuru*	*nuʔum*
3rd	*(r)aru*	*aⁿɽ~ar*	*ierⁿɽu*	*arⁿɽu*	*raʁ*	*raru*	*(ru)war*

Table 90: Independent Dual Pronouns. Neverver does not differentiate non-singular (dual vs. plural) forms in the independent pronouns.

The plural forms (Table 91) of Ninde are more like NCM pronouns, but Naati has clearly innovated on formerly more similar forms. All of the Naati plural pronouns bear a plural suffix *-lyl*, likely cognate with the second part of Ninde *wundil* 'all together' used as a post-pronominal quantifier. The final nasals of other languages are curiously missing in the Naati forms. In the other languages, dual forms appear to be derived from plural forms, but Naati could have regularized plural roots on analogy with the dual set. Once again, Ninde has a rather innovative form in *naʁa* 3PL; like the singular counterpar, a nominal *nV-* was prefixed to the third-person plural pronoun, but this could have been like Neverver and Neve'ei *ar* with a final vowel inserted, or it could have been more like Naati *rolyl* (cf. dual *rar* for what the form may have been without a round vowel context) and added to a root *ra. The final vowels of Ninde *kəte* 1PL.INCL, *naʁa* 3PL, and *nuwute* PLACE (an obligatory pronoun used generally in constructions for environmental conditions like *nuwute sousou* 'it is humid', but also as an object in *ŋor nuwute* 'oink') support affinity with the NCM forms. Final (likely devoiced) continuants and voiceless stops of the NCM source generally conditioned word-final vowel insertion.

	Naman	**Neverver**	**Avava**	**Neve'ei**	**Ninde**	**Naati**	**Nahavaq**
1st exclusive	*kamem*	*(i)nam~ (ᵑgu)mam*	*kopm*	*ᵑgemem*	*kamem*	*ŋgalyl*	*kamem*
1st inclusive	*(i)ᵑg(e~ə)t*	*(i)ᵑgit*	*ᵑgit*	*ᵑget*	*kəte*	*ndalyl~ ndolyl*	*niŋgin(d)*
2nd	*(i)ᵑg(e~ə)m*	*(i)ᵑgam*	*kam*	*ᵑgem*	*kamu*	*amulyl*	*ni,ŋgim*
3rd	*a͡ir*	*aⁿɽ~ar*	*ier*	*ar*	*naʁa*	*rolyl*	*ʔar*
PLACE	–	*kut*	*out*	–	*nuwute*	*naut*	*naut*

Table 91: Independent Plural Pronouns. The place pronouns are required as subjects or objects for some predicates relating to environmental conditions.

Once again, in the domain of independent pronouns, the lexical origins cannot be solidly linked to only NCM or SWB. The singular pronouns resemble SWB more closely, but the non-singular sets resemble NCM, and especially Naman and Neve'ei.

7.5.3 Renewal Cycles

Dependent pronouns, especially the plural forms, are often subject to rebracketing with or replacement by the independent pronouns. The attested Avava variant prefix for first-person exclusive plural *komat(i)-* incorporates the independent pronoun *kopm* and the corresponding bound pronoun *at(i)-*. Similarly, the second person singular form in Avava is *o-* or *oŋ-*; while the former could be an irregular realization of *u-*, the latter is identical to the independent pronoun. These two forms perhaps influenced each other.

An older layer of relationship suggests that independent pronouns and bound pronouns were initially not so distinct. Closely related languages differently derived the independent pronouns by prefixing either *i-* or *na-* to the same morphemes that provide the bound prefixes. These derivational prefixes could correspond to the third-person subject *i-* and the noun prefix *na-*. If this is the case, second-person singular *na-ku could have yielded Ninde *nuŋg*, Avava *oŋ*, and Neve'ei *ᵑgu*, whereas Neverver *(i)ox* would have opted for a non-nasal prefix and retained an oral *ku (with some inserted vowel). Perhaps both forms coexisted in the proto language, with *na- forms corresponding to true pronouns and *i-forms initially functioning as predicates.

In Ninde, there is great variation in rebracketed bound prefixes that apparently take on phonetic material from the independent pronouns that optionally precede them. One elderly speaker, Kaitipbuas Saaobal, can be heard on recordings (053 and 182) producing *kame-* for first-person exclusive verb forms; this is apparently a reduced form of *kamem e-* which approaches the level of erosion in the corresponding nonsingular root *ŋga-* of Naati. Generally, the second-person singular forms with an overt *nuŋg* could be *nuŋ(k) ku-* with reduction of the second vowel, and the second-person form is often heard in the post-verbal object position as *nuŋ*. The third-person singular *nije* is often reduced to *ni* before verbs and the irrealis *p-* may be realized as the syllable coda of the independent pronoun. The apparently great variation within Ninde is most likely an illusion of field work with many speakers across generations. With the exception of Neverver, the documentation of other languages was largely undertaken with very few speakers, sometimes as few as one.

7.6. Conclusions

Comparison of lexical systems rarely shows a simple story of inheritance in Ninde. In every system, there are similarities between Ninde and NCM and SWB in form, function, and grammatical patterning for individual forms, but these are rarely shared wholesale with just one set of languages. This could represent a situation where complexities of each system were creatively combined from the two main available sources.

A great deal of innovation or unique retention also sets Ninde apart from its relatives. While these may represent a yet to be identified contact language, it is just as likely that these are independent innovations. These rarely introduce complexity that is not shared with another language. Whatever the circumstances for these changes, it no doubt contributes to the general perception that Ninde is unlike other languages of Malekula. It is tempting to consider that language mixing and innovation may have been deliberate means of asserting a new and distinct identity of mixed-Ninde speakers.

Future work could address the relationship of Ninde with languages that have been named here as viable sources of borrowing (as distinct from the mixing process). Nesarian and Angavae are poorly documented and endangered, but they share a rare innovation in the numeral 'five' with Ninde. Based on the location of their traditional villages, they would have certainly been in contact with Ninde, especially if there had been a relatively direct trajectory from the regions of northern central Malekula to South West Bay by speakers of a source language. Ninde and Neve'ei also share many unusual lexical idiosyncracies, and incidentally both are complicated fits within the NCM language group.

Additionally, these semantic domains have a special importance to community members. They are often learned together in pairs, groups, or sequences, as they represent taxonomies. Linguistic practices observed in the Ninde-speaking community involve listing numbers from 1-10, or reciting pairs of relational antonyms from the kinship system. The work of this chapter also advances community interests in language maintenance and revitalization.

8 Conclusion

8.1 Main Findings

Ninde's status as a mixed language has been described by its speakers and the comparative method can shed light on the nature of that mixing. Furthermore, lexical systems may motivate perpetuation of semantic categories in the face of lexical innovation. The main source languages were most likely one that was in a language family with at least Naati and Nahavaq (SWB) and another that was in a language family with Neverver and Avava (NCM), bearing important similarities to Neve'ei and Naman (included here as NCM+), which lack some of the innovations shared by the other three NCM languages. Even without clarity about these relationships, Ninde can be characterized as a mixed language by comparing it with these two sources, even if NCM and SWB are certain to undergo reassessment as viable subgroups.

Although Ninde is included in Lynch's Western Malekula Linkage, there is more reason to include the languages of nothern Malekula in the same genetic and areal relationships. Prior work has discounted the innovation of linguolabial consonants as a diagnostic sound change, but this line of reasoning only considered feeding and blocking relationships in ordered sound changes. There is enough evidence that linguolabials assimilate and dissimilate to other coronal and labial consonants (including each other) to reassess the hypothesis of multiple innovations by contact. Ninde may share some consonant deletion patterns with V'ënen Taut, but this is more likely to represent chance parallel and independent innovation than far-flung areal relationships.

In terms of grammar, Ninde has several fundamental features in common with both of these sources. It shares with neighboring SWB languages many structures of negation, the general ordering of verbal prefixes, flexible nominalization of complex verbs, and a robust pattern of consonant mutation for realis-irrealis stem pairs (albeit one renewed by recent sound change). These features appear overlaid on perhaps older, frozen patterns shared with NCM+ languages: fusion of the third-person singular subject *i-* to monosyllabic verb roots, a generalization of the irrealis form *pe* 'go' to a general irrealis *p-*, aspect and mode constructions, determiner affixes on nouns, and irregular negative verb forms. Though the sources of this mix, commonly called Small Ninde and Big Ninde, would certainly have been related to each other and in contact, the result of that contact is a language with recognizable core properties from two different sources, which at times are blended in form, function, and grammar.

The phonology of most of Ninde's lexicon can be reconstructed with considerable success from models of NCM and SWB diachrony. The assumption of fundamentally exceptionless sound changes affected two subsets of the lexicon in parallel allows for the identification of individual words as either NCM or SWB in origin. Where there are exceptions, they can often be attributed to rebracketing changes linked to inflection that is in line with word-class, as well as variation attested in reduplicated forms. The substantial changes that have long obscured Ninde's diachronic phonology can be connected (presumably by areal diffusion) to yet a third set of languages centered on Lendamboi, with rampant devoicing of sonorous speech sounds. This phenomenon has altered Ninde's syllable structure and vowel quality and contributed a new phoneme: the pharyngealized lateral. The restructured phonological system, alongside bidirectional calques and doublets, have meant that the time depth of Ninde's divergence from other Malekula languages has been exaggerated, even as scholars have observed its shared innovation with disparate neighbors.

The same generalization holds for lexical systems. The way that semantic oppositions of kinship, geography, and time are structured generally favors other NCM languages, but it is often cognates with SWB languages that are evidently reassigned to new semantic oppositions. Blending of these systems has increased their complexity and Ninde-specific innovation has often rendered patterns (like formerly numbered days of the week) opaque. No single system is completely in line with simply just the NCM languages or the SWB languages, but they do make Ninde both familiar and eccentric in comparison with other Malekula languages.

8.2 Resulting Materials and Dissemination

The outcome of this work can be of use to various invested parties. Historical linguists may challenge the reconstructions offered here, both as a result of deeper understandings of phylogeny and contact, or in the face of new findings and orderings of change. Speakers of relevant languages may find historical hypotheses to be interesting or meaningful if they are made available in the form of etymological dictionaries. Some engaged in revitalization and language planning may opt to establish expected reflexes in modern language use, despite the inherent risk of innaccuracy. These expected forms can reconstitute the phonology of poorly documented forms or even recover unattested ones.

Corpus materials can be of use to community members and researchers, and will be available in at least two locations. The currently most accessible repository has been deposited with the Endangered Languages Archive (Crouch 2018). In accordance with the research permits we have held for work in Vanuatu, copies of all primary materials are also held by the Vanuatu Kaljoral Senta (Vanuatu Cultural Centre). We are committed to making these readily available in perpetuity.

For researchers, cognate sets and reconstructed proto forms will be made available in the standards of historical linguistics (Forkel et al. 2018). Efforts to standardize format have included the use of the International Phonetic Alphabet, restricting values to one per cell, including many information fields to allow for sorting along many parameters. The data and analyses, including word alignments, cognate judgments, reconstructed protos, and expected vs. attested reflexes, will all be published in the Cross-Linguistic Data Format (CLDF) specification on Zenodo.

Additionally, the software developed for this task will be available for other scholars to use on GitHub. This may allow for the deployment of diachronic sound change models en masse for entire lexicons. This could be a valuable tool for pedagogy, data exploration, and also dissemination of information about language families that respects the linguistic privacy of minoritized communities – for example, to demonstrate linguistic diversity to laypeople. This software can also be used alongside existing packages in Python like LingPy

For communities, predicted reflex lists may be of greater value. Though there is a potential for contamination of language use with inaccurate diachronic predictions, this can be mitigated in important ways. For forms that have multiple possible protos, due to insufficient comparative data in the face of mergers or unorderable sound changes, multiple expected reflexes can be linked in sets. Reconstitution of forms may be most effective in field work situations where a language is spoken, but much of the vocabulary lies at the boundaries of lexical recall. These forms can be disseminated in community orthographies wherever available, with the added benefit that it includes the lexicography that linguists have already undertaken. Font effects like shading can help to visually differentiate actual attested words from predicted ones. This is already being done for the Ninde data.

8.3 Future Direction

The work of this dissertation has uncovered areas for future development of auxiliary methods for computational reconstruction. Some of these, like large-scale reconstruction by computer and a focus on lexical systems, put communities first by addressing knowledge domains that are often priorities: vocabulary that is unlikely to come up in corpora and by elicitation, and that often serves taxonomic functions. Other innovative methods, like using decision trees to uncover complex interactions between assimilation and dissimilation, help to lend structure where sound change may appear to be fundamentally irregular. Like the findings they helped to uncover, all these methods face a long road to improvement and expansion.

The relationship of Neve'ei and Naman with Avava, Neverver, and Ninde also requires some work. While Neve'ei shares with Ninde unique innovations in its pronouns and kinship terms, it does not prove to be a powerful predictor of the lexicon at large. There are also superficial similarities in the two languages in debuccalization of both *s and *k, but these changes differ in terms of the phonetic environments in which they occur. With respect to NCM as a clade, Neve'ei lacks many of the innovations characterizing Avava, Neverver, and Ninde. Further work may be able to differentiate shared inheritance from language contact.

This work may serve to drive some discussion of other mixed languages that may exist in Vanuatu. Facing similar pressures to combine once-distinct communities, language mixing could generally have offered some way to negotiate linguistic relationships within combined communities with new shared identities, as Crowley (2016a) has suggested for Avava. Such work must make genealogical abstractions without committing to a model of singular linguistic parentage. Only with clear predictions about language change can exceptions be identified and characterized.

The tools developed for this dissertation also have much room for improvement. At present, extremely fine-grained rules are necessary to reduce computational complexity and generalize natural classes of phones and abstract environments for the reconstruction algorithm. Future development must incorporate phoneme inventories at the level of daughter languages and proto languages. Inventories will then be updated with each sound change on the basis of whether the rule is conditioned and what phonotactic environments are present in the lexical data. Additionally, parsing of phonetic sequences will be improved to represent multiple layers of morphophonological structure: syllable-, morpheme-, and word-boundaries. With these capabilities, sound changes can be formulated using more abstract phonological features, rather than specifying the range of anticipated input and output phones in the model.

REFERENCES

Arsenault, Paul. (2006, February). The adaptation of English alveolar stops in Telugu and Hindi. In *Handout from the Montreal-Ottawa-Toronto (MOT) Phonology Workshop* (Vol. 2, p. 2006).

Auer, Peter, and Hakimov, Nikolay. (2021). "From language mixing to fused lects: The process and its outcomes." *International Journal of Bilingualism* 25.2: 361-368.

Baerman, Matthew. (2014) "Suppletive kin term paradigms in the languages of New Guinea." *Linguistic Typology* 18.3: 413-448.

Bakker, Peter. (2015). Typology of mixed languages. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), The Cambridge Handbook of Linguistic Typology. Cambridge: Cambridge University Press.

Barbour, Julie. (2011). Contour segments in Neverver: Synchronic justifications. *Te Reo*, *54*, 3.

Barbour, Julie. (2012). *A grammar of Neverver* (Vol. 60). Walter de Gruyter.

Bedford, Stuart; Spriggs, Matthew; Regenvanu, Ralph; and Yona, Salkon. (2011). 19. Olfala Histri Wea i Stap Andanit long Graon. Archaeological Training Workshops in Vanuatu: A Profile, the Benefits, Spin-offs and Extraordinary Discoveries. *Working Together in Vanuatu*, 191.

Bedford, Stuart and Spriggs, Matthew (2014). The archaeology of Vanuatu: 3000 years of history across islands of ash and coral. *The Oxford handbook of prehistoric Oceania. Oxford University Press, Oxford. doi. org/10.1093/oxfordhb/9780199925070.013*, *15*.

Berez-Kroeker, Andrea L.; Gawne, Lauren; Smythe Kung, Susan; Kelly , Barbara F.; Heston , Tyler; Holton , Gary; Pulsifer , Peter; Beaver , David I.; Chelliah , Shobhana; Dubinsky , Stanley; Meier , Richard P.; Thieberger , Nick; Rice, Keren; and Woodbury, Anthony C. (2018). "Reproducible research in linguistics: A position statement on data citation and attribution in our field." *Linguistics* 56.1: 1-18.

Blevins, Juliette. (2018). Evolutionary phonology and the life cycle of voiceless sonorants. *Typological hierarchies in synchrony and diachrony*, 29-58.

Blust, Robert. (1996). "The Neogrammarian hypothesis and pandemic irregularity." *The comparative method reviewed: Regularity and irregularity in language change*: 135-56.

Blust, Robert (1999). Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In *Selected papers from the eighth international conference on Austronesian linguistics* (Vol. 1, pp. 31-94).

Blust, Robert (2013). *The Austronesian languages*. Asia-Pacific Linguistics, School of Culture, History and Language, College of Asia and the Pacific, The Australian National University.

Blust, Robert and Trussel, Stephen (2013). The Austronesian Comparative Dictionary: A Work in Progress. *Oceanic Linguistics* (Vol. 52, No. 2).

Blust, Robert, and Trussel, Stephen. (2020). "Austronesian comparative dictionary." Web resource. <https://www.trussel2.com/ACD/introduction.htm>

Bowern, Claire. (2006). Punctuated equilibrium and language change. *Encyclopedia of Language and Linguistics*, 286-289.

Bowern, Claire, and Koch, Harold. (2004). Eds. *Australian languages: Classification and the comparative method*. Vol. 249. John Benjamins Publishing.

Brayboy, Bryan McKinley Jones. (2021). "Tribal critical race theory: An origin story and future directions." *Handbook of critical race theory in education*. Routledge: 191-202.

Brotchie, Amanda. (2009). *Tirax grammar and narrative: an Oceanic language spoken on Malakula, North Central Vanuatu* (Doctoral dissertation).

Capell, Arthur (1954). *A linguistic survey of the south-west Pacific.* Nouméa: South Pacific Commission.

Capell, Arthur. (1962). Oceanic linguistics today. *Current anthropology*, *3*(4), 371-428.

Charpentier, Jean-Michel. (1982). *Atlas Linguistique Du Sud-Malakula-Linguistic Atlas of South Malakula (Vanuatu)* (Vol. 1). Peeters Publishers.

