

THE PHONOLOGY OF BETSIMISARAKA MALAGASY

by

Timothy O'Neill

A dissertation submitted to the Faculty of the University of
Delaware in partial fulfillment of the requirements for the degree of
the Doctor of Philosophy in Linguistics

Fall 2015

© Timothy O'Neill
All Rights Reserved

THE PHONOLOGY OF BETSIMISARAKA MALAGASY

by

Timothy O'Neill

Approved:

Benjamin Bruening, Ph.D.
Chair of the Department of Linguistics & Cognitive Science

Approved:

George H. Watson, Ph.D.
Dean of the College of Arts & Sciences

Approved:

Anne L. Ardis, Ph.D.
Interim Vice Provost for Graduate & Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Jeffrey Heinz, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Peter Cole, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Irene Vogel, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Karthik Durvasula, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

This project was funded by two grants: a University of Delaware Global Travel Award and NSF Graduate Research Fellowship #1123609. These institutions appreciated the genuine need for documentation of Betsimisaraka, and generously supported my project.

I would never have secured either grant without the guidance of my advisor Jeffrey Heinz, but of course his role went far beyond consulting on grant applications. As my first (and second!) phonology instructor, he inspired my original passion for—and gave me a solid foundation in—this fascinating field. He supervised my progress in Madagascar from afar, suggesting approaches to elicitation and analysis. His greatest contribution, though, was comments on all drafts of this dissertation. His suggestions ranged from background reading to whole new analyses, and always how to write more clearly and concisely.

All eight months in Madagascar, I had the companionship and invaluable guidance of my language assistant and friend Ludger Mamy. He taught my Betsimisaraka language lessons, aided in planning and executing the logistics of field trips, and consistently helped participants feel at ease and engaged. He worked through endless

hours of transcription and sometimes-repetitive questions with a passion for language—and a smile. The second assistant, Precilla V., joined the project late, but was nevertheless able to jump right in and adapt.

In Antananarivo, unofficial tour guides Merrill Baker and her husband Naunau Tata Midard led me through obtaining a visa extension and research permit—even taking care of later complications after I'd left the city. In Tamatave, the project would have been slow to start without Laura Tilghman and Ludo Razafindramazina, who found me lodging and consulted on how best to do fieldwork (and everything else!) in Madagascar. Ludo, a Betsimisaraka speaker, provided the initial recording of the language for me while I was still proposing this project. Along with Ludger, he continued to provide informal elicitation once I'd returned home.

Also back home, Delaware's entire Linguistics Department was invaluable, especially administrative assistant Jane Creswell and the professors on my committee, which (aside from Jeff Heinz) consisted of Peter Cole, Irene Vogel, and Delaware alumnus Karthik Durvasula. Jane helped me organize a mountain of receipts, and generally seemed to know whatever I needed to do to keep the university and the NSF satisfied. Peter was my first and only Field Linguistics professor, occasionally consulted on techniques while I was in Madagascar, and made key suggestions on vastly expanding the literature review. Irene

taught me in Sounds of the Worlds Languages, the fascinating class where I did the Betsimisaraka pilot study with Ludo; during the writing stage, she ever pushed me toward more professional phrasings, especially guiding the methodology section. Karthik was both my most detail-oriented editor and essential supplier of knowledge, off the top of his head, of additional relevant language data or approaches I could cite for a more well-rounded analysis.

Outside of Delaware, Ileana Paul, leader of the Malagasy-Canadian project to describe Malagasy linguistic diversity, included me in the project whenever possible and made connections for me in the Malagasy linguistics community. One of these connections was Jean-Lewis Botouhely, a dean at the University of Antsiranana and the official sponsor of this project in Madagascar.

Preliminary analyses were presented at Manchester Phonology Meeting 2013, Phonology 2013 at University of Massachusetts Amherst, and Dialectal Variation in Malagasy, the Malagasy-Canadian Project's culminating conference in Diego-Suarez, Madagascar in 2014. I am grateful to all colleague who attended my talks and post presentation, and whose questions suggested solutions to the trickiest problems. I also benefited from informal discussions with other colleagues, whether about fieldwork with Tim McKinnon, Malagasy linguistics with Eric Potsdam, Maria Polinsky, and Alexander Adelaar, or phonological analysis with Nadya Pincus, Gina Cooke, Joanna Zaleska,

and Michael Dow.

Having a son while writing my dissertation meant I always needed baby-sitters. Rachel Bristol, Mike Fink, Weipeng Wang, Sandy Abu Adas, Carolina and Pratyush Sharma, and of course all four grandparents: these fine people looked after the “Little Sir” while I either enjoyed responsibility-free time or, more often, while I worked. A not insignificant portion of this dissertation, in fact, was written over all-day sessions at the Sharma household while Carolina and her son Ishaan looked after mine.

Speaking of those grandparents, my parents, Tom and Kate O'Neill, sent me to Madagascar on my original trip as an undergraduate and, years later, when I had doubts about proceeding to graduate school, they talked me back into it. They continued to encourage me throughout graduate school (indeed, throughout my entire education), and contributed to this specific project at each stage, from transportation to the airport (with dinner!), to making the long trip to Madagascar for a Christmas visit. They were even kind enough to provide me with the very first and very last necessary copy-edits of this work. My parents- and sister-in-law, James, Lynda, and Melissa Kelder also gave key support during the writing stage of the project.

Last but not at all least, my wife KellyAnn O'Neill was there every step of the way. Thanks to the generosity of her aunt Marilyn Kelly, KellyAnn accompanied me for more than half my time in Madagascar,

participating in language lessons, navigating the culture, and talking through the first versions of many of the ideas presented here. When we were both back at home and I was writing, she kept me consistently on task, doing everything from inquiring after my progress to reading drafts of chapters as they were written. Along with KellyAnn (although he did not provide much in the way of proofreading), I am also grateful to our son John Charles O'Neill II, who came into our lives at the midpoint of this project. Together and separately, no matter how busy I was, KellyAnn and John Charles consistently brought me much-needed laughter, warmth, and love.

TABLE OF CONTENTS

LIST OF TABLES.....	xii
LIST OF FIGURES.....	xix
ABSTRACT	xx
Chapter	
1 INTRODUCTION AND MOTIVATION.....	1
1.1 The Need for a Study of Betsimisaraka Phonology....	4
1.1.1 The Need for a Study of Malagasy Phonology...	4
1.1.2 The Need for a Study of Malagasy Dialects.....	7
1.1.3 The Need for a Study of Betsimisaraka.....	10
1.1.4 Concluding Remarks on Motivation.....	11
1.2 Overview of the phonological grammar.....	12
2 METHODOLOGY	15
2.1 Data Collection Methodology.....	15
2.1.1 Advisors and Assistants.....	16
2.1.2 Pre-Field-Trip Language Lessons.....	17
2.1.3 Selection of the Field Site.....	18
2.1.4 Elicitation Preliminaries.....	21
2.1.5 Questionnaires and Elicitation.....	22
2.1.6 Transcription	24
2.1.7 Methodological Recommendations for Future Work.....	25
2.2 Theoretical Methodology: The Case for Ordered Rules	25
3 OVERVIEW OF MALAGASY PHONOLOGY.....	29
3.1 Genealogy of Malagasy.....	30
3.1.1 Sources of Malagasy.....	30

3.1.2 Relatedness of Malagasy Dialects.....	31
3.2 Segments of Malagasy.....	36
3.2.1 Consonants	36
3.2.2 Vowels	42
3.2.3 Sound-Based Divisions of Malagasy Dialects..	49
3.3 Syllable Structure.....	52
3.4 Morphology	54
3.4.1 Verbal Morphology.....	55
3.4.2 Compounding	57
3.4.3 Reduplication	59
3.5 Segmental Phonology.....	59
3.5.1 Consonant-Consonant Interactions.....	59
3.5.2 Alternations involving Vowels.....	68
3.6 Suprasegmental Phonology: Stress.....	75
3.6.1 Penultimate Stress and Final-Stress Exceptions	75
3.6.2 Antepenultimate Stress and Epenthetic Vowels	78
3.6.3 Potential-Coda Neutralization and Elision.....	86
3.7 Conclusions	95
4 SEGMENTS OF BETSIMISARAKA.....	99
4.1 Vowels	100
4.1.1 Phonemic and Allophonic Vowels.....	100
4.1.2 Vowels resulting from coalescence.....	104
4.2 Consonants.....	106
4.2.1 Distribution of Nasals.....	107
4.2.2 Obstruents of Betsimisaraka.....	110
4.2.3 Approximants of Betsimisaraka.....	112
4.3 Chapter Summary.....	113

5	STRESS AND STRESS-ASSIGNMENT IN BETSIMISARAKA...	114
5.1	Correlates of Stress in Betsimisaraka.....	115
5.2	Penultimate Stress Assignment and Antepenultimate Stress.....	122
5.3	Stress Assignment and Hiatus Resolution.....	126
5.4	Conclusion	130
6	SEGMENTAL ALTERNATIONS OF BETSIMISARAKA.....	132
6.1	Alternations affecting consonants.....	133
6.1.1	Nasal Assimilation.....	134
6.1.2	Consonant Cluster Reduction.....	138
6.1.3	Fortition	142
6.1.4	Lexically Triggered Voiceless Consonant Deletion	160
6.1.5	Variant Distribution of Root-Initial [v].....	164
6.1.6	Coda Deletion and Vowel Epenthesis.....	169
6.1.7	Coda Neutralization.....	174
6.1.8	Nasal Dissimilation.....	186
6.1.9	Distinctive Features for Consonants.....	190
6.2	Alternations Affecting Vowels.....	191
6.2.1	Stress-conditioned Alternations of High and Low Vowels.....	191
6.2.2	Backness Harmony of the 1S Suffixes.....	194
6.2.3	Harmony of Epenthetic Vowels in Betsimisaraka	202
6.2.4	Hiatus-Resolution Strategies.....	213
6.3	Final Rule Ordering.....	224
6.3.1	Crucial Rule Orderings: Consonants.....	224
6.3.2	Crucial Rule Orderings: Vowels.....	229
7	MAJOR CONTRIBUTIONS AND CONCLUSIONS.....	234
	REFERENCES.....	240
Appendix		
A	SWADESH LIST FOR BETSIMISARAKA MALAGASY.....	245
B	GLOSSARY OF ELICITED FORMS.....	248
C	IRB APPROVAL LETTERS.....	257

LIST OF TABLES

Table 3.1:	Malagasy words with non-Austronesian origins.....	31
Table 3.2:	A segmental inventory of the consonants of Official Malagasy..	36
Table 3.3:	Examples of palatalization.....	39
Table 3.4:	NC sequences of Malagasy, re-analyzed as pre-nasalized stops.....	40
Table 3.5:	Minimal pairs for regional Malagasy dialects, as compared to OM	41
Table 3.6:	Evidence for the derivation of dialectal /ŋ/ in Malagasy as a reflex of the same sound in Indonesian.....	41
Table 3.7:	Cognates in Merina (OM) and other dialects, where the former has /z/ and the latter has /j/.....	42
Table 3.8:	Phonological vowels of OM.....	42
Table 3.9:	Divergent pronunciations of <ia> in OM.....	46
Table 3.10:	Divergent pronunciations of <oa> in OM.....	47
Table 3.11:	Conversational pronunciations of <ao> and <ai> in OM.....	48
Table 3.12:	Consonant correspondences between different Malagasy dialects	51
Table 3.13:	Summary of transcriptions following a nasal-coda model of OM syllable structure vs. a pre-nasalized-stop interpretation.....	53
Table 3.14:	Active verbal morphology of OM.....	55
Table 3.15:	Nouns and verbs derived from the same root and SFP.....	56
Table 3.16:	Passive verbal morphology of OM.....	56
Table 3.17:	Passive and active verbal morphology evincing suppletion in root form.....	57
Table 3.18:	Four compounds of OM, as derived from their two roots.....	58

Table 3.19: Genitive compounds, all formed from ['tranu] ("house").....	58
Table 3.20: Nasal assimilation demonstrated by attachment of a nasal-final prefix.....	60
Table 3.21: Examples of fortition in compounds where the first root is nasal-final	63
Table 3.22: Examples of fortition where the reduplicant is nasal-final.....	63
Table 3.23: Alternations wherein a consonant appears in the suffixed form, but did not appear in the bare root.....	64
Table 3.24: Two additional alternations between bare and suffixed forms....	65
Table 3.25: Data from 3.23 and 3.24 featuring URs for the root that include the consonants heard only in the suffixed form.....	66
Table 3.26: Three examples of underlying vowel pairs surfacing as a single vowel.....	69
Table 3.27: Two examples of underlying vowel pairs surfacing as a pair of vowels	70
Table 3.28: Pairs of words that are pronounced generally the same due to vowel devoicing.....	71
Table 3.29: Vowel devoicing in the Taisaka dialect.....	72
Table 3.30: Different forms of the same verb that show stress-based alternation between [e] and [i].....	73
Table 3.31: Two- and three-syllable words exhibiting penultimate stress.....	75
Table 3.32: Borrowed words exhibiting final stress.....	76
Table 3.33: Three compounds that do not exhibit alternating stress.....	77
Table 3.34: Antepenultimately stressed words, all ending with weak syllables	78
Table 3.35: Two minimal pairs for stress.....	79
Table 3.36: A derivation table showing the ordered-rule analysis of antepenultimate stress and penultimate stress.....	80
Table 3.37: Taisaka Compounds wherein the first root is antepenultimately stressed and the second is consonant-initial.....	82

Table 3.38:	Antepenultimately stressed words in Sakalava, a Western dialect	83
Table 3.39:	Antepenultimately stressed words found in a survey of Northern dialects.....	85
Table 3.40:	Antepenultimately stressed roots and their corresponding passive forms, along with Malay cognates.....	87
Table 3.41:	Antepenultimately stressed roots with an older and newer passive form.....	88
Table 3.42:	Cases where Malay word coda [t] diverges to [r] in other Austronesian dialects and [tr̩] in Malagasy.....	89
Table 3.43:	Borrowings from English where the borrowed word ends in [t] in the original language, but ends with [tra̩] in OM.....	89
Table 3.44:	Borrowings from French where the borrowed word ends in an [r] or [l] in the original language, but ends with [tra̩] in OM.....	90
Table 3.45:	Exceptional cases where the root is antepenultimately stressed, ending in [ka], but the pre-passive consonant can be another consonant aside from the usual [h].....	90
Table 3.46:	Taisaka nominals and passives.....	91
Table 3.47:	Antepenultimately stressed words ending in [ŋa], from Northern dialects.....	92
Table 3.48:	Antepenultimately stressed and penultimately stressed versions of the same word, from the Northern dialects.....	94
Table 3.49:	Variants of the weak syllable [ŋa] in the Betsimisaraka of Vavatenina, and their OM variants.....	94
Table 4.1:	Phonemic vowels of Betsimisaraka.....	100
Table 4.2:	Active and passive forms of two verbs, showing disparate stress-conditioned behavior of /u/ and /o/.....	101
Table 4.3:	Disparate stress-conditioned behavior of /i/ and /e/.....	102
Table 4.4:	Surface realizations of /au/ and /ai/ sequences in Official and Betsimisaraka Malagasy.....	104
Table 4.6:	Formation of rising diphthongs in Betsimisaraka.....	105

Table 4.7:	Phonemic Consonants of Betsimisaraka and their articulations	107
Table 4.8:	Pairs of words that differ only by stress assignment and place features of the non-labial nasals.....	108
Table 4.9:	Minimal pairs for /n/ and /ŋ/ in Betsimisaraka.....	109
Table 4.10:	Different tense forms of verbs and adjectives employing the /n-/ past and /m-/ present tense prefixes.....	110
Table 5.1:	Betsimisaraka words exhibiting the usual, penultimate stress pattern	123
Table 5.2:	Betsimisaraka words exhibiting the frequent, antepenultimate stress pattern	124
Table 5.3:	Compounds wherein the first root is typically antepenultimately stressed, and the second is vowel-initial....	125
Table 5.4:	Surface stress in words featuring underlying hiatus.....	127
Table 5.5:	Active and imperative forms of verbs with vowel-final roots....	128
Table 6.1:	Words containing homorganic consonant clusters in Betsimisaraka	134
Table 6.2:	Passive and active forms of representative verbs for each place of articulation.....	135
Table 6.3:	Compounds using ['orunu] ("nose") as the first root.....	136
Table 6.4:	Passive and active forms for vowel-initial roots.....	137
Table 6.5:	Sample derivations of four verbs in the active present form...	138
Table 6.6:	Compounds wherein the first root is underlyingly consonant-final	139
Table 6.7:	Compounds consisting of an antepenultimately stressed word followed by a consonant-initial word.....	140
Table 6.8:	Derivation of compounds that undergo CC reduction or CC assimilation	141
Table 6.9:	Words containing the only acceptable CCs in Betsimisaraka	143

Table 6.10:	Compounds incorporating the genitive [n], which triggers fortition	144
Table 6.11:	Compounds wherein the nasal-final first root /homaN/ triggers fortition.....	145
Table 6.12:	Compounds wherein the nasal-final first root /enij/ triggers fortition.....	145
Table 6.13:	Summary of fortition of non-continuant segments in Betsimisaraka	146
Table 6.14:	Sample derivations of four compounds.....	153
Table 6.15:	Sample derivations of four compounds undergoing fortition...154	
Table 6.16:	Compounds wherein the first root has antepenultimate stress	155
Table 6.17:	Sample derivations of compounds undergoing fortition and CC Reduction.....	156
Table 6.18:	Phonemic Consonants of Betsimisaraka and their articulations	157
Table 6.19:	[+syllabic] segments of Betsimisaraka and their distinctive features	158
Table 6.20:	Coronals ([-dorsal -labial] segments) of Betsimisaraka and their distinctive features.....	159
Table 6.21:	Imperative and active forms of verbs that undergo fortition..160	
Table 6.22:	Imperative and active forms of verbs wherein the onset is absent in the active form.....	161
Table 6.23:	Sample derivations of active present verb forms.....	163
Table 6.24:	Imperative and active forms of verbs wherein the [v] onset of the imperative does not appear in the active.....	164
Table 6.25:	Sample derivation of [v]-initial roots.....	169
Table 6.26:	Active and Passive forms of verbs wherein the fricative that occurs just before the suffix in the passive form is entirely absent in the active.....	170
Table 6.27:	Sample derivations of active and passive forms of select verbs	173

Table 6.28:	Distinctive feature chart for [-syllabic +continuant] segments that are defined for [\pm strident].....	173
Table 6.29:	Active and Present forms of verbs wherein the active ends with the weak syllable [ŋa].....	175
Table 6.30:	Active and Present forms of verbs wherein the active ends with the weak syllables [ka] and [tʃa].....	177
Table 6.31:	Summary of the changes that occur when consonants are morpheme-final position.....	180
Table 6.32:	Sample derivations for the suite of rules that neutralize, then eliminate, codas through vowel epenthesis.....	183
Table 6.33:	Parallel derivation of two labial-fricative-final roots.....	184
Table 6.34:	Near-minimal pairs wherein the only aspects to change are the stress assignment and the place of the nasals.....	186
Table 6.35:	Compounds containing [...ŋVŋ...] strings.....	187
Table 6.36:	Derivation of monomorphemic “greens” and the compound translated here as “goose greens”.....	190
Table 6.37:	Distinctive features for [-syllabic] segments of Betsimisaraka	190
Table 6.38:	Stress-conditioned height alternation of non-low vowels.....	192
Table 6.39:	The root /dalov/ (“pass/drop by”) derived in the active and imperative	193
Table 6.40:	Distinctive features necessary so far for [+syllabic] segments	194
Table 6.41:	The behavior of 1S suffix “-ko” in both uses in Official Malagasy	195
Table 6.42:	The 1S passive suffix attached to high-vowel-final roots.....	196
Table 6.43:	Data from two Betsimisaraka dialects for 1S passive forms....	196
Table 6.44:	The behavior of the 1S passive suffix on some non-high-vowel-final forms in Betsimisaraka.....	197
Table 6.45:	Behavior of 1S possessive suffixes in Betsimisaraka.....	199
Table 6.46:	Optional Vowel harmony of Epenthetic Vowels.....	203

Figure 6.47: Possible pronunciations of a root that underlyingly has a dorsal stop as its word-final consonant.....	204
Table 6.48: Optional Vowel harmony of Epenthetic Vowels.....	206
Table 6.49: Four derivations that show the four grammars available to speakers of Betsimisaraka Malagasy.....	211
Table 6.50: Desyllabification of unstressed high vowels, in service of hiatus resolution.....	215
Table 6.51: Desyllabification of unstressed high vowels, in service of hiatus resolution.....	215
Table 6.52: Desyllabification of unstressed high vowels.....	216
Table 6.53: When an underlying form contains a VV sequence featuring an unstressed /a/, that vowel does not surface.....	218
Table 6.54: Formation of mid lax vowels as the surface realization of sequences of /a/ followed by high vowels.....	220
Table 6.55: Examples where the orthography suggests an underlying hiatus and the pronunciation includes a lax mid vowel.....	220
Table 6.56: Derivation showing the interaction of the three rules proposed in this subsection.....	222
Table 6.57: All distinctive features necessary for [-consonantal] segments	223
Table B.1: Correspondences between Betsimisaraka spelling and IPA.....	248

LIST OF FIGURES

Figure 1.1:	Map of Madagascar and surrounding areas.....	3
Figure 2.1:	Map of locations relevant to the project.....	19
Figure 3.1	Dialect map of Madagascar, including five major divisions.....	34
Figure 5.1	Spectrogram (in Praat) of ['botʃa] (“chubby”).....	116
Figure 5.2	Spectrograms (in Praat) of ['lalə] and [la'lane].....	118
Figure 5.3:	Spectrogram (in Praat) of [fufun'dzavatʃa] ("smell-of-something")	119
Figure 6.1:	Hasse Diagram, representing all crucial orderings in the consonant-affecting rules of Betsimisaraka Malagasy.....	228
Figure 6.2	Hasse Diagram, representing all crucial orderings in the vowel-affecting rules of Betsimisaraka Malagasy.....	232
Figure C.1:	IRB approval letter for 2011.....	258
Figure C.2:	IRB approval letter for 2012.....	259
Figure C.3:	IRB approval letter for 2013.....	258

ABSTRACT

This document constitutes the first phonological grammar Betsimisaraka Malagasy, a form of the Malagasy (Austronesian) language spoken in the island nation of Madagascar. Betsimisaraka specifically is the name of an ethnic group with approximately a million members living on the East Coast of the island, as well as the various dialects they speak. However, the forms of Malagasy, which are generally believed to be closely related dialects of the eighteen traditional ethnic groups, may actually constitute separate languages (Bouwer 2005, 2007). While generative phonological approaches to Malagasy do exist (Dziwirek 1989, Albro 2005), this is the first to focus on any variety of Betsimisaraka, a mostly undescribed variety in any area of theoretical linguistics, although see Kikusawa (2006, 2007), who focuses on syntax.

Beginning with the knowledge of previous descriptions of related dialects, mostly Official Malagasy (OM), known alternations were re-elicted in the Betsimisaraka town of Vavatenina in 2011 and 2012, and previously unanalyzed alternations were also noted and elicited from multiple speakers in the town. In the end, seventy speakers participated in the study.

All alternations observed during elicitation are analyzed using a set of nineteen ordered derivational rules—two of these are optional between and among speakers. Based on the processes these rules describe, this work also proposes a set of minimal features for the phonemes of the language. Throughout the analysis, alternations are first presented atheoretically and then one or more rules is proposed. This analysis is then compared to previous approaches to similar data in related dialects (usually OM). Many of the rules are in a counterbleeding relationship, indicating a high degree of opacity in the language.

Distinctions between Betsimisaraka and OM start with their inventories: Betsimisaraka has four sounds [o],[ŋ], [ɛ], and [ɔ] that are rarely or never used in OM, the first two of which are phonemic. Both dialects have word-final consonant neutralization processes, stress-conditioned vowel heightening, and palatalization of velar consonants, though the mechanics of the processes differ between the two varieties. Betsimisaraka also has processes unknown in OM, though they do occur in other dialects, such as word-final nasal deletion and copy-epenthesis, as well as two processes unattested in previous descriptions of any dialect of Malagasy: two similar but distinct allomorphies of first-person singular suffixes and nasal dissimilation.

Several aspects of the Malagasy spoken in Vavatenina suggest that it should be classified as Northern Betsimisaraka, based on

analysis of the data and previous dialectological findings: existence of phonemic /o/, lack of existence of phonemic /j/, and the aforementioned copy-epenthesis pattern. It is hoped that from the great distinction demonstrated between OM and Betsimisaraka, this work—along with other works of dialectology—can help put an end to the myth of Malagasy linguistic unity.

Chapter 1

INTRODUCTION AND MOTIVATION

Introduction and Overview

Betsimisaraka ([be,tsimi'saraka]) is one of the various regional forms of the Malagasy language spoken throughout the island of Madagascar. Like all forms of Malagasy, Betsimisaraka is an Austronesian language containing significant Bantu vocabulary, the result of a still-unclear settlement pattern of the island that has its roots in Indonesia. This work follows the custom of naming Malagasy dialects after the traditional ethnic groups who speak them, in this case the million or so members of the Betsimisaraka ethnic group, who inhabit the East Coast of the island. What follows is the first phonological grammar of this language variety: an inventory of the underlying sounds and a presentation of segmental and suprasegmental alternations that affect those sounds in given environments.

This work establishes the phonological system of Betsimisaraka as distinct from that of Official Malagasy (OM), the island's standard language. The analysis here includes the many aspects the two language varieties have in common, as well as all of the following

differences. In its inventory, Betsimisaraka has four sounds [o], [ŋ], [ɛ], and [ɔ] that are rarely or never used in OM, the first two of which are phonemic. Like OM, Betsimisaraka has an opaque word-final neutralization process, although the mechanics are different. Stress-conditioned vowel height alternations are also present but distinct in the two languages. Betsimisaraka also has three vowel harmony processes where OM has none, two of which comprise allomorphies of first-person singular suffixes, and the third of which results in copy epenthesis. This copy epenthesis pattern, along with existence of /o/ and /ŋ/ and an optional word-final nasal deletion process, will be shown to be phonological aspects of other non-official forms of Malagasy. Finally, Betsimisaraka has nasal dissimilation affecting surface realizations of its velar nasal phoneme, constituting a system that research for this project did not reveal any precedent in previous literature.

Fieldwork for this study was conducted in the town of Vavatenina, Madagascar; as shown on the map in Figure 1.1, below, this town is close to the center of the Betsimisaraka region. Elicitations were performed over several trips to Vavatenina in 2011 and 2012. Therefore, while this study will use the name of the ethnic group “Betsimisaraka” as a shorthand to refer to the language it describes, other authors might describe it more fully as South or North Betsimisaraka, depending where the region is split in half. It can be

said most accurately, then, to describe Vavatenina's Betsimisaraka, although from the data presented here, it will be shown to most likely ascribe to Northern Betsimisaraka. In one case, there is enough data to compare the Malagasy spoken in Vavatenina with Maroantsetra, an even more northern city in the Betsimisaraka region.

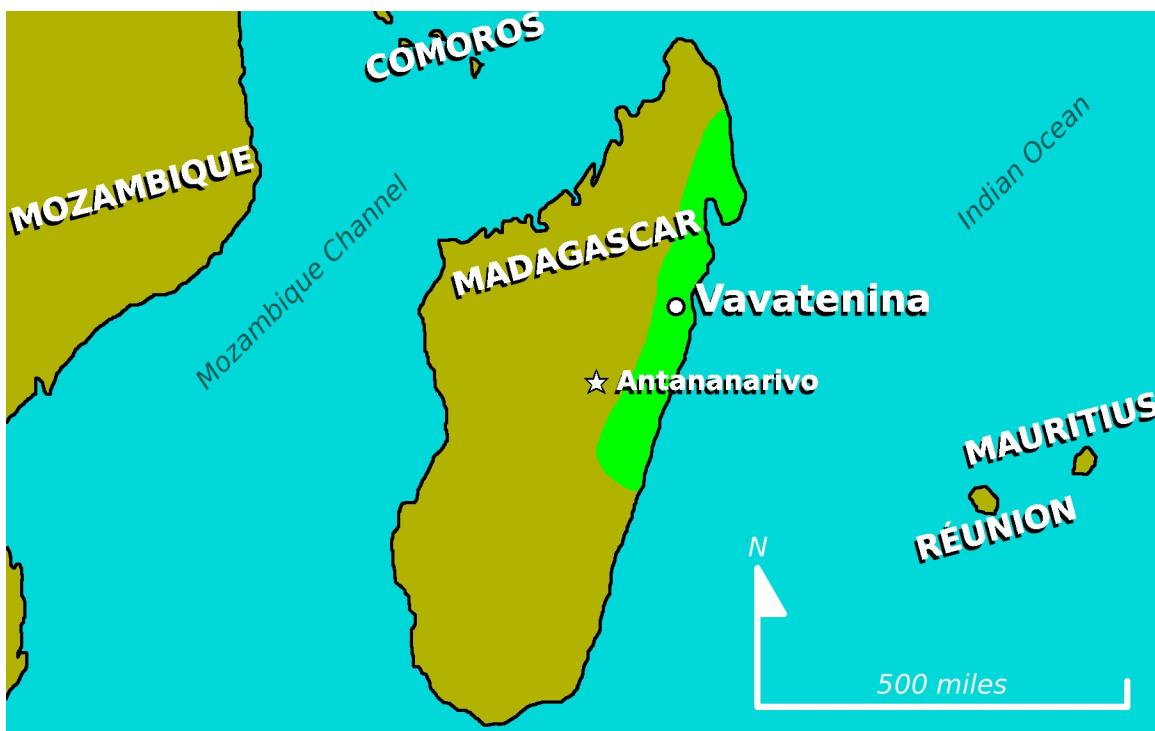


Figure 1.1: Map of Madagascar and surrounding areas, including mainland Africa to the west, the archipelago nation Comoros to the north, and to the east, the country of Mauritius and French overseas department Réunion. The approximate area associated with the Betsimisaraka ethnic group is highlighted on the East Coast, as is the country's Official-Malagasy-speaking capital Antananarivo and Vavatenina, where elicitations for this study took place. Like all maps include here, it was drawn for this work.

This chapter consists mainly of a theoretical justification for the project in Section 1.1. This constitutes an assessment of the quality

and quantity of previous work on Malagasy phonology¹, about Malagasy dialectology, and specifically about Betsimisaraka, concluding that is a valuable contribution to cover the phonology of this specific language variety. The section concludes, in 1.2, with an overview of the work as a whole.

1.1 The Need for a Study of Betsimisaraka Phonology

This section motivates the entirety of the mission of this work: to provide, for the first time, a comprehensive phonology of Betsimisaraka Malagasy. The subsections answer why the phonology of Malagasy deserves greater attention (1.1.1), why the language varieties of Malagasy, typically called dialects, are in need of study (1.1.2), and why Betsimisaraka in particular represents an important, somewhat overlooked area or research (1.1.3). These reasons are summarized in the final subsection (1.1.4).

1.1.1 The Need for a Study of Malagasy Phonology

The sounds of Malagasy have been at least touched on by every author attempting to describe the language, but this work is primarily concerned with generative phonology, an approach whereby surface forms are derived, along featural dimensions, from abstract underlying representations. This section, then, indicates which previous works are

¹ Discussion of the actual findings from the works on Malagasy phonology will be left to Chapter 3

most useful for phonological analysis by highlighting those that reference underlying forms and features in their analyses. The small number of these, though, will motivate the current phonological approach to the language.

Although earlier dictionaries provided some pronunciation advice, the first thorough description of the sounds of Malagasy comes from missionary David Griffiths (1854), describing the Merina² dialect that would eventually form the basis for OM. Although his approach is inconsistent³, Griffiths does offer observations akin to those of a modern linguist, pertaining to phonotactics and inventory: “Every word terminates with a vowel, and a vowel always intervenes between two consonants except the double and triple consonants” (p. 8), i.e. affricates and nasal-oral stop clusters. In his chapter on etymology, he describes, in purely segmental/orthographic terms, many morphophonological interactions alternations that will be covered here. Later, Garvey (1964) thoroughly describes the articulation of all OM sounds, but her analysis of morphophonological alternations speaks only of specific orthographic changes, with no generalization offered. Pedagogical grammars, such as Rajemisa-Raolison (1969), with the goal of teaching the language, describe all alternations segmentally.

² Griffiths calls it “Ankova,” actually the place name where the Hova, or Merina, people live.

³ Some more poetic diversions include a comparison to Welsh (p. 7) and stating “The natural softness of Malagasy excludes harsh, or guttural sounds” (p. 8).

Even following the establishment of generative phonology, some later analyses continue to approach alternations segmentally, with only occasional reference to features or underlying forms. Rakotofiringa (1981), looking at consonant alternations between suffixed and unsuffixed verbs, takes a featural, phonetic approach, employing statistical analysis as opposed to one based on underlying forms. In contrast, Erwin (1994) proposes URs to explain these same interactions, but takes no recourse to features. Similarly, in analyzing the many alternations that occur following attachment of the active suffix, Paul (1996) does not ignore features in her analysis, but her rules⁴ primarily employ segment sets, and URs are not mentioned.

Focusing more on prosody and morphology, both Keenan & Razafimamony (1995) and Keenan & Polinsky (1998) provide in-depth analysis of the morphology, but again speak of alternations segmentally or orthographically rather than featurally, following Rakotofiringa. (For example, when speaking of nasal assimilation, these articles refer to nasals converting to *<m>* before ** and *<p>*, but not of the non-orthographically marked conversion to [ŋ] in front of [g] and [k].) Most recently, Rajaoranimanana's *Modern Grammar* (2004), while never employing the word "phonology," includes a chapter on word-formation with a primarily segmental, but occasionally featural, analysis of morphophonology, with no mention of underlying forms.

⁴ Paul's are formalized as functions in Boolean logic, rather than as standard phonological rules.

This is not to say, of course, that the current work represents the first attempt at explaining a dialect of Malagasy using a UR-based and featural understanding of phonology. Although her short sketch does not cover all alternations, Dziwirek (1989) proposes URs and feature-based, ordered rules to explain alternations in Official Malagasy. Albro's (2005) dissertation, fully up-to-date with modern phonology, employs underlying forms and a set of ranked, feature-based constraints to explain all alternations of Official Malagasy under the rubric of Optimality Theory. None of this is to say that these two works (and others⁵), concerned with phonology itself, are inherently superior to those from linguists primarily concerned with related-but-distinct disciplines such as phonetics or morphology.

1.1.2 The Need for a Study of Malagasy Dialects

Generally speaking, linguists assume that each of the traditional eighteen ethnic groups of the island has its own dialect⁶, but due to OM's prestige, most studies (e.g. all of those mentioned in the previous subsection) have focused on it. There is no reason to believe, though, that OM is a representative dialect or that it is more closely related to the proto-Malagasy from which all dialects derive. It is derived instead

5 Two doctoral dissertations from the University of Paris cover the phonology of Malagasy, according to Dez (1991). These works, Builles (1987) and Andriantsolo (1989) were regrettably unavailable to the author of this work.

6 As will be further discussed in Chapter 3, this is not necessarily the case.

from the dialect of the Merina ethnic group, who inhabit the central plateau of the island.

Merina came to be the basis for OM for three primary reasons, according to Deschamps (1936). Firstly, due to favorable relations with the French in the pre-colonial era, they received assistance in their conquest and occupation of the island. Additionally, and certainly following from this first fact, Merina was the language on which the current, Latin-alphabet-based orthography⁷ was based. Finally, it was the mistake of several French colonial administrators to assume that there was one Malagasy language, synonymous with Merina, and thus this was the one used by the elite class in the capital. Even today, due to OM being the language of education, young people enjoy using words from this dialect to appear sophisticated (p 30).

Despite the dominance of OM, missionaries and, later, linguists have not ignored other language varieties Aside from those who looked only at Betsimisaraka, to be covered in the next subsection, dictionaries and/or grammars have been published on Takarana (Baron 1893), Sakalava (Dalmond 1842, Thomas-Fattier 1982), Taisaka (Deschamps 1936), and a group of several northern dialects (Velonandro 1983), along with many others unable to be consulted for this work. Additionally, authors Ferrand (1909), Dahl (1951), Dez

⁷ An earlier, Arabic-script-based orthography called *sorabe* was used on the Southeast coast in an earlier era, but was limited to a select class of astrologers and fell into disuse before the advent of the modern script.

(1963), and Adelaar (2013) cover dialectology as a whole, the first two in service of a wider comparison to other Austronesian languages.

Even following all of this previous work, some Malagasy—including even some linguists involved in dialectology⁸—generally believe that Official Malagasy is quite similar to the other regional varieties. In fact, the latter may actually be much more different from the former and from one another, according to Bouwer (2007). She suggests that inter-ethnic-group communication is conducted employing OM as a lingua franca. Her experiments suggest that dialects are not even always mutually intelligible—one key factor in distinguishing dialects from languages. Even if they are mutually comprehensible with better described varieties, though, this does not mean that they do not have distinct phonologies, so describing them would still be worthwhile. Even so, because the forms of the language associated with the traditional ethnic groups are still conventionally viewed as dialects, this work uses this term, or else some neutral term like “language variety,” when referring to them.

Additionally, a complete dialectology of the island could offer important evidence as to how Madagascar was settled (Adelaar 2013). With a full comparison of Malagasy language varieties, one might determine if the first Austronesians arrived straight from Indonesia,

⁸ As represented at the concluding conference for The Malagasy-Canadian Project, a coordinated effort to record variation in Malagasy. The author was present at this event, which took place in Diego-Suarez, Madagascar in June 2014.

fvia Africa, or via the Comoros Islands to the North. Furthermore, one might be better able to tell if Bantu contact happened in Africa, in the Comoros, or in Madagascar.

1.2.3 The Need for a Study of Betsimisaraka

As mentioned previously, Betsimisaraka is the second-largest ethnic group, with a million members, but one of the least described language varieties. Adelaar (2013) in particular seemed vehement about the need for description of Betsimisaraka, saying “there are various subdialects, and the dictionaries, which are several centuries old, are phonologically imprecise and tend to represent unsystematic and rudimentary information” (p 458). He also notes that the dialect is not uniform, being spoken over great distance along the East Coast, and again calls it “poorly understood” (p 468).

Dalmond's (1842) work apparently consists of a dictionary, somewhat mixing Betsimisaraka with Sakalava, and grammar, though only the dictionary was obtainable for this study. Dez (1960, 1992) has two, unfortunately unpublished works on the Southern Variety. Dez (1963) himself even specifically says that the sounds used in the language are very little explored (p 507). In the current era, Ritsuko Kikusawa did fieldwork in Vavatenina, the same town where research for this study was conducted, producing a “A Malagasy Text with Grammatical Notes” (2006), which includes six pages on the

phonological inventory and alternations, and a conference presentation whose relevant material was limited to Northern Betsimisaraka's syntax (2007).

Some older, difficult-to-obtain dictionaries, the occasional reference in larger dialectologies, and one scholar currently focusing on its syntax: this is the sum of linguistic approaches to Betsimisaraka. Inspired by this underdescription, two leading dialectologists of Malagasy, Dez and Adelaar, both urge further description of the language varieties used by the Betsimisaraka, with the former specifically mentioning the sounds.

1.2.4 Concluding Remarks on Motivation

The inspiration for this study should now be clear. Theoretical approaches concentrate on OM syntax, and when compiling complete grammars or studies in related fields of phonetics and morphology of OM, underlying forms and features are not typically employed, especially not together. Two previous works, a sketch by Dziwirek (1989) and a dissertation by Albro (2005) have used both of these powerful theoretical tools, but analyzed only OM. Work on any aspect of Betsimisaraka has been practically null, and where existent, has focused on the syntax with only scattered, non-comprehensive references to its sounds. This work, which offers a thorough phonological analysis of the language, fills an important descriptive

and theoretical gap.

Furthermore the findings presented here could promote the language identity of the Betsimisaraka and other groups. As mentioned, the prestigious and pervasive nature of OM could erode the identity of an ethnic group (especially one smaller than the Betsimisaraka). This work does not seek to demonstrate that Betsimisaraka is a distinct language or even dialect: it merely describes the distinctions, allowing the Malagasy and wider linguistics community to use it for that or any other purpose.

1.2 Overview of the phonological grammar

Chapter 2 covers the methods used to compile this phonological grammar, including both a thorough recounting of the study's data-collection and transcription methodology as well as a justification for the analytical methodology of the study, specifically the use of ordered rules (ultimately following Chomsky and Halle 1968) as opposed to a constraint-based, Optimality-Theoretic approach (proposed in Prince and Smolensky 1993).

Chapter 3 summarizes and distills the literature that was only reviewed in 1.1, resulting in a phonological description of OM, and also covers previous works that have focused on other dialects of Malagasy, including the few references to Betsimisaraka. By covering previous

literature on related dialects before delving into underdescribed Betsimisaraka, authors specializing in the various forms of Malagasy can inform this phonological grammar of Betsimisaraka. For this reason, and for a comparison that highlights the ways in which Betsimisaraka is distinct, references to OM and other dialects will continue throughout this work. Except when designated as such, though, all examples in the following chapters is Betsimisaraka data collected specifically for this work.

Chapter 4 presents the phonological inventory of Betsimisaraka, including articulatory features for all segments. The inventory includes two Betsimisaraka phonemes, the mid-height back vowel /o/ and the velar nasal /ŋ/, that are absent-to-marginal phonemes in OM:. This chapter additionally mentions certain allophones, such as the lax mid-height vowels, which also distinguish Betsimisaraka from OM.

Chapter 5 presents an overview of stress in the language, which does not necessarily differ from OM: it is assigned to the penultimate syllable, but stress can be antepenultimate (due to vowel epenthesis) or final (due to hiatus resolution). This chapter, then, sheds new light on Malagasy dialectology through an analysis that unequivocally calls upon penultimate stress to explain all the disparate stress patterns.

Analysis of segmental alternations comes in Chapter 6, which will re-examine the alternations covered in Chapter 3 using Betsimisaraka lexical items; each alternation motivates a rule or rules, and sample

derivations are offered. Chapter 6 presents many processes that do not occur in OM, including place dissimilation in nasals, optional deletion of underlyingly final nasals, and three distinct vowel harmony processes: two limited ones operating on different first-person singular suffixes, and one more robust, though optional, harmony affecting all epenthetic vowels. Just as crucially, certain alternations seen in OM, such as labialization and palatalization affecting following back consonants, are seen to occur in a different (though related) environment in Betsimisaraka.

Finally, Chapter 7 will conclude the phonological grammar by tying together the distinct factors of the language mentioned in Chapters 3 and 5, presenting Betsimisaraka as a distinct variety both from the previously studied Official Malagasy as well as other dialects, along with presenting select shortcomings and possible areas of future study in Betsimisaraka phonology and Malagasy dialectology in general.

Chapter 2

METHODOLOGY

Introduction and Overview

Data collection for this study did not greatly deviate from the standard methodology⁹ for describing an understudied form of a language: a thorough review of research on better described varieties, on-site language lessons, elicitation of the Swadesh list, then follow-up elicitation of word-forms based on phenomena occurring in the related dialects and those evident from the Swadesh list, language lessons, and other recorded speech. Section 2.1, then, proceeds roughly chronologically through the steps taken and concludes focusing on practical recommendations for future projects. Section 2.2 justifies the less standard analytical methodology used in this work, an analysis employing an ordered rules rather than an Optimality-Theoretic approach.

2.1 Data Collection Methodology

To achieve a phonological grammar, this project started by gathering advice and hiring a staff (2.1.1), then began with initial

⁹ Three works read before, and consistently consulted during, this project were Bowern (2008), Chelliah and de Reuse (2010), and Newman and Ratliff (2001).

language lessons (2.1.2). Once those had served their purpose, an appropriate field site was chosen (2.1.3) and subjects there were selected and introduced to the project (2.1.4). Finally, those participants were asked questions carefully selected to yield relevant responses (2.1.5), and those responses, along with introductory material, were transcribed (2.1.6). The positives and negatives from this procedure are summarized in 2.1.7.

2.1.1 Advisors and Assistants

Establishing this project was greatly facilitated by colleagues who had already been conducting (unrelated) research in Madagascar. These colleagues were able to advise on diverse subjects such as navigating the local bureaucracy (visas, permits, etc.), securing lodging, locating research assistants, determining subject payments, and picking a field site.

The formal staff hired for the project consisted, for most of the time in Madagascar, of one language assistant, Ludger M., who provided language lessons and general guide/interpreter services, and conducted and later supervised elicitation and transcription. Himself a member of the Betsimisaraka ethnic group—hailing from the northern city of Maroantsetra ([marwan'tsetʃə])—he was able to instruct the author in the relevant dialect and communicate with subjects in their native language. Additionally, he could participate in one-off

elicitations himself, which could at least be used to inform future questionnaires, and his transcriptions did not seem to alter the dialectal forms used by subjects. It cannot be overlooked that he also worked as a stand-up and sketch comedian in the region, indicating a passion for engaging people that made subjects easy to find.

In the last portion of the project, a second language assistant, Percila V., was hired, this one a native of Vavatenina and thus a native speaker of the Betsimisaraka described here. Firstly, this meant that in conversation with subjects, she naturally exhibited all the features of the language variety, inspiring them to employ their everyday speech. More importantly, though, she was able to provide data on alternations particular to Vavatenina during transcription, allowing fine-tuned scripts for further questionnaires in cases when Vavatenina and Maroantsetra Betsimisaraka differed.

In this section, “the investigators” denotes the author and the assistants.

2.1.2 Pre-Field-Trip Language Lessons

The first four weeks in Madagascar consisted of language lessons, which had both practical and direct applications to the project itself. This is to say, they were an almost necessary reality of living surrounded by and interacting with speakers of the language, but they also directly benefited the study by providing a basis for initial

conversation with, and elicitation from, informants: while never fluent enough to conduct 100% of elicitation, the author could follow the conversation and redirect it if necessary.

Crucially, as well, language lessons are a kind of undirected elicitation, so all of the words and phrases learned were potential data for the project, albeit from only one speaker and in a context not likely to inspire non-citation forms. As will be seen in 2.1.4, alternations observed in the language lessons were one important tool in constructing questionnaires for use at the field site.

2.1.3 Selection of the Field Site

Vavatenina, the eventual field-site chosen for this study, was ideal for the first description of Betsimisaraka phonology. This town is in the center of the Betsimisaraka region, and has a much more homogenous subject population than earlier sites considered. This is to say that a great majority of residents were born and raised in the area, thus all had very closely related phonologies. This section will refer to other cities in the Betsimisaraka region, so a map of that region is provided in Figure 2.1 below, providing much more detail than the previous map in Figure 1.1



Figure 2.1: Map of locations relevant to the project, with rivers and roads indicated. This map (with North to the right) shows most of the East Coast of the island of Madagascar. Relevant cities are marked: **Antananarivo**, the capital, where the author secured his research permit. Not a Betsimisaraka city. **Mahanoro**, near the southern edge of the Betsimisaraka region, where the author attended a Betsimisaraka cultural festival. **Tamatave**, site of the initial language lessons and most transcriptions for the project. (**Ivoloina**, the initial field site, is within the area marked for Tamatave.) **Vavatenina**, the field-site for the project and thus the town whose dialect is the focused subject of this grammar. **Maroantsetra**, where the primary language assistant was born and raised, before relocating to Tamatave. Generally speaking, the Betsimisaraka region spans from Mahanoro to Maroantsetra, reaching about twice as far inland as Vavatenina.

The initial field sites attempted for this project were Tamatave ([,tama'tav]¹⁰), the large city where the project began and where the author lived, and Ivoloina ([i,vo'lwina]), a village just to the north. Unfortunately, monolingual speakers of Betsimisaraka are difficult to

¹⁰This is the French name for the city, which partially accounts for the unusual final stress, as will be seen in 3.5 and 4.2. Its official Malagasy name, seen on modern maps, is “Toamasina” ([tuə'masin(ə)] or [tuə'masi(nj)])¹¹, though is seldom referred to as such by residents, who preferred the French form, sometimes written “Tamatavy,” following Malagasy orthography.

locate even in an outer neighborhood of cosmopolitan Tamatave, and an attempt to find a more reliably monolingual population in the smaller village also failed. In both locations, it was very unlikely that a given potential subject had grown up there.

For this reason, a new field site in Vavatenina ([,vava'tenina])¹¹ was selected, a town with an estimated population of 30,000 in 2001, a decade before the elicitations for this study. From there, the only vehicle access to the rest of the island is a single 22-mile, winding, pot-holed road to the coast. This may be one reason why the dozens of subjects selected for elicitation there were all born and raised either in Vavatenina or the surrounding villages. Once this field site was chosen, all elicitation was conducted there, though recall that this work will use “Betsimisaraka” to denote the variety of Malagasy that it describes.

It should be noted that, despite the move to Vavatenina for more access to monolingual speakers, the investigators did not actively seek them out. Instead, the language assistant was briefed on which characteristics to look for, but as long as a potential participant self-

¹¹ Although this grammar employs “Vavatenina,” the best known name for the town, the pronunciation used by subjects who were comfortably referring to their hometown would be spelled “Vavategniny” ([,vava'tenini]), using the <gn> digraph for [ŋ]. The former spelling is employed in all official documents and was often used by residents, but the local pronunciation actually happens to include several indicators that distinguish Betsimisaraka from OM. Comparing “Vavatenina” to “Vavategniny,” the latter has a velar nasal (analyzed in 4.2.1), a harmonic vowel in an antepenultimately stressed word (described in 6.2.3), and a non-velar nasal as the onset of a final syllable of an antepenultimately stressed word (examined in 6.1.8). Furthermore, an alternate pronunciation of the town, simply “Vavategny” ([,vava'tenji]), evinces a type of variant not found in OM (presented in 6.1.6).

identified as a member of the Betsimisaraka ethnic group, he or she was accepted as a participant. Because educated subjects were not filtered out, a great number—though fortunately not all—participants had had schooling in and were conversant in OM. Still, all seventy participants were told that the point of the study was to record the Betsimisaraka language as it is used every day in unofficial contexts, and that despite the presence of a Westerner asking questions, taking notes, and proffering a microphone, this was not a situation that demanded official speech.

2.1.4 Elicitation Preliminaries

Elicitation was generally conducted wherever subjects were discovered, usually in their yards or inside their homes, but occasionally on grass mats or stools by the side of the road. When subjects agreed to participate, they were shown how to put on the noise-canceling microphone headset, which was connected to the author's computer. Recordings were originally made using Praat, and later Audacity. All questions and prompts were on a pre-printed script.

Permissions were recorded, rather than using a signed form, to include participants at all levels of literacy. After experimenting with other systems, the investigators settled upon participants repeating after the language assistant, phrases such as “My name is ___ and I agree to be recorded for this study,” and “I understand these

recordings could appear in published works and on the Internet, to be seen and/or heard by anyone in the world.”

The procedure of acquiring demographic information was also an evolving process. The investigators began by asking six questions all about the subjects, their language, and how they used language with others, but, as these proved time-consuming to transcribe and not entirely relevant to the goal of a phonological grammar, they were eventually reduced to just two: “where did you grow up?” and “in what language were you educated?”

2.1.5 Questionnaires and Elicitation

Three main sources were used in the construction of questionnaires. Firstly, before arriving in Madagascar, the author was aware of many alternations already reported by other authors for Official Malagasy; some of the first stimuli were to investigate whether these alternations also existed in Betsimisaraka. Secondly, some of the phenomena that distinguish Betsimisaraka from OM became apparent during the initial month of language lessons, providing potential forms for elicitation. And finally, beginning with the initial field trip to construct a Swadesh List for Betsimisaraka¹², each trip revealed alternations to investigate on the next. The remainder of this section lays out how these scripts were written and which techniques were

¹²Included here as Appendix 1

employed to elicit given forms.

Questionnaires were typically constructed by the investigator himself, but were always checked with one of the assistants to be certain that they would work and that the instructions were clear. The Swadesh list could be obtained through charades, pictures, indicating things participants could see, and non-leading questions from the language assistant. Throughout the project, these methods were used whenever a single word in isolation was required. With the small exception noted below, every word participants heard from the investigators was in Betsimisaraka, with no English, French, or OM employed.

Alternations, though, were primarily explored using compound words, or else comparing different forms of verbs. In the case of a compound, both words would be elicited individually, as described above, and then subjects would be asked to combine them. In cases where an onset consonant of the second root was alternating, then, the investigators would have a recording of it in both environments. Verb forms ideally were elicited naturally: to hear the active form, investigators would charade or describe the verb, and then to hear the imperative, subjects were asked how they might tell someone else to do that. Admittedly, the language assistant would occasionally fall back on using the words “*actif*,” “*imperatif*” or “*passif*,” the French grammatical terms that educated Malagasy would recognize.

2.1.6 Transcription

The investigators scrupulously transcribed every utterance, from permissions to demographics to the actual, sometimes meandering answers to the questionnaires. The author streamlined the process by conducting initial transcriptions unsupervised, a step that became easier as the project went on. Later, the language assistants would correct and add to these rough transcriptions, glosses, and translations. In this way, every recording was transcribed in the traditional manner, with a Betsimisaraka transcription, a word-for-word gloss, and (in cases where the subject was speaking in sentences and not simply saying words) an idiomatic English translation. Occasionally, a line was appended with phonetic notes as well. All transcriptions were saved as text files.

Although full transcriptions of every utterance could be useful for future linguistic investigations, they were not necessary for the goal of constructing a phonological grammar. The author recommends doing some full transcription, but in the case of questionnaires seeking specific forms, keeping mostly to a chart-based organization where relevant answers can be easily compared without searching through vast amounts of unrelated transcription. This would save time on-site, allowing more investigation in the field, and while writing, because less time would be spent searching for specific words.

2.1.7 Methodological Recommendations for Future Work

The relevant methodology of the project has now been laid out, representing a procedure that was successful in producing the data necessary for a phonological grammar. Aspects of this study that would be useful to future researchers are tapping the knowledge of interdisciplinary colleagues, being sure to have at least one language assistant who speaks the exact dialect focused on, spending the first part of the project on language lessons, choosing a field-site with as homogenous a subject population as possible, and doing a first, unsupervised draft of transcriptions. Importantly, future projects can also improve upon the methodology used here by only asking focused, relevant demographic information from the start, conducting an efficient and systematic transcription so as to return to elicitation before too much time has passed, and (as not previously mentioned) attempting to actually write a significant portion of the analysis while elicitation is still underway, so as to spot potential new avenues of questioning inspired simply by inability to explain an alternation in words.

2.2 Theoretical Methodology: The Case for Ordered Rules

As stated in Section 1.1, this work follows the phonological tradition by analyzing sound patterns as the result of underlying forms converted to surface forms along featural dimensions. In the context of

modern phonology, these alterations were once commonly viewed as due to a language's grammar containing a set of ordered rules, as in Chomsky and Halle's *The Sound Pattern of English* (1968). Currently though, analysis is almost entirely done in the tradition of Optimality Theory, first proposed in Prince and Smolensky (1993), and OM has already been described in this framework, by Albro (2005).

Given the dominance of Optimality Theory, then, a justification is here provided for this work's choice to employ ordered rules in place of ranked, violable constraints. All justification for this stems from the central goal of this project, which is to inform the linguistic field and the public at large about the phonology of Betsimisaraka. To that end, it is argued that rules are more accessible, less subject to change as a result of theoretical debate, and therefore more enduring.

Before motivating this choice, though, it must be stressed that none of these points seek to demonstrate the superiority of rules at explaining phonology in general, just their appropriateness in the context of a phonological grammar. For example, one strength of OT is that a single constraint can explain trends in related and unrelated languages that would require many disparate, even dissimilar rules. The selection of an ordered-rule approach also does not preclude cases where OT provides a more accessible, less controversial, and more enduring explanation. See, for example, rule (19) in Section 6.2.4, where two simple constraints easily demonstrate why two vowels

cannot stand in sequence, whereas stating the alternation as due to a single rule results in a messy nomenclature that must mention the features of the remaining vowel twice.

Certainly, reasons like these have resulted in the popularity of OT, but its use as an analytical tool by the vast majority of phonologists does not mean they are not all familiar with rules as well. Many are first introduced to phonology using *SPE*-style rules, and textbooks on phonology reflect this (Hayes 2009, Odden 2013), so nearly every phonologist is familiar with this approach. Therefore, using rules maximizes the potential audience.