Clark, David. (2009). **Leo tuai: A comparative lexical study of North and Central Vanuatu languages*. Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Clements, J. Clancy; Amaral, Patrícia; and Luís, Ana R. (2011). "Spanish in contact with Portuguese: The case of Barranquenho." *The handbook of Hispanic sociolinguistics*: 395-417.

Codrington, Robert Henry (1885). *The Melanesian languages*. Clarendon Press.

Comrie, Bernard. (1985). *Tense*. Vol. 17. Cambridge university press.

Comrie, Bernard. (2000). Language contact, lexical borrowing, and semantic fields. *Studies in Slavic and General Linguistics*, *28*, 73-86.

Crowley, Terry. (1985). Common noun phrase marking in Proto-Oceanic. *Oceanic Linguistics*, *24*(1/2), 135-193.

Crowley, Terry. (1990). *Beach-la-Mar to Bislama: The emergence of a national language in Vanuatu*. Oxford University Press.

Crowley, Terry. (1998). A salvage sketch of Nāti (southwest Malakula, Vanuatu). *Pacific Linguistics. Series A. Occasional Papers*, *92*, 101-148.

Crowley, Terry. (2000). The language situation in Vanuatu. *Current Issues in Language Planning*, *1*(1), 47-132.

Crowley, Terry. (2001). The indigenous linguistic response to missionary authority in the Pacific. *Australian Journal of Linguistics* (Vol. 21, No. 2).

Crowley, Terry. (2002). *Serial Verbs in Oceanic: A Descriptive Typology*. New York: Oxford University Press.

Crowley, Terry and Lynch, John. (2006a). *The Avava language of Central Malakula (Vanuatu)*. Pacific linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Crowley, Terry and Lynch, John. (2006b). *Naman: A vanishing language of Malakula (Vanuatu)*. Pacific linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Crowley, Terry and Lynch, John. (2006c). *Nese: A diminishing speech variety of Northwest Malakula (Vanuatu)*. Pacific linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Crowley, Terry and Lynch, John. (2006d). *Tape: A declining language of Malakula (Vanuatu)*. Pacific linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Deacon, Arthur Bernard and Wedgwood, Camilla Hildegarde. (1934). *Malekula: a vanishing people in the New Hebrides*. Routledge.

Deacon, Arthur Bernard and Wedgwood, Camilla Hildegarde. (1934). Geometrical Drawings from Malekula and Other Islands of the New Hebrides. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland, 64*, 129-175. doi:10.2307/2843952

Dempwolff, Otto. (1938). *Vergleichende lautlehre des austronesischen Wortschatzes*. D. Reimer.

Dempwolff, Otto. (1969a). *Vergleichende lautlehre des austronesischen wortschatzes. 3. austronesisches wörterverzeichnis*. Kraus.

Dempwolff, Otto. (1969a). *Die Lautentsprechungen der indonesischen Lippenlaute in einigen anderen austronesischen Südseesprachen* (Vol. 2). Kraus Reprint.

Dench, Alan. (2000). Comparative reconstitution. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, 57-72.

Dimock, Laura G. (2009). A grammar of Nahavaq (Malakula, Vanuatu).

Dixon, Robert M.W. (1997). *The rise and fall of languages*. Cambridge University Press.

Dodd, Royce. (2014). *V'ënen Taut: Grammatical topics in the Big Nambas language of Malekula.* Hamilton, New Zealand: University of Waikato (masters thesis).

Donohue, Mark and Denham, Tim. (2008). "The language of Lapita: Vanuatu and an early Papuan presence in the Pacific." *Oceanic Linguistics* 47.2: 433-444.

Dwyer, Arienne M. (2006). "Ethics and practicalities of cooperative fieldwork and analysis." *Essentials of language documentation* 178.

Dyen, Isidore, and David F. Aberle. (1974). *Lexical reconstruction: The case of the Proto-Athapaskan kinship system*. Cambridge University Press.

Ersheidat, Ghayeth, and Hafsa Tahir (2020). "Genetic relationship among languages: An overview." International Journal of Language Education and Applied Linguistics: 17-27.

Fixico, Donald L. (1983). The Native American researcher: Another view of historical documents. *The Midwestern Archivist*, *8*(2), 5-15.

Forkel, Robert; List, Johann-Mattis; Greenhill, Simon J.; Rzymski, Christoph; Bank, Sebastian; Cysouw, Michael; Hammarström, Harald; Haspelmath, Martin; Kaiping, Gereon A.; and Gray, Russel. D. (2018). Cross-Linguistic Data Formats, advancing data sharing and re-use in comparative linguistics. *Scientific data*, *5*(1), 1-10.

Fox, Greg J. (1979a). *Big Nambas grammar*. Department of Linguistics, Research School of Pacific Studies, The Australian National University.

Fox, Greg J. (1979b). Big Nambas Custom Texts. *Journal de la Société des océanistes*, (vol. 35, no. 65) pp. 286-293.

Fox, Helen. (1996). An honorific sub-dialect used among Big Nambas women. *J. Lynch and Fa'afo Pat (eds) Oceanic Studies, Canberra: Pacific Linguistics C-133*, 367-73.

François, Alexandre. (2002). *Araki: A disappearing language of Vanuatu*. Vol. 522. Australian National University.

François, Alexandre. (2004). Reconstructing the Geocentric System of Proto-Oceanic. University of Hawai'i Press: *Oceanic Linguistics* (Vol. 43, No. 1), pp. 1-31.

François, Alexandre. (2005). Unraveling the history of the vowels of seventeen northern Vanuatu languages. *Oceanic Linguistics*, 443-504.

François, Alexandre. (2011). Where *R they all? The geography and history of *R-loss in Southern Oceanic languages. *Oceanic Linguistics*, *50*(1), 140-197.

François, Alexandre. (2017). Words from our Ancestors: The art of sung poetry in northern Vanuatu. Keynote address, *10th Conference On Oceanic Linguistics* (Cool10). Honiara, Solomon Islands.

François, Alexandre; Franjieh, Michael; Lacrampe, Sebastien; and Schnell, Stefan. (2015). "The languages of Vanuatu: Unity and diversity." Vol. 5. *Asia Pacific Linguistics Open Access.*

Golovko, Evgeniy V. "Language contact and group identity: The role of 'folk'linguistic engineering." *The mixed language debate: Theoretical and empirical advances* (2003): 177-208.

Cruz, José Gómez. (2017). "Estructuras morfosintácticas del tojo-ab'al en perspectiva comparativa: el caso de una lengua maya mixta." *Recurso impreso, recurso electrónico*.

Gray, Russell D.; Drummond, Alexei J. and Greenhill, Simon J. (2009). "Language phylogenies reveal expansion pulses and pauses in Pacific settlement." S*cience* 323.5913: 479-483.

Greenhill, S.J., Blust. R, and Gray, Russell D. (2008). The Austronesian Basic Vocabulary Database: From Bioinformatics to Lexomics. *Evolutionary Bioinformatics*, 4:271-283.

Grinevald, Colette. (2003). "Speakers and documentation of endangered languages." *Language documentation and description* 1: 52-72.

Ho, Tin Kam. (1995). "Random decision forests." *Proceedings of 3rd international conference on document analysis and recognition*. Vol. 1. IEEE.

Hocart, Arthur M. (1937). "Kinship systems." *Anthropos* H. 3./4: 545-551.

Hock, Hans Henrich. (1991). *Principles of historical linguistics*. Mouton de Gruyter.

Hyslop, Catriona. (2002). "The Lolovoli dialect of North-East Ambae, Vanuatu." *Canberra: Pacific Linguistics*.

Iseli, Jacqueline. (2018). "Deaf Ni-Vanuatu and their signs: A sociolinguistic study."

Jiao, Tianlong (2021). "Archaeology of Southeast China and the Search for an Austronesian Homeland." *Social Sciences in China* 42.1: 161-170.

Kaiseng, Edwel (2018). “Kastom Dance”. Item ninde122 in *Ninde documentation and orthographic design project*. Crouch, Caroline. (2018). London: SOAS, Endangered Languages Archive. URL: http://hdl.handle.net/2196/00-0000-0000-0013-8EBB-6

Kali, Plas (2018). “Star Song”. Item ninde134 in *Ninde documentation and orthographic design project*. Crouch, Caroline. (2018). London: SOAS, Endangered Languages Archive. URL: http://hdl.handle.net/2196/00-0000-0000-0013-8EE8-3

Kastow, Masing Venevil. (2013). *History, South West Bay, Malekula*. Unpublished manuscript.

Keesing, Roger M., and Fifiʔi, Jonathan. (1969). Kwaio word tabooing in its cultural context. *The Journal of the Polynesian Society*, *78*(2), 154-177.

Kimmerer, Robin (2013). *Braiding sweetgrass: Indigenous wisdom, scientific knowledge and the teachings of plants*. Milkweed editions.

Ladd, D. Robert; G. Roberts, Seán; and Dediu, Dan. (2015). "Correlational studies in typological and historical linguistics." *Annu. Rev. Linguist.* 1.1: 221-241.

Ladefoged, Peter and Maddieson, Ian. (1986). WPP, No. 64: Some of the Sounds of the World's Languages.

Larcom, Joan. (1983). Following Deacon. *Observers observed: Essays on ethnographic fieldwork*, *1*, 175.

Law, Danny. (2017). "Language mixing and genetic similarity: The case of Tojol-ab'al." *Diachronica* 34.1: 40-78.

Layard, John W. (1928). "Degree-Taking Rites in South West Bay." *Malekula (Journal R. Anthropology Inst. LVIII)*. London.

Leonard, Wesley Y. (2018). "Reflections on (de) colonialism in language documentation." *Reflections on language documentation 20 years after Himmelmann 1998*.

Letpen. (2018a). “Small Ninde villages”. Item ninde160 in *Ninde documentation and orthographic design project*. Crouch, Caroline. (2018). London: SOAS, Endangered Languages Archive. URL: https://elar.soas.ac.uk/Collection/MPI1035091.

Letpen. (2018b). “Big Ninde villages”. Item ninde161 in *Ninde documentation and orthographic design project*. Crouch, Caroline. (2018). London: SOAS, Endangered Languages Archive. URL: https://elar.soas.ac.uk/Collection/MPI1035091.

Letpen (lyricist). (2018c). “Sla‘ saxa paxndel wut amwiam”. Item ninde062 in *Ninde documentation and orthographic design project*. Crouch, Caroline. (2018). London: SOAS, Endangered Languages Archive. URL: https://elar.soas.ac.uk/Collection/MPI1035091.

Levinson, Stephen C. (1996). "Frames of reference and Molyneux’s question: Crosslinguistic evidence." *Language and space* 109: 169.

List, Johann-Mattis and Forkel, Robert (2021). LingPy. A Python library for historical linguistics. Version 2.6.9. URL: https://lingpy.org, DOI: https://zenodo.org/badge/latestdoi/5137/lingpy/lingpy. With contributions by Greenhill, Simon, Tresoldi, Tiago, Christoph Rzymski, Gereon Kaiping, Steven Moran, Peter Bouda, Johannes Dellert, Taraka Rama, Frank Nagel. Leizpig: Max Planck Institute for Evolutionary Anthropology.

Love, Mark. (2016). *Making komuniti in Vanuatu (and beyond)* (Doctoral dissertation, PhD thesis, University of Queensland).

Lynch, John (1981). Melanesian diversity and Polynesian homogeneity: The other sideof the coin. *Oceanic Linguistics* 20:95–129.

Lynch, John. (2000a). Reconstructing proto-oceanic stress. *Oceanic Linguistics* (2000): 53-82.

Lynch, John. (2000b). Linguistic subgrouping in Vanuatu and New Caledonia. In *Proceedings of the Second International Conference on Oceanic Linguistics (SICOL)* (Vol. 2, pp. 155-184).

Lynch, John. (2002). The Proto-Oceanic labiovelars: some new observations. *Oceanic Linguistics*, *41*(2), 310-362.

Lynch, John. (2003). Low vowel dissimilation in Vanuatu languages. *Oceanic Linguistics*, 359-406.

Lynch, John. (2005). The apicolabial shift in Nese. *Oceanic Linguistics*, *44*(2), 389-403.

Lynch, John. (2008). Liquid vocalization and loss in Central Vanuatu. *Oceanic Linguistics*, 294-315.

Lynch, John. (2009). At sixes and sevens: the development of numeral systems in Vanuatu and New Caledonia. *Canberra: Pacific Linguistics.*

Lynch, John. (2014). Unexpected final vowel retention in Malakula. *Open Linguistics 1.1.*

Lynch, John. (2016). Malakula internal subgrouping: Phonological evidence. *Oceanic Linguistics, 399-431.*

Lynch, John. (2018). Final consonants and the status of Proto-North-Central Vanuatu. *Language and Linguistics in Melanesia*, *36*, 14-24.

Lynch, John. (2019a). "The Bilabial-to-Linguolabial Shift in Southern Oceanic." *Oceanic Linguistics* 58.2: 292-323.

Lynch, John. (2019b). The phonological history of Naman, a Western Malakula language. *Journal of the Linguistic Society of Papua New Guinea* 37, 21-41.

Lynch, John and Brotchie, Amanda. (2010). Vowel loss in Tirax and the history of the apicolabial shift. *Oceanic Linguistics*, *49*(2), 369-388.

Lynch, John and Crowley, Terence. (2003). More reflexes of Proto-Oceanic *q in Vanuatu languages. *Oceanic linguistics*, *42*(1), 233-238.

Mangke Kaisul, Letin. (2018). "Before we wore clothes". Item ninde007 in *Ninde documentation and orthographic design project*. Crouch, Caroline. (2018). London: SOAS, Endangered Languages Archive. URL: http://hdl.handle.net/2196/00-0000-0000-0013-8ED1-6

Matisoo-Smith, Elizabeth A. (2015). "Tracking Austronesian expansion into the Pacific via the paper mulberry plant." *Proceedings of the National Academy of Sciences* 112.44: 13432-13433.

Matras, Yaron. (2003). "Mixed languages: Re-examining the structural prototype." *The mixed language debate: Theoretical and empirical advances*: 151-176.

Matras, Yaron and Bakker, Peter. (2003). The mixed language debate. *Berlin: Mouton de Gruyter*.

McCarter, Joe. (2012). "Variation, transmission, and maintenance of traditional ecological knowledge on Malekula Island, Vanuatu."

McCarter, Joe; Gavin, M. C.; Baereleo, S.; and Love, Mark. (2014). The challenges of maintaining indigenous ecological knowledge. *Ecology and Society*, *19*(3).

Meakins, Felicity. (2013). Mixed languages. *Contact languages: A comprehensive guide*, *6*, 159-228.

Moodley, Yoshan; Linz, Bodo; Yamaoka, Yoshio; Windsor, Helen M.; Breurec, Sebastien; Wu, Jeng-Yih; Maady, Ayas; Bernhöft, Steffie; Thiberge, Jean-Michel; Phuanukoonnon, Suparat; Jobb, Gangolf; Siba, Peter; Graham, David Y.; Marshall, Barry J.; and Achtman, Mark (2009). The peopling of the Pacific from a bacterial perspective. *Science*, *323*(5913), 527-530.

Mous, Maarten. (2003). "The making of a mixed language." *The Making of a Mixed Language*: 1-342.

Musgrave, Joan (2007). *A grammar of Neve'ei, Vanuatu*. Pacific linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Muysken, Pieter. (1997). "Code-switching processes: Alternation, insertion, congruent lexicalization." *Language choices: Conditions, constraints, and consequences* 20: 361-380.

Næss, Åshild (2013). From Austronesian Voice to Oceanic Transitivity: Äiwoo as the "Missing Link". University of Hawai'i Press: *Oceanic Linguistics* (Vol. 52, no. 1).

Olson, Kenneth; Reiman, Will; Sabio, F., and da Silva, F. (2009). The voiced linguolabial plosive in Kajoko. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society* (Vol. 45, No. 1, pp. 519-530). Chicago Linguistic Society.

Pawley, Andrew. (2005). "The chequered career of the Trans New Guinea hypothesis: Recent research and its implications." *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples*.

Pearce, Elizabeth. (2015). *A grammar of Unua* (Vol. 647). Walter de Gruyter.

Poletto, Cecilia. (2008). "Syntactic encoding of aspect in some Northern Italian dialects." *Syntactic Encoding of Aspect in Some Northern Italian Dialects*: 499-517.

Posth, Cosimo; Nägele, Kathrin; Colleran, Heidi; Valentin, Frédérique; Bedford, Stuart; Kami, Kaitip W.; Shing, Richard; Buckley, Hallie; Kinaston, Rebecca; Walworth, Mary; Clark, Geoffrey R.; Reepmeyer, Christian; Flexner, James; Maric, Tamara; Moser, Johannes; Gresky, Julia; Kiko, Lawrence; Robson, Kathryn J.; Auckland, Kathryn; Oppenheimer, Stephen J.; Hill, Adrian V. S.; Mentzer, Alexander J.; Zech, Jana; Petchey, Fiona; Roberts, Patrick; Jeong, Choongwon; Gray, Russell D.; Krause, Johannes; and Powell, Adam. (2018). Language continuity despite population replacement in Remote Oceania. Nature ecology & evolution, 2(4), 731-740.

Radden, Günter. (2003). The metaphor TIME AS SPACE across languages." *Zeitschrift für interkulturellen Fremdsprachenunterricht* 8.2.

Ray, Sidney H. (1926). *A comparative study of the Melanesian island languages*. Cambridge: Cambridge University Press.

Regenvanu, Ralph. (1999). Afterword: Vanuatu perspectives on research. *Oceania*, *70*(1), 98-100.

Rhodes, Richard (1977). "French Cree—A case of borrowing." Algonquian Papers-Archive 8.

Rice, Keren. (1994). Peripheral in consonants. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *39*(3), 191-216.

Ross, Malcolm D. (1998). Proto-Oceanic adjectival categories and their morphosyntax. *Oceanic linguistics*, 85-119.

Ross, Malcolm D. (2004). The Morphosyntactic Typology of Oceanic Languages. *Language and Linguistics* 5.2:491-541.