Furthermore, it is not the goal of this dissertation to promote one theory or another as to how to overcome opacity. As will be seen, especially in 6.3, opaque phenomena are prevalent in Betsimisaraka. Optimality Theory, as originally proposed, cannot explain opaque alternations. Certainly this gap has been filled many times over, for instance with Bacovic and Wilson (2007)'s targeted constraints approach, or Finley (2008)'s modifications to Turbidity Theory. One of these—or another of the many solutions proposed¹³—might have been employed here to explain the opacity involved in the language. Indeed, in Albro's (2005) analysis of OM, he employed a modified form of correspondence theory (based on Steriade 1998 and Beckman 1998), as a means to arguing for his computational approach. However, as

¹³ Summaries of the major approaches are given in Kager (1999, Chapter 9) and McCarthy (2007, Chapter 2).

stated above, the only objective of *this* grammar is a thorough analysis of the language, with no further theoretical agenda.

Following from the previous point, *SPE*-style rules have persisted for the almost half-century between that work's publication and the writing of the current text. OT, while no theoretical newcomer, has been continually updated in the quarter-century since its proposal with no consensus as to the strongest method of analyzing opacity. If one method were selected here, that method might fall out of favor and render moot the paramount goal of this work: to allow access to knowledge of a previously undescribed phonology.

Finally, the analysis here is of course not entirely incompatible with OT. All of the URs and features proposed could be used to analyze the alternation data offered here in any framework: a fellow phonologist wishing to review an old—or propose a new—method of analyzing opacity in OT could plug in the forms here as evidence. In other words, the empirical generalizations proposed here are largely theory-independent.

The data are presented clearly, with alternation data always presented first, followed by empirical generalizations and then an analysis using the most enduring, widely understood, and accessible method available.

Chapter 3

OVERVIEW OF MALAGASY PHONOLOGY

Introduction and Overview

The following description presents a summary of what has already been written about the phonology of Official Malagasy¹⁴, a dialect closely related to the Betsimisaraka dialect that is the focus of this grammar. Furthermore, although other dialects are much less well described, limited descriptions do exist, especially in the form of dictionaries. Therefore, this chapter also includes, wherever possible, alternations in non-Official forms of Malagasy, either indicated in the literature, or from dictionary searches initiated for this project. These forms include Betsimisaraka, but as noted in 1.1.3 this is one of the least-described language varieties, and this overview of related dialects constitutes the best literature review for the phenomena that were encountered in the Betsimisaraka of Vavatenina¹⁵.

Before the description of Malagasy phonology, this chapter offers a section (3.1) on the genealogy of Malagasy, both in the larger

¹⁴ Typically, these works do not call the language they describe “Official Malagasy,” either calling it simply “Malagasy,” for reasons that will become clear in 3.1, or specifying it as “Merina Malagasy,” the ethnic/regional dialect that formed the basis for, and is in some ways still synonymous with, Official Malagasy

¹⁵ Placing this chapter previous to the chapters on the Betsimisaraka is in no way to be taken as an endorsement of assumptions that unofficial varieties of Malagasy are derived from, or in any way later versions of, Official Malagasy.

Austronesian language group and in the relationships between language varieties in Madagascar. Describing the sound system begins with the phonemic consonants and vowels of the language (3.2), including discussion of complex segments such as affricates, pre-nasalized stops, and diphthongs, the nature of all of which is still in dispute. Section 3.3 introduces the morphological processes that will be used to demonstrate alternations in the sections to come. Section 3.4 analyzes the language's syllable structure. Segmental alternations are discussed in 3.5, followed by an examination of stress in 3.6, which also includes the proposal of underlying codas and the segmental alternations entailed thereby. A brief conclusion (Section 3.7) will highlight the major questions left unanswered by previous studies on official Malagasy.

3.1 Genealogy of Malagasy

3.1.1 Sources of Malagasy

The closest linguistic relatives of Malagasy are the other languages of the Southeast Barito (SEB) subgroup, part of the larger Malayo-Polynesian branch of the Austronesian language family. These include the languages of the Central Kalimantan provinces on the island of Borneo, thus whence the original Austronesian settlers are thought to have left.

Due to extended contact with the Bantu people of nearby

Southeast Africa during the settlement of the country, Malagasy has developed nominal and tense-marking prefixes, neither of which is a feature of its closely related languages in Indonesia (Adelaar 2013). It may also have obtained its stricter syllable structure (see Section 3.4) from Bantu, but it did not acquire tone, despite that “many if not most” East African Bantu languages exhibit this feature (Randriamasimanana 1999). Many Malagasy words, including most domestic animal names, have their origin in Bantu vocabulary, as well as in other languages Austronesians would have encountered on their journey from Borneo to Madagascar.

Table 3.1: Malagasy words with non-Austronesian origins, from Dahl (1951).

Malagasy		Non-Austronesian	
gloss	word	word	language
“cattle”	umbi*	ngombe	Bantu†
“dog”	ambua	mboa	Swahili
	alika	alaka	Sanskrit
“arm”	sandrit‡	sandhi¤	

* also [aumbi], as turned out to be the Betsimisaraka form.

† Dahl does not provide a specific Bantu language here.

‡ In Merina; also “trunk of the body” in Tanosy.

¤ This word, familiar to phonologists, means “connection” in the source language.

3.1.2 Relatedness of Malagasy Dialects

In motivating this project, Section 1.1.2 already presented Bouwer's (2007) analysis of the various language varieties of

Madagascar. In her view, at least some (she compared those from the Southwest to Official Malagasy) may qualify as separate languages, at least by lack of mutual intelligibility. As mentioned there, this is not necessarily the standard view.

The governmental regulatory body of the Malagasy language, the *Académie Malgache*, officially declared in 1973 that there was one Malagasy language, an official variety (OM) with various regional forms (Dez¹⁶ 1991 p 53). This despite that OM was, before this declaration and the social forces that motivated it, just a regional dialect itself (Dez 1963), albeit one native to an ethnic group that was, at least during the era of colonization, the most prestigious and powerful: the Merina¹⁷. This one-language-with-regional-dialects view of the linguistic situation in Madagascar leads to the opposite of Bouwer's view: that "the Malagasy language is unified¹⁸ across all the island of Madagascar, but it presents itself under the form of numerous local ways of speaking, all closely related between them" (Dez 1991 p 5). To be fair to authors who put forth this view, which also include Dahl (1955), it may be the case that, in the mid-twentieth century, when the Malagasy population was just five million and the national

16 Jacques Dez was, for many years, himself a member of the *Académie Malgache*.

17 It is an anachronism in this work, then, to refer to "OM" when discussing works by authors before 1973. This is done, though, for consistency and for brevity, and it can be inferred that those works referred to Merina, formerly also called "Hova."

18 This quotation was translated from French for this work; in this case, Dez used the word *unique*, which can mean "exceptional" as it does in English, but here likely carries its other meaning of "one, single."

infrastructure was in much better condition, the regional forms of Malagasy actually were more intercomprehensible than they are today (Bouwer 2005). That said, one need not go back so far to find this view in print: twenty years ago Rajaonarimana (1995 p 8), introduced his *Modern Grammar of the Malagasy Language* saying, in part, “Malagasy is characterized by a very weak dialectization (hardly accentuated dialectal differences) that never impede intercomprehension.”

Whether or not Dez (1963) appreciated the extent of the diversity of Malagasy, he was the first (Adelaar 2013) to attempt to describe its diversity. Looking at phonology, lexicon, and morphology of the described dialects, he concluded that there is a major division between those spoken in the Southwest and West, and those spoken in the Center and the East (Dez 1963), as delineated on the map in Figure 3.1 below. Some phonological data of the type he analyzed will be included in Section 3.2.3.

While most agree with the major groupings Dez made, counting the Southwest, West, Center, and East, along with the North, it is still controversial as to how the family tree should be drawn—for example, some scholars have placed the Southwest on its own branch in opposition to the other four (Adelaar 2013). One complication in making these divisions has been determining whether a form is borrowed or inherited from another dialect, and the influence of

prestigious dialects¹⁹. Indeed, the need for a better understanding of the relatedness of Malagasy languages/dialects was one motivation for this project and those like it, as stated in Section 1.1.2.



Figure 3.1 Dialect map of Madagascar, including the five major divisions originally proposed by Dez (1963). The map above is directly based off the one presented in Adelaar (2013), drawn by cartographer Chandra Jayasuriya, although does not include all dialects included there. Particularly in the Southeast, near Tanosy and Taisaka, there are a few more dialects.

Further complicating the construction of a family tree for Malagasy is that language varieties are often assumed to relate to the eighteen traditional ethnic groups of Madagascar. In reality, though, a member of one ethnic group may actually speak a dialect more similar to that of a nearby member of another group than to that of a more distant member of the same group (Adelaar 2013). This is especially

¹⁹Certainly OM is an influence throughout the country, but regionally, the dialects spoken by the larger ethnic groups like the Sakalava who inhabit most of the West Coast, can also influence the speech of members of less populous ethnic groups.

relevant to describing Betsimisaraka, the second-most-populous ethnic group, with a long north-south territory bordering that of many other groups, as shown in Figure 3.1. For this reason, it is common to divide Betsimisaraka into two varieties, Northern and Southern, and depending on the dialectologist, either or both varieties could be assigned to the Northern or Eastern subgroup. Finally, depending where the dividing line is drawn, Vavatenina (close to the center of the Betsimisaraka region as shown previously in Figure 1.1) could ascribe to either dialect. Findings from this project—as well as those more closely describing Betsimisaraka of the Southern and Northern extremes of that ethnic group's territory—seem to indicate that it should be grouped with Northern Betsimisaraka.

From those who have charted linguistic diversity in Madagascar, then, we know that Northern Betsimisaraka—depending on classification—belong to a larger group of Northern or Eastern Malagasy language varieties. Although scholars have not reached consensus on a family tree for language(s) in Malagasy, they are in agreement that, for a language on the East Coast, the most different dialect will likely be one used in the Southwest where it will be more similar to one from the Central grouping, such as Official Malagasy. It is a happy coincidence in preparing to describe Betsimisaraka in subsequent chapters, then, that OM is the best described.

3.2. Segments of Malagasy

3.2.1 Consonants

The table below indicates the best compromise of all inventories presented in grammars of OM. For the most part, grammars cited here agree on the transcriptions given below for OM, except for small disagreements on the place features of /t/ and /d/, the manner of /r/, and the entire description of the sounds here called /tr/ and /dr/. This section reports those as well as dialectal variation.

Table 3.2: A segmental inventory of the consonants of Official Malagasy.

	Bilabial	Labiodent.	Dent.	Alv.	Anterior	Velar	Glot.
Stops	p b		t d			k g	
Affricates				ts dz	tr dr		
Fricatives		f v		s z			h
Nasals	m			n			
Lateral				l			
Trill				r			

/h/ is not always heard in Malagasy²⁰. Rajaonarimanana (1995) called it “an etymological consonant” and says it “is almost never pronounced in the current, spoken language.” Even if it is elided in most contexts, /h/ is a phoneme of OM, and is most frequently pronounced intervocally (Dziwirek 1989). Crucially for its phonemic status, /h/ can be seen to neutralize to [k] under post-consonantal fortition (see 3.5.1) and underlyingly word-finally (see 3.6.3). Vowel-

²⁰ Interestingly, according to Ferrand (1909 p 83) the /h/ of Malay typically does not appear in Malagasy at any level of representation, whereas Malagasy /h/ is typically a reflex of Malay /k/. This is a case of diachronic lenition, though Malay /k/ also sometimes manifests as /k/ or /g/ in Malagasy, as does Malay /g/.

initial and vowel-final roots do not act this way.

Rajaonarimanana (1995) describes the sounds transcribed in Table 3.2 as /s/ and /z/ as having pronunciations situated between alveolar /s/ and /z/ and post-alveolar /ʃ/ and /ʒ/. They are listed only as alveolar above based on Dez (1963), who describes the OM sounds as alveolar²¹ and the (other) regional dialect equivalents as post-alveolar. The same is confirmed by Deschamps (1936) in the Taisaka dialect²², at least for the voiceless segment. In Tanosy, they are alveolar except before /i/, where they are post-alveolar²³. In OM, the stops /t, d, n/ are pronounced dentally before /i, e, a/ but as alveolar before /o, u/ (Thomas-Fattier 1982), so are listed under both places of articulation.

The trill /r/ occurs as such when spoken by those of the Merina ethnic group, but may occur as the tap [ɾ] in fast speech (Albro 2005). According to Thomas-Fattier (1982), its distribution is contextual: [r] word-initially and [ɾ] elsewhere.

The sounds written in the standard writing system as <ts> and <j> are recognized unanimously as the affricates /ts, dz/ in OM—transcribed in some works cited here as /t̪s, d̪z/ or as /t̪s, d̪z/. Comparing these affricates to the anterior sounds /tr̪, dr̪/, Rajaonarimanana says the

21 In one report that runs contrary to all others, Deschamps seems to say that the Merina pronunciation of <s> is a palatal or even velar fricative. More specifically, he says that in Taisaka (the dialect he is focusing on) <s> is pronounced like French “ch,” i.e. [ʃ], whereas in Merina it is like the German “ch,” i.e. [ç] or [x].

22 According to the division in Figure 3.1, an Eastern dialect like Betsimisaraka.

23 Because Antanosy is where the French first encountered Malagasy, and due to vowel devoicing, the French word for “Malagasy” is “Malgache” [mal'gaʃ].

former employ the tip of the tongue while the latter employ the blade. Not all dialects have these affricates, employing instead /s/ or /t/; Deschamps (1936) uses the example of cognates to OM [a'tsimu] (“South”): Taisaka [a'simu] and Bara [a'timu].

The identity of /tr/ and /dr/ is still controversial, thus their slightly ambiguous listing above as “anterior.” These segments, written <tr> and <dr>, are sometimes analyzed as affricates of their constituent segments (/tr/ and /dr/) (Keenan & Razafimamonjy 1996 *inter alia*) and occasionally as the affricates /tʃ/ and /dʒ/ (Keenan & Polinsky 1998). Albro (2004) splits the difference in transcribing <tr> as the post-alveolar affricate /tʃ/ but <dr> as the alveolar affricate /dʒ/. The Francophone author Rajaonarimanana (1995) says these are pronounced close to English “tree” and “drive,” i.e. as [tʃu] and [dʒu], but transcribes them (like Erwin 1996) simply as [tr] and [dr].

One reason for these differences may be speakers applying the phonemes of their dialects to the lexicon of Official Malagasy. For example, the descriptions Keenan & Polinsky (1998) used, /tʃ/ and /dʒ/, are also used for Thomas-Fattier (1982) in his description of Taisaka; as will be seen in Chapter 4, this is also how they are pronounced in Betsimisaraka. Furthermore, both Dez (1963 p 515) and Thomas-Fattier (1982) say that the alveolar and anterior affricates are in the process of merging in some dialects. According to the latter, three dialects from

the South (Betsileo, Tandroy, Tanosy) have [ts] for both <tr> and <ts>²⁴. He proposes that dialects like Taisaka and Betsimisaraka are a transition between the OM, rhotic-like pronunciation and the one that has merged with [ts].

Some other sounds occur only as allophonic variants, such as the palatalized dorsals [k^j] and [g^j], while [h^j] surfaces as [j] due to OM tendency not to pronounce /h/. Examples of each of these sounds are seen in the table below.

Table 3.3: Examples of palatalization, from Rajaonarimanana (1995)

orthography	pronunciation	gloss
<isika>	[i'sik ^j a]	1P (inclusive)
<migoka>	[mi'g ^j uka]	“drink” (pres. act.)
<mihafy>	[mi'jafi]	“endure” (pres. act.)

Palatalization occurs following the high front vowel /i/, as in monomorphemic /alika/[alik^ja] (“dog”) or bimorphemic /vadi-ku/[vadik^ju] (“my spouse”). The process does not demand that the triggering /i/ be pronounced as a full vowel; as long as it is there underlyingly, it will trigger palatalization, as in /aingana/[a^jng^jana].

Note that in the last example, the /g/ palatalizes even after a nasal; this is used as evidence by those analyzing NC sequences as one pre-nasalized stop, transcribing the example above instead as

²⁴ He does not specifically say if he believes this includes the voiced affricates, although Dez (1963 p 513) says some Eastern dialects have /z/ where OM has /dz/.

[a^ŋgjana]. Pre-nasalized stops, listed below, are not generally proposed as phonemic consonants, operating instead as allophones of the non-pre-nasalized stops when those stops follow an underlying nasal. Further discussion of these segments is suspended until the section on syllable structure, which provides a main argument for their existence.

Table 3.4: NC sequences of Malagasy, re-analyzed as pre-nasalized stops

Bilabial	Dental	Alveolar	Anterior	Velar
^m p ^m b	ⁿ t ⁿ d	ⁿ ts ⁿ dz	ⁿ tr̪ ⁿ dr̪	ⁿ k ⁿ g

For fully nasal stops, Table 3.2 lists only /n/ and /m/, although OM certainly has [ŋ] on the surface as an allophone of /n/ or /m/ by nasal assimilation (see Section 3.5.1), as in [maŋka] (“mango”) and [maŋkahala] (“hate” V) (Dziwirek 1989). OM, though, is actually exceptional among Malagasy dialects for not having a phonemic velar nasal²⁵, which typically (as in the palatalization of velars seen above) surfaces as [ɲ] following /i/ and /e/ (Dez 1963 p 511). It is even observable in writing, where it is variously transcribed as <ñ>, <ň>, or <gn>, the last of which is employed in this project. Minimal pairs from Southern Betsimisaraka below justify the phonemic status of the third nasal.

²⁵ Deschamps (1936) actually goes even further, saying that “the Merina spoken in the countryside”—so perhaps within the Merina territory, but not the capital—does feature phonemic /ŋ/, just as most other dialects do.

Table 3.5: Minimal pairs for regional Malagasy dialects, as compared to OM, from Dez (1963 p 512). All pairs occur at least in the Southern Betsimisaraka Dez was describing (thus the column labels), although some occur in other dialects as well. In all these cases, the lack of a velar nasal in OM results in homophony, with the same utterance corresponding to both meanings. (Glosses were translated from French for this work.)

SB	gloss	SB	gloss	OM homophone
'aŋana	"greens"	'anana	"possession"	'anana
'orana	"rain"	'oranja	"shrimp"	'urana
vi'nani	"conjecture"	vi'naŋji	"mouth" (river)	'vinani
ma'niri	"increase"	ma'ŋiri	"desire"	'maniri

The presence of a phonemic velar nasal is another connection to fellow Malayo-Polynesian languages. For the majority of Malagasy dialects that have it, it frequently appears in the same position in Indonesian cognates, as can be seen in Table 3.6 below.

Table 3.6: Evidence for the derivation of dialectal /ŋ/ in Malagasy as a reflex of the same sound in Indonesian. OM, without a velar nasal, has alveolar nasals in the same position in cognates. Data are from Dez (1963) and Deschamps (1936).

Indonesian	OM	Dialectal Malagasy		gloss
		word	dialect	
balanja	vi'lani	va'laŋji	Sakalava, Vezo	"cooking pot"
daŋav	'tranu	'traŋu	all	"house"
lenat'	'lena	'leŋa	Southern Betsimisaraka	"wet"
laŋit	'lanit̪ra	'laŋit̪ra	Taisaka	"sky"

OM does not have any phonemic glides, altho as will be seen in the following Sections 3.2.2 and 3.5.2, diphthongs present in the

surface inventory of the language do incorporate offglides and perhaps onglides in their articulation. Indeed, no dialect seems to have phonemic /w/, but at least some seem to have /j/, occurring in the same environment where OM cognates have either /dʒ/ (Dez 1963) or /z/²⁶ (Adelaar 2013). The table below shows examples of the latter.

Table 3.7: Cognates in Merina (OM) and other dialects, where the former has /z/ and the latter has /j/. Data are from Dez (1963 p 513), who does not say from which dialect he draws the evidence in the left columns, just “a certain number of dialects.”

Merina	dialectal	gloss	Merina	Vorimo	gloss
'aiza	'aja	“where”	'hazu	'haju	“wood”
'izi	'iji	3S	'zanaka	'janaka	“child”

3.2.2 Vowels

The four-vowel phonological inventory of OM is as shown below.

Table 3.8: Phonological vowels of OM. /u/ is typically high but may surface as [o] in certain environments. /a/ (perhaps technically /ə/) is typically cited as being low without further description.

	front	back
high	i	u
mid	e	
low	a	

/o/ is conspicuous in its absence, although this is somewhat counteracted by Dez's assertion that /u/ is pronounced closer to [o] in

²⁶ Adelaar cites South Betsimisaraka as one language where this occurs, but data found for this project align with the Merina columns in Table 3.7, not the dialectal. This could be one indication that Vavatenina speaks Northern Betsimisaraka.

the vicinity of non-high vowels and [u] elsewhere. OM may also have an extremely marginal /o/ phoneme, used for some recent²⁷ loanwords and the exclamation of surprise, pronounced simply [o]. Although apparently not as prolific as /ŋ/, it is not unusual for Malagasy dialects to have /o/ as its own, fifth vowel phoneme. Dez lists Northern Betsimisaraka, Tsimihety, Bezanozano, and perhaps Betsileo and Bara as dialects with /o/. As will be made explicit in Section 4.1.1, data for this study also revealed phonemic /o/, further suggesting that Vavatenina is in the Northern Betsimisaraka region. As one example, Dez says that OM ['vula] (“money”) is ['vola]²⁸ in Northern Betsimisaraka, a finding backed up by data found for this project.

To understand another case where [o] can be heard in OM, we must confront the controversial subject of the VVs and diphthongs of OM, on which no two grammars seem to agree. We will begin discussing diphthongs in general, and return to the subject of [o]. Since different sources use the term “rising diphthong” differently, this grammar always explicitly refers to either rising sonority (e.g. [jɑ] but not [aⱼ], as in Dziwirek 1989) or falling sonority, which will mean rising tongue height (e.g. [aⱼ] but not [jɑ], as in Albro 2005). Further complicating matters is that, while a diphthong is a single vowel (with

²⁷ Older loanwords would approximate other languages' /o/ with vowel clusters pronounced as diphthongs, e.g. /paumi/ for “apple,” from French “pomme.”

²⁸ Either because the designers of the alphabet had dialects like this in mind, or due to the aforementioned contextual lowering of the non-low back vowel, /u/ is actually represented by <o> in the official orthography. Betsimisaraka and other dialects that have a distinct /o/, then, are left with <ô> as its official representation.

two places of articulation), because most authors describe them with their orthographical two-vowel representation, it can be difficult to discern what pronunciation is proposed.

Erwin (1996) provides the smallest diphthong inventory, saying that “traditionally” there are only two, which he calls “ai” and “au.” Indeed, all grammars consulted for this project agree that these two sequences at least can be pronounced as diphthongs. Their surface pronunciation is with falling-sonority, transcribed in Dziwirek (1989 *inter alia*) as [a^j] and [a^w]. She employs glides to show which underlying vowel has decreased in sonority and superscripts because syllable structure that does not allow glide-consonant clusters.

Contrasting Erwin's minimalist set, Rajaonarimanana (1995) recognizes “seven ‘diphthongs’ in Malagasy,” listing <ai, au, ui, eu, ia, oa, io> although he is actually talking about “the pronunciation of the sounds of letters,”²⁹ so his list denotes the language's pronounceable digraphs. These seven, then, can be interpreted as permitted underlying hiatus, only for four of which does Rajaonarimanana provide a diphthong pronunciation. /ui/, he says, is pronounced [u^j], and he also lists /eu/ and /iu/ as diphthongs, never actually stating that they are pronounced [e^w] and [i^w]. Like Erwin, he gives [a^w]³⁰ as the surface realization of /au/ as, as in /lalau/ → [lala^w] (“game”). Note that even

29 This may explain why he places quotation marks around “diphthongs.”

30 Although he does not provide a transcription, he cites “cow” as a rhyme.

counting all seven, this is five short of the twelve possible VV combinations where the two vowels are different. These five missing digraphs, *<ea, ei, ae, ue, ie>*, all start or end with *<e>*, by far the least occurring vowel in OM. These sequences, then, could simply represent a lexical gap, although note that Albro (2005) actually proposes the constraint *|ei|. Furthermore, Dez (1963) says all underlying hiatus can surface as diphthongs, and specifically mentions both *<ea, ie>*, although this does not preclude Albro's analysis.

What is known about true diphthongs in OM, then, is that most that occur in the language are the falling-sonority/rising-tongue-height type, mostly the two low-high pairs that underlyingly are /ai, au/, as well as possibly high-high /iu, ui/. The existence of rising-sonority diphthongs [ja, wa] is controversial, although authors do cite forms that contain /ia, ua/. As described by Rajaonarimanana (1995), though, these VVs surface as a monophthongs, at least in normal conversation. The pronunciation of /ia/ on the surface is either [e], a coalescence of the two vowels into one with place features in between the two, or simply [i]. Rajaonarimanana explains this disparity by saying that /ia/ at the beginning of the word is pronounced [e], whereas if it has stress assigned, it is pronounced [i], as seen in the examples in Table 3.9.

Table 3.9 Divergent pronunciations of <ia> in OM, from Rajaonarimanana (1995 p 14-15). According to his analysis, these vowels are pronounced [e] at the beginning of words, and [i] if stress is assigned to them.

<ia> prounced [e]			<ia> prounced [i]		
orthography	SR	gloss	ortho.	SR	gloss
<i ^a nao>	e'na ^w	2S	<vilia>	vi'li	"plate"
<d ⁱ a>	'de	3S	<masiaka>	ma'sika	"cruel"
<m ⁱ anavaratra>	m ^e na'vara ^t ra	"go north"	<niavi ^a ny>	ne'vini	"origin"

Rajaronarimanana's proposal does not explain /dia/, where stress is assigned to the /i/, but this could be accounted for by positing that, in choosing a hiatus resolution strategy, stress assignment is not as powerful a trigger for /a/-deletion as the beginning of the word is a trigger for coalescence. Of the examples on the left, one has primary stress, one has secondary, and one does not have stress, so it follows that stress is not a factor for cases where /ia/ are the first vowels of the word. Because, as will be analyzed in Section 3.6, stress is assigned to the penultimate vowel, all the words on the right have stress assigned to /i/, not /a/, although Rajaonarimanana does not say whether this makes a difference.

Word position is similarly important in Rajaonarimanana's analysis of the usual pronunciation of underlying /ua/, which also undergoes coalescence or /a/-deletion, depending on word-position, as in the Table 3.10 below.

Table 3.10 Divergent pronunciations of <oa> in OM, from Rajaonarimanana (1995 p 14-15). According to his analysis, these vowels are pronounced [u] at the end of words, and [o] in other cases.

<oa> prounced [o]			<oa> prounced [u]		
orthography	SR	gloss	ortho.	SR	gloss
<toaka>	['toka]	"liquor"	<soa>	su	"good"
<ahoana>	[a'hona]	"how"	<tokoa>	tu'ku	"truly"

Although the distinction is not quite the same, it is similar: hiatus-resolution by low-vowel-elision occurs toward the end of the word, although in this case this is the more important factor, since the /ua/ in <soa> are both the first *and* the last vowels. It is possible with an analysis involving more examples, a more systematic analysis of the environment for coalescence vs. deletion could be possible.

Before entirely leaving discussion of Rajaonarimanana's examination of hiatus resolution in discourse forms, let us look at the most common diphthongs, pronunciations of /ai/ and /au/. Although these are usually the falling-sonority diphthongs [a^j] and [a^w], Rajaonarimanana says these can also undergo coalescence, pronounced as the mid-vowel compromise-place-feature vowels, [e]³¹ and [o], as in the examples below:

31 Kikusawa (2006), working on the same variety of Betsimisaraka covered in Chapters 3-5, identified the coalesced vowel as [ɛ], not [e], as did this author. She went further, though, in saying /ɛ/ was underlying, a reflex of OM /ai/, while the account in Sections 4.1.2 and 6.2.4 is synchronic, not diachronic.

Table 3.11 Conversational pronunciations of <ao> and <ai> in OM, from Rajaonarimanana (1995 p 14-15). Although the usual pronunciation of these is the diphthongs [a^w] and [a^j], in fast speech occasionally the monophthongs are heard.