Ross, Malcolm. (2005). "Pronouns as a preliminary diagnostic for grouping Papuan languages." *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples* 572: 15-65.

Ross, Malcolm D. (2014). "Reconstructing the history of languages in northwest New Britain: Inheritance and contact." *Journal of Historical Linguistics* 4.1: 84-132.

Ross, Malcolm D.; Pawley, Andrew; and Osmond, Meredith. (1998). *The lexicon of Proto Oceanic: Material culture* (Vol. 1). *Canberra: Pacific Linguistics.*

Ross, Malcolm D.; Pawley, Andrew; and Osmond, Meredith. (2007). *The Lexicon of Proto-Oceanic: The Culture and Environment of Ancestral Oceanic Society. The physical environment* (Vol. 2). ANU E Press.

Ross, Malcolm D.; Pawley, Andrew; and Osmond, Meredith. (2008). The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society: Plants (Vol. 3). *Canberra: Pacific Linguistics.*

Ross, Malcolm D.; Pawley, Andrew; and Osmond, Meredith. (2011). The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society: Animals (Vol. 4). *Canberra: Pacific Linguistics.*

Ross, Malcolm D.; Pawley, Andrew; and Osmond, Meredith. (2016). The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society, People: Body and mind (Vol. 5). *Canberra: Asia-Pacific Linguistics.*

Rothe-Wulf, Annelie, Sieghard Beller, and Andrea Bender. (2015). "Temporal frames of reference in three Germanic languages: Individual consistency, interindividual consensus, and cross-linguistic variability." *Quarterly Journal of Experimental Psychology* 68.5: 917-939.

Saao Baal, Luwi. (2018). "When we came down from the bush". Item ninde030 in *Ninde documentation and orthographic design project*. Crouch, Caroline. (2018). London:

SOAS, Endangered Languages Archive. URL: https://elar.soas.ac.uk/Collection/MPI1035091.

Shafer, Robert. (1965). *Was New Guinea the graveyard of 100 south Asian and Pacific cultures?*. Centre International de Dialectologie Générale de l'Université Catholique de Louvain.

Shaw, Jason A., and Kawahara, Shigeto. (2018). "The lingual articulation of devoiced /u/ in Tokyo Japanese." *Journal of Phonetics 66: 100-119.*

Shimelman, Aviva; Heggarty, Paul; Abete, Giovanni; Wägerle, Laura; Dërmaku-Appelganz, Darja; and Walworth, Mary. (2019). Vanuatu Voices: Malakula. <https://vanuatuvoices.com>

Smith, Linda Tuhiwai. (2021). *Decolonizing methodologies: Research and indigenous peoples*. Bloomsbury Publishing.

Takau, Lana Grelyn. (2016). "A grammar of Nese." Dissertation. University of Newcastle, Australia.

Thieberger, Nick. (2017). Missionary-induced language change, on the trail of the conditional in Nafsan, central Vanuatu.

Thieberger, Nick, and Ballard, Chris. (2008). "Daniel Macdonald and the" Compromise Literary Dialect" in Efate, Central Vanuatu." *Oceanic linguistics* 47.2: 365-382.

Thomason, Sarah G. (1997). "A typology of contact languages." *Creole Language Library* 19: 71-90.

Thomason, Sarah G. (2009). "How to establish substratum interference." *Senri ethnological studies* 75: 319-328.

Thomason, Sara., and Terrence Kaufman. (1988). "Contact-Induced language change: An analytic framework." *Language Contact, Creolization and Genetic Linguistics*: 35-64.

Tryon, D. T. (1976). *New Hebrides languages: An internal classification* (No. 50). Dept. of Linguistics, Research School of Pacific Studies, Australian National University.

Versteegh, Kees. (2017). "The myth of the mixed languages." *Advances in Maltese linguistics* 21: 245.

Walworth, Mary. (2017). "Reo Rapa: a Polynesian contact language." *Journal of Language Contact* 10.1: 98-141.

Watkins, Calvert. (1962). *Indo-European origins of the Celtic verb* (Vol. 1). Dublin Institute for Advanced Studies.

Wurm, S. A. (1982). *Papuan languages of Oceania*. Tübingen: Gunter Narr Verlag.

Zinken, Jörg. (2015). *Temporal frames of reference*. Institut für Deutsche Sprache, Bibliothek.

APPENDIX

How to Read the Sound Change Formulations

The following sound change formulations are written more or less as is common practice, with the following additional conventions:

Line breaks represent a new step in the diachronic sequence of sound changes (but not all chronological relationships are crucially ordered).

Commas separate features of the same segments within a sequence of input, output, or conditioning phones.

Pipes (|) separate ordered phones and natural classes in a sequence.

Curly brackets group together multiple sound changes that target the same abstract structure (e.g., position in a syllable) or natural classes of input phones – this is only important for backwards application, where the generation of an earlier stage of a lexical item could preclude an alternative sound change from being applied.

Asterisks introduce hyphen-separated ranges (minimum-maximum) for the number of phones that are allowed to match in sequence. An asterisk followed by 0 represents any number, or none.

Sequence-final rules with a composite tilde and arrow (~>) generate two variants representing free variation.

In what follows, the sound changes as supplied to the algorithm are presented first for North Central Malekula (NCM) languages in alphabetical order: Avava, Nahavaq, and Ninde; then, for the South West Bay (SWB) languages in reverse alphabetical order: Ninde, Nahavaq, Naati.

Following the sound change models, several of the reconstructed proto forms with few or no competing alternatives are presented with the relevant lexical reflexes and a brief explanation of how they fit the model. First, PNCM forms are listed, then PSWB forms are listed.

Proto Forms and Sound Changes (as supplied to the algorithm)

Sound Changes in Avava

V,unstressed>V,stressed/_C*0-2|V|#

V,unstressed>V,stressed/V,unstressed|C*1-2_C*1-2|#

V,unstressed>V,stressed/#|C*0_C*0|#

[i]~>0/C_#

[l:]>[^{n}d]

[k]|[a]|[k]>[k:]/_[a]

[l]|[i]|[l]>[^{n}d]/_V,front

C,prenasalized,bilabial,plosive>C,prenasalized,bilabial,trill,-secondary/_#

[a]>[u]/_C*0|C,rounded

0>[i]/#_C*1-2|V*1-2|C*1|#

[k]>[$^{\text{ŋ}}$g]/[n]|V_

[k]>[x]/V_V

[x]>[k]/V,-high_#

{[x]>[j]/#|[i]_V,low

[x]>[j]/#|[i]_V,upper-mid}

[i]>0/V|C*0|V|C_C|V,stressed

[i]>0/V|C_C|V,stressed|C*0|V

[^{n}s]>[n]/C_

[^{m}b]>bilabial,trill/_[u]

[$^{\text{ŋ}}$g]>[ŋ]/_#

[n]|[i]>0/_C,plosive,prenasalized|V|C*0|V

[e]>[i]/_C|[o]

[i]>0/#_

[e]|[i]>[i]

[a]>[e]/_C|[e]

[e]>[a]/C,bilabial,labialized_

[e]>[i]/[e]|C_

[e]>[i]/_C|[i]

V,low,front,unrounded,lax,unstressed>[o]/_C,bilabial,labialized

V,high,front,unrounded,tense,unstressed>[u]/_C,bilabial,labialized

[x]>0/[i]_#

[^{n}d] >[t]/V_#

[ŋ]>[m]/[p]_

C,long>C,short

[u]>[y]/V,unrounded|C,unrounded_C,unrounded

{[e]>0/V|C,-nasal*1-2_C*1-2|V,stressed

[o]>0/V|C,-nasal*1-2_C,-nasal*1-2|V,stressed}

[k]>[$^{\text{ŋ}}$g]/#|[n]|V_

[ʦ]>[t]/_#

[e]|[i]>[e]

[e]>[i]/_C*0-1|[i]

[e]>[i]/_fricative|#

[e]>[i]/_[l]|#

[a]>[e]/_C*1-1|[i]

[e]>[i]/_[x]|[i]

[a]>[e]/_C*0|[e]

[e]>[o]/_[w]|[e]|#

[e]>[o]/_C|V,rounded

[e]>V,low,front,unrounded,lax/_C|V,low,front,unrounded

{[o]|[x]|[o]>[oː]

[a]|[x]|[a]>[aː]

[e]|[x]|[e]>[eː]

[i]|[x]|[i]>[iː]

[u]|[x]|[u]>[uː]}

{[x]>0/V_V

[x]>0/#_V

[x]>0/_C}

[k]>0/V_C*0|V

{V,high,back,rounded,unstressed,tense|[^{n}s]>[m]/bilabial_#

V,high,back,rounded,unstressed,tense|[^{n}s]>[m]/bilabial_C

V,high,back,rounded,unstressed,tense|[n]>[m]/C,bilabial_#

V,high,front,unrounded,tense|[n]>[m]/V|C,bilabial_#}

[n]>[m]/C,bilabial_

[ŋ]>[m]/C,bilabial_

[^{n}s]>[s]/V,high,front_

[s]>[h]/V_#

[y]>[e]/_[h]

[s]>[h]/V_[i]|#

[x]>[h]/V_#

[β]>[p]/_C,-lateral,-trill

[β]>[p]/_#

[β^{w}]>[p]/_C,-lateral,-trill

[β^{w}]>[p]/_#

[x]>0/C_V

[^{n}s]>[s]

{[n]|[a]>0/#_C*1-2|V|C*1-2|V*1-2

[n]|[a]>0/#_[^{ŋ}g]|V|C*0|V

[n]|[a]>0/#_C*1-2|V|C*0|V

[n]|[aː]>[a]/#_

[n]|[o]>0/#_C*1-2|V|C*0-2|V*1-2

[n]|[o]>0/#_C*1-2|V|C*0-2|V

[n]|[i]>0/#_C*1-2|[i]|C*0|V

[n]|[i]>0/#_C*1-2|[iː]

[n]|[i]>0/C*1-2_[^{n}d]

[n]|[e]>0/#_C*1-2|V,front|C*0|V

[n][i]>0/plosive|[n]_prenasalized|V|C*0|V

[n]>0/#_V|C*0|V|C*0|#}

[i]>0/V|C_[s]|V

[i]>0/V|C_C|V

V,unstressed>0/V|C_C|V,stressed

[e͡u]>[o]/_#

[e͡u]>[uː]/_C

[w]|[u]>[o]/_#

V,stressed>V,unstressed

0>[w]/V,rounded_[a]

[^{m}b]>[m]/V,high_#

C,bilabial,labialized>bilabial,-secondary/_#

C,bilabial,labialized~>bilabial,-secondary

0~>[i]/V_#

[^{ŋ}g]~>[k]/#_

[i]|[i]>[iː]

Sound Changes in Neverver

V,unstressed>V,stressed/_C*0-2|V,unstressed|#

V,unstressed>V,stressed/_C*1-2|#

{V,low,front,unrounded,lax,unstressed>0/#|C,oral_C|V,low

V,low,front,unrounded,lax,unstressed>0/V|C,oral_C*1-2|V,stressed

[e]>0/C,-nasal,short_C*1-2|V,stressed

V,low,front,unrounded,lax,unstressed>0/[m]_C*1-2|V,stressed

V,low,front,unrounded,lax,unstressed>0/V|C,oral_C|V

[a]>0/C,plosive,-nasal_C,approximant|V,stressed

[e]>0/#|C,-nasal,short_C|V

[i]>0/#|C,-nasal,short_C|V,high

[o]>0/#|C,-nasal,short_C|V,-high

[o]>0/V|C,-nasal,short_C|V,-high

[u]>0/#|C,-nasal,short_C|V,high

[a]>0/#|[m]_C|V,low

[e]>0/#|[m]_C|V

[i]>0/#|[m]_C|V,high

[o]>0/#|[m]_C|V,-high

[u]>0/#|[m]_C|V,high}

[p]|[p]>[pː]/_V

[k]|[k]>[kː]/_V

[m]|[m]>[mː]/_V

[r]|[r]>[rː]/_V

[s]|[s]>[sː]/_V

[t]|[t]>[tː]/_V

[n]|[n]>[nː]/_V

[l]|[l]>[lː]/_V

[e͡u]>[o]/_C

{[u]|[n]>0/V|C*1-2_#

[i]>0/C,oral,plosive_[n]|#}

[k]>[t]/_[n]

[k]>[ɣ]/voiced_voiced

[x]>[ɣ]/_voiced

[k]>[ɣ]/#_V

[o]|[x]>[ɣ]|[o]/V|C,-long,-approximant_#

[i]>0/bilabial,labialized|[e]_C

[^{m}b]>bilabial,trill/_[u]

[^{m}b^{w}]>bilabial,trill,-secondary/_[u]

[^{m}b^{w}]>bilabial,trill,-secondary/_#

C,prenasalized,plosive>C,nasal,plosive/V_C

[a]>[e]/_C|[e]

[e]>[i]/#|[n]_C*1-2|[e]

[e]>[i]/#|[n]_C*1-2|V,upper-mid,front,unrounded

V,low,front,unrounded,lax,unstressed>[e]/#|[n]_C*1-2|V,low,front,unrounded

[a]>[u]/_C*0|C,rounded

V,high,front,unrounded,tense,unstressed>[u]/C,rounded_

[i]>[u]/[w]_

V,upper-mid,front,unrounded,unstressed,tense>[o]/C,rounded_

[e]>0/#|C*0|V*1-2|C_C|V

C,plosive,oral>C,voiced,prenasalized,plosive/C,nasal_

[u]>[w]/_V

[ʦ]>[s]/_#

V,upper-mid,front,unrounded,tense,unstressed>[o]/[w]_

V,upper-mid,front,unrounded,tense,unstressed>[u]/_C,rounded

[a]>[i]/_[w]|[o]

[w]>0/[i]_[o]

0>[w]/V,rounded_V,unrounded

[p]>[ɸ]/_#

[β]>[ɸ]/_#

[β^{w}]>[ɸ]/_#

[t]>[^{n}d]/C,nasal_

[s]>[^{n}s]/C,nasal_

[u]>0/V|C_C|V

[n]|[w]>[nː]

C,bilabial,labialized>C,bilabial,-secondary/_V,-front

[ᵐʙ]>[ᵐʙ̥]/_#

[w]>0/C,alveolar_

C,fricative>C,fricative,voiceless/_#

[ᵑg]>[ᵑk]/_#

[ᵑk]~>[ŋ]/_#

C,bilabial,labialized~>C,bilabial,-secondary

V,stressed>V,unstressed

[ɸ]~>[p]|[ɸ]/V_#

Sound Changes in Ninde (NCM)

[ᵑg]>[k]/C,-nasal|V_V

[i]>0/C_[u]

0>[i]/#_C,-nasal*1-2|V*0|C*0|#

0>[i]/#_[m]|V*0|C*0|#

0>[i]/#_[ŋ]|V*0|C*0|#

0>[h]/#_[i]

0>[o]/V,rounded|[ᵑg]_#

0>[a]/[ᵑg]_#

V,high,back,rounded,tense>V,high,front,rounded,tense/C,unrounded_C,unrounded|C,-rounded

V,high,back,rounded,tense>V,high,front,rounded,tense/C,unrounded,-velar_C,unrounded|#

[ᵑg]>[k]/C,-nasal|V_#

[x]>[k]/_#

[x]>[k]/_C

{[k]>[ᵑg]/C,nasal|V*0_

[k]>[ʔ]/C,-velar|[a]_

[k]>[ʔ]/C,-velar|[o]_

[k]>[ʔ]/_[i]

[k]>[x]/V,low_V,low}

{[k:]>[k]|[o]|[ʔ]/_V,upper-mid,back,rounded

[k:]>[k]|[a]|[ʔ]/_V,low

C,voiceless,plosive,long>C,plosive,short}

V,front,upper-mid,unrounded>V,back,rounded/rounded_

[e͡u]>[e]/_#

[e]|[i]>[e͡i]

[e͡i]>[i]/C_C

C,bilabial,labialized>C,bilabial,-secondary/_V,upper-mid,rounded

{[e]>[a]/_C,bilabial,labialized

[i]>[u]/_C,rounded}

[i]>[u]/_C|[u]

[β]|[o]>0/[a]_C

[s]>[h]/V_#

0>[o]/[u]|[h]_#

0>[e]/V,high,front,tense|[h]_#

{[a]|[x]|V,low,front,unrounded,lax>[ˈe̥]/_C

[a]|[x]|V,low,front,unrounded,lax>[e̥͡i̥]/_#

[o]|[x]|V,low,front,unrounded,lax>[e̥]

[u]|[x]>[ˈi]/_#}

C,voiceless,plosive,-glottal>C,voiced,prenasalized,plosive/[n]|V_V,unstressed

V>voiceless/C,voiceless_

[l]>voiceless,lateral/V,voiceless_

[l]>voiceless,lateral/_V,voiceless

[l]>voiceless,lateral/[a]_#

V,upper-mid,back,rounded,tense>V,low,front,unrounded,lax/V,unrounded|C*1-2_[ʔ]

V,upper-mid,front,unrounded,tense>V,high,back,rounded/#|[n]_C|V,high,back,rounded

[r]>C,voiceless,alveolar,trill/_V,voiceless

[w]>0/C,alveolar_

[s]>[h]/V_V

0>[e]/V,front,-low|C,alveolar_#

0>[o]/#|[h]|V,voiceless|C*1-2|V,high,back,rounded|C,trill_#

0>[a]/V,voiceless|C,alveolar,-prenasalized,-affricate,-nasal_#

0>[o]/V,high,back,rounded,voiceless|C_#

0>[o]/#|[h]|V,voiceless|C*1-2|V,-high|C,rounded_#

[^{n}s]>[s]

[y]>[0]/V|[l]_C|V

[i]>[0]/V|[l]_C|V

C,prenasalized>C,oral,voiceless/V,voiceless_

V,high,front,unrounded,tense>0/V|[s]_C,bilabial|V

V,high,front,rounded,tense>0/V|[s]_C,bilabial|V

V,high,back,rounded,tense>0/V|[s]_C,bilabial|V

[ʦ]>[s]

[t]||[r]>[^{n}ɽ]

bilabial,trill>bilabial,plosive

[^{ŋ}g]>[ŋ]||[g]

[^{m}b]>[m]||[b]

[n]>[m]/C,bilabial

V,low,front,unrounded,lax>V,front,upper-mid,unrounded,tense/V,front,-low|C*0_C,-rounded*0|V,-low

V,low,front,unrounded,lax>V,front,upper-mid,unrounded,tense/V,front,-low|C,-uvular*0_[s]*0|#

{[i]>0/C_#

[u]>0/C_#}

C,long>C,short

0>[w]/V_V,high,back,rounded

V,high,back,rounded,tense>V,high,front,rounded,tense/C,unrounded,-velar_C,unrounded|#

[i]>[ə]/[i]|C_C|-front

[i]>[ə]/[i]|C_C|#

[βʷ]>[w]

[β]>[w]/_V,back,rounded

[e]>[y]/_C|V,rounded

[r]>[ʁ]

[r̥]>[ʁ]

[ᶯɽ]>[r]

0>[a]/V,voiceless|[t]_#

0>[a]/[a]|[t]_#

V,high,back,rounded>V,upper-mid,back,rounded,tense/_[ʁ]

V,high,front,unrounded,long>[i]|[a]/_[ʁ]

V,high,front,unrounded,short>[a]/_[ʁ]

V,high,back,rounded>V,upper-mid,back,rounded,tense/_[ɮ]

V,high,back,rounded>V,upper-mid,back,rounded,tense/[ʁ]_

V,high,front,rounded>V,low,front,unrounded,lax/_[ʁ]

V,central,mid>V,low,front,unrounded,lax/_[ʁ]

V,upper-mid,front,unrounded,tense>V,low,front,unrounded,lax/_C,voiceless,lateral

V,low,front,unrounded,lax>V,upper-mid,back,rounded,tense/V,rounded|C_#

V,low,front,unrounded>V,upper-mid,back,rounded,tense/C,rounded_#

V,low,front,unrounded>V,upper-mid,front,unrounded,tense/_#

V,low,front,unrounded,lax>V,upper-mid,front,unrounded,tense/V|C*0_#

V,low,front,unrounded,lax>V,upper-mid,front,unrounded,tense/C,-labialized,-glottal_C,-labialized,-glottal|V,-low

V,upper-mid,front,unrounded,tense>V,low,front,unrounded,lax/_C,lateral,voiceless

V,low,front,unrounded,lax>V,upper-mid,front,unrounded,tense/C,-labialized,-glottal_C,-labialized,-glottal|V,upper-mid,front,unrounded,tense

V,upper-mid,front,unrounded,tense>V,low,front,unrounded,lax/C,lateral,voiceless_C*1-2

V,upper-mid,front,unrounded,tense>V,low,front,unrounded,lax/[ʁ]_

V,upper-mid,front,unrounded,tense>V,low,front,unrounded,lax/_[ʁ]

V,low,front,unrounded,lax>V,high,front,unrounded,tense/_C,-prenasalized|V,high,back,rounded

V,low,front,unrounded,lax>V,high,front,unrounded,tense/_[h]|V,low,front,unrounded

V>voiceless/_[h]

V>voiceless/[h]_

[β]>0/V,stressed_V

[e]>0/C_[y]

[o]>0/C_[y]

[e]>0/_[e]

[l]>[l̥]/_V,voiceless

[l]>[l̥]/V,voiceless_

[l̥]>[l]/_#

[l̥]>voiced,lateral,fricative

lateral,approximant,long>voiced,lateral,fricative

[n̥]>[n]

[h]>[j]/V,front,-low_V

[h]>[w]/V,rounded_V

[x]>[j]/V,front,-low_V

0>[w]/V,rounded_V

0>[w]/V_V,rounded

C,bilabial,rounded>-secondary/_V,-front

C,bilabial,rounded>-secondary/_[a]

C,bilabial,rounded>-secondary/_#

[^{n}d]>[n]/_#

[h]>0/_V

[h]>0/V_

[x]>[w]/V_V,rounded

{[x]>0/_V

[x]>0/V_}

[o]>[u]/_C|[i]

0>[j]/V,high,front,unrounded_V

V,voiceless>V,voiced

[ʔ]>[k]/_V,high

[y]>[u]/_[k]|[u]

[i]>0/#_

[u]>0/#_

[y]>0/#_

[a]>0/#_C,uvular

[o]>0/#_C,uvular

C,nasal,voiceless>voiced

V,front,-low,tense,unrounded>V,back,tense,rounded/_C,rounded

V,front,low,lax,unrounded>V,upper-mid,back,tense,rounded/C,rounded_[l]*0|#

V,upper-mid,front,tense,unrounded>[o]/V,back,rounded|C*1-2_

V,upper-mid,front,tense,unrounded>[o]/V,back,rounded|C*1-2_#

[i]>[u]/[w]_

V,stressed>V,unstressed

[y]~>[ə]/C,back_C

[a]~>[ə]/C,back_C

[i]~>[ə]/C,back_C

[o]~>[ə]/C,back_C|-rounded

[u]~>[ə]/C,back_C

[y]~>[ə]/C_C,back

[a]~>[ə]/C_C,back

[i]~>[ə]/C_C,back

[o]~>[ə]/C_C,back|-rounded

[u]~>[ə]/C_C,back

[ŋ]|[e]~>0/V_#

[ŋ]|[o]~>0/[o]_#

[m]|[o]~>[m^w]|[e]/V_#

[p]|[o]~>[p^w]|[e]/V_#

[b]|[o]~>[b^w]|[e]/V_#

[ə]~>[u]/_C|V,rounded

Sound Changes in Ninde (SWB)

V,long>V,short

[e]>[y]/_C|[u]|#

[u]>[e]/[e]|C,unrounded_[h]

[e]>0/[e]|C_C|V

[e]>[o]/_C|[u]

[e]>[a]/_[m]*0-1|C,rounded

[e]>[a]/_C,retroflex

[a]>[e]/_C|[e]|#

[a]>[e]/_C|[e]|[l]|#

[u]>0/[y]|C_#

[e͡i]>[o͡i]/[w]_#

[e͡i]>[e]/C_#

[e͡u]>[o]/_#

[e͡u]>[u]/_#

{[u]>0/V|C_C|V

[i]>0/V|C_C|V}

[ʔ]>[k]/C_

[ʔ]>[k]/V,front,-low_#

[ʔ]>[k]/#_

[ʔ]>[k]/_V,front,monophthong

[u]>[i]/_#

[m]|[bʷ]>[pʷ]/[k]|V*0_

V,front,rounded>V,back,rounded|_C0|V,back,rounded

[β]>[βʷ]/_V,back,rounded

[βʷ]>[w]

[β]>0/V,rounded_V,rounded

C,bilabial,rounded>C,bilabial,-secondary/_#

[ø]|[y]>[y]

0>[a]/[ᶯɽ]_#

0>[e]/V,front,-low|C,alveolar_#

[a]>[i]/_[h]|[a]

[h]>[ʔʰ]/[h]|V0|C0|V0_#

[ʔ]>[h]/_V0|C0|V0|[ʔʰ]|#

[ʔʰ]>[ʔ]

0>[h]|[i]/#_C,-nasal*0|V|C*1-2|#

V,voiced>V,voiceless/_[h]

V,voiced>V,voiceless/[h]_

[l]>[ɮ]/V,voiceless_

[l]>[ɮ]/_V,voiceless

{0>[e]/V,front|C,alveolar_#

0>[o]/V,high,back,rounded|C_#,

0>[o]/C,rounded_#}

[e]>[a]/[ɮ]_

[e]>[a]/_[ɮ]

[u]|[w]|[e]>[o]

[u]>[y]/C,unrounded_C,unrounded

[i]>[ə]/[i]|C_C

[a]>[o]/V,rounded|C_#

[a]>[o]/C,rounded_#

[u]>[o]/C,rounded_C|#

[u]>[o]/C,rounded_C*2-3

[u]>[o]/_C|[o]

[u]>[o]/_C|[o͡i]

0>[w]/#_[u]

[a]>[e]/_#

[e]>[u]/_[w]

[a]>[o]/_[w]

[e]>[y]/_C|V,rounded

C,bilabial,rounded>unrounded/_V,rounded

C,bilabial,rounded>unrounded/_[a]

[r]>[ʁ]

[ɳ]|[ɽ]>[r]

[h]>[w]/_[u]

[u]>[o]/_[ʁ]

[u]>[o]/[ʁ]_

[u]>0/C,rounded_V

[ø]>[o]

[a]>[i]/_C|[u]

[y]>[a]/_[ʁ]

[ə]>[a]/_[ʁ]

[y]>[a]/[ʁ]_

[ə]>[a]/[ʁ]_

V,voiceless>V,voiced

[h]>0/_V

[h]>0/V_

0>[w]/V,rounded_V

0>[w]/V_V,rounded

0>[j]/V,front,-low_V

0>[j]/V_V,front,-low

[e]>[i]/_[j]

[e]>[i]/_C|[i]

[y]>[u]/_#

[y]|[y]>[y]

[u]|[y]>[u]

[y]>[e]/_#

[i]>0/#_

C,nasal,voiceless>voiced

C,bilabial,rounded>-secondary/_V,rounded

[a]~>[ə]/C,back_C

[i]~>[ə]/C,back_C

[o]~>[ə]/C,back_C

[u]~>[ə]/C,back_C

[a]~>[ə]/C_C,back

[i]~>[ə]/C_C,back

[o]~>[ə]/C_C,back

[u]~>[ə]/C_C,back

[ŋ]|[e]~>0/V_#

[ŋ]|[o]~>0/[o]_#

Sound Changes in Nahavaq

[ʔ]>0/#_V

[e͡i]>[i]/_C*0-1|[e͡i]

[e͡i]>[i]/_C*0-1|[i]

[e͡i]>[e]|[j]

[a͡i]>[a]|[j]

[o͡i]>[o]|[i]

[a]>[e]/C,bilabial,-secondary_

[a]>[e]/C,nonlateral,-secondary_[h]|#

[a]>[e]/C,nonlateral,-secondary_[p]|#

[a]>[e]/V,high_C*0|#

[a]>[e]/_C*0|[i]|C

[a]>[e]/V,high_

{[a]>[e]/_fricative|#

[a]>[e]/_#}

{[e]>[i]/_C,-secondary*1-2|V,high

[e]>[i]/_C,-secondary*1-2|[e]

[e]>[i]/_C,-secondary*1-2|[e͡i]

[e]>[i]/_C,-secondary*1-2|[e͡u]}

[o]>[u]/_C,glottal|[u]

[e]>[i]/#|[n]_C,-secondary*1-2|V,high

{[u]>0/V|C,oral_C,plosive,nasal|V

[u]>0/V|C,oral_C,nasal|C,plosive|V}

[e]>[u]/C,bilabial,labialized_

[e]>[u]/_[m]*0-1|C,bilabial,labialized

[e]>[a]/_C*0|[a]|C*1-2|V

[e]>[a]/#|C,nasal|C*0_C*0|[a]

[a]|[a]>[a]

0>[w]/V,rounded_V

[u]>0/C_#

[u]>[o]/[u]|C*1-2_#

[β]|[i]|[β]>[β]_V

[o͡u]>[o]|[w]

[a͡i]>[a]|[j]

[o͡i]>[o]|[j]

[a͡u]>[a]|[w]

[e͡u]>[e]|[w]/_#

[e͡u]>[o]|[w]

[i]>[j]/V_#

[e]>[o]/_C*0|[o]

[ø]>[o]/_C*0|[o]

[ø]>[o]/_#

[ø]>[e]

{[y]>[u]/_[n]|#

[n]>0/V_#}

{[n]|[e]>0/#_C*1-2|[e]

[n]|[a]>0/#_C*1-2|[a]

[n]|[i]>0/#_C*1-2|[i]}

[e]>[a]/_[w]|[a]

[e]>[i]/#|[n]_[m]*0-1|C,bilabial,labialized|[i]

[ɳ]|[ɽ]>[n]|[d]/_#

[ɳ]|[ɽ]>[n]|[d]|[r]

{[n]|[d]>[n]/_#

[m]|[b]>[m]/_#

[ŋ]|[g]>[ŋ]/_#}

[n]|[i]>0/#_[n]|[d]

[i]>[j]/_V

0>[j]/V_[i]

[e]>[u]/[i]|C*0-1|C,labialized_C*1-2|[u]

0>[j]/V,front_V

V,front,rounded>V,back,rounded

[β]>[p]/_#

[β]>[p]/_C

[βʷ]>[pʷ]/_#

[βʷ]>[p]/_C

Sound Changes in Naati

0>[w]/V_V,rounded

[ʔ]>0/#_V

[h]>0/[ʔ]_V

[e]>[o]/[w]_

[a]>[o]/_[l]|[u]

[a]>[u]/_[m]*0-1|C,labialized*1|[u]|#

[u]>[o]/_#

[e]|[βʷ]|[u]>[e͡u]

[e]>[i]/_C*0|[e͡i]

[e]>[i]/_C*0|[e͡u]

[e]>[i]/_C*0|[e]

[e]>[i]/_C*0|[o]

[i]>0/V|[t]_[m]|V

[e]>[ø]/_C*0|[ø]

[e]>[ø]/[m]*0-1|C,labialized_

[e]>[ø]/_[m]*0-1|C,labialized

[a]>[e]/[e]|C*1-2_C*0|V|C

[a]>[ø]/_[m]*0-1|C,labialized

[a]>[ø]/_C,bilabial|[a]

[ø]>[a]/C*1-2|V*1-2|C*1-2_[m]*0-1|C,labialized

[e]>[i]/_C|[a]

[h]>0/[ʔ]|V*0_

[a]>[u]/_[m]*0-1|C,labialized|[o]

[u]>[y]/C,-secondary_C,-secondary

[i]>[o͡i]/C,labialized,bilabial_#

[ø]>[u]/_[m]*0-1|C,labialized,bilabial|[o͡i]

{[u]>[y]/[e]|C*0-1_

[u]>[y]/_C*0-1|[y]}

[y]>[i]/C,velar,plosive_

[y]>[i]/_C,velar,plosive

V,rounded,back>V,rounded,front/_[ʔ]

[y]>[ø]/C,bilabial,labialized_C,glottal

C,bilabial,labialized>C,bilabial,-secondary/_V,rounded

C,bilabial,labialized>C,bilabial,-secondary/_[o͡i]

C,bilabial,labialized>C,bilabial,-secondary/_[a]

[o]>[a]/_[ʔ]

[k]>[ʔ]/_#

[β^{w}]>[p^{w}]/_#

[β^{w}]>[p^{w}]/_V|C*1-2|V|C,voiceless,bilabial,plosive

[i]>[y]/C,bilabial_C,alveolar|C

[i]>[y]/C,bilabial_C,alveolar|#

[i]>[y]/[l]_C,alveolar|C

[i]>[y]/[l]_C,alveolar|#

[i]>[u]/_[m]*0-1|rounded

[e]>[y]/C,bilabial_C,alveolar|C

[e]>[y]/C,bilabial_C,alveolar|#

[e]>[y]/[l]_C,alveolar|C

[e]>[y]/[l]_C,alveolar|#

[ø]>[y]/C,bilabial_C,alveolar|C

[ø]>[y]/C,bilabial_C,alveolar|#

[ø]>[y]/[l]_C,alveolar|C

[ø]>[y]/[l]_C,alveolar|#

[y]>[i]/_C*0|[e]

[ʔ]>0/[ø]_#

C,bilabial,labialized>-secondary/_#

[a]>[e]/_[ŋ]|#

[n]>0/V_#

[u]>[o]/_[n]|#

{[a]>[o]/_C*0|[o]|[n]|#

[a]>[o]/_fricative|[u]

[a]>[o]/_[ʔ]|[u]}

[h]>0/V_V||[ʔ]

[a]||[a]>[aː]

[j]>0/#_V

[β]>[p]/_[s]

[p]>[β]/V_V

[ɳ]||[ɽ]>[t]||[r]/C,nasal_

[n]||[d]>[t]/C,nasal_

[ŋ]||[g]>[k]/C,nasal_

[m]||[b]>[p]/C,nasal_

{[n]||[a]>[n]||[ø]/#_[m]*0-1|C,bilabial,labialized

[n]||[e]>[n]||[u]/#_rounded

[n]||[e]>[n]||[i]/#_C*0|V,high}

[n]>0/V_C,nasal|oral,plosive

[ŋ]>0/V_C,nasal|oral,plosive

[e]||[ʔ]||[u]>[o]||[ʔ]||[u]

[a͡i]>[e͡i]/_#

[e]>[eː]/_[ŋ]||[g]

V,short>V,long/_C,alveolar,nonlateral,oral|V

V,long>V,short/_[s]

[m]||[u]||[w]>[mʷ]

[o]>[o͡i]/_[j]

0>[a]/[ʔ]_C

Excerpts from the proto lexicons reconstructed for PNCM and PSWB

Proto NCM reconstructions

***(jox)jox ‘vomitus’** : Probably cognate with Avava *jok-ian* (exp. *jojok*). Expected Neverver *joɣjox*, Ninde *joʔjoʔ*.

***at- ‘3.PL.SUBJ.REAL’** : Avava *at-*, Neverver *at- ~ at-i-*. Expected Ninde *ate*.

***aut ‘ashore, edge’** : Neverver *aut*, Ninde *o͡ute ~ ewute*. Expected Avava *aut*.

***ilaŋal ‘ten’** : Avava *laŋal*, Ninde *ɮaŋal*. Expected Neverver *ilŋal*.

***ilim ‘five’** : Neverver *i-lim*. For Avava *ilim*, inserted initial [i] not explained. Probably cognate with Ninde *<se>lme* (exp. *ɮim*).