<ao> pronounced [o]			<ai> pronounced [e]		
orthography	SR	gloss	ortho.	SR	gloss
<laoka>	['loka]	"meat" ³²	<manaikitra>	manekitra	"bite"
<misaotra>	[misotra]	"thank"	<akaiky>	a'keki	"near"

Coalescence may be a newer phenomenon. Dez (1963 p 509) looks at some nineteenth-century transcriptions of <toaka> and <loaka> and finds that the French speakers who transcribed them clearly heard two vowels. Either way, it exists today, including in non-OM varieties, though the result may be different depending on dialect. For example, OM <taulana> ("bone") is pronounced ['tolanya] in Southern Betsimisaraka (Dez 1963) and ['tula]³³ in Taisaka (Deschamps 1936). Further examples of coalescence in Southern Betsimisaraka from Dez (1963) include [teza] for <taiza> ("looking after") and as in Table 3.10, [toka] for <toaka> ("alcohol"). Other dialects also employ the hiatus resolution strategy where the low vowel is deleted, as in Taisaka ['miti] for OM <mainti> ("black"), although note that the /n/ is also missing.

When considering coalescence vowels, then, they seem to occur

³²For the Merina (and other ethnic groups), a meal is typically made up of rice, light broth, and <laoka>, which is typically meat, including fish, but can also be made of beans or other ingredients. Another translation could be "main course."

³³Due to OM's use of <o> for [u], and some authors tendency to employ both phonetic and orthographic glyphs, it is possible Deschamps (19360 meant ['tola]. For information about why the final syllable is missing here, see Section 3.6.3.

only in conversational speech, sometimes taking the place of diphthongs, as in Table 3.11, and in other cases being the surface pronunciations of hiatus whose usual pronunciation is undetermined. In any case, coalesced vowels may represent even a careful pronunciation in some dialects.

Not considered thus far is how underlying hiatus is pronounced when both underlying vowels surface as monophthongs. Dez (1963 p 518) proposes that, at least for Southern Betsimisaraka, an excrecent glide may be inserted between the vowels to ease the transition. /eo/ and /ea/ strings, then, can be pronounced [ejo] and [eja], although Dez says that the [e] can be so weak that the pronunciation may be closer to [jo] or [ja]. Looking at Northern Betsimisaraka, Kikusawa (2006) found a similar result for underlying /ia/ and /ua/, with surface pronunciations [ija] and [uwa].

To remain more neutral on the subject of how hiatus is resolved, SRs given in examples given in this work will display underlying hiatus unless its exact pronunciation is relevant to the discussion.

3.2.3 Sound-Based Divisions of Malagasy Dialects

Since Dez (1963), Malagasy dialects have traditionally been divided based on their reflex of two historical strings: *li and *ti. A major piece of evidence for segregating SW dialects (e.g. S. Sakalava, see Figure 3.1) from all others is that for these strings they have [li]

and [ti], while OM and others (including Betsimisaraka as described here) use [di] and [tsi], although a minority of the latter use [si] instead. In Dez's analysis, Anosy (spoken in the southeast corner) is the only dialect on one side of the [li] vs. [di] isogloss, but the other side of [ti] vs. [tsi]/[si]. According to Adelaar (2013), though, if this was previously true, it may be no longer, as modern findings show it having both [di] and [tsi], like other eastern languages.

One example of corresponding lexical items with either /d/ or /l/ is OM /vadi/ ("spouse") (also found in many other dialects, apparently including Betsimisaraka), which in some Southern dialects (Bara, Antandroy, and Tanosy) is /vali/. This type of cognate is usually seen where OM has an intervocalic /d/; another example is /madiu/ ("clean") as opposed to Southern /maliu/. Some dialects, such as southeastern Taisaka, may have both pronunciations, although typically this dialect has the OM-matching forms (Deschamps 1936).

Dez (1963) originally included Takarana—the furthest north dialect—with the western and southwestern varieties. Now, though, Adelaar (2013) proposes that either because more information is available or because the dialect itself has shifted under influence from OM and neighboring dialects, it is more appropriate to group it with the Central/East group. Conversely, its neighboring dialect to the southwest Tsimihety may similarly either have shifted toward, or else always have been part of, the Western and Southwestern [li]/[ti] group.

Due to this uncertainty, Northern dialects are not included in the consonant correspondence chart below.

Table 3.12 Consonant correspondences between different Malagasy dialects. Included is a column of the Proto-Malay-Polynesian roots that are the progenitors of all the forms given. Most data is from Adelaar (2013 p 462), except the Betsimisaraka words, which were gathered for this project. (Adelaar does provide some tentative, limited Betsimisaraka data, not included.)

PMP	Merina (Central)	Betsimisaraka (East)	Sakalava (West)	Tandroy (SW)	gloss
* l ima	dimi	dimi	lime	lime	“five”
* kulit	huditra	huditri	hulitsi	hulitse	“skin”
* putiq	futsi	fotsi	futi	futi	“white”
*ha- timur ³⁴	atsimu	atsimu	a(n)timu	atimu	“south”

Aside from the correspondences presented in the table, notice as well that some dialects may have optional word-medial nasal codas—or else optionally sub a fully oral stop for its pre-nasalized counterpart. These differences result in some homonymy, e.g. OM [futsi] (“white”) and [funtsi] “ravanala” (a type of palm tree) can both be [futsi] in Antaisaka (Deschamps 1936).

Dez (1963) believes that the division of dialects based on [l]/[d] and [t]/[ts] occurred on Madagascar, but Simon (1988) presents evidence that these distinction already exist among the various South East Barito varieties in Central and South Kalimantan. They thus may

³⁴ Adelaar gives the meaning “dry monsoon” for the PMP word, which is the root of Indonesian “timur” (“east”) as well as “south” for Malagasy. Similarly, the PMP *habaRat, for wet monsoon, gives Malay its word “barat,” meaning “west.” Adelaar (p.c.) suggests this is because these are the directions these particular winds come from in these different countries.

have existed as a variation at the time of migrations, which eventually manifested as an on-island split, or (reconciling this argument with Dez's proposal) it is not impossible that the same language change happened twice (Adelaar 2013). In any case, regardless of their historical roots, these divisions remain a useful first partition in sorting out the dialectology of Malagasy. The following section returns to describing the phonology of OM, with dialectal data where available.

3.3 Syllable Structure

Only simplex onsets are permitted, though complex onsets do marginally occur in some loan words, such as [frantsai] (from the French word for "French"). Codas are a bit more controversial, though no author posits the existence of complex codas. Authors disagree, though, on whether nasal codas are allowed in some positions, or whether no codas are allowed and seemingly nasal codas are just part of the onset.

If nasal codas are allowed (Dziwirek 1989, Albro 2005), words would be syllabified between nasals and following consonants, as in [mam.ba] ("crocodile"), and [kin.ta.na] ("star"). Under this analysis, the syllable type is (C)V(N). The nasal is not allowed in coda position word-finally³⁵, and—as shown in discussion of hiatus resolution—vowels

³⁵ Albro (2005), alone, posits that nasals are allowed word-finally, citing all antepenultimately stressed words ending in [na] as being penultimately stressed words ending in [n]. Although that is not the analysis used here, it follows from the devoicing of final vowels in that scenario.

may be unable to start any but the first syllable.

As discussed briefly in 2.1.1, syllable-final nasals can also be interpreted as the first element of a pre-nasalized stop, complex segments with a nasal and oral component, indicated by writing the nasal component as a superscript before the oral stop or affricate.. Erwin (1996) motivates this assertion by the lack of word-coda nasals and also cites some lexical items that have a nasal and an oral element to their onsets, such as [ngu.li] (“numb from cold”) and [ndzu.la] (“cross-eyed”); these represent exceptions in Dziwirek's model. Syllable-type is thus simplified to (C)V, assigning Dziwirek's (1989) analyzed nasal codas to the following onset.

Table 3.13: Summary of transcriptions following a nasal-coda model of OM syllable structure (Dziwirek 1989 *inter alia*), vs. a pre-nasalized-stop interpretation (Erwin 1996 *inter alia*)

nasal-coda	pre-nasalization	gloss
mam.ba	ma. ^m ba	“crocodile”
kin.ta.na	ki. ⁿ ta.na	“star”
ngu.li	ⁿ gu.li	“numb from cold”
ndzu.la	ⁿ dzu.la	“cross-eyed”

It is not within the realm of this project to decide between these two hypotheses on theoretical terms. This analysis will continue to discuss both options when relevant, with an appreciation for the parsimony argument for pre-nasalization while not ignoring the practicality of typesetting the nasal-coda interpretation. When not

discussing the two hypotheses, examples will be transcribed as if the nasal and oral consonant were separate segments.

Due to vowel devoicing (to be discussed in Section 3.5.2), some words may sound as if they *do* have codas, in that the devoiced vowel's onset consonant may sound as if it is a coda consonant to the previous vowel. Devoicing is the traditional analysis for why these vowels are not heard, perhaps due to the strict syllable structure seen elsewhere in the language. One may hear an utterance that sounds like [tranku], derived from /tranu+ku/ ("my house"), but one sign that the voiceless [u] is still there is that it prevents nasal assimilation: one does not hear *[trəŋku]. Another possible analysis given these data is that the vowel is deleted and that rule ordering explains the surface opacity. This work follows the traditional analysis that the vowels are devoiced without offering phonetic evidence³⁶ either way.

3.4 Morphology

This section looks at some morphological processes of Official Malagasy. It is not meant to be taken as an exhaustive look at all ways words can be formed, but rather to look at three common ones—verbal prefixes and suffixes, compounding, and reduplication—that will be very useful in observing phonological alternations. These three will be used later in this chapter, as well as extensively in Chapter 5, since

³⁶For those who would like to see some spectrograms of devoiced vowels, some are offered in examining stress in Betsimisaraka, in Figures 5.1, 5.2, and 5.3.

little-to-no morphological differences were found between Betsimisaraka and OM. Questions pertaining to these three processes were used on every field visit for this study, and through analyzing them, evidence came to light for every phonological process covered.

3.4.1 Verbal Morphology

According to Dziwirek (1989), Malagasy active verb formation combines three morphemes. Note that only tense is marked: there is no inflectional subject agreement.

- 1 A tense prefix: /n-/ (past), /m-/ (present), or /h-/ (future)
- 2 A stem-forming prefix (SFP) with no independent meaning, but which can indicate transitivity vs. intransitivity
- 3 A root, which has a semantic association but no defined meaning

Expanding on item 3, the verbs for “tattle” and “say” in Table 3.13 below both involve speaking and both employ the root /laza/. The SFP, though, has no semantic content, having instead a limited syntactic role; for example, verbs that take /an-/ as their SFP are far more likely to be transitive.

Table 3.14: Active verbal morphology of OM, following Dziwirek (1989). Shown are three verbs with /laza/ as their root, with disparate but semantically linked meanings. In the SR for “tattle,” the onset of /laza/ has undergone fortition (see 3.5.1) following /n/.

tense prefix	SFP	root	SR	gloss
m-	i-	laza	milaza	“say” (present)
n-	an-		nandaza	“tattle” (past)
h-	anjka-		haŋkalaza	“celebrate” (future)

Use of SFPs is not limited to verbs, and is carried over into nouns³⁷ that are derived from those same roots, as in Table 3.14.

Table 3.15: Nouns and verbs derived from the same root and SFP, using different prefixes. The prefix pronounced /p-/ is written <mp>, but as cited by Dziwirek (1989) *inter alia*, the <m> is just an orthographic convention.

prefix	SFP	root	SR	gloss
m-	i-	laza	misutru	“drink” (verb)
f-			fisutru	“drink” (noun)
p-			pisutru	“drinker” (noun)

In forming the passive, the vast majority of verbs employ a bare root³⁸ without the SFP, as in the examples below.

Table 3.16: Passive verbal morphology of OM, following Dziwirek (1989). It is there proposed that consonant-initial roots employ /nu-/ and /hu-/ as tense prefixes while vowel-initial ones use the /n-/ and /h-/ familiar from active verb morphology. Personal suffixes (1S /-ku/, 2S /-nau/, etc.) mark the agent subject of the verb; if it is unspecified, the agentless subject is /-na/.

prefix	root	suffixes		SR	gloss
		passive	personal		
nu-	fafa	-a	-na	nufafana	“was swept”
Ø	sutru	-i	-ku	sutruiku	“is drunk by me”
h-	aruv	-a	-na	haruvana	“will be protected”

³⁷ Nelson (2012) says “prefix f- that appears on deverbal nouns is not a nominalizer, but is rather part of the verbal complex.”

³⁸ Which of course raises the question as to how to distinguish passives, for example, for “say” and for “tattle.” Dziwirek does not cover this, although data collected for this study showed that, in Betsimisaraka, passive verbs can use the SFP, but still be distinguished from the active by using the separate set of passive tense prefixes (shown in Table 3.14 below) as opposed to the similar but disparate set used in the active (shown in Table 3.12 above).

The selection of passive suffix /-a/ vs. /-i/ is not connected to the choice of /an-/ vs. /i-/ (SFPs, above), and Dziwirek (1989), along with Keenan & Razafimamonjy (1996), states that it is lexically determined by the root. Tense prefixes used on the passive differ from the active in the lack of a pronounced present tense prefix and in the /Cu-/ allomorph which surfaces before consonant-initial roots.

Dziwirek (1989) also identified instances of irregularity, with slightly distinct forms used in the active and passive, as shown below:

Table 3.17: Irregular passive and active verbal morphology, after Dziwirek (1989). [hanimba] (“will ruin”) also evinces a voiceless-consonant deletion after /an-/ (see 2.4.1).

prefixes		root	suffixes		SR	translation
tense	SFP		psv.	actor		
h-	an-	simba			hanimba	“will ruin”
hu-		sumba	-i	-na	husumbaina	“will be ruined”
n-	ian-	anatr̪			nianatr̪a	“studied”
n-		enatr̪	-i	-ku	nenatriku	“was studied by me”

The vowel-quality changes above have not been analyzed as systematic, although again, more examples of this type may be necessary to see the triggering factors.

3.4.2 Compounding

Compounding can occur between various parts of speech; two examples in Table 3.17 are noun-noun, but “bless” is verb-noun.

Table 3.18: Four compounds of OM, as derived from their two roots, with examples from Martin (2005). Stress is indicated below as an indication that it is assigned to root words, not the final compound. Further discussion of stress is found in Section 3.6.1. The compound form for “bless” also exhibits opaque fortition of the onset of the second root, analyzed in Section 3.6.2.

Root 1		Root 2		Compound	
SR	gloss	SR	gloss	SR	gloss
'ba	“stocking”	ki'raru	“shoe”	,baki'raru	“sock”
ta'kela	“sheet”	'bi	“metal”	ta,kela'bi	“sheet metal”
'vavu	“heart”	'fu	“mouth”	,vavu'fu	“solar plexus”
'mitsu	“blow”	'ranu	“water”	,mitsu'd̪ranu ³⁹	“bless”

Genitive expressions—identified as such (and not as “compounds” per se) by Keenan & Polinsky (1998)—are used to show relationships between nouns or between noun heads and other words. These differ from the two-morpheme compounds in Table 3.17 by incorporating the genitive morpheme /n/ (similar to “of”), inserted between the two words, as shown in the example below in Table 3.18.

Table 3.19: Genitive compounds from Keenan & Polinsky (1998), all formed from the head ['tranu] (“house”). The form for “barn” displays nasal assimilation and the form for “stable” evinces fortition, both of which are covered in Section 3.4.1.

head		second word		compound	
SR	gloss	SR	gloss	SR	gloss
'tranu	“house”	an'd̪riana	“lord”	,tranun'an'd̪riana	“lord's house”
		'bibi	“animal”	,tranum'bibi	“barn”
		,sua'vali	“horse”	,tranun,tsua'vali	“stable”

³⁹The [r]/[d̪r] alternation is indicative of fortition, analyzed in 3.4.1, indicating that the UR of the first root may be underlyingly consonant-final.

3.4.3 Reduplication

Official Malagasy only has one reduplication pattern. Its semantic effect is similar to applying the English suffix “-ish” to adjectives, though it can also be applied to nouns with a meaning of “X-like thing/substance,” or to give verbs a casual or less intense nuance. According to Martin (2005), the reduplicant is selected based on metrical structure, targeting the final full foot such that the reduplicated form of [a'lika] (“dog”) is [a,lika'lika]. When the last full foot is not word-final, like some loanwords and compounds (see 3.6.1), the reduplicant is infix-like, preceding the last syllable. Examples of this type from Martin (2005) include [,suku,suku'la], from [,suku'la]] (“chocolate”) and [,zavu,zavu'ka], from [,zuvu'ka]] (“avocado”).

Certain words do not contain a full foot. This happens if the word is only one syllable long or for two-syllable words with final stress. Examples of the latter include onomatopoeia, demonstratives, and loanwords. In these cases, according to Keenan & Polinsky (1998), the final stressed syllable reduplicates, as in [,lu'lu], from [lu] (“rotten”) and [vu,vu'vu], from [,vu'vu] (“bark” V).

3.5 Segmental Phonology

3.5.1 Consonant-Consonant Interactions

This section covers the interactions between consonants, starting with the related processes of nasal assimilation and post-nasal

fortition, along with the deletion of voiceless sounds post-nasally. Concluding the section is a discussion of cases where suffixed forms have a wide variety of consonants not heard in corresponding roots; two hypotheses for this are weighed and one wherein these consonants are underlyingly present at the end of the root is selected.

Nasal place assimilation occurs when a nasal-final morpheme is concatenated before one that is consonant-initial. In the derived form, the nasal will possess the place features of the following oral consonant. This can be analyzed using the SFP /an-/, identified in 3.4.1 as the transitive active verbal prefix. As a frequent prefix, it occurs before all segments of Malagasy, surfacing as /am-/, /an-/, or /anŋ-/ depending on the place of the next morpheme's initial consonant.

Table 3.20: Nasal assimilation demonstrated by attachment of a nasal-final prefix, including a case where the nasal surfaces faithfully due to a vowel-initial root. Examples come from Paul (1996), but the transcriptions follow Dziwirek (1989) and Albro (2004) in showing assimilation at the velar place or articulation, which is not noted in the orthography of Malagasy.

root	active verb		gloss (verb)
	nasal assimilation	pre-nasalization	
aluka	manaluka	manaluka	“protect”
beda	mambeda	ma ^m beda	“grab”
disu	mandisu	ma ⁿ disu	“mislead”
gedza	mangedza	ma ⁿ gedza	“squeeze”

In a pre-nasalization analysis, the nasal segment itself does not have to change place. Rather, because the prefix has a nasal coda in a

language that does not allow them, this segment either (a) becomes the onset of the next vowel, the first segment of the root, or (b) is deleted for being a coda, but crucially its nasal feature remains. This feature attaches to the following consonant, surfacing as pre-nasalization on that segment. Only this [+nasal] feature remains of the segment, the place features lost with the segment itself.

Another alternation occurring at the prefix-root boundary is deletion of voiceless consonants. Unlike the nasal assimilation seen above, though, this process occurs uniquely at this morphological boundary. Although forms like [ampi] (“enough”), with a voiceless consonant following a nasal, are not uncommon, the same type of sequence is forbidden if (and only if) the nasal in question is that of the SFP /an-/ . Thus we see the underlying form /man+taitra/ (“startle”) surface as [mañaitra].

This prefix is so common, though, that this restricted post nasal voiceless consonant deletion is an important process in the language⁴⁰. If the deleted consonant is labial, then its place features are preserved as those of the nasal, for example /man+fihina/ surfaces as [mamihina] (“hug”)⁴¹. In a nasal-assimilation model, this creates opacity, with the nasal assimilating and then the consonant to which it assimilated being

⁴⁰ Nor is this process unique to Malagasy. The cognate Indonesian suffix /mən-/ has the same properties (Urbach 1987). Pater (1999) and Blust (2004) examine the related processes in several Austronesian languages.

⁴¹ If the deleted consonant is velar, i.e. /h/ or /k/, /n/ does not assimilate because [ŋ] cannot be a syllable onset in OM. As will be seen in 6.1.1, it can, and does, surface as velar in Betsimisaraka.

deleted, exceptionally, only in cases of the SFP /an-/. In a pre-nasalized stop analysis, there is no opacity: exceptionally, the nasal in the SFP /an-/ will *not* form complex segments with voiceless consonants, and forms instead a simplex segment with the place features of the stop applied to the nasal from the prefix.

Voiced sounds are not deleted after the SFP /an-/, but they do undergo fortition. This is a general phonological process occurring to all post-nasal consonants and is not limited to interactions with any single morpheme. Alveolar fricative /z/ surfaces as the associated affricate [dʒ] post-nasally, while /v/, without an associated affricate, surfaces as [b]. Also following nasals, liquids /l/ and /r/ surface as obstruents [d] and [ɖ].

Post-nasal fortition can be observed by attaching the SFP /an-/—compare [zavuna] (“cloud”) with its verbal form [mandzavuna]—but it is actually not the best trigger for this alternation, as voiceless sounds are deleted after it. In compounding and other forms of morphological concatenation though, voiceless sounds can follow /n/, and they all undergo fortition, as shown in Table 3.21 below.

Table 3.21 Examples of fortition in compounds where the first root is nasal-final, from Rajemisa-Raolison (1969 p 9-11). For clarity the first roots (but not the second) is listed in its UR, although in their surface forms they undergo epenthesis because word codas are not allowed, as is further described in Section 3.6.2.

root1 UR	gloss	root2	gloss	compound
/havan/	“relative”	lavitra	“far”	havandavitra
/ravin/	“leaf”	ramiari	<i>Datura alba</i> (flower)	ravin d ramiari
/lalan/	“road”	vauvau	“new”	lalambauvau
/fufun/ ⁴²	“smell”	havana	“relative”	fufunjkavana
		fanafudi	“remedy”	fufump p anafudi
/mihinan/	“eat”	saundzu	“taro”	mihinantsaundzu

Finally, fortition can also be heard in reduplicated forms. In these cases, because the final word includes both the original word onset and its copy, with one still word-initial and the other following a nasal, both the fortified and non-fortified alternant of the consonant can be heard in the same word.

Table 3.22 Examples of fortition where the reduplicant is nasal-final. As above, SRs of the word are vowel-final due to epenthesis. Examples are from Rajemisa-Raolison (1969 p 17-19).

gloss	UR	SR	reduplicated (SR)
“hoarse”	/farin/	farina	farimparina
“deep”	/lalin/	lalina	lalindalina
“virtue”	/hasin/	hasina	hasinkasina
“tired”	/vizan/	vizana	vizambizana

42 ['fufuna] literally means smell, but in compounds generally indicates a less formal version of the second root. A [fufun]'kavana], then, is a distant relative, just like a [havandavitra], and a [fufumpanafudi] would be a placebo. Not included above is its compound with ['vadi] (“spouse”): [fufum'badi], meaning “fiancé(e)”

In a pre-nasalization model, pre-nasalized segments are limited to those that share the stopped oral airflow of nasal stops, thus stops and affricates. The inventory, then, does not include pre-nasalized consonants *[^mv] or [ⁿl]; when these would be derived, grammatical [^mb] or [ⁿd], respectively, surface in their stead. This need not be motivated purely by the language's inventory: when the nasal attaches to next syllable, it contributes not only its nasality, but also its stop feature, such that when /v/ absorbs the nasality and stop features of the deleted nasal, it surfaces as [^mb].

Nasal assimilation, fortition, and the lexically limited voiceless consonant deletion process are all consonant-consonant interactions, and have been analyzed at the prefix-root boundary. The last alternation to be discussed in this subsection concerns consonants that appear in suffixed forms that do not surface in unsuffixed ones. These consonants will ultimately be analyzed as underlyingly root-final, although we will consider the hypothesis that they are inserted.

Table 3.23: Alternations wherein a consonant (emboldened below) appears in the suffixed form not in the bare root. It is not part of the suffix. Data comes from Rakotofiringa (1981 p. 22-24), who does not cite the meaning of the suffixed forms, though based on Dziwirek (1989), the first two appear to be passives and the third an imperative.

gloss (root)	bare root	suffix	suffixed form
“surround”	[funu]	/-ina/	[funus <u>ina</u>]
“burning”	[tsulu]	/-ana/	[tsuluf <u>ana</u>]
“love”	[tia]	/-u/	[tia <u>v</u> u]

There are three possible hypotheses to explain these consonants, although no author proposes that they are part of the suffix. This leaves two: that they are inserted to avoid hiatus or that they are part of the root. The former is discussed in the following paragraphs, though this analysis ultimately employs the latter.

The consonant-insertion hypothesis is put forward by Rakotofiringa (1981) and followed by Keenan & Polinsky (1998) and Rajaonarimanana (1995). In this analysis, it intrudes to prevent hiatus, but the choice of consonant is not entirely predictable. Rakotofiringa (1981) proposes many factors to predict it, including consulting Indonesian cognates, as shown in Table 3.24 below, although he does not describe the mechanism by which these consonants are accessed by modern speakers of Malagasy.

Table 3.24: Two alternations between bare and suffixed (apparently passive) forms, with Indonesian cognates offered by Rakotofiringa (1981) to explain what he calls the intrusive consonant.

gloss (root)	bare root	suffixed form	Indonesian cognate
“flee”	fuli	fulisina	pulet
“fire”	sali	salazana	t’alaʒ

Rakotofiringa, though, also offers a synchronic explanation, determining some “intrusive consonants” by dissimilation, as [-s-] or [-z-] are more likely candidates than [-f-] and [-v-] when there are labials in the root. He also links dissimilation to the place features of

the closest root vowel, including a statistical likelihood of co-occurrence between front vowels and alveolar intrusive consonants, leaving front labial fricatives to associate with the back vowels.

An alternative hypothesis is offered by Dziwirek (1989) and Albro (2005), and is considered, along with Rakotofiringa's proposal, by Rajaonarimanana (1995). In this analysis, the cognates cited by Rakotofiringa in Table 3.24 are not related just to the suffixed forms, but to the underlying form of the root. In OM, the codas are deleted because the syllable structure does not permit them, as seen in 3.3. When those same consonants followed immediately by a vowel-initial suffix, though, they are preserved and serve as onsets thereto. The data from Tables 3.23 and 3.24 is reinterpreted below following from this alternative hypothesis.

Table 3.25: Data from 3.23 and 3.24 featuring URs for the root that include the consonants heard only in the suffixed form. In the bare root, these consonants are deleted because syllable structure does not allow word-final codas. Roots for “flee” and “fire” can now be seen as more closely related to their Indonesian cognates, in which language the codas are allowed.

gloss (root)	bare root	root UR	suffix	suffixed form
“surround”	[funu]	/funus/	/-ina/	[funusina]
“burning”	[tsulu]	/tsuluf/	/-ana/	[tsulufana]
“love”	[tia]	/tiav/	/-u/	[tiavu]
“flee”	[fuli]	/fulis/	/-ina/	[fulisina]
“fire”	[sali]	/saliz/	/-ana/	[salazana]

Aside from Indonesian word codas, Rakotofiringa (1981) proposes

multiple factors within the bare root that offer a better-than-chance prediction of the consonant seen in the suffixes form. There is no reason to believe, though, that the underlying coda cannot simply be learned from the suffixed form⁴³ and dropped in the active. With a synchronic underlying coda analysis, none of the Rakotofiringa's statistical strategies are necessary—perhaps they reflect basic tendencies in morpheme construction—and the UR is revealed to be a closer cognate to its Indonesian equivalent. Furthermore, as will be seen in Section 3.6.2 and 3.6.3, there is other evidence of word-final codas, altho they are resolved through neutralization and epenthesis, not deletion. That subsection even presents evidence that nasals can also be deleted when they represent potential codas, in other dialects.

Four consonant alternations were examined in this section. In the first, the active verbal prefix was seen as /n/ before vowels, but if a labial consonant followed it, the nasal was instead [m]; this was analyzed as due to either nasal assimilation or, in a prenasalization analysis, concatenation of the features of two segments. Following that same verbal prefix (and only that prefix), voiceless consonants heard as the onset of the root were not present in the prefixed form, an alternation here analyzed as due to lexically defined deletion of voiceless consonants. A third CC alternation entailed a limited distribution of consonants following nasals, which in a nasal-coda

43 This is especially true given Malagasy's tendency to employ the passive form much more frequently than, for example, in English.

model was motivated by fortition, but in a prenasalization schema was due to a limited inventory of pre-nasalized consonants. Finally, both “thematic consonant” and coda deletion hypotheses were considered to explain cases where suffixed forms exhibited root-final consonants not heard in bare root forms. An analysis wherein word-final codas are deleted was selected here as more theoretically robust.