***ioᵑg ‘you’** : Avava *oŋ*, Neverver *ox~iox*. Probably cognate with Ninde *nuŋg* (exp. *wəŋgo~woŋgo*).

***itl ‘three’** : Neverver *i-tl*, Ninde *tl*. For Avava *itl*, inserted initial [i] not explained.

***jal ‘sing’** : Avava reduplicated root *jal*. Expected Neverver *jal*, Ninde *jal*.

***jaᵑg ‘born’** : Avava *jaŋ*, Neverver *jaŋ*. Expected Ninde *jaŋge~jəŋge*.

***jox ‘vomit’** : For Avava *jox*, fricative final [x] not explained. Probably cognate with Ninde *joʔjoʔ* (exp. *jaʔ*). Expected Neverver *jox*.

***kˈamem ‘we, us,’** : Neverver *inam ~ nam-*. For Ninde *kamem*, vowel height of [a] not explained. Probably cognate with Avava *kopm* (exp. *kimim*).

***kːan ‘eat’** : Avava *-kan*, Neverver *kːan*, Ninde *kaʔan*.

***k:is 'peel (fruit)'** : Avava *kih*, Neverver *k:is*. Probably cognate with Ninde *ŋgi<lim>* (exp. *kije~kəje*).

***k:on 'bitter'** : Avava *-kon*, Neverver *k:on*, Ninde *koʔone*.

***kala-n 'nephew'** : Neverver *xalan*. For Ninde *kala-*, approximant [l] not explained; deleted final *n not explained. Expected Avava *kalan*.

***ka^{m}b^{w}at 'ancestor spirit and cultural hero'** : Avava *ka^{m}bat* ~ *^{ŋ}ga^{m}bat*. For Ninde *kapat*, deleted final *e not explained. Expected Neverver *k^{m}b^{w}at~k^{m}bat*.

***kil 'dig up'** : Avava *il*, Neverver *ɣil*. Probably cognate with Ninde *kil* (exp. *kəɮe~kiɮe*).

***ku- '2SG.SUBJ.REAL'** : Ninde *ku-*. Probably cognate with Neverver *ku-* (exp. *ɣu*). Expected Avava *u*.

***kut 'where, if'** : For Neverver *kut*, plosive initial [k] not explained. Probably cognate with Ninde *wut* (exp. *kəto~kuto*). Expected Avava *ut*.

***l:ak 'marry'** : Ninde *ɮaʔ*. Probably cognate with Avava *^{n}dak^{n}daka* (exp. *^{n}dak*). Expected Neverver *l:ak*.

***l:eβex 'pour out, tip out'** : For Neverver *l:iβix*, vowel height of [i] not explained; vowel height of [i] not explained. Probably cognate with Avava *^{n}deβeh* (exp. *^{n}diβik*). Expected Ninde *leβek*.

***lak 'in secret, in hiding'** : Probably cognate with Avava *(la)lak(i)* (exp. *lak*), Ninde *ɮaʔ(ɮaʔ)* (exp. *ɮaʔ*). Expected Neverver *lak*.

***lap:an 'under'** : Avava *lapan*, Neverver *lap:an*, Ninde *βe-* ~ *βene*.

***laxau 'step over (obstacle)'** : For Neverver *lɣa*, deleted final *u not explained. Probably cognate with Ninde *ɮo͡u* (exp. *lejowu*). Expected Avava *la:ui~la:u*.

***laβ 'get'** : Avava *lap*, Neverver *laɸ*. Probably cognate with Ninde *lip* (exp. *ɮaβ*).

***laβ 'plant'** : Avava *lap*. For Ninde *ɮap*, plosive final [p] not explained. Expected Neverver *lapɸ~laɸ*.

***leb 'give birth'** : For Avava *lem*, nasal final [m] not explained; For Neverver *le^{m}p*, prenasalized final [^{m}p] not explained. Expected Ninde *ɮab*.

***leŋ 'remove husk'** : Avava *leŋ*, Neverver *leŋ*. Expected Ninde *ɮəŋ~ɮaŋ*.

***leβ 'big'** : Probably cognate with Avava *<leβ>-lep* (exp. *lip*), Neverver *la^{m}p* (exp. *lepɸ~leɸ*), Ninde *<t>lepe* (exp. *ɮaβ*).

***lile 'near, nearby, soon'** : Neverver *lile*. Probably cognate with Avava *lile* (exp. *^{n}dei~^{n}de*). Expected Ninde *lile*.

***liŋ 'leave'** : For Ninde *liŋ*, approximant initial [l] not explained. Probably cognate with Avava *liŋ-liŋ(^{m}bitep OBJ)* (exp. *liŋ*), Neverver *liŋ-liŋ* (exp. *liŋ*).

***liŋ 'ROOT IMPL. BY leave, put'** : Avava reduplicated root *liŋ*, Neverver reduplicated root *liŋ*. Expected Ninde *ɮəŋ~ɮiŋ*.

***lit 'yellow'** : Avava *lit*. Probably cognate with Ninde *lit* (exp. *ɮite*). Expected Neverver *lit*.

***liβ^{w}oxˈa-n 'in the middle of, between'** : Avava *lupa-n*, Ninde *liβe-ne*. Expected Neverver *liβɣan~liβ^{w}ɣan*.

***liβʷoxat 'night'** : Avava *lupat*, Neverver *liβɣat*, Ninde *liβete*.

***lualu 'vomit'** : For Avava *lualu*, deleted *w not explained. Probably cognate with Ninde *liluwo* (exp. *luwol*). Expected Neverver *lalu*.

***lue 'shoot at (with a bow and arrow)'** : For Avava *lu*, deleted final *e not explained; For Ninde *luwo*, approximant initial [l] not explained. Probably cognate with Neverver 's reduplicated root *lu* (exp. *lo*).

***lulumʷ 'sweet'** : Neverver *l:um*. For Avava *lum*, deleted initial *l not explained; deleted *u not explained; For Ninde *lum*, deleted initial *l not explained; deleted *u not explained.

***lumus 'wash'** : Avava *-lumuh*, Neverver *lmus*. Probably cognate with Ninde *lumus* (exp. *lumyje*).

***luxlux 'stay behind, wait'** : Neverver *luɣlux*. For Avava *luluk*, velar final [k] not explained; plosive final [k] not explained. Expected Ninde *lyklyk~ləklək*.

***m:ap 'heavy'** : Avava *-map*. For Neverver *m:ap*, deleted final *ɸ not explained. Expected Ninde *map*.

***ma 'HAB'** : For Avava *ma-VERB*, inserted final [V] not explained. Expected Neverver *ma*, Ninde *me*.

***malamal 'naked'** : For Avava *malamal*, inserted [a] not explained; For Neverver *malmal*, vocalic [a] not explained; vowel quality of [l] not explained; vowel place of [l] not explained; vowel manner of [l] not explained; vowel height of [l] not explained; vowel backness of [l] not explained; vowel stress of [l] not explained; vowel aspiration of [l] not explained; vowel laterality of [l] not explained; vowel syllabicity of [l] not explained. Probably cognate with Ninde *maɮemaɮe* (exp. *malamal*).

***malu 'come out, emerge, leave'** : Avava *ma<lu>*. Expected Neverver *malu*, Ninde *mal*.

***m^{w}em^{w}e 'father'** : For Avava *mom (^{n}de-)*, inserted [^{n}d] not explained; vowel height of final [e] not explained. Probably cognate with Neverver *mama* (exp. *memo~m^{w}emo*). Expected Ninde *momo~mom^{w}e*.

***marak 'jump, (of water) splash'** : Avava *marak*, Ninde *maʁaʔ*. Expected Neverver *mrak*.

***maran 'tomorrow'** : Avava *maran (nan)*, Neverver *maran*, Ninde *maʁan*.

***mas 'cooked'** : Avava *mah*. Expected Neverver *mas*, Ninde *ma*.

***mas 'spear'** : Avava *-mah*. Probably cognate with Ninde *namase* (exp. *ma*). Expected Neverver *mas*.

***matak 'fear, be afraid of'** : Avava *matak*. For Ninde *metaʔ*, vowel height of [e] not explained. Expected Neverver *m^{n}dak*.

***mat͡s 'be dead, die'** : Avava *-mat*, Neverver *mas*, Ninde *mes*.

***matur 'sleep'** : For Avava *-matur*, vowel backness of [u] not explained; For Neverver *matur*, inserted [a] not explained; oral [t] not explained. Probably cognate with Ninde *mitoʁ* (exp. *metəʁa~metaʁa*).

***maur 'live'** : Avava *maur*, Neverver *maur*. For Ninde *mowoʁ*, vowel quality of [o] not explained; vowel height of [o] not explained; vowel backness of [o] not explained.

***^{m}bak^{m}bak 'hide'** : Avava *^{m}ba^{m}bak*. For Neverver *^{m}baɣ^{m}bax*, fricative final [x] not explained. Expected Ninde *mbaʔmbaʔ*.

***^{m}bakur 'tamanu (Bis: nambakura)'** : For Avava *^{m}bakur*, inserted [k] not explained; vowel backness of [u] not explained. Probably cognate with Ninde *nəmbəkoʁo* (exp. *mbaʔoʁo*). Expected Neverver *^{m}baɣur*.

***^{m}ba^{n}d 'owl (Tyto alba)'** : Ninde *niβinmban*. Expected Avava *^{m}bat*, Neverver *^{m}ba^{n}d*.

***^{m}ba^{ŋ}g 'seizure'** : Relationship with Ninde *nanmbaŋge* (exp. *paŋge~pəŋge*) uncertain. Expected Avava *^{m}baŋ*, Neverver *^{m}baŋ~^{m}ba^{ŋ}k*.

***^{m}bar '(of eyes) blind'** : Avava *^{m}bar*, Neverver *^{m}bar*. Expected Ninde *pəʁ~paʁ*.

***^{m}bel 'chase (s.t.)'** : Neverver *^{m}bel*. For Avava *^{m}bel*, vowel height of [e] not explained. Expected Ninde *pele*.

***^{m}ber 'ROOT IMPL. BY be long, long'** : Avava *-^{m}ber*, Neverver reduplicated root *^{m}ber*. Expected Ninde *pəʁa~paʁa*.

***^{m}bor '(of ears) deaf'** : Avava *^{m}bor*. Probably cognate with Ninde *mboʁo͡i* (exp. *poʁ*). Expected Neverver *^{m}bor*.

***^{m}b^{w}ˈir 'break, win'** : Avava *^{m}b^{w}ir*, Neverver *^{m}bir*. Probably cognate with Ninde *βaʁ* (exp. *pəʁa~paʁa*).

***^{m}b^{w}erˈakin 'true, real, really'** : Avava *^{m}bariːn ~ ^{m}b^{w}iriːn*, Neverver *βratn*, Ninde *-motne~mosne*.

***^{m}ʙum 'grandfather'** : Avava *^{m}ʙum*, Neverver *^{m}bu^{m}bu~^{m}ʙu^{m}ʙu*. Probably cognate with Ninde *<tə>mbum* (exp. *ʙ̥ym*).

***^{m}ʙun ‘be full’** : Avava -*^{m}ʙun*. Probably cognate with Ninde *mbun* (exp. *ʙ̥yne*). Expected Neverver *^{m}ʙun*.

***^{m}ʙur ‘swell’** : Avava *^{m}ʙur*, Neverver *^{m}ʙur*. Probably cognate with Ninde *mboʁo* (exp. *ʙ̥aʁa~ʙ̥əʁa*).

***^{m}ʙut ‘stop crying’** : Avava *^{m}ʙut*, Neverver *^{m}ʙut*. Probably cognate with Ninde *mbut* (exp. *ʙ̥yte*).

***^{m}ʙuxut ‘inside’** : Avava *^{m}ʙu:t*. For Neverver *^{m}ʙuɣut*, inserted [u] not explained. Expected Ninde *mbuwuto*.

***^{m}ʙ^{w}ut ‘step’** : Avava -*^{m}ʙut*. Probably cognate with Ninde 's reduplicated root *mbəte* (exp. *ʙ̥ut*). Expected Neverver *^{m}ʙut*.

***me(re)res(al) ‘light (in weight)’** : For Ninde *ma<ʁa>ʁas*, deleted *a not explained; deleted final *ʁ not explained. Probably cognate with Avava *meres<al>* (exp. *mirirsal*). Expected Neverver *mrersal*.

***meler ‘(be) clear, transparent’** : Neverver *m-ler*. For Avava *meler*, vowel height of [e] not explained; vowel height of [e] not explained. Expected Ninde *melaʁa~meləʁa*.

***melikis ‘green, blue’** : For Avava *melih*, vowel height of [e] not explained; vowel length of [i] not explained; For Ninde *mel<ə>kise*, inserted final [ə] not explained. Expected Neverver *mliɣis*.

***mem ‘dry’** : Ninde *mem*. For Avava *mim*, vowel height of [i] not explained. Probably cognate with Neverver *m:as* (exp. *mem*).

***mesax ‘sick, ill’** : Neverver *m^{n}sax*. For Ninde *mijaʔ*, vowel height of [i] not explained. Expected Avava *masak*.

***mial 'red'** : Avava *mial*, Neverver *mial*. For Ninde *mijaɮe*, fricative [ɮ] not explained; inserted final [e] not explained.

***min 'drink'** : Avava reduplicated root *min*. For Ninde *min*, deleted final *e not explained. Expected Neverver *min*.

***mitutak 'fear ($Deacon)'** : For Avava *mitutak*, inserted [u] not explained; For Neverver *mtutax*, oral [t] not explained; fricative final [x] not explained. Expected Ninde *mututaʔ*.

***mol 'rest (NTR?)'** : Avava *mol*, Neverver *mol*. Expected Ninde *mol*.

***mo^{ɳ}ɽo^{ŋ}g 'be hungry'** : Ninde *moroŋgo*. For Avava *morot*, alveolar [r] not explained; trill [r] not explained; oral [r] not explained; alveolar final [t] not explained; oral final [t] not explained. Expected Neverver *m^{ɳ}ɽo^{ŋ}k~m^{ɳ}ɽoŋ*.

***mul 'skin shed by a snake'** : Neverver *mul*. For Avava *mulu-n*, inserted [u] not explained; inserted final [n] not explained. Expected Ninde *myle*.

***muluxul 'round'** : Avava *muluːl*. Expected Neverver *mluɣul*, Ninde *muluwuɮo*.

***m^{w}as 'must'** : Neverver *mas*. Expected Avava *mah~m^{w}ah*, Ninde *m^{w}o*.

***m^{w}en 'sweat'** : Neverver *men*, Ninde *mone*. Expected Avava *m^{w}an~man*.

***m^{w}iir 'left'** : Avava *m^{w}iːr*. For Ninde *mijaʁ*, plain initial [m] not explained; deleted final *a not explained. Probably cognate with Neverver *mer* (exp. *mwir~m^{w}wir*).

***nˈaβu^{n}s 'banana'** : Neverver *naβu^{n}s~naβus*, Ninde *nys~nesu-*. Probably cognate with Avava *apm* (exp. *aβeh*).

***naka ‘tree’** : Avava *a*^{ŋ}*gai* ~ *a*^{ŋ}*ga*, Neverver *naɣa*. Probably cognate with Ninde *ne͡i* (exp. *naʔe*).

***nakerwe ‘large root’** : For Avava ^{ŋ}*gerwe-*, vowel height of [e] not explained. Expected Neverver *niɣero*, Ninde *naʔaʁo~naʔuʁo~naʔəʁo*.

***nalaŋ ‘sky’** : Avava *alaŋ*, Neverver *nelaŋ* ~ *nilaŋ*. For Ninde *naɮaŋ*, fricative [ɮ] not explained.

***namaliŋ ‘bed’** : Neverver *nemaliŋ~nimaliŋ*. For Avava *maliŋ*, vowel height of [a] not explained. Expected Ninde *nemeliŋ~nemeləŋ*.

***naman ‘bird’** : Avava *aman*, Neverver *neman*, Ninde *nemen*.

***nama^{n}sir ‘t.o. banana’** : For Avava *masir*, vowel height of [a] not explained. Probably cognate with Ninde *nəməsaʁ* (exp. *namesəʁa~namesaʁa*). Expected Neverver *nam*^{n}*sir*.

***namap ‘level place’** : Avava *amap*. For Ninde *nemep*, vowel height of [e] not explained; vowel height of [e] not explained. Expected Neverver *nemapɸ~nemaɸ*.

***namar ‘hunger’** : Avava *(a)mar*, Neverver *nemar*. For Ninde *^nemaʁa*, vowel height of [e] not explained; inserted final [a] not explained.

***na^{m}bak ‘turtle’** : Avava *a*^{m}*bak*, Ninde *nambaʔ*. Expected Neverver *ne*^{m}*bak*.

***na^{m}bal ‘swamp harrier (Circus approximans) (Bis: hokdei)’** : Avava *a*^{m}*bal*. Probably cognate with Ninde *nembaɮe* (exp. *nambal*). Expected Neverver *ne*^{m}*bal*.

***na^{m}bala ‘sow, pig’** : Avava ^{m}*bala*. Probably cognate with Ninde *nambaɮa͡i* (exp. *nambele*). Expected Neverver *ne*^{m}*bala*.

***na^{m}balaka ‘cottonwood’** : For Ninde *nambaɮaŋgo*, inserted final [a] not explained. Probably cognate with Avava *^{m}balaka* (exp. *^{m}bala:i~^{m}bala:*). Expected Neverver *nemlaɣa*.

***na^{m}ba^{ŋ}g ‘banyan’** : Avava *a^{m}baŋ*, Neverver *ne-^{m}baŋ*. For Ninde *nəmbaŋge*, vowel quality of [ə] not explained; vowel height of [ə] not explained; vowel backness of [ə] not explained.

***na^{m}bari ‘dream’** : For Avava *^{m}bari*, vowel height of [a] not explained. Probably cognate with Ninde *nəmbʁyje* (exp. *nambaʁ~nambəʁ*). Expected Neverver *ne^{m}bari*.

***na^{m}barox ‘tooth’** : For Avava *^{m}boroh*, vowel quality of [o] not explained; vowel height of [o] not explained; vowel backness of [o] not explained; rounded [o] not explained; glottal final [h] not explained; fricative final [h] not explained. Probably cognate with Ninde *nəmbaʁi* (exp. *nambəʁəʔ~nambaʁəʔ~nambəʁaʔ~nambaʁaʔ*). Expected Neverver *ne^{m}barɣo*.

***na^{m}ʙuas ‘male pig’** : Neverver *ni^{m}ʙwas ~ ni^{m}ʙuas*. For Avava *a^{m}ʙuah*, inserted initial [a] not explained; deleted *w not explained. Probably cognate with Ninde *nəmbuwas* (exp. *nambuwo*).

***na^{m}ʙuŋ ‘day’** : For Avava *a^{m}ʙuŋ*, vowel backness of [u] not explained. Expected Neverver *na^{m}ʙuŋ*, Ninde *nambyŋ~nambəŋ*.

***na^{m}ʙ^{w}ut(a) ‘footprint’** : Avava *^{m}ʙuta-*. For Neverver *na-^{m}ʙut*, inserted [u] not explained; deleted final *a not explained. Probably cognate with Ninde *nembətembəte* (exp. *na^{m}buto*).

***nani ‘coconut’** : Neverver *nani*. Probably cognate with Avava *(n)ani* (exp. *eni~enii~en*), Ninde *nimyt* (exp. *nan*).

***na^{n}sa^{n}s ‘croton’** : Neverver *ne^{n}sa^{n}s*. For Avava *asah*, glottal final [h] not explained. Expected Ninde *nasas*.

***na^{n}siβilen 'hair (of head)'** : Avava *siβlen*. Expected Neverver *na^{n}siβilen*, Ninde *nasβilene*.

***na^{ɳ}ɽar 't.o. tree'** : Avava *a^{ɳ}ɽar*. For Ninde *narara*, alveolar [r] not explained; trill [r] not explained; inserted final [a] not explained. Expected Neverver *ne^{ɳ}ɽar*.

***na^{ŋ}galat 'nettle'** : Avava *^{ŋ}galat*, Neverver *na^{ŋ}glat~ni^{ŋ}glat~ne^{ŋ}glat*. For Ninde *naŋgaɮate*,.

***na^{ŋ}goβi 'egg (N)'** : For Avava *oβi-*, deleted initial *^{ŋ}g not explained; deleted final *i not explained; For Neverver *ni^{ŋ}goβin*, vowel height of [i] not explained; inserted final [n] not explained. Expected Ninde *nəŋgoβ~naŋgoβ*.

***napa^{ŋ}gala^{m}b 't.o. banyan'** : For Neverver *nepaŋalam*, deleted *^{ŋ}g not explained; inserted final [a] not explained. Expected Avava *pa^{ŋ}gla^{m}ʙ*, Ninde *nambaʔaɮamb*.

***napap 'penis wrapper'** : Avava *apap*. Expected Neverver *nepaɸ~neparɸ*, Ninde *nambap*.

***nar͡eu-n 'leaf'** : Neverver *noron*. For Avava *aruː-*, deleted final *n not explained. Probably cognate with Ninde *nuwoʁo-* (exp. *nəʁywyne~nuʁywyne~naʁəwyne~nəʁəwyne~naʁywyne~nəʁuwyne~naʁuwyne*).

***narukum 'crab'** : For Ninde *noʁkum*, vowel quality of [o] not explained; vowel height of [o] not explained; vowel backness of [o] not explained; rounded [o] not explained; deleted *ə not explained; deleted final *o not explained. Probably cognate with Avava *arum* (exp. *ryum*). Expected Neverver *narɣum*.

***narup^{w} 'garden'** : Avava *arup*. Probably cognate with Ninde *noʁpo* (exp. *nuʁop~naʁop~nəʁop*). Expected Neverver *narup~narup^{w}.*

***nasal ‘road, path’** : Avava *asal*. For Ninde *naɮe*, deleted *i not explained; deleted *j not explained. Expected Neverver *nesal*.

***nasix ‘kingfisher’** : Neverver *nasix*. Probably cognate with Avava *asik* (exp. *eh~ehii~ehi*), Ninde *nes<um>* (exp. *naik*).

***natal ‘vine, rope, cable, cord’** : Avava *atal*. Probably cognate with Ninde *netel* (exp. *na*^{n}*dal*). Expected Neverver *netal*.

***natan ‘land, earth, ground, dirt, soil’** : Avava *(a)tan*, Ninde *netene*. Expected Neverver *netan*.

***naur ‘lobster, prawn’** : Neverver *naur*. For Avava *our*, vowel quality of initial [o] not explained; vowel height of initial [o] not explained; vowel backness of initial [o] not explained; rounded initial [o] not explained. Probably cognate with Ninde *nuwoʁa<ʔ͡ai>* (exp. *nowoʁ*).

***naus ‘rain’** : Avava *auh*, Neverver *naus*. Probably cognate with Ninde *nuwo* (exp. *nowuwo*).

***nawal ‘hole’** : Avava *awal*. For Ninde *nowol*, vowel quality of [o] not explained; vowel height of [o] not explained; vowel backness of [o] not explained; rounded [o] not explained. Expected Neverver *nuwal*.

***nawa^{ŋ}g ‘boat’** : Avava *awaŋ*, Neverver *nuwa*^{ŋ}*k~nuwaŋ*. Probably cognate with Ninde *nowoŋgo* (exp. *nawaŋge~nawəŋge*).

***naxamal ‘house’** : Avava *amal*, Neverver *naɣmal ~ naɣamal*. For Ninde *nemel*, vowel height of [e] not explained.

***naxa^{m}ʙ 'fire'** : For Neverver *naɣa^{m}ʙ̥*, vowel height of [a] not explained. Probably cognate with Avava *a:^{m}ʙ* (exp. *a^{m}ʙ*), Ninde *nemb* (exp. *neʙ̥*).

***naxa^{m}ʙu-n ''** : Probably cognate with Avava *e^{m}ʙu-* (exp. *a:^{m}ʙyn*), Neverver *naɣa^{m}ʙun* (exp. *naɣ^{m}ʙ̥*), Ninde *ne͡imbu-* (exp. *neʙ̥yne*).

***naxa^{n}dl 'earwax'** : Avava *a^{n}dl*. Probably cognate with Ninde *ni^{n}dle* (exp. *netl*). Expected Neverver *neɣanl*.

***naxari 'victory leaf'** : For Avava *a:ri*, vowel height of initial [a:] not explained; For Neverver *naɣari*, vowel height of [a] not explained; For Ninde *naʁi*, inserted final [i] not explained.

***naxas-n 'jaw.3SG.POSS'** : Avava *as-n*, Ninde *nesne-*. Expected Neverver *neɣasn*.

***naxati^{m}b^{w}el 'dragon plum (Dracontomelon vitiense)'** : For Avava *ati^{m}bol*, inserted [i] not explained; vowel quality of [o] not explained; vowel height of [o] not explained; vowel backness of [o] not explained; rounded [o] not explained. Probably cognate with Ninde *naʁu* (exp. *netimbol*). Expected Neverver *naxti^{m}b^{w}el~naxti^{m}bel*.

***naxaβax 'yam mound'** : For Avava *apak*, plosive [p] not explained; For Neverver *naɣaβax*, vowel height of [a] not explained; inserted [a] not explained. Expected Ninde *neaʔ*.

***naxaβix 'Malay apple'** : For Neverver *naɣaβix*, inserted [a] not explained. Probably cognate with Avava *aβik* (exp. *e:βi~e:βii~e:β*), Ninde *neβyke* (exp. *neik*).

***naxaβux 't.o. tree'** : Probably cognate with Avava *aβik* (exp. *a:βyh*), Neverver *niβɣo* (exp. *naɣβux*), Ninde *neβyke* (exp. *nowək~nowyk*).

***naxut ‘louse’** : Neverver *naɣut ~ noɣut*. For Avava *aut*, vowel backness of [u] not explained. Probably cognate with Ninde *nuŋgut* (exp. *nejuto*).

***naβan(an) ‘fruit’** : Avava *βanan*, Neverver *neβan*. Expected Ninde *naβanan*.

***naβat ‘stone’** : Neverver *neβat*. For Avava *(n)aβat*, inserted initial [n] not explained; For Ninde *neβet*, deleted final *e not explained.

***naβul ‘moon’** : Neverver *naβyl ~ naβul*. Probably cognate with Avava *a^{m}basi* (exp. *aβyl*), Ninde *neβe* (exp. *neβyle*).

***naβunul ‘maggot’** : For Avava *βunul*, vowel backness of [u] not explained. Probably cognate with Ninde *nanwuɮo* (exp. *nowunyle*). Expected Neverver *naβnul*.

***naβ^{w}ilaŋ ‘fly (insect)’** : Avava *β^{w}ilaŋ*. Probably cognate with Ninde *nawaɮ͡ai* (exp. *nowulaŋ~nowuləŋ*). Expected Neverver *naβ^{w}laŋ~naβlaŋ*.

***^{n}da ‘PERF’** : Avava *^{n}da*, Ninde *!paʔ ~ mbaʔ*. Expected Neverver *^{n}da*.

***^{n}dam^w ‘shout, yell, call out, (of dog) bark’** : Avava *^{n}dam*. For Ninde *^{n}dam^{w}e*, prenasalized initial [^{n}d] not explained. Expected Neverver *^{n}dam~^{n}dam^w*.

***^{n}dan ‘set (of the sun), set’** : Avava *^{n}dan*, Neverver *^{n}dan*, Ninde *^{n}den*.

***^{n}daŋ ‘breach (a dam)’** : Neverver *^{n}daŋ*. Expected Avava *^{n}daŋ*, Ninde *taŋ~təŋ*.

***^{n}das ‘go down’** : Avava *-^{n}dah*, Neverver *^{n}das*. Expected Ninde *ta*.

***^{n}deŋ ‘remove, take out’** : Avava *^{n}deŋ*, Neverver *^{n}daŋ~^{n}deŋ*. Probably cognate with Ninde *^{n}diŋ* (exp. *teŋ*).

***^{n}deŋ ‘slow’** : Neverver *^{n}deŋ*. Relationship with Ninde *^{n}daŋ^{n}daŋ* (exp. *teŋ*) uncertain. Expected Avava *^{n}deŋ*.

***^{n}dilim^{w} ‘swallow’** : Avava *-^{n}dilim*, Ninde *^{n}dlum* ~ *^{n}dlim*. For Neverver *^{n}dlom*, vowel quality of [o] not explained; vowel height of [o] not explained; vowel backness of [o] not explained; rounded [o] not explained.

***^{n}doŋon ‘count’** : Avava *^{n}doŋon*, Neverver *^{n}doŋon*. Expected Ninde *^{n}doŋon*.

***^{n}dor (+ *-kon) ‘burp’** : Avava *^{n}dor^{n}dorkon*. Probably cognate with Neverver *^{n}dor* (exp. *^{n}dorɣon*), Ninde *taʁpako* (exp. *^{n}doʁkon~^{n}dəʁkon*).

***ne-maur-ian ‘life’** : Avava *maur-ian*, Neverver *ne-maur-ian*, Ninde *<nu>mowoʁ-ijene*.

***ne-mes-ian ‘funeral’** : Neverver *ni-mas-ian*. Expected Avava *misian*, Ninde *nemejijan*.

***nele ‘voice’** : Ninde *nele*. Probably cognate with Avava *ele-* (exp. *ili~il~ilii*). Expected Neverver *nile*.

***neleme ‘tongue’** : For Ninde *nelime-*, vowel height of [i] not explained. Probably cognate with Avava *leme-n* (exp. *limi~lim~limii*), Neverver *neleme-n* (exp. *nilme*).

***neleso ‘scrotum and/or testicles’** : Avava *lese-~liso-*. Expected Neverver *nilso*, Ninde *neɮyjo*.

***nematˈa ‘eye’** : Ninde *nemete-,* Avava *mata-n*, for Neverver *nim^{n}dan*, vowel height of [i] not explained.

***nematˈa-nal ‘watch, clock, time’** : Avava *matanal*. Probably cognate with Ninde *nimytnaɮe* (exp. *nematanal*). Expected Neverver *nem^{n}danal*.

***ne^{m}bˈatu- ‘head’** : Avava *^{m}bat(u)- ~ ^{m}bat (nan)*, Neverver *ne^{m}bat-~ni^{m}bat*. Probably cognate with Ninde *nəmbute-* (exp. *nembytu*).

***ne^{m}baun ‘knee’** : Neverver *ne^{m}baun*. For Avava *^{m}bau-*, deleted final *n not explained. Expected Ninde *nembywun*.

***ne^{m}bet ‘breadfruit’** : Neverver *ni^{m}bet*. Probably cognate with Avava *(e)^{m}bet* (exp. *i^{m}bit*), Ninde *nimbytep* (exp. *nembete*).

***ne^{m}b^{w}ilok ‘kava’** : Probably cognate with Avava *^{m}b^{w}ilo<lo>k* (exp. *nu^{m}b^{w}lok~nu^{m}blok*), Ninde *nəmbiloʔ<ore>* (exp. *nə^{m}b^{w}əlaʔ~na^{m}b^{w}ilaʔ~na^{m}b^{w}əlaʔ~nə^{m}b^{w}ilaʔ*). Expected Neverver *nu^{m}b^{w}lok~nu^{m}blok*.

***nemelakas ‘cold’** : Avava *melekit ~ malah*, Ninde *nemaɮaʔ*. Expected Neverver *nimelɣas*.

***nemele^{m}biko^{n}s ‘t.o. skink commonly found under stones’** : Ninde *nemelembəkos*. Relationship with Avava *mele^{m}bieh* (exp. *mil^{m}bos*) uncertain. Expected Neverver *nimle^{m}biɣo^{n}s*.

***nemeta-liu ‘door’** : For Neverver *nim^{n}dali*, deleted final *e not explained. Probably cognate with Avava *matali* (exp. *mitliu~mitliui*), Ninde *nimitlu* (exp. *nemetal*).

***nemi^{n}da^{ŋ}g ‘t.o. tree’** : Avava *mi^{n}daŋ*, Neverver *nemi^{n}daŋ*. Probably cognate with Ninde *nəm^{n}daŋge* (exp. *nemi^{n}dəŋge~nemi^{n}daŋge*).

***nem^{w}at ‘snake’** : Avava *am^{w}at*, Ninde *namate*. For Neverver *nemat*, vowel quality of [e] not explained; vowel height of [e] not explained; vowel backness of [e] not explained; unrounded [e] not explained.

***nem^{w}et 'knife'** : Neverver *nim^{w}et*. Probably cognate with Avava *(n)em^{w}et* (exp. *amat~am^{w}at*), Ninde *nemes͡ei* (exp. *nymot*).

***ne^{n}dam^{w}at 'peace'** : For Avava *^{n}damat*, vowel quality of [a] not explained; vowel height of [a] not explained; vowel backness of [a] not explained; unrounded [a] not explained; For Ninde *netamate*, oral [t] not explained. Expected Neverver *ne^{n}dam^{w}at~ne^{n}damat*.

***ne^{n}di^{ŋ}giaβ^{w} 'reef slug'** : For Avava *^{n}dikiap*, oral [k] not explained; vowel quality of [a] not explained; vowel height of [a] not explained; vowel backness of [a] not explained; unrounded [a] not explained; For Ninde *na^{n}diktikj͡au*, vowel height of [a] not explained; inserted final [k] not explained. Expected Neverver *ne^{n}di^{ŋ}giapɸ~ne^{n}di^{ŋ}giaɸ*.

***ne^{n}doŋ 'mangrove.swamp'** : Avava *o^{n}doŋ*, Neverver *ni^{n}doŋ*, Ninde *ne^{n}doŋo.*

***ne^{n}siβir 'coconut lorikeet/lory (Triglossus haematodus)'** : Avava *siβir*, Ninde *nesβaʁa*. Expected Neverver *ne^{n}siβir*.

***ne^{ɳ}ɽe 'blood'** : Avava *e^{ɳ}ɽe ~ e^{ɳ}ɽei (nan)*, Neverver *ni^{ɳ}ɽe ~ ne^{ɳ}ɽe*. For Ninde *na-re*, vowel height of [a] not explained.

***ne^{ŋ}geritaŋ 't.o. tree with yellow fragrant flowers'** : For Avava *^{ŋ}geritaŋ*, vowel height of [e] not explained; inserted [i] not explained. Expected Neverver *ni^{ŋ}gritaŋ*, Ninde *neŋgəʁətaŋ~neŋgaʁitaŋ~neŋgəʁətəŋ~neŋgaʁətaŋ~neŋgəʁitaŋ~neŋgəʁitəŋ*.

***nep^{w}eno- 'face'** : For Avava *ipno-*, deleted initial *n not explained; vowel backness of [i] not explained; unrounded [i] not explained. Probably cognate with Ninde *nopmo-* (exp. *nembono*). Expected Neverver *nip^{w}no~nipno*.

***nes:am^{w}u ‘broom’** : Ninde *nesam*. For Neverver *ni-s:amu*, vowel height of [i] not explained. Expected Avava *sam^{w}ui~samu~sam^{w}u~samui*.

***net'e͡u ‘chicken’** : Avava *oto ~ otoi*. For Ninde *nete*, oral [t] not explained. Expected Neverver *net'e͡u*.

***netamat͡s ‘devil’** : Avava *tamat*, Neverver *netmas*. Probably cognate with Ninde *netemes* (exp. *ne^{n}damas*).

***netaxal ‘spider’** : For Ninde *netaƺe*, vowel height of [a] not explained. Probably cognate with Avava *lala* (exp. *ata:l*). Expected Neverver *netɣal*.

***netuswe ‘ocean’** : For Avava *tuswe*, vowel backness of [u] not explained; For Neverver *netusu*, vowel height of final [u] not explained. Probably cognate with Ninde *netes* (exp. *nu^{n}duwo*).

***neβ ‘(of fire) burn down to embers’** : For Avava *nep*, vowel height of [e] not explained. Expected Neverver *neɸ~nepɸ*, Ninde *neβ*.

***neβara-n ‘hand’** : Neverver *neβran*. For Avava *(a)βara-*, inserted initial [a] not explained; deleted final *n not explained; For Ninde *neβaʁa-*, deleted final *n not explained.

***neβe ‘ray’** : For Ninde *neβe͡i*, vowel height of final [e͡i] not explained. Probably cognate with Avava *eβe* (exp. *iβi~iβii~iβ*). Expected Neverver *niβe*.