3.5.2 Alternations involving vowels

Along with other regional dialects, Betsimisaraka has more alternations affecting vowels than OM, in that it has vowel harmony. Because understanding which vowels are epenthesized entails an understanding of antepenultimate stress, discussion of the harmony exhibited by several Northern dialects is suspended until Section 3.6.2; harmony in Betsimisaraka specifically is in Section 6.2.3. OM, though, has at least three vowel alternations, which will be considered here. Although already covered when considering vowel inventories, some notes are offered on the controversial matter of hiatus resolution strategies. This is followed by a look at vowel devoicing in Official Malagasy, and finally the subsection concludes with an analysis of vowel weakening (by changing place).

As recounted in 3.2.2, scholars can only agree on very few points as to how VV sequences are realized, finding common ground only in that the usual realization of /au/ and /ai/ are (given some

transcriptional leeway) [a^w] and [aⁱ]. Some authors propose other possible diphthongs, and Rajaonarimanana (1995) lists several monophthong results of underlying⁴⁴ VVs, including two resulting from concatenation, [o] from /au/ or /ua/ and [e] from /ai/ and /ia/.

Rajaonarimanana's spelling-based approach is useful, but a complete analysis also incorporates interactions when a morpheme-final vowel meets a morpheme-initial one. Keenan & Razafimamonjy (1996) look at reduplication as well as pairs of words that cross word- and inter-compound-morpheme- boundaries, finding that in all three cases hiatus resolution depends greatly on stress⁴⁵. Unstressed /a/, Keenan & Razafimamonjy say (p. 39-40), elides pre-vocally, as will a vowel the same quality as the following vowel, as in the data below.

Table 3.26: Three examples of underlying vowel pairs surfacing as a single vowel. In the first two cases, [a]s are deleted in the resulting expression because they are unstressed [a] preceding another vowel. In the third expression, involving reduplication, the first vowel again elides, this time because it matches the vowel that follows it. Data is from Keenan & Razafimamonjy (1996).

morph1	gloss	morph2	gloss	expression	gloss
'uluna	person	'efatra	four	,ulu'nefatra	"four people"
'tapaka	"broken"	'elatra	"wing"	,tapa'kelatra	"has a broken wing"
'ivi	"spit"	'ivi	"spit"	'ivivi	"spittle"

44 Rajaonarimanana is actually considering pronunciations of orthography, which are here interpreted as SRs of URs. Some Adjustment is necessary for the IPA, e.g. replacing <o> with /u/ and word-final <y> with /i/, and cases where the author drops /h/ or devoiced vowels (see later in this subsection) are reinstated here as, while likely true, irrelevant to hiatus resolution.

45 This insight informed discussion of stress assignment in evaluating Rajaonarimanana (1995) in 3.2.2

If both vowels are stressed, as in the two-word expressions⁴⁶ below in Table 3.27, the two vowels both surface, protected from deletion by stress. The data in Table 3.26 represents how /a/ and matching vowels were the most likely to elide, but below, an [a.a] sequence is allowed to surface because both are stressed.

Table 3.27: Two examples of underlying vowel pairs surfacing as a pair of vowels, as both are assigned stress. Data is from Keenan & Razafimamonjy (1996). (Final stress words are imperatives; a phonological motivation for final-stress imperatives is offered for Betsimisaraka in Section 5.3.)

word 1	gloss	word 2	gloss	expression	gloss
,maŋka'tu	"obey"	'uluna	"person"	,maŋka,tu 'uluna	"obey people"
man'da	"refuse"	'azi	3S	man,da 'azi	"refuse him"

Both Keenan & Razafimamonjy (1996) and then Rajaonarimana (2004) discuss VV pairs where the unstressed vowel is not pronounced in favor of the stressed one. Depending on how close their transcriptions are, these could represent examples of the wider phenomenon of devoicing, a process that occurs to most unstressed vowels, depending on register. In all but the most formal, careful speech, though, is /a/ susceptible to devoicing, manifesting sometimes in as little as a release of the previous stop (Dziwirek 1989), though, as noted in Section 3.3, nasal assimilations and other phonological interactions still respect it as a segment. At the end of a word, /a/ will

⁴⁶ While it is true that all of Keenan & Razafimamonjy (1996)'s 'V'V data occur across word boundaries, this is due to the difficulty of finding final stress in any variety of Malagasy (see the Section 5.3), and note that "four people" from Table 3.26 was also across the word boundary but still evinces elision.

sometimes surface as a [ə] rather than becoming completely devoiced, as in /mandika/[mandikə] (“transgress”). Malagasy's other vowels are also seen to undergo devoicing: the UR /an+sulu+ku/ (“change” psv 1S) surfaces as [,anul'uku]⁴⁷ in careful speech, but can also be heard in discourse almost as [,an'luk]⁴⁸.

Because of the difficulty of perceiving place features on a voiceless vowel, certain word pairs (such as those below) will be indistinguishable in casual speech.

Table 3.28: Pairs of words that are pronounced generally the same due to vowel devoicing, as provided by Keenan & Razafimamonjy (1995), along with ambiguous transcription encompassing both, with no devoiced vowels transcribed. In [ma'nen.nə], the transcription indicates something like a surface geminate.

ambiguous transcription	word 1		word 2	
	citation	gloss	citation	gloss
'ent.na	'entana	“packages”	'entina	“carried”
am.'pang	am'panga	“accusation”	am'pangu	“rice crust” ⁴⁹
ma'nen.nə	ma'nenuña	“weaves”	ma'nenina	“regret”

In discussing vowel devoicing in the Taisaka dialect, Deschamps (1936) states that, in cases of antepenultimate stress (see the next

47 /s/ deletes (as analyzed in 3.5.1), because it is voiceless following the SFP /an-/

48 Note that, if it were not for the devoiced vowel between /n/ and /l/, the form would surface as *[anduk] due to fortition. As noted in 3.3, it would not be difficult, though, using rule-ordering or another multi-level approach, to account for this as fortition occurring previous to vowel deletion.

49 Rice is the staple food of Madagascar, and for many ethnic groups (including Merina and Betsimisaraka) accompanies three meals a day. It is typically cooked in a small cauldron and allowed to burn to the sides. The traditional beverage during and after meals is to boil water in the crust to absorb the nutty flavor of toasted rice. It is even often available complimentarily at restaurants.

subsection) the final vowel is generally devoiced, and that the penultimate /i/ and /u/ are somewhat devoiced while /a/ is not very pronounced. He says that it is due to this tendency that one occasionally finds two different careful pronunciations of the same word, as in [furⁱtri] or [fur^utri], both meaning “folded.” Some examples of careful and more conversational pronunciations from his research are presented in the table below.

Table 3.29: Vowel devoicing in the Taisaka dialect, as shown by careful and conversational pronunciations of the same words. Data is from Deschamps (1936 p 14), who uses “(a)” to denote a partial pronunciation of /a/, as opposed to total devoicing of /i, u/.

pronunciation		gloss
careful	casual	
'variki	'vark	“lemur”
'alut ^{tri}	'altʃ	“shelter”
'angat ^{tri}	'ang(a)tʃ	“phantom”
'lunaki	'lun(a)k	“chief”

Because of the difficulty of determining, without in-depth phonetic analysis, whether a vowel has been deleted or devoiced, aside from the general trends mentioned so far, vowel devoicing will not be further analyzed here. Because so much depends on the CV (or CV(N) except for final syllables) syllable template of Malagasy, though, it must be mentioned because, like [am'pang] (either “accusation” or “rice crust”?) in Table 3.28, those investigating the language will likely hear (or believe they hear) word-final codas and question the validity

of this important assertion. All previous authors, though, tell us that those vowels are in some sense there⁵⁰, just unvoiced.

The last vowel-based alternation to be analyzed here, like hiatus resolution or vowel-devoicing, takes into account which syllables are stressed and which are not. Because (as will soon be seen in Section 3.6) stress is assigned counting from the end of the word, adding suffixes will often change where the stress falls on the root. This means that, by looking at different words in a verb paradigm, vowel quality can change along with stress, as in the table below.

Table 3.30: Different forms of the same verb that show stress-based alternation between [e] and [i]. Adapted directly from Albro (2005 p. 228). Consonants that appear only in suffixed forms (e.g. the [s] in “kill”) were covered in the previous subsection; consonants that alternate when the root is suffixed (e.g. [k]/[h] in “sit”) will be covered in Section 5.6.3.

gloss	unsuffixed	suffixed
“kill”	maha'fati	a,hafa'teṣana
“please” (V)	ma,hafi'narīṭra	,aha,fina'retana
“sit”	mi'peṭraka	mipe'ṭraha
“look at”	mi'dzeri	,idze'rena

From the forms of “kill” and “please,” where a stressed [e] in the suffixed forms alternates with an unstressed [i] in the unsuffixed, a simple distribution suggests itself: [e] occurs in stressed syllables and

⁵⁰ One notable exception to this is Albro (2005), who (as previously mentioned) has an alternative analysis of antepenultimately stressed words with what are typically analyzed as nasal-onset final syllables. In his analysis there is no final vowel, the words are penultimately stressed, and these nasals are codas of the previous syllable. This analysis could contribute to why he accepts a CVN syllable template.

[i] appears in unstressed ones. Given that other examples have [i] in stressed syllables, Albro (2005) decides that all emboldened vowels in Table 3.29 are underlyingly /e/, surfacing as [i] in some unstressed environments. From “sit” and “look at,” though, we see that /e/ surfaces faithfully in other unstressed syllables. Albro makes the generalization that /e/ will only be raised following the main stress of the word, which holds true for the data in Table 3.29. It also recalls the pattern wherein /a/ was also raised to be pronounced [ə] in final syllables, from earlier in this subsection.

This pattern is not unique to OM, although the version that will be seen for Betsimisaraka (in Sections 4.1.1 and 6.2.1) is more general, with [i] alternating for [e] in all unstressed syllables and an identical distribution for [o] and [u] due to the presence of phonemic /o/. At the other end of the spectrum, Adelaar (2013) reports that [e] surfaces in unstressed syllables in Sakalava and other dialects.

Vowel-based alternations in Official Malagasy, then, are all related to the assignment of stress. When two vowels form hiatus, if only one is pronounced it is invariably the stressed one; when a vowel does not have stress, it can be devoiced; and in OM, in unstressed syllables following the main stress of the word, the low vowels /e/ and /a/ are raised to the higher vowels [i] and [ə], although the [a]/[ə] is not mandatory.

3.6 Suprasegmental Phonology: Stress

The final data examined in this chapter will be for stress, whose assignment has been hinted at throughout explorations of segmental alternations. When previous authors looked at stress, the presumably listened for the same acoustic correlates as Ferrand (1909): longer duration (p 241-242) and higher pitch (p 243-244).

The first subsection (5.6.1) presents the usual case of penultimate stress along with some exceptions that result in final stress. Section 5.6.2 shows how penultimate stress assignment can result in antepenultimate stress, and the nature of the final vowels in those cases. Finally, 5.6.3 presents the changes that can occur in the final consonants of antepenultimately stressed words.

3.6.1 Penultimate Stress and Final-Stress Exceptions

In Malagasy, unmarked stress assignment is to the penultimate syllable, resulting in the majority of two- and three-syllable words (and, with polymorphemic words, four- and up as well) exhibiting this pattern, as in the examples in the table below.

Table 3.31: Two- and three-syllable words exhibiting penultimate stress. These examples are from Martin (2005).

2σ word	gloss	3σ word	gloss	3σ word	gloss
1. 'futsi	"white"	u'mali	"yesterday"	ru'nunu	"milk"
2. 'vadi	"spouse"	ka'rama	"salary"	ta'nura	"young"
3. 'lela	"tongue"	a'tsimu	"South"	an'kizi	"children"

Final stress, though, is not unknown in the language. According to Martin (2005), for example, one-syllable words like [fe] ("thigh") and [^mba^j³⁷] ("step aside") are stressed, and some demonstratives are lexically defined as having final stress, such as [i'ti] ("this") and [i'za^j⁵¹] ("that"). Final stress is often found on imperatives as well; Thomas-Fattier (1982) cites the past and imperative of "wash," ['sasa] and [sa'sa], as minimal pairs for stress. These can be analyzed as an exception in this verbal category, but Section 5.3 presents the case, at least for Betsimisaraka, that the final vowels are actually the result of hiatus resolution between the penultimate, stressed vowel and the final vowel of the imperative suffix.

Many loanwords, too, have final stress due to the stress pattern of the native language, especially since the final-stress language French is such a prolific lexifier of Malagasy.

Table 3.32: Borrowed words exhibiting final stress, as they were in the language from which they were borrowed, from Martin (2005).

word	gloss	source
di'te	"tea"	"du thé" [dy'te] (Fr.)
zu'ma	"Friday"	"جمعة" [ʒami'ʕa] (Arabic)
suku'la	"chocolate"	"chocolat" [ʃoko'la] (Fr.)
laki'le	"key"	"la clé" [la'kle] (Fr.)
zavu'ka	"avocado"	"les avocats" [lezavo'ka] (Fr.)
sari'bu	"charcoal"	"charbon" [ʃaʁ'bɔ̃] (Fr.)

⁵¹[^mbai] and [i'zai] are transcribed as diphthongs here following Martin's transcription. If stress is assigned to the penultimate vowel, and not syllable (which is tantamount to saying stress assignment occurs following diphthong creation), then both can be seen as having, at some level, penultimate stress.

Another set of surface forms that do not match the penultimate-stress pattern are compound words. In forms where they do exhibit the expected stress pattern for Malagasy, with penultimate stress and secondary stress on alternating syllables leftward, it is because their second constituent root has an even number of syllables: consider the combination of [‘tranu] (“house”) and [‘bibi] (“animal”) to render [,tranum‘bibi] (“barn”). Given the great number of three-syllable words in the language, though, it can be easily demonstrated that stress is assigned to each morpheme individually in compounding, and a two-syllable second root just happens to trigger a pattern identical to the one for non-compounds. In the cases shown below, the stress pattern of compounds is that of two separate words, except that there is more emphasis (primary stress) on the stress of the second root.

Table 3.33: Two compounds that do not exhibit alternating stress, demonstrating that stress is assigned to each root individually.

root1	gloss	root2	gloss	compound	gloss
‘teni	“word”	a’tsimu	“south”	,tenina’tsimu	“South language”
pa’isu	“peach”	va’zaha	“foreign”	pa’isumba’zaha	“plum”

The two roots have joined together at some level of the prosodic hierarchy, but are still distinct at the level at which stressed syllables are marked. The unusual stress pattern seen in the compounds of Table 3.33 is used by Martin (2005) to motivate another unusual stress pattern: the one seen in loanwords such as [,laki’le] (“key”) in Table 3.32. These, he posits, could be analyzed by speakers as two bound

roots ([*'laki*] and [*'le*]) forming a compound to explain their unfamiliar stress pattern. This analysis is parsimonious in its elimination of many lexical exceptions—if not all: presumably demonstratives like [*i'ti*] (“this”) could also be split into separate morphemes⁵². It also demands, though, that the speaker store many meaningless morphemes (or novel uses for existing morphemes): approximately two for each loanword. Selecting his analysis here would actually only affect the count of underlying morphemes in the language, a list of which is not a goal of this project, so we leave both a lexical-exception and a bound-morpheme analysis open to future study.

3.6.2 Antepenultimate Stress and Epenthetic Vowels

Antepenultimate stress is actually a fairly common phenomenon in Malagasy languages. In OM, the group of words exhibiting this kind of stress all end with one of what are traditionally called the so-called “weak” syllables, [na], [tra], and [ka], which are said to be ignored by stress assignment (Keenan & Razafimamonjy 1996, *inter alia*). Stress, then, ends up assigned to the antepenult.

Table 3.34: Antepenultimately stressed words, all ending with weak syllables [tra], [ka], and [na], from various sources.

word	gloss	word	gloss	word	gloss
'vuhitra	“hill”	'tapaka	“broken”	'kintana	“star”
'sarutra	“difficult”	'satruka	“hat”	'uluna	“person”

52 As with imperatives, they also might be said to be, for example, underlyingly /itii/, with stress assigned to the penultimate /i/ and the final /i/ then deleted.

However, no author proposes that these three syllables are universally ignored in stress assignment; indeed, words ending [-na] form minimal pairs for stress, as seen in Table 3.33. Keenan & Razafimamonjy (2005) explain penultimately stressed “weak syllable” words (like [ta'nana] and [la'lana] below) by defining the group of pseudo-weak, listing seventeen forms and using ellipses to indicate that the list is incomplete.

Table 3.35: Two minimal pairs for stress, which can be found cited in almost every Malagasy grammar. Given that the last syllable [na] is one of the “weak” syllables, in this case the penultimately stressed forms are unexpected.

Penultimate Stress		Antepenultimate Stress	
word	gloss	word	gloss
ta'nana	“village”	'tanana	“hand”
la'lana	“law”	'lalana	“road”

In this model, there are three ways stress is assigned to words: typically, stress is penultimate, but it is antepenultimate if final syllable is in the “weak” category, except if the word is lexically defined as “pseudo-weak,” in which case it is penultimately stressed after all. Although this does explain the data in Tables 3.31, 3.34, and 3.35, a more parsimonious analysis is possible if the surface form of these words is not taken as the form to which stress is assigned.

Albro (2005) posits that all antepenultimately stressed words are underlyingly consonant-final. Recall in 3.5.1, an analysis of root-final consonant deletion was chosen over pre-suffix “thematic consonant”

insertion. This allows for root-final consonants, but the reader may recall that the only consonants seen deleting are /s, z, f, v/. According to Albro (2005) analysis of OM, [ka], [\widehat{tra}], and [na] are the surface realizations of all other word-final consonants, which neutralize to one of [k, $\widehat{t}r$, n] before undergoing vowel epenthesis. There are then two resolution strategies for potential word-final codas: the non-dorsal⁵³ fricatives /s, z, f, v/ delete, while the rest undergo epenthesis; both strategies serve to prevent word codas. Crucially, though, in the latter case, stress is assigned as if the epenthetic vowel were not there. This can be done derivationally, as below, or by marking epenthetic [a] as invisible to stress assignment.

Table 3.36: A derivation table showing an ordered-rule analysis of antepenultimate stress and penultimate stress on words ending in “weak” and “pseudo-weak” syllables. If stress were allowed to be assigned to epenthesized words, then the forms would be entirely homophonous.

Gloss	“hand”	“village”	“road”	“law”
UR	/tanan/	/tanana/	/lalan/	/lalana/
Stress Assignment	/'tanan/	/ta'nana/	/'lalan/	/la'lana/
Epenthesis	/'tanana/	---	/'lalana/	---
SR	['tanana]	[ta'nana]	['lalana]	[la'lana]

Either way, this analysis is more parsimonious because it does not require a separate stress-assignment rule based on the identity of the last syllable, nor lexical listings of exception to that separate rule.

53 Thus, not /h/. Albro's (2005) identifies these not as non-dorsal, though, but as strident. This will be crucial in motivating why some seeming root-final /f/s delete while others become weak syllables: he concludes that strident /f/ deletes and non-strident /ɸ/, an abstract phoneme that usually surfaces as [f], neutralizes.

It does, however, require an additional process of epenthesis, but this process is motivated by maintaining syllable structure, whereas it is not certain what would motivate a separate, “weak-syllable” stress assignment rule, or the many lexical exceptions to that rule.

An epenthesis analysis of the final vowel of antepenultimately stressed words also aids in comparing dialects. For example, where OM has [-ka] and [-tra], Taisaka has [-ki] and [-*tri*], some southern varieties have [-ke] and [-tse], and Bara has [-ki] and [-tsi]⁵⁴ (Deschamps 1936). (Due to elision of potential nasal codas, as seen at the end of this subsection, it is less useful to compare -na across dialects.) Under a “weak syllable” analysis, it might be difficult to motivate why antepenultimately stressed words change their final syllables’ vowels from dialect to dialect, while others do not. If “weak syllable” words are interpreted as underlyingly consonant-final, though, each variety simply has a different epenthesis rule to resolve word codas: OM and others epenthesize [a], while Bara and Taisaka use [i] and the Southern varieties analyzed by Deschamps employ [e].

Knowing that the Taisaka epenthetic vowel is [i], we can now appreciate some further evidence that antepenultimately stressed words have underlying potential codas. Recall that when discussing assimilation and fortition in 3.5.1, we considered compounds with a first root that typically ends in weak [na]. In the compound, though,

⁵⁴ According to Deschamps (1936), you hardly hear the vowel at all in Taisaka, you hear it more in OM, and most of all in Southern forms.

that [a] was not seen, with the /n/ assimilating to, and triggering fortition in, the following consonant. We did not cover cases where the weak syllable was [ka] or [tra], but as can be observed in the data below (adjusting for the Taisaka weak syllables [ki] or [tri]) the potential coda in the first root deletes, but still opaquely triggers fortition in the following consonant.

Table 3.37: Compounds in Taisaka wherein the first root is antepenultimately stressed and the second is consonant-initial. The underlying potential coda of the first root is deleted but, if the initial consonant of the second root is a non-continuant, it undergoes fortition. From Deschamps (1936).

root1	gloss	root2	gloss	compound	gloss
'angatri	“spirit”	velu	“living”	angabelu	“ghost”
'afaki	“free from”	'baraka	“honor”	afa'baraka	“dishonored”

Crucially, fortition is *not* triggered in compounds if the first root does have a weak syllable to “drop.” Underlying /tr/ and /k/ codas⁵⁵, then, behave identically to /n/ in the same scenario, although unlike /n/ they cannot surface because NC are the only grammatical clusters. Having more thoroughly established the presence of underlying potential codas, let us return to examining the dialectology of Malagasy coda resolution strategies.

When Ferrand (1909) describes Northern Sakalava, he says it is like the Southern varieties mentioned above in having [-tse], [-tsi], or

⁵⁵As will be shown in the next subsection, the ultimate underlying coda is one of /r, t, h/ in most cases, which neutralizes to this smaller set of consonants.

perhaps [-tsa] in place of OM [- $\widehat{\text{tra}}$]. A different, slightly more complex pattern emerges from a look through Thomas-Fattier's (1982) dictionary of Sakalava: in antepenultimately stressed words, the final syllable begins with one of [k, $\widehat{\text{tr}}$, n] and the final two vowels are of the same quality. More important than the examples listed below, though, is that the search of Thomas-Fattier's dictionary revealed a lack of forms where the vowels do *not* match.

Table 3.38 Antepenultimately stressed words found in Thomas-Fattier's (1982) dictionary of Sakalava, a Western dialect. These words all have matching pairs of final vowels, and none were found that did not. They are interpreted here as evidence of copy-enthesis due to potential word-codas.

word	gloss	word	gloss
[u]-final			
ma'hevuku	"transpire"	ma' η atuju	"approach"
man'tumbuku	"harass"	man'drukutru	"grate" (V)
ma'meluju	"support"	a'maluju	"eel"
[i]-final			
ma'maliki	"give back"	ma'iziki	"somber"
'aljni	"night"	man'drafitri	"do woodwork"
ma'madiki	"restart"	ma'mandriki	"trap-hunt"
[a]-final			
maha'sega η a	"rejoice"	ma'manaka	"have children"
ma'lamat $\widehat{\text{ra}}$	"slip" (v)	ma'nora $\widehat{\text{ra}}$	"write"
ma'minta η a	"fish with a line"	man'dahat $\widehat{\text{ra}}$	"arrange"

Without modern terminology, Ferrand (1909) comes close to analyzing forms like those seen above as underlyingly consonant final

and then undergoing epenthesis of a vowel matching the root's final vowel for quality, called copy epenthesis. Recall that we saw in Table 3.24 of Section 3.5.1 that the PMP root can in some ways replicate what we now analyze as the UR. Ferrand, therefore, proposes a diachronic analysis wherein what was formerly the Malay coda has taken on a vowel, and that that vowel is a copy of the previous. He cites an even earlier work (Baron 1893 p 57), which says:

The rule seems to be that the last vowel, at any rate after *k* and *tr*, is the same as the preceding one... Large numbers of such words therefore in the dictionary require correction, since they have been changed by those who have collected the words to harmonize with the Hova [i.e. OM before OM] form.

Velonandro (1983) shares Baron's suspicions about prescriptivist lexicographers, saying "Certain [editors] had a tendency to 'Merinize' [i.e. make more like OM] the words they were noting." Velonandro's project is, in part, a compilation of words from many sources, all covering a set of Northern dialects, mostly Takarana, but also Betsimisaraka, and some Betsileo and Sakalava. A search of antepenultimately stressed words in this wordlist reveals both apparent [a] epenthesis and copy epenthesis.

Given the number of dialects Velonandro (1983) covers, though, prescriptivism is not the only explanation for the [a]-final forms in the left column below. One of the dialects might have the same coda-resolution strategy as OM. Additionally, as will be seen to be the case

for Betsimisaraka in 6.2.3, copy epenthesis might be optional in some of the dialects—or all if the list, as intended, is not interpreted as exhaustive. The presence of multiple dialects might also explain potential minimal pairs for epenthetic vowels, such as ['felika] (“greens”) vs. ['feliki] (“movements of the tail”).

Table 3.39 Antepenultimately stressed words found in Velonandro's (1982) survey of Northern dialects, including Takarana, Betsimisaraka, Betsileo, and Sakalava. Forms on the left have [a] as their final vowel while those on the right have a vowel that matches the previous. Underlined forms also had an [a]-epenthetic variant.

[a]-epenthesis		copy-epenthesis	
word	gloss	word	gloss
'fasina	“sand”	'vurutru	“scrotum”
am'pendzika	“newcomer”	mam'positri	“fall over”
an'tsinika	“thingamajig”	ba'barini	“blackish leather”
'tsitsika	“curse” (N)	be'soruku	“thumbtack”
'endrina	(type of lemur)	buru'buruku	“caterpillar”
'vovuna	“roof”	mi'duduku	“get wet”
mi'bitsuka	“scatter”	mamorum'poruŋu	“crumple”
'esika	“blockage”	maŋetsiki	“hunt”

Given the data in the literature, then, it is likely that several Northern dialects exhibit copy epenthesis, including Takarana, Northern Sakalava, and Northern Betsimisaraka. As will be seen in 6.2.3, the Betsimisaraka spoken in Vavatenina exhibits copy epenthesis. As has been well analyzed, OM and other Central dialects epenthesize [a], and In some southern dialects the vowel epenthesized is invariably [i] or

[e], depending on dialect, although they also tend to have [ts]⁵⁶ in place of [tr] in “weak syllables.” The only pattern not seen, then, is [u]-epenthesis, though further explorations of Malagasy dialectology may reveal a language variety of this type.

3.6.3 Potential Coda Neutralization and Elision

In our discussion of Malagasy “weak syllables,” we have just learned that the final vowel is epenthetic, with its quality dependent on dialect, and we have said that the onset consonants are limited to [k], [tr]/[ts], and [n]/[ŋ]. The previous subsection motivated an analysis where antepenultimately stressed words are underlyingly consonant-final, though it has not explicitly been stated those underlying final consonants are, for OM, [k, tr, n]. In fact, although a hint to the potential coda can be seen in looking at Malay cognates, they can be seen synchronically by adding the passive suffix, as shown below. Ferrand's (1909 p 213-215) own interpretation of the data below is that the consonant seen before the passive suffix is an infix, although (like Rakotofiringa 1981) he does not propose how it is derived except by stating that it is a reflex of the Malay coda.

⁵⁶ Although this appears to be the standard analysis of the reflex of PMP final *t (see the following subsection), Adelaar (2013) cautions that in some cases this conclusion may be based on early grammar- and dictionary-writers being unclear on which symbol to use for which affricate. It may then be necessary to check these dialects again; if some of them invariably have [tr] as the final consonant of antepenultimately stressed words, they have either since merged with the Central and Eastern dialects, or were transcribed incorrectly in earlier fieldwork.

Table 3.40 Antepenultimately stressed roots and their corresponding passive forms, along with Malay cognates, from Ferrand (1909 p 211). The author identifies the emboldened consonant as an infix related to the Malay word coda, but this work analyzes it as the Malagasy underlying word coda.