***neβelu^{n}s ‘fill.laplap’** : For Avava *iβleh*, deleted initial *β not explained; inserted [β] not explained. Probably cognate with Neverver *βles* (exp. *neβlu^{n}s*). Expected Ninde *neβelse*.

***neβe^{m}ʙun 't.o. bird (green)'** : Avava *nep^{m}ʙun*. Probably cognate with Neverver *neβe^{m}ʙun* (exp. *neβ^{m}ʙ̥*), Ninde *nembin* (exp. *neβembyne*).

***neβereβeras 'hornet'** : For Ninde *neβaʙaβaʙas*,. Probably cognate with Neverver *leβriβras* (exp. *niβreβras*). Expected Avava *βirβirah*.

***neβilːak '®'** : Neverver *ni-βilax~ni-βilːax*. For Ninde *(nevilaʔ)*, labiodental [v] not explained. Expected Avava *βi^{n}dak*.

***neβles 'earth oven'** : For Avava *iβleh*, vowel height of initial [i] not explained; vowel height of [e] not explained. Probably cognate with Ninde *nalse* (exp. *neβɮe*). Expected Neverver *niβles*.

***neβule 'feather'** : Probably cognate with Avava *^{m}ʙule-* (exp. *βyle~βylei*), Neverver *niβilun* (exp. *neβle*), Ninde *nele-* (exp. *nuwulo*).

***neβ^{w}ene^{m}ʙ^{w}us 'bamboo'** : Probably cognate with Avava *β^{w}ini^{m}ʙuh* (exp. *β^{w}en^{m}ʙ^{w}uh~βen^{m}ʙuh*), Neverver *nu^{m}bu* (exp. *niβon^{m}ʙus*), Ninde *nanambuwo* (exp. *nywona^{m}buwo*).

***neβ^{w}enm^{w}e^{n}seu 'star'** : For Neverver *niβinːim^{n}so*, vowel quality of [i] not explained; vowel height of [i] not explained; vowel backness of [i] not explained; unrounded [i] not explained; deleted final *n not explained. Probably cognate with Avava *β^{w}inmese* (exp. *βanmeseui~β^{w}anm^{w}eseu~β^{w}anm^{w}eseui~βanmeseu*), Ninde *nanmysi* (exp. *nywonmosowu*).

***ni-teβ-teβ-ian 'growth'** : Neverver *ni-teβ-teβ-ian*, Ninde *nititipijene*. Expected Avava *niteptiβian*.

***niki + *melis 'redfish'** : For Avava *iki melih*, inserted initial [i] not explained; inserted [e] not explained; For Ninde *nikimijaɮe*, inserted final [m] not explained. Expected Neverver *niɣimelis*.

***nikim^{w}i 'chin?¤chin'** : Avava *kim^{w}i-* ~ *^{ŋ}gim^{w}i-*. Probably cognate with Ninde *nukum¤nesi-* (exp. *nəkum~nukum~nəkəm~nukəm*). Expected Neverver *niɣimu*.

***ni^{m}bi^{ŋ}g 'giant turban shell'** : Avava *i^{m}biŋ*. For Ninde *nəmbəŋge*, vowel quality of [ə] not explained; vowel height of [ə] not explained; vowel backness of [ə] not explained. Expected Neverver *ni^{m}biŋ~ni^{m}bi^{ŋ}k*.

***ni^{m}b^{w}isian 'language, talk, utterance, message, word, story'** : Avava *^{m}b^{w}is-ian*. Expected Neverver *ni^{m}bsian~ni^{m}b^{w}sian*, Ninde *nu^{m}b^{w}ijijan~nə^{m}b^{w}ijijan~nə^{m}b^{w}əjijan~nu^{m}b^{w}əjijan*.

***ni^{m}ʙus- 'tail'** : For Ninde *nimbyse-*, alveolar [s] not explained; fricative [s] not explained. Probably cognate with Avava *^{m}ʙus-* (exp. *i^{m}ʙeh*). Expected Neverver *ni^{m}ʙus*.

***ni^{m}ʙ^{w}ia^{ŋ}g 'swamp taro'** : For Neverver *ni^{m}ʙuaŋ*, vocalic [u] not explained; For Ninde *nəmbijaŋge*, inserted [m] not explained; oral [b] not explained; plain [b] not explained. Probably cognate with Avava *^{m}ʙuaŋ* (exp. *nu^{m}ʙiaŋ~nu^{m}ʙ^{w}iaŋ*).

***nimoxmok 'woman'** : For Neverver *(ni)moɣmox*,. Probably cognate with Avava *momok* (exp. *nimmok*). Expected Ninde *nimeʔmoʔ*.

***nimukut 'person'** : Neverver *nimɣut*. For Ninde *nymuŋgut*, vowel backness of [y] not explained. Probably cognate with Avava *muːt* (exp. *nimyut*).

***ni^{n}deliŋa-n 'ear'** : Neverver *ni^{n}dliŋan*. For Ninde *ni^{n}diliŋ a-*, vowel height of final [i] not explained. Probably cognate with Avava *^{n}diliŋa-* (exp. *^{n}delŋan*).

***ni^{n}dueβ 'canoe tree'** : For Avava *^{n}duep*, vowel height of [e] not explained. Expected Neverver *ni^{n}deɸ~ni^{n}depɸ*, Ninde *nu^{n}duwoβ*.

***nixi^{n}sa 'name'** : For Avava *isa-*, vowel length of initial [i] not explained; For Neverver *niɣi^{n}san*, inserted final [n] not explained. Relationship with Ninde *niki(ja)* (exp. *nijise*) uncertain.

***niβni 't.o. tree (Macaranga sp.) (Bis: navenue)'** : Avava *ipmi*. Expected Neverver *niβni*, Ninde *niβn*.

***niβ^{w}isoxo 'flesh'** : For Neverver *niβisxon*, inserted final [i] not explained. Probably cognate with Avava *ipso-* (exp. *upso:~upso:i*). Relationship with Ninde *nowoʔ-* (exp. *nuwujowo*) uncertain.

***noe^{ŋ}g 'island cabbage (Abelmoschus manihot)'** : For Avava *oweŋ*, inserted [w] not explained. Expected Neverver *nowe^{ŋ}k~noweŋ*, Ninde *nowəŋgo~nowoŋgo*.

***no^{m}bolo 'large, (big)'** : Avava *^{m}bolo*. Expected Neverver *nomlo*, Ninde *nombolo*.

***no^{m}boŋˈo- 'mouth'** : Avava *^{m}boŋo-*, Neverver *nomŋo-*, Ninde *nombo-ŋo*.

***nomoxo^{m}b 'skink'** : For Avava *omo:^{m}b*, plosive final [^{m}b] not explained. Probably cognate with Ninde *nemaʔamb* (exp. *nomowop*). Expected Neverver *nomoɣo^{m}b*.

***nono^{n}si- 'nose'** : For Avava *onos(i)-*, inserted initial [o] not explained; deleted final *i not explained. Probably cognate with Ninde *nuŋgun* (exp. *nonosi*). Expected Neverver *nono^{n}si*.

***nowe 'water, river'** : Avava *owe*. For Ninde *nowo͡i*, vowel quality of final [o͡i] not explained; vowel height of final [o͡i] not explained; vowel backness of final [o͡i] not explained; unrounded final [o͡i] not explained. Probably cognate with Neverver *nio* (exp. *nowo*).

***^{n}sal '(be) sick'** : Avava *-sal*, Neverver *^{n}sal*. Expected Ninde *seɮe*.

***^{n}sax 'be here'** : Neverver *^{n}sax*. Expected Avava *sak*, Ninde *saʔ*.

***^{n}saβi 'sneeze'** : Neverver *^{n}saβi*. For Avava *saβi*, vowel height of [a] not explained; deleted final *i not explained. Probably cognate with Ninde *seβije* (exp. *saβ*).

***^{n}sem^{w} 'chew'** : Avava *sem*, Neverver *^{n}sem*, Ninde *sam^{w}e*.

***^{n}sil 'go torch-fishing, go fishing at night with burning torches'** : Avava *-sil*. Probably cognate with Ninde *sle* (exp. *siɮe*). Expected Neverver *^{n}sil*.

***^{n}si^{m}ber 'reach'** : Avava *-si^{m}ber*. For Neverver *s^{m}ber*, oral initial [s] not explained; deleted *i not explained; For Ninde *spaʁ*, deleted *i not explained; deleted final *a not explained.

***^{n}sin '(of the sun) shine'** : For Ninde *sne*, deleted *i not explained. Expected Avava *sin*, Neverver *^{n}sin*.

***^{n}su^{m}b^{w} 'sit, sit down, be seated'** : Ninde *spo*. For Avava *-sum*, plosive final [m] not explained; nasal final [m] not explained. Expected Neverver *^{n}su^{m}ʙ̥*.

***nulu 'arrow'** : Neverver *nu-lu*. For Avava *ulu<kai>*, inserted [k] not explained; inserted [a] not explained. Expected Ninde *nyl*.

***nunwu- 'spirit'** : Avava *uno-*. For Neverver *nun:u<n>*, inserted final [n] not explained. Expected Ninde *nyn*.

***nuu^{n}d 'Palolo worm (Palola, Eunice viridis)'** : Ninde *nuwun*. Expected Avava *uut*, Neverver *n:u^{n}d*.

***nuus 'penis'** : Neverver *nus-*, Ninde *nuwuse-*, alveolar fricative [s] and unrounded final [e] not explained. Expected Avava *uuh*.

***ᶯɽoxᶯɽox ‘bend over’** : Probably cognate with Avava *rorok*, Ninde *roʔroʔ*. Expected Neverver *ᶯɽoɣᶯɽoɣ*.

***ŋar-ŋar ‘breathe’** : Avava *ŋarŋar*. For Ninde *ŋaʁaŋaʁa*, inserted medial [a] not explained; inserted final [a] not explained. Expected Neverver *ŋarŋar*.

***ŋaβ ‘pant’** : Avava *ŋap*. Probably cognate with Ninde *ŋaβeŋaβe* (exp. *ŋaβ~ŋəβ*). Expected Neverver *ŋapɸ~ŋaɸ*.

***ᵑgal ‘be stuck’** : Avava *kal*, Neverver *ᵑgal*. Probably cognate with Ninde *ŋgaɮe-saʁa* (exp. *kəl~kal*).

***ᵑgaras ‘drag fire with rakes through (garden site) while burning off’** : Avava *karah*. Expected Neverver *ᵑgaras*, Ninde *ŋgəʁa~ŋgaʁa*.

***ᵑgawats ‘cross?’** : Avava *kawat*, Neverver *ᵑgwas*. For Ninde *ŋgowus*, vowel qualities are both unexpected.

***ᵑgis ‘squeeze’** : Avava *ᵑgih*, Neverver reduplicated root *ᵑgis*. Expected Ninde *kije~kəje*.

***ᵑgit ‘we, us’** : Avava *ᵑgit*, Neverver *ᵑgit~iᵑgit*, Ninde *kəte*.

***ᵑgor ‘block’** : Avava *-kor*, Neverver *kor ~ ᵑgor(SIC)*. Probably cognate with Ninde *ŋgoʁo* (exp. *koʁ*).

***ᵑgun ‘bend’** : Avava *ᵑgun*, Neverver *ᵑgun*, Ninde *ŋgun*.

***ᵑgunᵑgun ‘crouch’** : Ninde *ŋgunŋgun*. Probably cognate with Neverver *ᵑgunᵑgun* (exp. *ᵑgunᵑk~ᵑgunŋ*). Expected Avava *ᵑgunᵑgun~kunᵑgun*.

***ŋiŋ 'smile'** : For Avava *ŋiŋi*, inserted final [i] not explained; For Neverver *ŋis*, final [s] not explained; For Ninde *niŋ*, alveolar initial [n] not explained.

***pːenox~*venox 'steal'** : Avava *pinok*, Ninde *<βene>βenaʔ*, Neverver *panɣo* (exp. *pːenoɣvnox*).

***pusel 'FRUST'** : Neverver *pusel*. For Ninde *pijaɮe*, vowel backness of [i] not explained; unrounded [i] not explained; palatal [j] not explained. Probably cognate with Avava *^{m}biβil* (exp. *pusil*).

***rːik(in) 'throw'** : Ninde *ʁətn ~ ʁkin*. Relationship with Neverver *rːik* (exp. *rːitn*) uncertain. Expected Avava *riːn*.

***rak 'clear ground'** : Avava *-rak*, Ninde *ʁaʔ*. For Neverver *rax*, fricative final [x] not explained.

***ran 'dawn'** : Avava *ran*, Neverver *ran*. Expected Ninde *ʁən~ʁan*.

***rer 'hot'** : For Avava *rar*, vowel height of [a] not explained. Expected Neverver *rer*, Ninde *ʁaʁa~ʁəʁa*.

***riβik 'good'** : Avava *-riβik*. For Neverver *irβix*, inserted initial [i] not explained; fricative final [x] not explained. Expected Ninde *ʁəβək~ʁiβək*.

***riβ^{w} 'run'** : Neverver *riɸ*, Ninde *ʁop.* For Avava *-rup*, vowel quality of [u] not explained.

***roŋ '(want,feel/), ear'** : For Avava *ro-ro*, deleted final *ŋ not explained; For Neverver *rot*, alveolar final [t] not explained; For Ninde *ʁo(ŋe)*, inserted final [e] not explained.

***roŋ- 'know'** : Neverver *roŋ<il>*. Probably cognate with Avava *-rokut* (exp. *roŋ*), Ninde *raŋgəle* (exp. *ʁoŋ*).

***sax 'ascend'** : Avava *-sak*, Neverver *sax*, Ninde *jaʔ*.

***saβ^{w} ‘dance’** : Avava *-sap*, Neverver *saɸ*, Ninde *jawo*.

***saβ^{w} ‘four’** : Probably cognate with Avava *iβat* (exp. *sap*), Neverver *i-βas* (exp. *sapɸ~saɸ*), Ninde *βes* (exp. *jawo*).

***seβ^{w} ‘cough’** : For Avava *sep*, vowel height of [e] not explained. Probably cognate with Neverver *^{n}soβ* (exp. *seɸ~sepɸ*). Expected Ninde *jawo*.

***sian ‘pregnant’** : Avava *sian*, Neverver *sian*, Ninde *<ʁome>sijene*.

***su^{m}b^{w}aʦ ‘join’** : Avava *su^{m}bat*. For Neverver *s^{m}bas<m>*, deleted *u not explained; inserted final [m] not explained. Expected Ninde *supas*.

***sup ‘shave, epilate, scrape out’** : Avava *sup*. Probably cognate with Ninde *wop* (exp. *jyp*). Expected Neverver *supɸ~suɸ*.

***sup:ak ‘nearly, almost, close, near’** : Avava *supak*, Neverver *sup:ax ~ sup:ak*. Expected Ninde *supaʔ*.

***ta^{ŋ}g ‘cry ($Deacon)’** : For Neverver *ta^{ŋ}g*, prenasalized final [^{ŋ}g] not explained. Expected Avava *taŋ*, Ninde *teke*.

***taβ^{w}ak ‘explode, detonate, blow up, (of lighning) thunder, sprout (of a seed)’** : Ninde *tawaʔ*. For Avava *-taβak*, vowel quality of [a] not explained. Expected Neverver *tβ^{w}ak~tβak*.

***teβ ‘sprout’** : Neverver *teɸ*. Probably cognate with Avava *-tep<tep>* (exp. *tip*). Expected Ninde *teβ*.

***titini^{n}s ‘play’** : For Avava *titinih*, inserted [i] not explained. Expected Neverver *t:ini^{n}s*, Ninde *titinise*.

***tiβin ‘bury’** : Avava *tipm*, Neverver *tβin*. Expected Ninde *tiβine*.

***tntn ‘roast’** : Avava *-tn-tn*, Neverver *tn*, Ninde *tn*.

***toroβ ‘jump’** : Avava *torop*. For Neverver *troβ*,. Expected Ninde *toʁoβ*.

***tuan ‘some’** : Avava *tuan*, Neverver *tuan*, Ninde *tuwane~*^{n}*duwane*.

***tur ‘stand up’** : Avava -√*tur*, Neverver *tur*. Probably cognate with Ninde *toʁ* (exp. *təʁa~taʁa*).

***tus ‘write’** : For Avava, reduplicated root *tih*, vowel backness of [i] not explained; unrounded [i] not explained; For Neverver *tos*, vowel height of [o] not explained. Expected Ninde *tyje*.

***tux ‘strike’** : Neverver *tux*, Ninde *tək*. Expected Avava *tuh*.

***tuxtux ‘hammer’** : Neverver *tux-tux*, Ninde *təktək*. Expected Avava *tutuh*.

***wel ‘lever (V)’** : Neverver *wel*. Expected Avava *wil*, Ninde *wol*.

***xan ‘eat’** : Avava -√*jan*, Neverver *ɣan*, Ninde *jen*.

***xaʦ ‘bite’** : Avava *-jat*, Ninde *jes*. For Neverver *xas*,.

***xavax ‘plant (yams)’** : For Neverver *ɣaβax*, inserted [a] not explained; bilabial [β] not explained. Expected Avava *avak*, Ninde *avaʔ*.

***xo ‘scrape out’** : Neverver *ɣo*. Expected Avava *joi~jo*, Ninde *jo*.

***xup ‘blow’** : Avava *up*. For Neverver *βuɸ*, bilabial initial [β] not explained. Expected Ninde *wupo~wup*^{w}*e*.

***βaŋ 'ROOT IMPL. BY be alight (of fire)'** : Neverver reduplicated root *βaŋ*. Expected Avava *βaŋ*, Ninde *βəŋ~βaŋ*.

***βere 'outside of the fenced residential compound, in exile, banished, shunned'** : Probably cognate with Ninde *βaʁ<kije>* (exp. *βaʁa~βəʁa*). Expected Avava *βir~βirii~βiri*, Neverver *βre*.

***βi 'be'** : Avava *-βi*. Probably cognate with Ninde *βi* (exp. *β*). Expected Neverver *βi*.

***βiar 'Viar.TOPOS'** : Avava *βiar*, Neverver *^βiar*. Expected Ninde *βijəʁ~βijaʁ*.

***βitum^{w}as 'laugh, laugh at'** : Ninde *βitumo ~ βitimo*. Expected Avava *βitm^{w}ah~βitmah*, Neverver *βtumas~βtum^{w}as*.

***βoβ^w 'rain'** : Avava *-βop*. For Ninde *wuwo*, vowel height of [u] not explained. Expected Neverver *βoɸ~βopɸ*.

***βul 'go shopping'** : Probably cognate with Neverver *βulβul* (exp. *βul*). Expected Avava *βul*, Ninde *βyle*.

***βu^{n}s 'ROOT IMPL. BY axe (to) (pieces) (many) (things)'** : Ninde reduplicated root *βyse*. Expected Avava *βus*, Neverver *βu^{n}s*.

***βuru^{n}s 'grab (hold of)'** : For Avava *-βuruh*, glottal final [h] not explained. Expected Neverver *βru^{n}s*, Ninde *woʁəse~woʁyse~wəʁəse*.

***β^{w}axβ^{w}ax 'be short, shorten'** : Neverver *βaɣβax*, Ninde *waʔwaʔ*. Expected Avava *β^{w}aβ^{w}ak~βaβak*.

***β^{w}el(em) 'come'** : Neverver *βlem*. For Avava *-β^{w}el*, deleted *e not explained; deleted final *m not explained. Probably cognate with Ninde *wul* (exp. *wolom*).

***β^w^er 'say'** : Neverver *βer,* Ninde <*ti*>*woʁ.* For Avava *βa*, deleted final *r not explained.

***β^w^ol 'buy'** : Ninde *wol*. For Neverver *βul*, vowel height of [u] not explained. Relationship with Avava *-leh* (exp. *β^w^ol~βol*) uncertain.

Proto SWB reconstructions

***amel ''** : Ninde *emel,* Naati *amel.* Expected Nahavaq *amel*.

***haʔ 'climb'** : Ninde *jaʔ*, Naati *haʔ*, Nahavaq *haʔ*.

***i 3SG.SUBJ.REAL**: Ninde *je ~ (i) ~ jet*, Nahavaq *i-*. Expected Naati *i*.

***kamem 'we, us, '** : Ninde *kamem*, Nahavaq *kamem*. Probably cognate with Naati *ŋgalyl* (exp. *kamem*).

***kinaŋg 'I, me'** : Naati *kinaŋg*. Probably cognate with Ninde *kəne* (exp. *kənaŋg~kinəŋg~kinaŋg~kənəŋg*), Nahavaq *kinaŋg ~ kinaŋ* (exp. *kinaŋ*).

***lamb^w^um 'yesterday'** : Nahavaq *lamb^w^um*. For Naati *lambum*, vowel height of [a] not explained; unrounded [a] not explained. Relationship with Ninde *ɮombun* (exp. *lambomo*) uncertain.

***latan 'down'** : For Ninde *ɮata-ne*, fricative initial [ɮ] not explained; For Naati *latan*, vowel length of [a] not explained; inserted final [n] not explained. Probably cognate with Nahavaq *leten* (exp. *lata*).

***larap 'afternoon, evening'** : For Ninde *ɮa-ʁap*, fricative initial [ɮ] not explained; For Nahavaq *larap*, vowel height of [a] not explained. Probably cognate with Naati *reßreß* (exp. *la:rap*).

***lip 'take'** : Naati *lip*, Nahavaq *lip*. For Ninde *lip*, approximant initial [l] not explained.

***lumus 'wash'** : For Naati *lumus*, vowel backness of [u] not explained; vowel backness of [u] not explained. Probably cognate with Ninde *lumus* (exp. *lymyso*), Nahavaq *lumʷus~lumus* (exp. *lumus*).

***luʔluʔ 'hide '** : Nahavaq *luʔluʔ*. Probably cognate with Ninde *ɮaʔ(ɮaʔ)* (exp. *lyʔlyʔo*), Naati *luʔluʔ* (exp. *lyʔalyʔ*).

***mamal 'red'** : Naati *mømal*. For Nahavaq *mamal*, vowel height of [a] not explained; vowel height of [a] not explained. Probably cognate with Ninde *mijaɮe* (exp. *mamale*).

***mata 'eye'** : For Naati *mata-*, vowel length of [a] not explained; For Nahavaq *mete-*, vowel height of [e] not explained. Probably cognate with Ninde *nemete-* (exp. *mate*).

***mataʔ 'fear'** : For Ninde *metaʔ*, vowel height of [e] not explained; For Naati *mataʔ*, vowel length of [a] not explained; For Nahavaq *mataʔ*, vowel height of [a] not explained.

***mbarap 'long'** : Naati *mba:rap*. For Ninde *paʁap*, deleted initial *m not explained; For Nahavaq *mbarap*, vowel height of [a] not explained; vowel height of [a] not explained.

***mboŋo 'mouth'** : Naati *mboŋo-*, Nahavaq *mboŋo-*. For Ninde *nombo-ŋo*, inserted initial [n] not explained; inserted [o] not explained.

***mes 'dead, die '** : For Ninde *mes*, deleted final *e not explained; For Naati *mes*, vowel height of [e] not explained; unrounded [e] not explained. Probably cognate with Nahavaq *(mis)mes* (exp. *mes*).

***mete-n-nal 'hour'** : For Nahavaq *mete-n-nal*, vowel height of [e] not explained. Probably cognate with Ninde *nimytnaɮe* (exp. *metennale*). Expected Naati *mi:tennal*.

***min 'drink'** : For Ninde *min*, deleted final *e not explained; For Naati *myn*, inserted final [n] not explained; For Nahavaq, reduplicated root *min*, inserted final [n] not explained.

***møβys 'white'** : Naati *møβys*, Nahavaq *meβus*. For Ninde *mys*, deleted final *e not explained.

***namʷui 'earthquake'** : For Ninde *nəmʷi{mʷi}*, inserted [mʷ] not explained; inserted final [i] not explained; For Naati *namui*, vowel height of [a] not explained; unrounded [a] not explained; For Nahavaq *na-mʷuj ~ mʷi*, palatal [j] not explained; inserted [mʷ] not explained.

***nan 'of'** : For Ninde *nene*, vowel height of [e] not explained. Probably cognate with Naati *nan ~ nin* (exp. *na*). Expected Nahavaq *na*.

***naʔut 'lice'** : Nahavaq *naʔut*. For Naati *noʔut*, vowel quality of [o] not explained; vowel height of [o] not explained; vowel backness of [o] not explained; rounded [o] not explained; vowel backness of [u] not explained. Probably cognate with Ninde *nuŋgut* (exp. *naʔyto*).

***ndelŋa 'ear'** : Naati *ndelŋa-*, Nahavaq *ndilŋe-*. Probably cognate with Ninde *niⁿdiliŋ a-* (exp. *ndelŋe*).

***ndipʷ 'heavy '** : Nahavaq reduplicated root *ndipʷ*. For Naati *ndip*, vowel backness of [i] not explained; unrounded [i] not explained. Expected Ninde *ndip*.

***n͡eilit 'yard'** : For Ninde *nolte*, vowel quality of [o] not explained; vowel height of [o] not explained; vowel backness of [o] not explained; rounded [o] not explained; deleted *i not explained; For Naati *n͡eilit*, unrounded [i] not explained; For Nahavaq *nilit*, inserted initial [n] not explained; inserted [i] not explained.

***nejum^{w} ‘house’** : For Naati *nijum*, vowel backness of [u] not explained; For Nahavaq *nejum*^{w}, vowel height of [e] not explained. Expected Ninde *nojymo*.

***nelambut ‘rat’** : For Naati *nelembut*,. Probably cognate with Ninde *noɮombute* (exp. *nelambyto*), Nahavaq *nalamb*^{w}*ut~nalambut* (exp. *lambut*).

***nelip^{w} ‘bullet’** : For Naati *nilip*, vowel backness of [i] not explained; unrounded [i] not explained; For Nahavaq *nelip*^{w}, inserted initial [n] not explained; inserted [e] not explained. Expected Ninde *nilip*.

***nembuŋ ‘day’** : Nahavaq *ni-mbuŋ*. For Naati *nembuŋ*, vowel height of [e] not explained; vowel backness of [u] not explained; rounded [u] not explained. Expected Ninde *nembyŋo*.

***nemb^{w}i ‘song’** : Ninde *namb*^{w}*i*, Naati *numbo͡i*. For Nahavaq *nimb*^{w}*i*, vowel backness of [i] not explained; unrounded [i] not explained.

***nemen ‘bird’** : For Ninde *nemen*, deleted final *e not explained; For Naati *nimen*, vowel height of [e] not explained; unrounded [e] not explained; inserted final [n] not explained; For Nahavaq *ne-men*, vowel height of [e] not explained; inserted final [n] not explained.

***neɳɽ͡ei ‘blood’** : Ninde *na-re*, Naati *niɳɽ͡ei*. Probably cognate with Nahavaq *ne-ɳɽej ~ ne-ndej* (exp. *ndrej*).

***neŋ͡ei ‘native almond’** : Nahavaq *niŋej*. For Naati *neŋ͡ei*, vowel height of [e] not explained. Probably cognate with Ninde *nəŋi* (exp. *ne~neŋe*).

***netel ‘rope’** : For Ninde *netel*, deleted final *e not explained; For Naati *nitel*, vowel length of [i] not explained; For Nahavaq *netel*, vowel height of [e] not explained.

***netes ‘sea’** : For Ninde *netes*, deleted final *e not explained; For Naati *nites*, vowel length of [i] not explained; For Nahavaq *ne-tes*, vowel height of [e] not explained.

***net͡eu ‘chicken’** : For Naati *nit͡eu*, vowel length of [i] not explained; For Nahavaq *ne-tew*, vowel height of [e] not explained. Probably cognate with Ninde *nete* (exp. *nyto*).

***†luej ‘water.LOC’** : Nahavaq *luwej*. Probably cognate with Ninde *now͡oi* (exp. *luwij*), Naati *nuw͡ei* (exp. *luej*).

***newuh ‘rain’** : For Naati *nuwuh*, vowel backness of [u] not explained; For Nahavaq *newuh*, vowel height of [e] not explained. Probably cognate with Ninde *nuwo* (exp. *nowuwo*).

***neʔeha ‘name’** : For Naati *neʔia-*, vowel height of [e] not explained. Probably cognate with Nahavaq *ne-ʔehe- ~ ne-ʔhe-* (exp. *ʔihe*). Relationship with Ninde *niki(ja)* (exp. *neʔa*) uncertain.

***neβaran ‘hand ’** : For Ninde *neβaʁa-*, deleted *n not explained; deleted final *e not explained; For Naati *niβara-*, vowel height of [a] not explained; vowel length of [a] not explained. Probably cognate with Nahavaq *βara-* (exp. *niβera*).

***neβet ‘stone’** : For Ninde *neβet*, deleted final *e not explained; For Naati *niβet*, vowel height of [e] not explained; unrounded [e] not explained; For Nahavaq *neβet*, vowel height of [e] not explained.

***βi ‘weave’** : Nahavaq *βi*. Probably cognate with Ninde *βiβi* (exp. *βi*), Naati *βiβ͡ei* (exp. *βi*).

***nimbʷuas ‘pig’** : Nahavaq *ni-mbʷuwes*. Probably cognate with Ninde *nəmbuwas* (exp. *nimbʷəse~nimbʷase*), Naati *nimbuas ~ nømbuas* (exp. *nimbuas*).

***nimetu 'coconut'** : For Ninde *nimyt*, deleted final *e not explained; For Naati *nime:tu*, vowel height of [e:] not explained; vowel height of final [u] not explained. Probably cognate with Nahavaq *nimetu* (exp. *mit*).

***nimomoʔ 'woman'** : Nahavaq *nimomoʔ*. Probably cognate with Naati *nimomoʔ* (exp. *numomø*). Expected Ninde *nimomoʔ*.

***neten 'ground'** : Ninde *netene*. For Nahavaq *neten*, vowel height of [e] not explained; inserted final [n] not explained. Probably cognate with Naati *nitan* (exp. *ni:te*).

***n͡ou 'cloth-like fiber at top of coconut tree'** : Naati *n͡ou*. Expected Ninde *n͡ou*, Nahavaq *now*.

***noʔo 'face'** : Nahavaq *noʔo-*. For Naati *noʔo-*, vowel backness of [o] not explained. Expected Ninde *noʔo*.

***noʔolsin naʔ͡ai 'belt around waist made of bark for holding penis wrapper'** : For Naati *noʔolsin naʔ͡ai*, vowel backness of [o] not explained. Expected Ninde *noʔolsinnaʔ͡ai*, Nahavaq *noʔolsinnaʔaj*.

***nuhyhy 'breast'** : Nahavaq *nu-huhu-*. Probably cognate with Ninde *nu-* (exp. *nuwywu*), Naati *nehuhu-* (exp. *nyhyhy*).

***rar 'sore'** : Naati *rar*, Nahavaq *rar*. For Ninde *ʙaʙa*, vowel height of final [a] not explained.

***raʔ 'work'** : Ninde *ʙaʔ*, Naati *raʔ*. For Nahavaq *(ra)raʔ*, inserted initial [r] not explained; inserted [a] not explained.

***roŋhur 'know '** : Nahavaq *roŋhur*. For Naati *roŋhur*, vowel backness of [u] not explained. Expected Ninde *ʙoŋəʙo~ʙəŋuʙo~ʙoŋuʙo~ʙəŋəʙo*.

***rop^w 'run'** : Ninde *ʁop*, Naati *rop*, Nahavaq *rop^w*.

***sem^{w}e 'chew'** : For Ninde *sam^{w}e*, vowel height of [a] not explained; For Nahavaq *sem^w*, vowel quality of [e] not explained; vowel height of [e] not explained; vowel backness of [e] not explained; unrounded [e] not explained. Expected Naati *sumø*.

***seʔer 'lost, missing '** : For Naati *seʔer*, vowel height of [e] not explained; For Nahavaq *seʔer*, vowel height of [e] not explained. Expected Ninde *seʔeʁe*.

***sian 'pregnant '** : For Naati *sian*, inserted final [n] not explained. Probably cognate with Ninde *{ʁome}sijene* (exp. *sijane*). Expected Nahavaq *sje*.

***sinmb^{w}uŋ 'forget '** : Nahavaq *sinmb^{w}uŋ*. For Naati *simbuŋ*, bilabial [m] not explained. Probably cognate with Ninde *sumbu* (exp. *sinmboŋo~sinmbəŋo~sinmbo*).

***sut 'some '** : Nahavaq *sut*. For Naati *sut*, vowel backness of [u] not explained. Expected Ninde *syto*.

***taris 'be (somewhere), '** : Probably cognate with Ninde *toʁ* (exp. *təʁəse~taʁise~təʁise~taʁəse*). Expected Naati *ta:ris*, Nahavaq *tiris*.

***ta͡tai 'father'** : Ninde *ta͡tai*, Nahavaq *tataj*. Probably cognate with Naati *ta͡tai* (exp. *ta:t͡ei*).

***taʔu (meaning unknown)** : Expected Ninde *taʔi*, Naati *taʔo*, Nahavaq *taʔ*.

***tembtemb 'defecate'** : Ninde *tembtemb*. For Naati *tembtemb*, vowel height of [e] not explained. Expected Nahavaq *tembtem*.

***teŋ 'cry'** : Naati *teŋ*. Probably cognate with Nahavaq *(tiŋ)teŋ* (exp. *teŋ*). Expected Ninde *teŋ*.

***toʔ 'stay, be '** : Probably cognate with Naati *toʔ* (exp. *tø*). Expected Ninde *toʔ*, Nahavaq *toʔ*.

***turtur 'sew'** : Nahavaq *turtur*. For Ninde *taʁtaʁ*, deleted final *o not explained; For Naati *turtur*, vowel backness of [u] not explained; vowel backness of [u] not explained.

***ty 'put'** : Ninde *tu*, Nahavaq *tu*. Expected Naati *ty*.

***wup^w 'blow'** : Naati *wup*. Probably cognate with Nahavaq *wup^w* ~ *hup^w* (exp. *wup^w*). Expected Ninde *wopo*.

***ʔaʔan 'eat '** : For Ninde *kaʔan*, deleted final *e not explained; For Naati *ʔaʔan*, inserted initial [ʔ] not explained; inserted final [n] not explained; For Nahavaq *ʔa-ʔan*, inserted initial [ʔ] not explained; inserted final [n] not explained.

***ʔil 'dig'** : For Ninde *kil*, deleted final *e not explained; For Naati *ʔil*, inserted initial [ʔ] not explained; For Nahavaq *ʔil*, inserted initial [ʔ] not explained.

***iβøs 'four'** : Nahavaq *iβes*. For Naati *iβøs*, vowel height of [ø] not explained. Probably cognate with Ninde *βes* (exp. *βose*).

***ʔy '2SG.SUBJ.REAL'** : Ninde *ku-*, Nahavaq *u-*. Expected Naati *y*.

***βaŋas 'speak (v)'** : Naati *βaŋas*. Probably cognate with Nahavaq *(βi)βaŋas* (exp. *βiŋes*). Expected Ninde *βəŋase~βəŋəse~βaŋase~βaŋəse*.

***β͡ei 'go'** : Ninde *βe*, Nahavaq *βej*. Expected Naati *β͡ei*.

***βin 'COP'** : Nahavaq *β(i)*. Probably cognate with Ninde *βi* (exp. *βine*), Naati *βin* (exp. *βy*).

***βiʔis 'lie, tell a lie'** : Naati *βiʔis*, Nahavaq *βiʔis*. Expected Ninde *βiʔəse*.

***βʷul 'buy'** : Nahavaq *βʷul*. For Ninde *wol*, deleted final *o not explained. Expected Naati *βul*.

***βysar 'outside ' :** Naati *βysar*. Probably cognate with Nahavaq *βusar~βʷisar* (exp. *βusar*). Relationship with Ninde *misaʁa* (exp. *βysaʁe~βysəʁe*) uncertain.