Malay	Malagasy		gloss
	root	passive	
surat	'sura ^{t̪} ra	su'rata ⁿ a	"write"
tulis	'surit̪ra	su'ritana	"write"
bibir	'vivit̪ra	vi'vitana	"edge, side"
tiyup	'tsiu ^{t̪} ra	tsi'u ^f ina	"breathe"
kikil	'kikit̪ra	ki'kirina	"gnaw"
lipat	'lefet̪ra	le'ferina	"fold"

In Ferrand (1909)'s useful analysis of 300 antepenultimately stressed words where the final consonant is [tr], he finds that the alternant before the passive suffix is typically [r] (200) and [t] (80), plus 10 more cases where it can be either. The rest were other consonants, including 7 instances of [f]; these last might be analyzed by Albro (2005) as SRs of an abstract phoneme /ɸ/, as will be explained further in 6.1.7. For this section, it suffices to say that, unlike /f/, /ɸ/ is not deleted word-finally.

Although Ferrand's work significantly predates generative phonology, the data he collected and organized can now be viewed as instances of word-final codas /r, t, ɸ/ that are preserved when suffixed, for example, in the passive. This same set of codas all undergo neutralization to [tr] word-finally, but the trigger for neutralization is

not evident because it is then followed by the appropriate epenthesis process for the dialect.

This synchronic analysis accounts for alternations like those seen in Table 3.39 without recourse to the Malay roots. One consequence, then, is that the speaker can only determine the underlying word coda by hearing the suffixed form. One might expect, then, that (despite that Malagasy tends to employ the passive with more frequency than English), when speakers posit a UR, they might occasionally employ the most frequent underlying coda /r/ in place of a less frequent one like /t/. Ferrand (1909) shows data for which this is the case, comparing some then-modern forms to the ones found in the very first manuscripts of Malagasy⁵⁷.

Table 3.41 Antepenultimately stressed roots with an older and newer passive form. The pre-passive-suffix consonant has changed from [t] to [r]. Data is from Ferrand (1909 p 219), who notes that the former is closer to the word-coda of the Malay cognates.

active	passive		gloss
	original	“modern” (1909)	
'elat̪ra	e'latina	e'larina	“wing”
'havit̪ra	ha'vetina	ha'verina	“hook”
'vezat̪ra	ve'zatina	ve'zarina	“open”

Continuing in his diachronic analysis, Ferrand (1909 p 219) points out that having [tr̪] for Malay [t] is not so unusual, considering cases where other Austronesian languages have [r] for Malay [t].

⁵⁷ Transcribed in the 15th century using the Arabic-based “Sorabe” alphabet.

Table 3.42 Cases where Malay word coda [t] diverges to [r] in other Austronesian dialects and [tr̪] in Malagasy. Data is from Ferrand (1909 p 219); forms in the “Malay” column may be from another form of Malay-Indonesian, e.g. Batak.

Malay	other Austronesian		Malagasy	gloss
	word	language		
senjat	sigir	Tagal	senitra	“stinger”
lalat	laler	Javanese	lalitra	“fly” (n)
pusat	puser		fuitra	“navel”
laut	lauer	Bulusch	alautra	“sea”
kait	kawer		havitra	“fog/cloud”
tumit	tumir		tumitra	“talon”
lilit	kaliliri	Bugui	lilitra	“twist”
gigit	gegel	Sundanese	kekitra	“bite”

Furthermore, Ferrand (1909) presents cases where Malagasy employs [tr̪] as the equivalent of word-final [t] when borrowing English words, and of word-final [r] and [l] in French. These are important for an argument that word-final /r, t/ neutralize to [tr̪] word-finally, because these cases do not rely on suffixes forms to show that they are word-final: they are clear cases where a Malagasy hears one of word-final [t, r] and pronounces the same as [tr̪].

Table 3.43 Borrowings from English where the borrowed word ends in a [t] in the original language, but ends with [tr̪] in OM. From Ferrand (1909).

English	Malagasy	English	Malagasy
rabbit	ra'bitr̪u	cabinet	kabi'netr̪a
bracket (punc.)	bu'raketr̪a	sergeant	sari'zenitr̪a
support	sa'poritr̪a	carbonate	karibu'netr̪a

Table 3.44 Borrowings from French where the borrowed word ends in an [r] or [l] in the original language, but ends with [tra] in OM. From Ferrand (1909).

French	Malagasy	gloss	French	Malagasy	gloss
'sab ^r	'sabat ^{ra}	“saber”	e'pɛgl	'pingit ^{ra}	“pin”
dy'pwav ^r	dipu'avat ^{ra}	“pepper”	la'tabl	la'tabat ^{ra}	“table”

Significantly, Ferrand also provides data (not given here) where no other coda is pronounced [tr̚] when the word is borrowed.

Ferrand conducted a similar analysis of “weak” [ka], finding 600 roots, all of which have [h] before the passive and [k] as the word-coda of related Malay forms. 40 of these also allow [f, t] before the passive in Malagasy, but none had [k] in this position. Cases where non-[h] pre-suffix consonants are allowed tend to be a reflex of the Malay coda, as shown below.

Table 3.45 Exceptional cases where the root is antepenultimately stressed, ending in [ka], but the pre-passive consonant can be another consonant aside from the usual [h]. These are a reflex of their Malay cognates. From Ferrand (1909 p 188).

Malay	Malagasy		gloss
	radical	passive	
tiyup	'tsiu ka	tsi'u fina	“breathe”
dʒelap	'le laka	le'lafina	“lick”
tahit	'tar ika	ta'ritina	“drag”

Again, these pre-passive consonants are here interpreted as underlying word codas that undergo neutralization to [k] word-finally,

before undergoing epenthesis. These same potential codas surface faithfully when suffixed, although in this case it is more difficult to say why /f, t/ should not neutralize to [tr̪] as seen above. Albro (2005) again employs abstract phonemes, identifying /f/-that-neutralizes-to-[k] as [β], as opposed to /f/-that-neutralizes-to-[tr̪], which is /ɸ/ and /f/-that-elides, which is actually /f/. The form for “drag,” not analyzed by Albro, can be treated as a lexical exception.

As Ferrand (1909)'s analysis depends on looking at older manuscripts written in the Southeast, as well as then-modern Merina (not yet OM) Malagasy, there is no reason to think that what we are now analyzing as neutralization is restricted to OM. Indeed, Deschamps (1936) presents similar data for Taisaka, where the epenthetic vowel, as shown in the previous subsection, is [i] instead of [a]. In his data, we see root-final /r/ and /t/ neutralizing to [tr̪], /h/ neutralizing to [k], and /v/ being deleted, as was previously examined for OM in Section 3.5.1.

Table 3.46 Taisaka nominals and passives, including three cases where the passive suffix reveals the consonant that is neutralized in the nominal. In the final row, the passive suffix reveals a consonant that was deleted in the nominal. From Deschamps (1936 p 15).

nominal	passive	gloss
'sambutri	sam'buri	“seize”
'suratri	su'rata	“write”
'petraki	pe'trahi	“sit”
'tsizu	tsi'zuvi	“perceive”

Ferrand's (1909) systematic approach to what is here analyzed as neutralization concludes by examining "weak" [na]. Almost all have, as their pre-suffix consonant in the passive, [n], although 15 had /m/, suggesting all nasals neutralize to aleveolar word-finally, pre-vowel-epenthesis. A single case where the revealed potential coda is /v/ can be treated as exceptional.

In some Northern dialects, the result of word-final neutralization for nasals is not [n], but [ŋ]; given its inventory, this is not an option for OM. Although Velonandro (1983) does not give corresponding suffix forms, at this point antepenultimate stress should constitute sufficient evidence for underlying potential codas.

Table 3.47 Antepenultimately stressed words ending in [ŋa], from Northern dialects. Data is from Velonandro (1983), who collected data from Takarana, Betsimisaraka, and other language varieties. Note that the form for "truth" exhibits the copy epenthesis described in the previous subsection, and thus /ŋ/ appears instead as its palatalized allophone [ŋ].

word	gloss
ma'ŋairanja	"measure"
a'naranja	"name"
an'ketipni	"true/truth"
ma'ŋaranja	"fish (by hand)"
'bedanja	"act of landing"

Although Velonandro (1983) mostly presents antepenultimately stressed words where [ŋ] (or [ŋ]) is the final consonant, his wordlist is not without instances of "weak" [na], for example [anka'ranana] (type

of wood) and ['bedana] (post-rain mud). These may be influenced by OM or other dialects; interestingly, he has two translations for “dull, worn down”: [dumuna] and ['dumuŋu], one pure OM form with [n] as the result of nasal neutralization and [a] as the epenthetic vowel, and one dialectal form neutralization to [ŋ] and copy epenthesis. The lack of compromise forms, i.e. those that exhibit one dialectal markers for Northern varieties but not the other, suggests that speakers are aware, on some level, of what rules go with what dialect.

We have just seen that nasal consonants have two different results of neutralization, depending on dialect. What is here analyzed as underlying potentially word-final nasals also have an alternate form in some Malagasy dialects where the nasal is elided, a strategy that was previously used to analyze resolution of certain fricative codas in 3.5.1. Deschamps (1936) indicates that OM ['vuruna] (“bird”), from Malay “burung,” is [vuru] in Taisaka, saying similar variants are found in Taimoro, Sakalava, Betsileo, Betsimisaraka, and especially Bara⁵⁸. Ferrand (1909 p 197) says that, depending on dialect, words with nasal “weak” syllables in OM can be nasal-final in other dialects, or the entire syllable can be missing, with or without nasalization on the previous vowel. Indeed, both weak-syllable and the elided versions are found in Velonandro's (1983) survey of Northern dialects, and Kikusawa (2006)

58 Deschamps (1936) points out that, in all these dialects, when neutralization is used in place of elision to resolve nasal codas, the surface nasal is [ŋ], not [n], suggesting the former is a less capable coda.

indicates that all of Ferrand's possibilities are found in the Betsimisaraka of Vavatenina. As will be shown in Section 6.2.3, this finding is supported by fieldwork in Vavatenina done for this project.

Table 3.48 Antepenultimately stressed and penultimately stressed versions of the same word, from the Northern dialects surveyed by Velonandro (1983). Its UR can be analyzed as having a potential coda, resolved by neutralization for the former and nasal elision in the latter.

antepenultimate stress (neutralization)	penultimate stress (elision)	gloss
'alina	'ali	"night"
'embaña	'emba	"wing," "birdflight"

Table 3.49 Variants of the weak syllable [ŋa] in the Betsimisaraka of Vavatenina, and their OM variants, from Kikusawa (2006 p 10).

Betsimisaraka	OM	gloss
'anjaranja, 'anjaranj, 'anjarā	anara	"name"
sa'hoŋa, sa'hoŋ, sa'hō	sahuna	(type of frog)

Dez (1963 p 517) agrees that of words without weak-syllable [na] a variant often occurs wherein the [na] is entirely absent. He notes that in Southern Betsimisaraka, OM ['funtuna]'s cognate is so ['fontū]⁵⁹, or even ['fōtū]⁶⁰ without an equivalent final syllable and *with* the now-final

59 This form also exhibits /o/, phonemic in Betsimisaraka and other regional dialects.

60 Dez (1963) wonders if some transcriptions exhibiting nasalized vowels in place of nasal codas are due to his fellow French speakers applying their own phonotactics to Malagasy. He is thus uncertain how to interpret, for example, ['vita], one author's transcriptions of the Bara dialect cognate of OM ['vinta] (destiny). Did it lose the /n/ diachronically, or does it lose it synchronically and the author missed the nasalization? Difficulty in hearing nasalization is especially the case, he posits, for nasalized /i/, which does not exist in French, a reference point for almost all of those investigating Malagasy.

[u] nasalized. There is even some evidence that, diachronically at least, other potential codas can be completely dropped: Deschamps (1936) reports that OM ['haru] (“mix”), ['haruka] (“searching”/ “scouring”), and ['haruna] (“basket”) are all pronounced ['haru] in Taisaka.

Although the phenomenon has been described diversely by more than a century of authors working in different traditions, the data presented here support an analysis where Malay words (C)V(C)V often have a Malagasy cognate, '(C)V(C)V, where the final consonant and vowel are narrowly specified by the dialect. The Malagasy UR, though, will be /(C)V(C)/, and while the final C may still be different from its Malay cognate, its possible features are not nearly so restricted. This underlying final consonant will be revealed in cases where vowel-initial suffixes are added; while previous non-generative approaches analyzed it as an “intrusive consonant” or an “infix” and not an underlying potential coda, they did note the featural connection to the original Malay root.

3.7 Conclusions

The dozen or so previous studies of OM phonology agree on many factors, especially those that can be based on analyses of analogous alternations in other Austronesian languages, such as those surrounding Indonesian /mən-/ prefixation. The inventory count is

generally agreed upon as well, though the surface realization of certain segments, principally the ones written <tr> and <dr>, remains in question. This is especially true when considering their pronunciations in (non-OM) regional dialects, where they may even have merged with non-anterior affricates written <ts> and <j>. Still, there are two larger controversies to consider when examining any dialect of Malagasy.

The first concerns syllable structure, which is certainly restricted in many ways. Still, consensus proves elusive as to whether nasals are allowed as codas or if seemingly coda nasals are actually instances of prenasalization. Intriguingly, the two fully phonological treatments of the subject, Dziwirek (1989) and Albro (2005) support, or at least opt for, the coda-nasal analysis while Keenan's two morphophonological articles (Keenan & Razafimamonjy 1995; Keenan & Polinsky 1998), plus those that cite them, continually promote the temptingly parsimonious pre-nasalization account. One might think, since the justification for pre-nasalization is almost entirely phonological, that the reverse would be true. Deciding between these models is beyond the scope of this project, and thus this grammar follows its fellow phonological treatments in positing the existence of nasal codas.

The second controversy is over whether roots can underlyingly be consonant-final. The difference is, again, between phonologists and other linguists, although in this case the phonologists propose the phonologically motivated analysis. Phonologists Dziwirek (1989) and, in

a fuller account, Albro (2005), both take advantage of the idea of underlying potential codas to explain many surface irregularities: consonants that appear between a root and its suffix that do not appear on the bare root, the limited inventory of onsets to “weak” syllables, and the very existence of antepenultimate stress. Keenan & Polinsky (1998) *inter alia* propose that the consonants between roots and suffixes are “intrusive” (following Rakotofiringa 1981), and that weak roots must be lexically marked. Seeing how none of those steps are necessary, an analysis allowing root-final consonants underlyingly should be adopted for any and all future studies of Malagasy phonology, and will be shown to be useful in Betsimisaraka in Chapters 5 and 6 as well.

Work on non-OM dialects and on dialectology itself points to various markers to look at when analyzing an under-described language variety like Betsimisaraka. Many dialects were said to have phonemic /o/, and apparently all but OM have /ŋ/. Cognates of Malay forms that have *li and *ti strings should be examined to see if the dialect uses faithful [li] and [ti], like Southwestern and some Western dialects, or novel [di] and [tsi]/[si], like Central, Eastern, and some Northern dialects. Some dialects were said to not allow nasal codas even at the syllabic level, although researchers should take care that they are not applying biases from their own language—even if that language is another dialect of Malagasy!

Finally, those describing Malagasy dialects should pay particular attention to how the “weak syllables” are pronounced. The onset consonants that (here analyzed as the results of opaquely word-final neutralization) vary between dialects. They might be [k, \widehat{tr} , n], like OM; alternatively they might sub [ts] for [\widehat{tr}] or—if they have the phoneme—[ŋ] for [n]. The vowels of weak syllables might universally be one of [a, i, e], or might be a copy of the previous vowel, here analyzed as the final underlying vowel of the root. And in any case, because of vowel devoicing, the vowel quality of the vowel might be nearly impossible to discern.

Chapter 4

SEGMENTS OF BETSIMISARAKA

Introduction and Overview

According to Betsimisaraka folklore, during a conflict between the Betsimisaraka and Merina people, a single sentence was used to determine if someone claiming to be Betsimisaraka actually was.

Akôho lahy magnèno ambinagnin' ny Managnàra.
chicken male sing delta gen. Mananara
"The rooster sings in the Mananara delta."

This sentence, written above in the Betsimisaraka orthography used by the primary language assistant, contains both of the phonemes that Betsimisaraka has (emboldened in the example above), but that Official Malagasy lacks, to be used as a shibboleth to out possible Merina interlopers: the fifth phonemic vowel /o/, written <ô>, and a third phonemic nasal, /ŋ/, written with the digraph <gn>. A Merina speaker would likely pronounce either the <ô> as [u] or one of the <gn>s as [n], respectively, as in his own dialect, and find himself with at least some serious explaining to do.

This chapter delves into the sounds of Betsimisaraka, offering both a phonemic inventory and an explanation of contextually determined allophonic pronunciations. Where possible, minimal pairs

are provided. While Northern and Southern varieties of Betsimisaraka have several lexical differences, and some small changes in rules and rule ordering, the phonemic inventories of the dialects are identical.

4.1 Vowels

4.1.1 Phonemic and Allophonic Vowels

Unlike Standard Malagasy, Betsimisaraka has five phonemic vowels, consisting of the five cardinal vowels, as indicated below. As is typical with this arrangement of vowels, the non-low back vowels are pronounced with lip rounding, and there is no low-vowel distinction for frontness.

Table 4.1: Phonemic vowels of Betsimisaraka

	front	back
high	i	u
mid	e	o
low	a ⁶¹	

OM's distribution of [o] sounds was said, in section 3.1.2, to be limited to coalescence of certain /au/ and /ua/ sequences, an interjection (“ô!”), and some loanwords. Instances of [o] in Betsimisaraka, though, are due to the existence of an /o/ phoneme; its

⁶¹In employing the IPA symbol [a], technically reserved for low front vowels, this approach does not purport that the vowel is phonetically front. It is more likely (given alternations seen in Subsection 6.2.2), in fact, to phonetically be the back vowel [ɑ], but as there is no front/back distinction, the more typographically advantageous glyph was chosen for this work.

phonemic status can be motivated entirely⁶² by its stress-conditioned alternation with [u], distinct from the behavior of the separate phoneme /u/, as indicated in Table 4.2. There, because stress is assigned from right to left, adding a suffix assigns stress to different vowels in the same root, showing how the same underlying vowel reacts with and without stress.

Table 4.2: Active and passive forms of two verbs, showing disparate stress-conditioned behavior of /u/ and /o/. Both vowels surface faithfully in stressed syllables, but as [u] in unstressed ones.

Gloss	Active	Imperative
“assemble (for a meeting)”	[mi' vuri]	[miv <u>u</u> 'riə]
“pass”/“drop by”	[man'dalu]	[,mandə'l <u>ovə</u>]

If [o] were not a phoneme and just the form /u/ takes in stressed syllables, then it should do so in “assemble.” The same would be true of /o/ were the phoneme and [u] its unstressed allophone.

At first glance, this distribution recalls the stress-conditioned distribution of *front* vowels seen in Section 3.4.1, wherein, from the same root, a mid vowel ([e]) appeared in stressed syllables while a high vowel ([i]) appeared in unstressed. Betsimisaraka's front vowels, though, have a more general stress-conditioned alternation: recall that

62 In Betsimisaraka, no minimal pairs could be found for back round vowels /o/ and /u/. When questioned on possible minimal pairs, knowledge of official Malagasy often stood in the way. For example, [voluna] (“moon”) contrasts in the minds of Betsimisaraka speakers with [vuluna], the OM pronunciation of the word. Even working with those who did not speak OM, when subjects were asked about a list of words featuring /o/, but with /u/ in the place of /o/, they would invariably respond either that the form existed, but was the “Merina” (i.e. OM) pronunciation, or else did not exist if the form was a lexical item unique to Betsimisaraka.

in OM, pretonic unstressed mid-vowels maintained their quality. Betsimisaraka /e/, like fellow mid-vowel /o/, does not surface faithfully in unstressed syllables, as heard in examples in Table 4.3 below. Formalization of height alternation will be left to Section 6.2.1.

Table 4.3: Disparate stress-conditioned behavior of /i/ and /e/, wherein both vowels surface faithfully in stressed syllables, but as [i] in unstressed ones. This is a more robust pattern than the only-after-primary-stress raising seen in OM in Table 3.17.

Gloss	Active	Imperative
“cut”	[ma'ndidi]	[di'diə]
“look”	[mi'dzeri]	[dzi'revə]

Stress-based alternation between mid and high vowels, then, is characteristic of Betsimisaraka. Note in both Table 4.2 and 4.3, that another vowel, [ə], only appears in unstressed syllables, which was presented in Section 3.4.2 as a distribution of OM as well. In the examples above, [ə] is heard primarily as the realization of the imperative suffix, but note also the imperative form of the verb “drop by,” [,mandə'lovə], where [ə] follows the [d] in the root, a position where [a] is heard in the active [man'dalu]. This alternation, with [a] (phonetically lower) in the stressed positions and [ə] (phonetically higher) in the unstressed, is similar to the [o]/[u] and [e]/[i] distributions, with two important differences. First, the mid/high alternations represent neutralization, whereas Betsimisaraka has no evidence of phonemic */ə/, which is to say, no word where [ə] occurs in

stressed position. The second difference is that, no matter how careful the pronunciation, neither *[man'dalo] nor *[mi'dzere] (ungrammatical forms with mid vowels in unstressed positions) are heard, whereas [mandalova] (with [da] rather than [də] as the second syllable) is a careful but otherwise unremarkable alternative to [mandə'lovə].

From examining the alternations in verb paradigms, then, Betsimisaraka's five-way distinction stressed syllable vowel pronunciations emerges, allowing [i, e, a, o, u], whereas unstressed syllables permit just three, either [i, a, u] or [i, ə, u].

Perhaps due to the prestigious nature of OM, Betsimisaraka speakers are well aware of their distinct pronunciations. From OM, Betsimisaraka gets the convention of writing /u/ as <o>, relegating /o/ to the diacritic-enhanced <ô>. Furthermore, only surface [o] is written <ô>; unstressed /o/, pronounced [u], is written as <o>. [mandəlovə], then, is written <mandalôva>, but [mən'dalu] is universally <mandalo>, never *<mandalô>.

One reason for this may be that Malagasy is written with optional diacritics marking stress, as in <tanàna> ("hand") and <tànanana> ("town"), although the second form is rarely seen, since, with "weak" <na> as the final syllable, antepenultimate stress is the expected case. Diacritics, then, are expected only in stressed syllables, which is the only place /u/ is pronounced [o]. It is not entirely surprising for a neutralizing phonological rule to be more noticeable to speakers, and

thus recorded in the orthography⁶³, which could motivate why the same practice follows for orthographic <e> and <i>, writing the latter for /e/ except when it is pronounced [i].

4.1.2 Vowels resulting from coalescence

Another distinctive aspect of Betsimisaraka is the presence of the vowels [ɛ] and [ɔ], which will be analyzed here as realizations of underlying /ai/ and /au/ sequences⁶⁴. As noted in Section 3.4.2, these sequences are pronounced [a^j] and [a^w] in OM, as shown below for two words with the same UR in each dialect:

Table 4.4: Surface realizations of /au/ and /ai/ sequences in Official and Betsimisaraka Malagasy.

gloss	standard orthography	pronunciation	
		OM	Betsimisaraka
2S	<anao>	[a'naw]	[a'nɔ̄]
“where”	<haiza>	['ajza]	['(h)ɛza]

From just the evidence above, one might suggest that Betsimisaraka has phonemes */ɛ/ and */ɔ/, and that the coalescence

63 This is also true, to some extent, in English nasal assimilation. As English has only two or three phonemic nasals, we only show nasal assimilation in writing when it neutralizes to a phoneme, such that <impossible> uses the grapheme <m>, while <infelicitous> and <incredible> (pronounced with [m] and [n], respectively), the nasals are written <n>.

64 The author never learned to reliably hear the difference between instances of tense and lax mid vowels in non-citation forms. Conversely, the project's language assistants had no trouble distinguishing instances of [a'tɔ̄] (/atau/, a passive form of “do”) and [a'to] (/ato/, one of a few words corresponding to English “here”). When pronounced in isolation, the former obviously had the lax vowel, but the trouble emerged when transcribing participants' natural speech.

analysis is just a result of an OM-centric Betsimisaraka analysis, yet such an analysis does not follow from alternation evidence presented in Section 6.2.4, specifically Table 6.53. If some mid vowels are due to coalescence, there is thus no reason to posit phonemes */ɛ/ and */ɔ/. A treatment of the rewrite rules that render mid-lax vowels from underlying sequences can also be found in section 6.2.4.

The mid-lax vowels [ɛ] and [ɔ] replace the most common OM diphthongs [a^j] and [a^w], but diphthongs are still found in Betsimisaraka. The presence of rising-sonority diphthongs was controversial in OM (see 3.4.2), but they are present in Betsimisaraka, at least in some registers. When non-low vowels precede stressed [a], they can join to become a diphthong wherein the first vowel becomes non-syllabic, becoming an on-glide. Like /a/ reducing to [ə], this alternation typically occurs, but is not required, especially not in citation forms. Examples of this disparity are offered in Table 4.6, below.

Table 4.5 Formation of rising diphthongs in Betsimisaraka. To be clear, in the usual gerund column, [ʷa] and [jɑ] are rising-sonority diphthongs, not coarticulation on the onset consonant.

gloss	active	gerund	
		citation	usual
“wilt”	mə'lazu	fandazu' a ŋa	fandə'zʷaŋə
“break”/ “read”	mə'maki	famaki' a ŋa	fama'kʷaŋə

Other underlying instances of hiatus result in diphthongs

following a similar procedure: the unstressed vowels become non-syllabic and converts to the onglide or offglide portion of the diphthong, with the stressed vowel staying on as the syllable nucleus. This is true of sequences of two high vowels as well, as when [miudi] (“go home,” active) is pronounced [mjudi]. It is less certain what happens to mid vowels: either they are raised because they are in unstressed syllables (as we saw in section 4.1) and then converted to the same glides as high vowels, or else are converted to mid glides, which were mistranscribed as high glides due to the author's unfamiliarity with the latter. More in-depth discussion of the surface representation of unstressed mid vowels in hiatus resolution requires recourse to formalization and is thus suspended until Section 6.2.4.

Throughout this work, examples will employ VV sequences like those shown in the citation forms in Table 4.6, as it is always possible to pronounce them that way and how exactly the vowels are pronounced will not usually be relevant to the discussion at hand. Diphthongs are transcribed more narrowly above to give a full accounting of vowel sounds in the language, and they will be transcribed as such in later chapters when relevant, for example, when a full accounting of hiatus resolution strategies is given in 6.2.4.

4.2 Consonants

The distribution of consonants in Betsimisaraka is almost

identical to that found in OM, with some variation. The major differences, expanded in the sub-sections to follow, are the presence of the velar nasal as a phoneme of the language and the pronunciation of the phonemes spelled <tr> and <dr> in the shared orthography, which are here identified as the post-alveolar fricatives familiar to English speakers, but are analyzed many different ways, both in OM and across dialects, as recounted in 3.1.1; this study employed retroflex affricates in Chapter 3.

Table 4.7 Phonemic consonants of Betsimisaraka and their articulations. By convention, voiced sounds are to the right of their voiceless counterparts. Bilabial fricatives in parentheses indicate that these are abstract consonants. Unless affected by other processes, they ultimately neutralize to the labiodental fricatives [f, v]. Justification for these abstract segments is provided in sections 6.5 and 6.7, but as they are not “sounds” of Betsimisaraka, no further discussion will be provided about them here.

	Bilab.	Labiodent.	Dental	Alveolar	Post-Alv.	Velar	Glottal
Stop	p b		t d			k g	
Affricate				ts dz	tʃ dʒ		
Fricative	(ϕ β)	f v		s z			h
Nasal	m		n			ŋ	
Lateral				l			
Rhotic				r			

4.2.1 Distribution of Nasals

The most significant difference is the existence of an additional phoneme, the velar nasal /ŋ/, which literate Betsimisaraka speakers are aware of as a distinct feature of their dialect. One subject was

heard to say, without being questioned on the matter:

ilay	ana	môka	misy	tréma
the	2S	[pause word]	there's	(diacritic) ⁶⁵
"That's 'ana' with a <i>tréma</i> ."				

This subject, after using his dialectal word for "you" in the previous sentence, felt a need to point out that he had not said [ana] that time, but rather [aña]. The reason he refers to it as having had a *tréma* is that Malagasy's Official Orthography uses not <gn> but <ñ>, seen, for example, in the alternate name for the Southern town of Ft. Dauphin on some maps, <Tôlañaro>⁶⁶.

While bilabial and alveolar nasals can be found word-initially or word-internally, the velar nasal is only found word-internally. Furthermore, given word pairs such as in 4.8, it may be tempting to hypothesize that [ŋ] is a stress-conditioned allophone of /n/.

Table 4.8 Pairs of words that differ only by stress assignment and place features of the non-labial nasals.

penultimate stress	gloss	antepenultimate stress	gloss
ta'naja	"town"	'tanjana	"hand"
ma'rini	"near"	'mariŋi	"correct"

In the lexically unrelated forms above, the words are the same but for the placement of stress and the place of the nasals. By the way

⁶⁵The diacritic *tréma* is used in French to mark two vowels pronounced separately, such as in the French word <Noël> ("Christmas"), pronounced [no'ɛl]. In Betsimisaraka its function is simply to visually distinguish <ñ> from <n>.

⁶⁶Another unofficial, but popular option, is to use <ñ>, as seen in the works of Jacques Dez (1963, 1991, *inter alia.*) and other linguists.

place features alter with stress, one might posit stress-conditioned [n]/[ŋ] allophony. For “town” and “hand,” for example, one could hypothesize that nasals become dorsal following stressed vowels. Unfortunately for this analysis, the reverse is true of “near” vs. “correct”: /i/ is followed by [n] when stressed, but [ŋ] when unstressed. Therefore, rather than positing a system wherein one stressed vowel causes the following nasal to shift *to* alveolar while a different vowel causes the following nasal to shift *away from* alveolar, an analysis wherein a velar nasal exists as a phoneme of Betsimisaraka is adopted.

Additionally, the following minimal pairs may occur for some speakers of Betsimisaraka, although it should be noted that these do not come from fieldwork in Vavatenina, but from a colleague with knowledge of Betsimisaraka.

Table 4.9: Minimal pairs for /n/ and /ŋ/ in Betsimisaraka. These forms were provided by Jeannot Fils Ranaivoson (p.c), a Malagasy phonetician familiar with Betsimisaraka, and subsequently checked with the primary language assistant, so there is no guarantee that they are used in Vavatenina.

word featuring /n/	gloss	word /ŋ/	gloss
'lani	“expend”	'lanji	“goal” or “stench” ⁶⁷
'reni	“mother”	'renji	hear (psv.)

Certain neutralization alternations between [n] and [ŋ], also contribute to justifying the phonemic status of /ŋ/. These alternations

⁶⁷These two glosses are not intended as two meanings of the same word. Rather, the first gloss was supplied by Ranaivoson, and the second meaning by the study's primary language assistant, who was asked to confirm these pairs.

will be analyzed in Chapter 5: Like OM, Betsimisaraka has a nasal-assimilation process (6.1.1) and a distribution whereby all nasals neutralize to [ŋ] as the onset of the final syllables of antepenultimately stressed words, (6.1.7). Unlike OM, though, Betsimisaraka also has a nasal-dissimilation process across vowels (6.1.8).

To establish that /n/ and /m/ are distinct phonemes, one need only compare the past and present tense of any active verb, or any adjective, because (as presented in 3.2.1), /m-/ is the present prefix and /n-/ is the past.

Table 4.10 Different tense forms of verbs and adjectives employing the /n-/ past and /m-/ present tense prefixes.

gloss	verb	
	past	present
“see”	na'hita	ma'hita
“go home”	ni'odi	mi'odi
“rain”	na'noranja	ma'noranja

4.2.2 Obstruents of Betsimisaraka

Most obstruents are the same as in OM. Among labials, stops are pronounced bilabially while fricatives are labiodental, although this only distinguishes them phonetically and does not prevent, for example [b]/[v] or [f]/[p] alternations, as analyzed in 6.1.4. As in OM, non-anterior coronal obstruents have dental stops [t̪, d̪, n̪] (although the diacritics will only be employed here when relevant) and alveolar

affricates and fricatives [ts, dz, s, z]⁶⁸. A minimal pair that demonstrates the phonemic status of oral vs. nasal stops is [an'tudi] ("egg") vs. [an'tuni] ("reason").

However, Betsimisaraka's anterior obstruents—the ones written with the digraphs <tr> and <dr>—are postalveolar affricates, rather than the retroflex affricates employed for OM in Chapter 3 (See Section 3.1.1). For Betsimisaraka speakers, though, pronunciation of these sounds invariably consisted of a stop followed by a release with the characteristic spectrum of a post-alveolar fricative.

As mentioned in 3.1.1, Dez 1963 presented evidence that retroflex and alveolar affricates might be merging in certain dialects. Aside from their disparate pronunciations, and their different behavior vis-à-vis fortition (Section 6.1.3), at least one minimal pair exists for the voiced set of affricates: [gidʒu] ("monkey" or similar creature) vs. [gidzu] ("path").

Distribution of Betsimisaraka /h/ is similar to that of OM, in that it is not always pronounced⁶⁹, although before stressed vowels it is typically heard. Just as in OM, though, the existence of /h/ is plain from

68 The feature [+anterior] distinguishes coronals into two sets, the [+anterior] post-alveolar and retroflex sounds, i.e. [ʃ, ʒ, ʂ, ʐ] from [-anterior] dentals and alveolars, i.e. [θ, ð, s, z]. This feature will turn out to be important to the distinctive feature analysis in Chapter 6.

69 In fact, due to the author's unreliability in distinguishing [o] from [ɔ] (See 4.1.2), it was difficult for him to discern when subjects were saying /akohu/ ("chicken") and when they were saying /akau/ ("there"). Like in the former example of /atau/ and /ato/, neither language assistant had any difficulty whatsoever, even when the words were played in isolation.

when it surfaces as its allophone [k], for example as the final consonant of an antepenultimately stressed word (See 6.1.7), or due to post-nasal fortition (See 6.1.3). These pronunciations of /h/ as [k] are important in establishing its phonemic status, as it cannot be analyzed as simply a writing convention, as has been the case for some studies of OM.

4.2.3. Approximants of Betsimisaraka

In Section 4.1, the glides [j] and [w] were proposed to occur as allophones of corresponding front and back vowels, existing as on- or off-glides on surface diphthongs. No data exists providing any support for their existence as phonemes⁷⁰ or even as independent segments, nor is there evidence of any other glides.

Like OM, Betsimisaraka has a rhotic and a lateral liquid, with only one minimal pair found for this study: [oruju] (“nose”) vs. [oluŋju] (“person”). Additionally, although /l/ neutralizes to [d] under fortition (Section 6.1.3), the minimal pairs ['fadi] (“taboo”) vs. ['fali] (“happy”) and [mi'vidi] (“buy”) vs. [mi'vili] (“turn”) establish the phonemic distinction between /d/ and /l/.

When the rhotic is not word-initial it is almost always pronounced [r]. The trill [r] is heard from time to time, but given that the

⁷⁰This is not to dispute Dez's claim of finding /j/ as a reflex of /z/ in Southern Betsimisaraka. He was studying a variety in the far South of the region, closer to Mahanoro on the map in Figure 1.1.

prestigious dialect uses a trill, it is not surprising to hear of some free, if limited, variation. Still, for ease of transcription, [r] will be employed in the system used in this work

4.3. Chapter Summary

This chapter has now set the stage to begin examining the alternations that occur in Betsimisaraka, beginning with suprasegmental ones in Chapter 5 and moving back into the fully segmental realm in Chapter 6. Chapter 6 will also adapt the phonetic features proposed in 4.1 and 4.7 to assign phonological distinctive features to all the phonemes described here. This chapter has confirmed remarks about Betsimisaraka by previous researchers, in that it has two phonemes /ŋ, o/ not present in OM and two results of coalescence [ɔ, ε] not heard there, as well as distinct pronunciations /tʃ, dʒ/ for its anterior affricates.

Chapter 5

STRESS AND STRESS-ASSIGNMENT IN BETSIMISARAKA

Introduction and Overview

Stress in Betsimisaraka can be understood as a bundle of co-occurring acoustic markers and segmental features that tend to occur on the penultimate syllable of morphological words, although several segmental interactions will result in surface stress falling on the final or antepenultimate syllable.

The indicators of stress will be examined in 5.1, starting with acoustic features and moving on to segmental ones. Section 5.2 presents the most common stress assignments on surface forms: penultimate and antepenultimate, the latter of which can be analyzed as due to an epenthesis process that derivationally follows stress assignment. Cases of final stress are examined in Section 4.3, and explained as resulting from various hiatus resolution strategies that eliminate final syllables by way of merger or elision.

Because in so many cases stress appears to be assigned to a non-surface form of the word, it will be impossible to discuss any non-penultimate stress assignment without some discussion of the segmental alternations from Chapter 6. Every effort will be made to

provide data wherein only the relevant information is shown, but the reader will understand that Betsimisaraka is a living language and inevitably some forms offered will display other than the relevant alternations.

5.1 Correlates of Stress in Betsimisaraka

The phonetic correlates of stress in Betsimisaraka are those that are typically employed cross-lingusitically, which is to say: longer syllables, higher pitch and higher amplitude pronunciations, along with lack of vowel reduction. Some spectrograms of Betsimisaraka words are offered below to illustrate the indicators speakers employ to mark stressed vowels. However, this is just one small section of a larger work decidedly in the realm of phonology, and it does not attempt to offer an in-depth acoustic phonetic analysis of stress in the language.

The criteria for selecting the sound files for these visualizations was non-scientific, so these examples should not be taken as clear examples of the signals of stress and not statistically average. These were, rather, pronunciations that the language assistants found typical, and that contain evidence of the correlates of stress found throughout the study. For a straightforward example, consider the following spectrogram of a recording of the word written <bôtra> (“chubby”), pronounced ['botʃa] or ['botʃə].

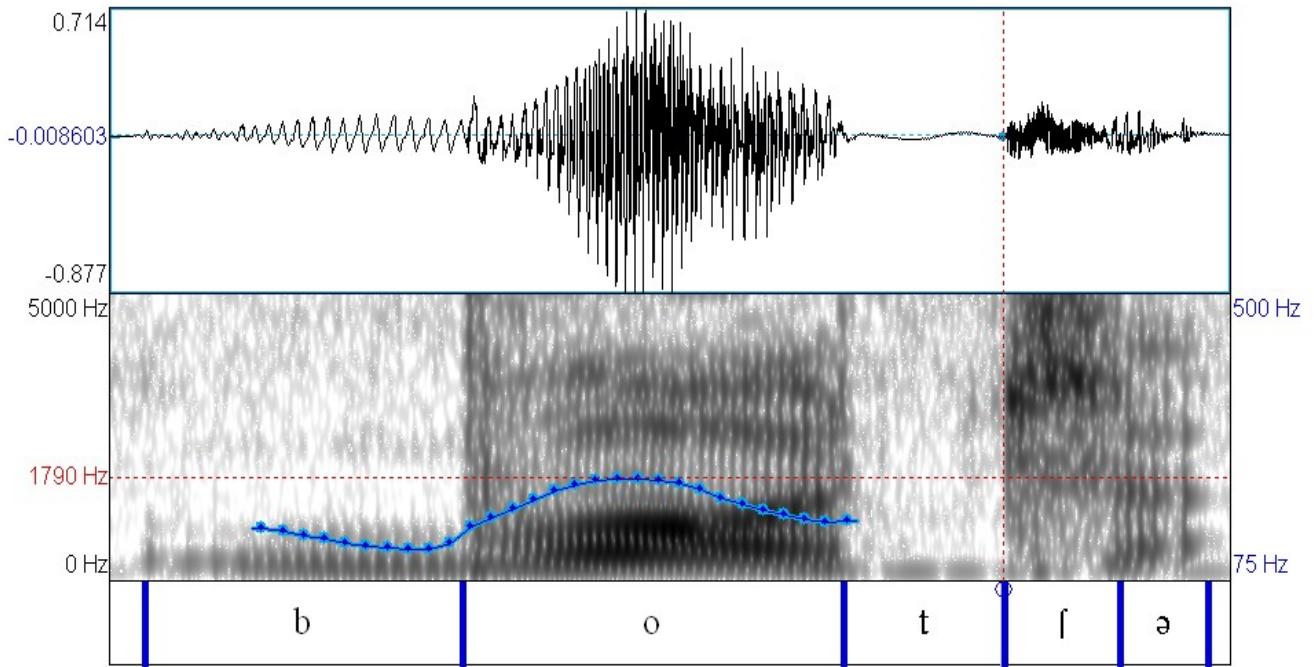


Figure 5.1 Spectrogram (in Praat) of ['botʃə] ("chubby"), showing increased length and amplitude, as well as high frequency, on the stressed vowel, along with full place value of a mid vowel. The unstressed vowel has shorter length, lower amplitude, no discernible pitch, and is reduced to [ə]. (The stop and fricative portions of the affricate are marked separately above, but this is not to indicate that they are not part of a single phonological segment.)

In this example, the period of high energy and characteristic vowel formants signifying the [o] vowel lasts several times longer than the one for the [ə], and pitch tracking reveals that, while the speaker increased in pitch until the mid-point of the [o], Praat was unable to locate any pitch whatsoever for the [ə]. In this example, then, what we are calling the stressed syllable is marked by longer duration, as well as higher amplitude and pitch.

Aside from these prosodic features, though, stress is also indicated by segmental alternations between different vowel heights.

The word ['botʃa] ("chubby") was chosen to demonstrate the correlates of stress because its stressed vowel is the mid vowel [o], a vowel that never occurs in unstressed syllables, as previously discussed in 4.1.1. Recall that, due to suffixation that alters where stress is assigned, a syllable that would have the vowel [o] or [e] when stressed would be instead pronounced [u] or [i] when unstressed. That the speaker could pronounce a mid-vowel in the first syllable of ['botʃa], then, is another signifier of what this grammar will be referring to as stress⁷¹. Similarly, the final vowel, although still optionally pronounced [a] even in unstressed syllables, can be pronounced [ə] here, an option not open to low vowels in stressed syllables, even the word-final ones seen in Section 5.3.

Before proceeding to a more complex example, let us examine a contrastive pair. Although this study produced no recordings of minimal pairs for stress, some similar words still exhibit disparate stress patterns, which can be used to similar effect. The terms that will be used in the figure below are <lalàna> ("law"), pronounced [la'lana], and <lalagna> ("road"), pronounced in careful speech ['lalaŋa] but otherwise simply as ['lala]. It would have been ideal to have a pronunciation wherein the final syllable was not dropped⁷², but this process was so

⁷¹ ['botʃa] was never heard with any stress-shifting suffixes, but several examples of verbs exhibiting [o]/[u] alternations are offered in Section 6.2.1.

⁷² As will shortly be seen in Section 5.2, this antepenultimately stressed word is indicative of an underlying potential coda. 6.1.7 formalizes the optional deletion of potential nasal codas.

common that no such recordings could be found of this common word.

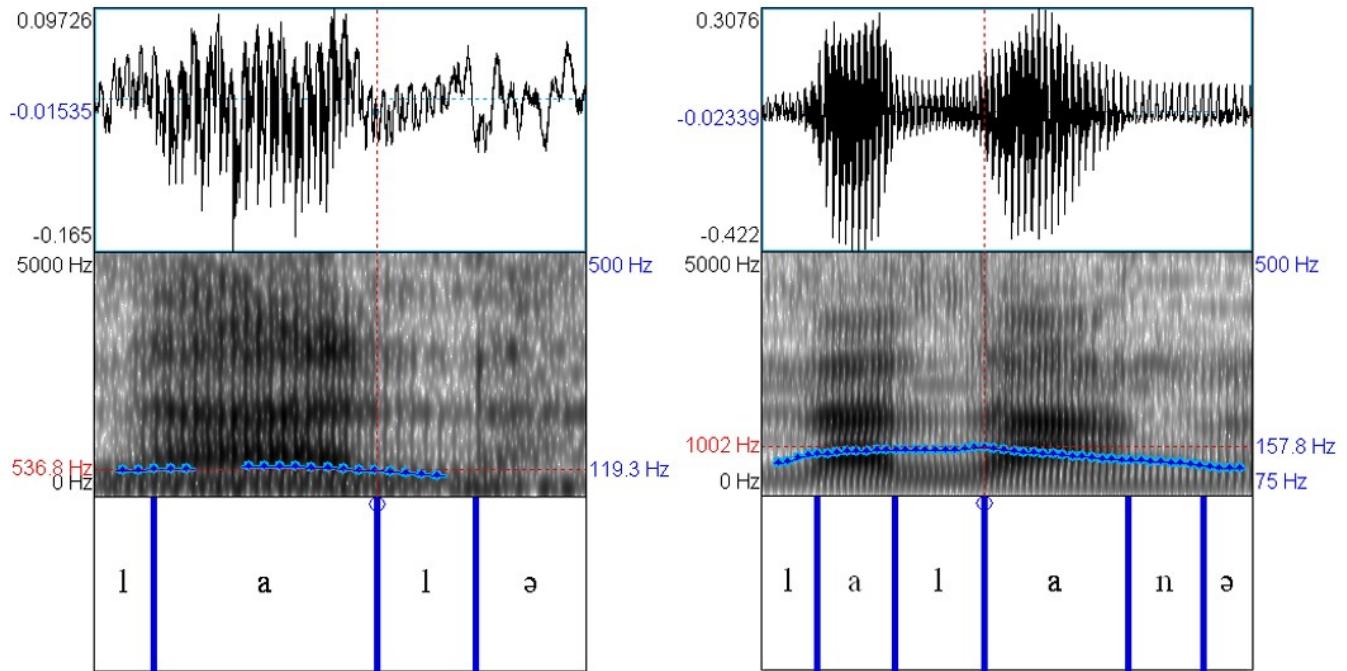


Figure 5.2 Spectrograms (in Praat) of ['lalə] ("road") and [la'lənə] ("law"), showing increased relative duration of stressed vowels. For ['lalə], pitch and amplitude also seem to be useful indicators, while this cannot be stated with confidence for [la'lənə], where a glance shows little difference.

Concentrating on just the [lala] sequences the two words have in common, we can see again that duration is a valuable acoustic correlate of stress, as the vowel we have identified as stressed is longer than the unstressed. The spectrogram for ['lalə] shows that the [a], as compared to the [ə], has pitch, higher amplitude, greater duration, and is not reduced, exhibiting all four correlates of stress. In [la'lənə], though, while the second [a] differs from the [ə] by the former having far greater duration, higher amplitude, and lack of reduction, the distinction between the two unreduced vowels is more

nuanced. Certainly the stressed vowel is longer, but neither the pitch or the amplitude appear to be significantly higher. Duration, then, may be the best indicator for stress.

In Betsimisaraka, complete vowel-devoicing⁷³ can also be characteristic of unstressed vowels. In the figure below, a narrower transcription would mark two of the unstressed vowels as devoiced, to wit [fufun'dzavətʃə]. For those two syllables, while the consonants of two are clear, there is little-to-no voicing bar or vowel formants for the corresponding vowels.

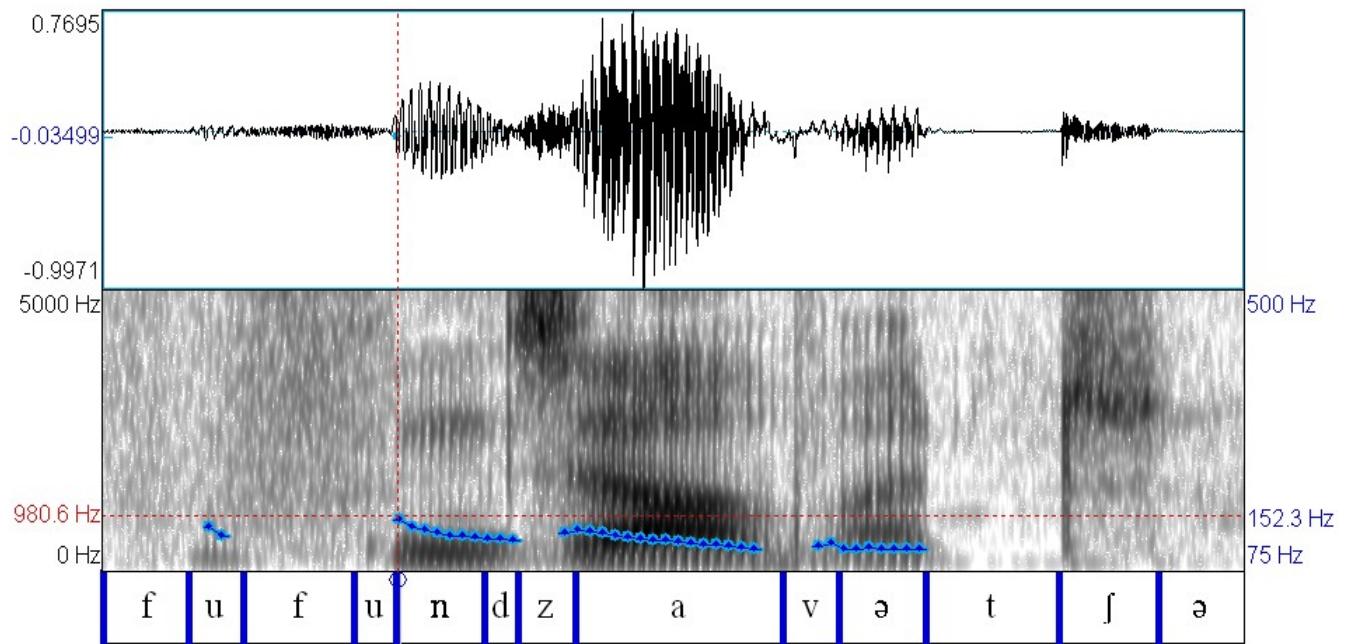


Figure 5.3: Spectrogram (in Praat) of [fufun'dzavətʃə] (“smell-of-something”), showing the previously mentioned increased relative vowel length, pitch, and amplitude for stressed vowels, as well as vowel reduction (the two [ə]s) and devoicing (the [u]s) for unstressed vowels.

⁷³As discussed in Section 3.4.2, vowel devoicing is the traditional analysis for OM, although given Figure 5.3, complete vowel deletion cannot be entirely ruled out. Determining exactly which process is occurring here, though, is a phonetic concern outside the realm of this phonological grammar

The longest-duration, highest amplitude vowel in this compound, and thus the stressed one, is the [a]. This [a] is also stressed in the original second root, ['zavatʃa] ("thing"), as is the first [u] in ['fufuna] ("smell"), although in the compound the latter only gets secondary stress. As in Figure 4.1, this vowel also has a higher pitch than the others (though, interestingly, apparently not higher than the preceding nasal). The following, unstressed [ə] (the penultimate syllable), then, has a lower pitch and is less than half the length of the stressed [a]. The second [ə] (the final syllable) has some light vowel formants, but no pitch or voicing whatsoever, although its onset, the affricate [tʃ], has all its acoustic indicators present. This lack of voicing is another indicator that vowel in question is unstressed.

Similarly, while of the [u]s, the first has a voicing bar and visible pitch, the second barely has either. Both vowels, though, are much shorter in duration than the stressed [a], and neither is the primary stress-bearer for the word. This grammar, then, will consider the distinction between the two [u]s as one in which the first bears *secondary* stress, while the second is not marked for stress at all.

One type of vowel reduction exhibited by unstressed vowels is part of hiatus resolution: non-low vowels desyllabify to on- or offglides of the following/previous vowel, while low ones delete. This process cannot be analyzed phonetically as it depends on underlying and surface forms of words, but is mentioned here so as to have a

complete understanding of factors used to determine stress.

Diphthongization was mentioned briefly in 4.1.2; section 5.3 explores these processes as they pertain to stress assignment, and a featural analysis appears in 6.4.2.

In this grammar, then, when a vowel is said to be stressed ([+stress] when speaking in terms of features), it is at least longer, along with being possibly louder and of higher pitch than one that is not. It is also able to host a greater distinction of vowel quality (allowing the mid vowels [e] and [o]), will never reduce to [ə], will always never lose its syllabacy or delete to resolve hiatus, and will never devoiced. A vowel with secondary stress will have higher pitch than an unstressed one, will not be totally devoiced, and will also be unaffected by hiatus resolution. Finally, then, an unstressed vowel is short, low-pitch, quiet, reduced (optionally in the case of [ə]), and additionally can become a glide or (for low vowels) delete altogether due to hiatus.

The purely prosodic features of length, amplitude, pitch will not be referenced further in this grammar, and optional reduction to [ə] will be largely ignored as well. Diphthongization will only be marked if relevant to the analysis. The reduction of non-low vowels, though, is a segmental neutralization (both /e/ and /i/ are pronounced [i] in unstressed syllables while /o/ and /u/ reduce to [u]), so vowel raising will continue to be represented in all transcriptions, as (it was noted in

Section 4.1.1) they are in Betsimisaraka orthography.

5.2 **Penultimate and Antepenultimate Stress**

Now that the factors identifying stress have been inventoried, the stress-assignment system can be analyzed. The factors unified under the stress umbrella will be shown to occur almost entirely on the penultimate or antepenultimate syllables of words. The latter case will be seen to only occur only if the final syllable begins with one of [k, tʃ, ŋ]⁷⁴, similar to the “weak” syllables of OM. These resulting analysis—developed independently here but following the one used in for OM in section 3.5—posits that all antepenultimate stress is due to underlyingly consonant-final forms, which receive penultimate stress, but also an epenthetic vowel ignored by stress assignment. Both penultimately and antepenultimately stressed words, then, will have primary stress assigned only to the penultimate underlying vowel.

The usual case for Betsimisaraka words is for emphasis to be heard on the penultimate syllable, as shown in the representative forms below.

⁷⁴ Or, very occasionally, [n], due either to influence of OM or to an alternation that will be analyzed as nasal dissimilation in 6.1.8. This exception will apply throughout this chapter.

Table 5.1 Betsimisaraka words exhibiting the usual, penultimate stress pattern. Note that the vowels [e] and [o], mentioned in Section 5.1 as being segmental markers of stress, only appear in the second-to-last syllable.

word	gloss	word	gloss
[mi'roŋku]	"close" (adj)	['ampi]	"enough"
[am'boni]	"above"	['aŋgi]	"father"
[ma'venti]	"big"	['fun̩ʃi]	"many"
[a'kun̩dʒu]	"banana"	['hena]	"meat"
[an'deha]	"go"	['andʒu]	"day"

Still, it is not at all uncommon for words, like those in Table 5.2 below, to be heard exhibiting antepenultimate stress. In a similar distribution to, but distinct pattern from, OM, these antepenultimately stressed words can only occur if the final syllable begins with one of three segments: either the affricate [tʃ], the oral stop [k], or the nasal stop [ŋ]⁷⁵. Another factor linking antepenultimately stressed words is that the final two, unstressed vowels are the same quality; pronunciations where the final vowel of antepenultimately stressed words can always be [a] are optional, as well be explained in 6.2.3.

⁷⁵These consonants can also be onsets to the final syllables of penultimately stressed words (note [mi'roŋku] and ['fun̩ʃi], in Table 5.1), so they alone cannot explain antepenultimately stressed words. It is true that both of these words also end with VNCV string instead of VCV, but recall that the previous chapter did prevent minimal pairs for penultimate and antepenultimate stress.

Table 5.2 Betsimisaraka words exhibiting the frequent, antepenultimate stress pattern. The words below are organized into columns depending on the onset of the final syllable, the only three allowed in antepenultimately stressed words.

word	gloss	word	gloss	word	gloss
'huditʃi	"skin"	'lavaka	"hole"	'tadiŋi	"ear"
'luatʃa	"surplus"	ma'diniki	"small"	mi'adaŋa	"slow"
'menaŋa	"shy"	mi'araka	"mix" (verb)	a'ŋaraŋa	"name"
'ohatʃa	"example"	mi'petʃaka	"sit"/"stay"	u'vianja	"when"
zavatʃa	"thing"	ta'ranaka	"descendent"	'oruŋu	"nose"

Aside from the limited inventory of consonants and vowels in their final syllables, antepenultimately stressed words also group together in their behavior when occurring as the first root of the compound. When the second root is consonant-initial, a normally antepenultimately stressed first root's final syllable will be entirely absent (see Table 6.6 in the next chapter). If the onset of the final syllable is nasal, though, the consonant will occur and the vowel will not. As can be seen in Table 5.3 below, though, the consonant onset of the final syllable of antepenultimately stressed words will always surface if the second root is vowel-initial.

Table 5.3: Compounds wherein (at least) the first root is typically antepenultimately stressed, and the second root is vowel-initial. In all cases, the final vowel heard when the first root is spoken by itself is not pronounced in the compound form. Exceptionally for this grammar, secondary stress is marked below, and (like in Malagasy orthography systems), a hyphen is employed to mark the morphological word boundary between the two roots.

Root 1	gloss	Root 2	gloss	Compound	gloss
'lavaka	"hole"	'oruŋu	"nose"	,lava'k-oruŋu	"nostril"
'huditʃi	"skin"	a'umbi	"cow"	,hudi'tʃ-a'umbi	"cow leather"
'tandʒuka	"horn"			,tandʒuk-a'umbi	"cow-horn"
'tungutʃu	"ear"	'oluŋu	"person"	,tungu'tʃ-oluŋu	"human ear"
'oruŋu	"nose"			,oru'ŋ-oluŋu	"human nose"
'helatʃa	"wing"	a'kohu	"chicken"	,orun-a'kohu	"chicken beak"
				,helatʃ-a'kohu	"chicken-wing"

Given just the data in Table 5.2 and 5.3, we have two possible hypotheses to explain the behavior of the first roots. Firstly, they could be exceptionally assigned antepenultimate stress, all end with one of three consonants and then a copy of the previous vowel by coincidence, and have that vowel deleted in compounds, perhaps due to hiatus resolution⁷⁶. More systematically, though, all of these factors can be explained by the underlying form of the word actually being closer to the one seen in the compounds in Table 5.3 above. They are thus underlyingly consonant-final and only underlying vowels are counted when assigning penultimate stress. This work selects the second analysis, in part because stress assignment can be interpreted

⁷⁶Indeed, deletion will be analyzed as one hiatus resolution strategy in 6.2.4, though it is not the case that the first vowel is systematically deleted.

as universally penultimate, but more importantly, the limited distribution of consonants and vowels can be interpreted, respectively, as due to word-final neutralization (Section 6.1.7) and epenthesis followed by optional harmony (6.2.3).

5.3 Stress Assignment and Hiatus Resolution

In Betsimisaraka underlying forms, hiatus is frequent, despite that on the surface, underlying vowel pairs are frequently resolved through diphthongization, coalescence, and deletion. The coalescence of /ai/ and /au/ to [ɛ] and [ɔ], as well as the creation of various rising and falling diphthongs, was presented in 4.1.2 to distinguish the surface inventory from the underlying one. Like the vowel epenthesis alternation in 5.2, though, words that have [ɛ], [ɔ], or diphthongs do not necessarily exhibit expected stress patterns. This will be analyzed in this chapter as due to stress being assigned to the underlyingly penultimate vowel, even if on the surface the only trace of that vowel is certain features of the *final* vowel.

Table 5.4: Surface stress in words featuring underlying hiatus. URs provided are deduced from the surface presence of lax-mid vowels or diphthongs, neither of which are elements of the phonemic inventory of Betsimisaraka. They are emboldened in the chart below. Superscript semivowels represent on- or off-glides off main vowels, not labialization or palatalization.

Penultimate Stress			Final Stress		
UR	SR	gloss	UR	SR	gloss
/araiki/	[a'ɾɛki]	one	/mahai/	[ma'hɛ]	be able to
/aiza/	['ɛza]	where	/zareu/	[za'rew]	3P
/dziabi/	['dzjabi]	all	/anau/	[a'nɔ]	2S
/tavuangi/	[ta'vʷaŋgi]	bottle	/akau/	[a'kɔ]	there ⁷⁷

Although URs are provided above for a fuller picture, focus first on the SRs. From just these two columns, one might conclude that the penultimately stressed words have stress assigned as normal, while those with final stress would be lexical exceptions. Such a hypothesis would have to somehow account for the presence of either diphthongs or lax-mid vowels in all of these final-stress forms. If we accept the URs offered, though, we can postulate a stress-assignment process that is assigned as if both hiatus-resolution-motivated diphthongization and coalescence had not occurred, perhaps because these processes derivationally follow stress assignment, as will be proposed in section 6.2.4.

Telltale markers of hiatus resolution like diphthongs and lax mid

⁷⁷ Like other Austronesian languages, Betsimisaraka has an extremely rich set of demonstrative pronouns denoting not just distance from the speaker (like English “here” vs. “there”), but also distance from the listener and visibility, along with other factors. [a'kɔ] appears to refer to a location distant from, but perceptible to, the speaker, and adjacent to the listener.

vowels are not always available to justify aberrant final stress. A similar tactic, though, can be employed to justify final stress as resulting from penultimate stress assignment, even in wordforms that do not contain these markers, as seen below.

Table 5.5: Active and imperative forms of verbs with vowel-final roots. If the final vowel of the root is [u] or [i], the imperative takes the /-a/ suffix and is pronounced with penultimate stress, but if the final vowel is [a], there is no possible pronunciation for *[aa] and the word is pronounced with final stress.

Penultimate-Stress Imperative			Final-Stress Imperative		
gloss	active	imperative	gloss	active	imperative
“blow”	mi'sut <u>ju</u>	su't <u>ʃ</u> ua	“see”	ma'hita	hi'ta
“murder”	ma'munu	mamu'nua	“allow”	magnam'bela	ambi'la
“dig”	maŋ'gadi	maŋga'dia	“hunt”	mi'haza	ha'za
“breed”	mi'umpi	um'pia	“cross”	miam'pita	ampi'ta
“believe”	'minu	mi'nua	“lift”	mi'bata	ba'ta

The analysis in the rest of the section will depend on the presence or absence of suffixes⁷⁸, as stress is assigned from the right word edge. In the left columns, active forms end in [u] or [i] while imperative forms all end with [a], following the vowel that had been the final segment of the active form. While both actives and imperatives have the expected penultimate stress, the stressed syllable in the imperative is one to the right of its position in the active

⁷⁸ The reader may recall the prefixes present mostly in the active forms as the SFP of OM (analyzed in 3.2.1) with the present-tense marker [m]. The morphology of imperatives is more complex, varying between verbs and between and among speakers, although representative forms are supplied. Some imperative forms in Table 5.5 can be interpreted as employing this prefix cluster (“murder,” “believe,” and “dig”), while others do not, or do so only optionally, in the imperative.

form. This is due to that additional vowel in the imperative, interpreted here as the imperative suffix /-a/.

In the right columns of Table 5.5, the pattern is similar, with a rightward shift of stress from active to imperative verb forms, but for these verbs, both the active and the imperative end in [a], and the imperative forms have final stress. Following from the hypothesis of an imperative suffix /-a/, though, this can be understood as yet another hiatus resolution strategy that occurs after stress assignment⁷⁹. The surface-final syllables of imperatives like [hi'ta] or [ambi'la], then, are underlyingly penultimate, but the ungrammatical *[ta.a] or *[la.a] strings are resolved by pronouncing just a single instance of the vowel⁸⁰.

In this section, then, we have seen how, if two vowels occur in sequence underlying at the right word boundary, the surface form of the word will have final stress. This is because stress will be assigned

79 One might well ask, then why the imperatives in the left columns do not *also* trigger deletion of the /-a/ suffix; after all, ['u.a] and ['i.a] also represent hiatus. Indeed, forms without the [a]s are certainly heard, but hiatus resolution is only *necessarily* triggered if the vowels are the same, if the unstressed vowel can become a glide, or if the vowels can coalesce into a lax mid vowel. ['ua] and ['ia]—along with, at least theoretically, ['ea] and ['oa]—are acceptable strings in carefully pronounced Betsimisaraka.

80 One intriguing, alternate interpretation is available, recalling that length was the most reliable acoustic correlate of stress: perhaps both [a]s *do* surface, and ['aa] strings were simply mistaken for instances of ['a]. Following this analysis, the forms for “hunt” would be [mi'haza] and ?[ha'zaa]. One might expect that ['zaa] from the passive to exhibit longer duration than the ['ha] in the active. Additionally, to the extent that pitch and intensity are reliable correlates of stress, one would expect these quantities to decrease over the duration of ['zaa]. Answering this question is additionally complicated by devoicing of final syllables, which may itself be interpreted as deletion.

to the penultimate vowel, which will merge with the final vowel. This merging can connote diphthongization, coalescence to a lax mid vowel, or deletion, but in any case will surface as a single vowel which displays the stress assigned to the underlyingly penultimate one.

5.4 Conclusion

Stress in Betsimisaraka can be understood as a bundle of acoustic markers and segmental features applied to the underlyingly penultimate vowel of a word. Increased vowel length is the most reliable, while higher pitch and amplitude are also factors.

Betsimisaraka also distinguishes itself from the better known Official Malagasy changes in vowel quality by allowing [o], but only in stressed syllables, while restricting [e] to stressed ones. Like OM, though, Betsimisaraka optionally reduces [a] to [ə] in unstressed syllables, as well as reducing unstressed high vowels to glides and deleting unstressed low vowels, both to resolve hiatus.

Like in OM, Betsimisaraka stress assignment can be said to be uniformly to the penultimate vowel as long as segmental alternations are allowed to alter the structure of words that have already had their stress assigned. In one such alternation, word-coda-triggered vowel epenthesis results in antepenultimate stress, although the identity of the epenthesized vowel is not always [a] as it is in OM. In the other set of alternations, the various processes triggered by hiatus resolution

merge two final vowels into one, which then exhibits the previously penultimate stress. This phonological grammar opts for a rule ordering with early stress assignment both for uniformity of penultimate stress assignment and to account for the segmental alternations discussed.

Chapter 6

SEGMENTAL ALTERNATIONS OF BETSIMISARAKA

Introduction and Overview

This chapter presents a comprehensive overview of segmental alternations of Betsimisaraka, showcasing the phonological differences of this understudied dialect of Malagasy. Some of the alternations are analogous to those described in Chapter 3 as occurring in OM and other dialects, but when presented in this chapter, all primary examples will be from Betsimisaraka. All alternations will be presented starting fresh with the data collected for this project, with no assumed knowledge of previous work in Malagasy languages. That said, this analysis is built upon analysis of previous analyses, especially the generative approaches in Dziwirek (1989), Paul (1996), and Albro (2005), and references to literature and to discussion thereof in Chapter 3 will occur throughout.

Each alternation will first be presented atheoretically, to the extent possible, with data showing representative forms. This will be followed by a generalization, and then a formalization of that generalization as a rewrite rule. Analysis will conclude with sample derivations, including relevant interactions with previously mentioned

rules. Finally, as the alternations are motivated, the distinctive feature set for the language will be continually defined. Section 6.1 presents alternations affecting consonants while 6.2 covers vowels; each section will conclude with a summary of the rules and their ordering; a final feature matrix for relevant segments will be presented toward the end of each chapter. Section 6.3 presents a summary of all rules from previous sections and their ordering.

6.1 Alternations affecting consonants

Betsimisaraka consonant alternations are similar to those of Official Malagasy: nasal assimilation, consonant cluster reduction, post-nasal fortition, lexically restricted voiceless consonant deletion, and the neutralization and deletion processes that occur to underlying codas. Additionally, aside from neutralization of underlying word-final consonants, the same rules proposed here could explain the alternations discussed in Chapter 3.

The only consonant alternation that is entirely unique to Betsimisaraka is nasal dissimilation, occurring when [ŋ] heads two subsequent syllables. This last process obviously could not have been said to occur in OM due to the latter's lack of an underlying velar nasal. And while /ŋ/ does exist in all non-OM forms of Malagasy, nasal dissimilation has never before been proposed. An analysis of this process concludes Section 6.1.

6.1.1 Nasal Assimilation

Consonant clusters wherein the first consonant is nasal⁸¹ are the only ones allowed in Betsimisaraka, the same pattern from OM, as described in Section 3.3. Furthermore, as is common across world languages, the cluster is always homorganic, i.e. with the two consonants pronounced at the same place of articulation. Examples of Betsimisaraka clusters at each place of articulation are provided below, although the best evidence of this distribution is the total lack of heterorganic clusters to be found in the language.

Table 6.1: Words containing homorganic consonant clusters in Betsimisaraka. Transcription here is more narrow than usual for this chapter to show the coarticulation found in dental and post-alveolar places of articulation.

labial	dental	alveolar	post-alveolar	velar
[amboni] “above”	[anðeha] “go”	[matandzaka] “strong”	[akundʒu] “banana”	[aŋgi] “father”
[ampi] “enough”	[mavent̪i] “big”	[tsi maɪnts̪i] “must”	[funt̪i] “many”	[miroŋku] “close” (adj)

Additionally, aside from the apparently pre-existing homorganic clusters above, the language can be observed to alter the place of a nasal that forms part of a cluster to assimilate to that of the following consonant. This assimilation occurs in polymorphemic words wherein a nasal-final morpheme is concatenated in front of a consonant-initial

⁸¹This distribution assumes that nasal-consonant pronunciations are nasal-oral clusters, not pre-nasalized stops. As decided in 4.2.1, this work is formalized using the former analysis, but as mentioned there, the alternations can be analyzed using pre-nasalized stops as well.

root. Alternations of this nature employing the nasal-final active verb prefix /maŋ-/⁸² can be seen in the Table 6.2 below.

Table 6.2: Passive and active forms of representative verbs for each place of articulation⁸³, with the invariably homorganic NC clusters (emboldened below) in the active forms. Narrow transcription is used below as well.

articulation	passive-3S	active-present	gloss
labial	babi'ani	ma m' babi	"carry" / "support"
dental	didi'ani	ma n' didi	"cut"
alveolar	dzu'ritini	ma n'dz urutſu	"pull"
velar	gadi'ani	ma ŋ'g adi	"dig"

The consonants at the beginning of the passive are also present in the active, but preceded by a segment sequence [mN-], wherein N is a nasal of the same place of articulation as the consonant. Even though the nasal is part of the prefix, then, its place or articulation can be seen to be determined by the onset of the root. Nor is this type of distribution limited to /mN-/.

The alternation can also as easily be heard using compounds, although the first root will have to be selected so that it is underlyingly nasal-final. Recall from section 5.2 that antepenultimately stressed

82 As in OM, presented in Section 3.2.1, this string is made up of the present-tense prefix /m-/ (as opposed to past /n-/ and future /h-/ plus the active "stem-forming prefix" (Dziwirek 1989) /aŋ-/. The bimorphemic nature of this "prefix" will only be mentioned in this chapter when relevant.

83 No verb roots (which is to say, no passives or imperatives) were found that began with [dʒ]. Although verbs were encountered with active forms beginning [mandʒ...], the passive (thus the root) always began /r/, with the surface [dʒ] the result of fortition, as seen in Section 6.1.3. Furthermore, we would not expect [tʃ] following /maŋ-/ because voiceless sounds are deleted following this prefix, as accounted for in Section 6.1.4.

words have consonant-final URs that have undergone vowel epenthesis. The SR ['orunu] ("nose")⁸⁴, then, has an underlying nasal coda, and when used in compounds, epenthesis will not be triggered, meaning if the second root is vowel initial, nasal assimilation will occur.

Table 6.3: Compounds using ['orunu] ("nose") as the first root and various animals as the second, providing nasal assimilation data for a root with underlying potential coda.

root 2	gloss	compound (animal's nose)
[pisu]	"cat"	orum'pisu
[tʃandʒaka]	"tenrec" ⁸⁵	orun'tʃandʒaka
[gisi]	"goose"	orunj'gisi

From the data in Tables 6.2 and 6.3, we can extrapolate a language-wide generalization whereby nasals surface at different places of articulation based on the following consonant. This generalization, which is true for OM as well, can be accounted for by an assimilation rule, in standard featural notation and prose below:

- 1) Consonant Cluster Place Assimilation (“CC Assimilation”)**
 $[-\text{syllabic}] \rightarrow [\alpha\text{place}] / [-\text{syllabic } \alpha\text{place}]$
 “When consonant clusters occur, the first consonant will adopt the place features of the second”

It would not affect the theory developed in this work—or the language as a whole—if the rule above used [+nasal] as its target feature rather than [-syllabic]. As was mentioned at the beginning of

⁸⁴ With alternate forms ['orü] or ['oru], with processes described in Section 6.1.6.

⁸⁵ A small hedgehog-like mammal found in Madagascar, of the family Tenrecidae.

the chapter, all CCs in Betsimisaraka, as in OM, have a nasal as their first consonant. The rule above employs [-syllabic], then, by theoretical reasoning: it is the [-syllabic] feature that the segments have in common that causes one to assimilate to the other, not the [+nasal] feature of the first segment. Due to the restrictions on CCs, it would be impossible to prove which feature is the crucial one for the target, and so the most general rule possible was chosen.

Knowing that the place features of final nasals are determined by the following consonant, the underlying place features of /məN-/ might be determined by examining this segment's behavior when the following root is vowel-initial.

Table 6.4: Passive and active forms for vowel-initial roots.

passive-3S	active-present	gloss
i'saini	ma'ŋisa	"count"
am'barani	maŋam'bara	"tell"
a'miani	maŋa'mia	"give"

In the cases above, where the passive (and thus the root) is vowel-initial, the nasal coda of the prefix is velar, suggesting that this is the underlying value of the nasal⁸⁶. Although data collected for this study produced no /u/- or /o/-initial root that takes /maŋ-/ in the active, there is no reason to suggest that the place of the nasal is due to the

⁸⁶ Unless all vowels are analyzed as marked for the feature [+dorsal], as Sagey (1986) suggests. This is not generally agreed upon though: Clements (1991) asserts that only back vowels are [+dorsal], in which case we would not expect [ma'ŋisa] above.

features of the following vowel. Knowing the UR of the active prefix and the nature of the assimilation rule, we can now show some derivations using this rule.

Table 6.5: Sample derivations of four verbs in the active present form, undergoing CC Assimilation.

gloss	“give”	“carry/support”	“cut”	“dig”
UR	/maŋ-amia/	/maŋ-babi/	/maŋ-didi/	/maŋ-gadi/
CC Assim.	---	/mam-babi/	/man-didi/	/maŋ-gadi/
SR	[maŋamia]	[mambabi]	[mandidi]	[maŋgadi]

CC Assimilation makes use of two features, although [αplace] is a stand-in for the place features that will be formalized in Section 6.1.3. The only feature to be defined by this rule, then, is [\pm syllabic], for which the vowels [i, e, a, o, u] are [+syllabic], and thus irrelevant to CC Assimilation, while all consonants (listed in Table 4.7) are [-syllabic], and will trigger and/or undergo CC Assimilation.

Using data from the active verb prefix and from compounding, this section offers an explanation for one restriction on CCs in Betsimisaraka, that the first syllable must be a nasal homorganic to the following consonant. The next section analyzes the result when, due to concatenation, a non-nasal appears as the first consonant in a cluster.

6.1.2 Consonant Cluster Reduction

In the previous subsection, underlyingly nasal-final words had

their codas preserved when used as the first root of a compound, owing to Betsimisaraka phonotactics allowing internal NC clusters. CC clusters (where the first segment is oral) are not found⁸⁷, but through compounding they do occur underlyingly. ['oruŋu] ("nose") was assumed to be underlyingly consonant-final due to its antepenultimate stress pattern, but recall from section 5.2 that onsets of the so-called weak, final syllables also include [tʃ] and [k]. Just like [ŋ], these onsets will occur, without their nuclei, in compounds where the second root is vowel-initial.

Table 6.6: Compounds wherein the first root is underlyingly consonant-final and the second is the vowel-initial ['oluŋu] ("person").

first root	gloss	compound with ['oluŋu]	gloss
'sarutʃu	"difficult"	saru'tʃoluŋu	"difficult person"
'oruŋu	"nose"	oru'ŋoluŋu	"human nose"
'tapaka	"cut"	tapa'koluŋu	"cut someone"

The alternations above are consistent with an epenthesis account for the antepenultimately stress. The URs of the first roots above are identical to their SRs without their final vowels, and epenthesis was not triggered in the compounds because the first root was not word-final⁸⁸,

⁸⁷ As in OM, they are permitted in certain loanwords, e.g. [frantsai] "French" or [krismasi] "Christmas."

⁸⁸ As mentioned in 5.2, an analysis wherein the final vowels of the first root are part of the UR, and are subsequently deleted due to hiatus resolution seems possible given just the data here, but will not account, for example, for the alternations in Table 6.5, e.g. why the [u] at the end of [oluŋu] ("nose") should not appear in [oluŋgisi] ("goose's nose").

and the underlying coda is pronounced as the onset of the second root's first syllable. In this way, a syllable structure that only allows nasal codas, and only word-internally, is maintained.

In compounds where the second root is *consonant-initial* and the first has antepenultimate stress when on its own, the first root is missing both the nucleus *and* onset of its final syllable.

Table 6.7: Compounds consisting of an antepenultimately stressed (underlyingly consonant-final) word followed by a consonant-initial word.

root 1	gloss	root 2	gloss	compound	gloss
'sarutʃu	“difficult”	'koru	“course”	saru'koru	“hard class”
		'tani	“earth”	saru'tani	“difficult soil”
'tapaka	“cut”	'tsihi	“mat”	tapa'tsihi	“half-mat”
		'pati	“pasta”	tapa'pati	“small noodles”

Given that the final vowels of both ['sarutʃu] and ['tapaka] result from epenthesis, the underlying forms of these words are, respectively, /sarutʃ/ and /tapak/. When the second word of the compound is consonant-initial, the result is an ill-formed cluster wherein the first consonant is non-nasal. As occurs in the compounds in Table 6.7, Betsimisaraka dispenses with these ungrammatical clusters by a structure preservation rule deleting the first consonant of a cluster if it is not a nasal⁸⁹.

89 Elicitation to produce the forms seen in Table 6.7 was inspired by Keenan & Razafimamony's (1995) analysis of the same compounds in OM, but their analysis depended an underlying “weak syllables” interpretation of the data, with the vowel or the whole syllable deleting. Under an underlying-root-final-consonant analysis, CC Reduction in OM occurs exactly as here proposed for Betsimisaraka.

2) Consonant Cluster Reduction (“CC Reduction”)

[-syllabic -nasal] → Ø / [-syllabic]

“Oral consonants are deleted when occurring before other consonants.”

There is no strict need to order CC Reduction with respect to CC Assimilation. As mentioned above, the latter rule employs [-syllabic], rather than [+nasal] as its target, but even if CC Assimilation were to precede CC Reduction, the unnecessarily assimilated consonant would be deleted. For example, /tapak-tsihi/ (“cut-mat”) would have a medial form ?/tapat-tsihi/, but CC Reduction would reduce /ts/ to simply [ts], producing the attested [tapatsihi]. For purely aesthetic reasons, one might wish to avoid an ordering that mandates consonants destined to be deleted must first change their place. Still, this rule ordering produces the desired grammar, and more importantly will be mandated due to an alternation described in 6.1.5, so is the ordering in the sample derivations below of the two rules formalized so far.

Table 6.8: Derivation of compounds that undergo CC reduction or CC assimilation.

glosses	“nose” + “cat”	“difficult” + “person”	“cut” + “pasta”
UR	/orun-j-pisu/	/sarutʃ-oluŋu/	/tapak-pati/
CC Assimilation	/orum-pisu/	---	/tapap-pati/
CC Reduction	---	---	/tapa-pati/
SR	[orumpisu]	[sarutʃolunu]	[tapapati]

The features employed in CC Reduction were [\pm syllabic], which

has already been defined, and [\pm nasal]. The rule targets the non-nasal consonants, thus leaving out the vowels and the nasal consonants /m, n, η . Although the rule would not be affected by the existence of nasals vowels, none have been proposed for Betsimisaraka, leaving the [+syllabic +nasal] quadrant empty.

Sections 6.1.1 and 6.1.2 have described alternations that affect the first consonant of a cluster. In the two subsections to follow, we will investigate changes that affect the second one.

6.1.3. Fortition

In Section 3.4.1, Official Malagasy was shown to neutralize consonants to a restricted set of non-continuants post-nasally. In the following section, this process will be examined in-depth using Betsimisaraka examples, beginning by isolating the nasal-final forms that trigger the alternation. Alternations will be demonstrated, followed by an articulatory motivation for the phenomenon. The analysis concludes with a rule-based analysis of the interaction, and a derivation of examples, interacting with the nasal-assimilation rule above. Finally, fortition data contributes to a nearly full formalization of the distinctive features of consonants.

As defined in the previous section, consonant clusters in Betsimisaraka are restricted to those wherein the first consonant is a nasal, homorganic to the following consonant, but another look at Table

6.1, repeated below, suggests that there are restrictions on the second consonant as well:

Table 6.9: Words containing the only acceptable CCs in Betsimisaraka, wherein the second consonant is a stop or an affricate.

labial	dental	alveolar	post-alveolar	velar
[amboni] “above”	[andeha] “go”	[matandzaka] “strong”	[akundžu] “banana”	[aŋgi] “father”
[ampi] “enough”	[maventi] “big”	[tsi maintsi] “must”	[funtʃi] “many”	[mironjku] “close” (adj)

Every non-continuant can be found post-nasally, producing the ten grammatical CC clusters, as emboldened in Table 6.9. The five fricatives ([f, v, s, z, h]) and the two liquids ([l, r]) are never the second consonant in a cluster. As the remainder of this subsection will show, if morphological processes place one of the seven continuants in this position, it undergoes fortition, resulting in one of the ten CCs allowed.

To trigger fortition, examples below employ the genitive morpheme⁹⁰, identified in isolation or in slow speech as either [na] or [ni]⁹¹. Its function can often be glossed in English as “of,” such that [tʃanu na rabe] is “house of Rabe” (or “Rabe's house”). For our purposes here, though, the genitive morpheme is important because it

⁹⁰Section 6.1.1 made use both of compounding and verbal prefixing, but, as will be analyzed in 6.1.4, the latter cannot reliably be used to observe fortition as it triggers deletion of following voiceless consonants.

⁹¹Written <ny>, [ni] is also often a definite article in Betsimisaraka, though when used in this way, it may represent a borrowing from Official Malagasy. Native Betsmisaraka typically prefers simply [i] as definite article, itself an article in OM, albeit with a much more limited distribution.

is typically pronounced simply [n]. In the example above, “Rabe’s house” would usually be pronounced [tʃanundʒabe]. Not only is the above [na] reduced to [n], but the onset of [rabe] has been fortified from the liquid [r] to the affricate [dʒ]. Equipped with a morpheme that is typically just [n], we can observe fortition by creating compounds wherein the second word begins with a liquid or fricative, as below.

Table 6.10: Compounds incorporating the genitive [n], which triggers fortition.

root 1	gloss	root 2	gloss	compound	gloss
'vua	“seed”	'lalana	“road”	vuan'dalana	“souvenir” ⁹²
		'hazu	“tree”	vuan'kazu	“fruit”
'kilo	“kilogram”	'vari	“rice”	kilom'bari	“a kilo of rice”
		zari'ko	“bean”	kilon <u>d</u> zari'ko	“a kilo of beans”

Before analyzing the alternations above, let us add to our data set alterations involving a slightly different type of compound, both to see that the process is not limited to interaction with the genitive morpheme [n] and to capture the behavior of the liquids and fricatives not covered in Table 6.10. To that end, consider compounds wherein the first root is underlyingly nasal-final, as in Table 6.5 above.

The Betsimisaraka word for “eat” or “consume,” ['homanya]⁹³. As for /oluŋu/ in Section 6.1.2, based on the antepenultimate stress

⁹²A <voandalana> is, specifically, a present you get for someone who did not get to go on the trip with you. These are an expected part of Betsimisaraka culture.

⁹³Often pronounced ['homã] or ['homa]

pattern, we can assume an underlyingly nasal-final root: /homan/. As before, the underlying coda can be seen to trigger fortition when the following root begins with a liquid or fricative.

Table 6.11: Compounds wherein the nasal-final first root /homaN/ triggers fortition in the initial consonant of the second root.

root 2	gloss	compound	gloss
['lɔku]	“fish”	[homan' dɔ ku]	“eat fish”
['ranu]	“water”	[homan' dʒ anu]	“drink water”
['sira]	“salt”	[homan' tʃ ira]	“eat salt”
['hena]	“meat”	[homan' k ena]	“eat meat”

Exhibiting the same fortition-triggering pattern as ['homanya], wherein the underlying nasal coda can be deleted or epenthesized, is ['enijŋi] (“six”)⁹⁴.

Table 6.12: Compounds wherein the nasal-final first root /enij/ triggers fortition in the initial consonant of the second root.

root 2	gloss	compound	gloss
['fulu]	“ten”	[enim' p ulu]	“sixty”
['zatu]	“hundred”	[enin' d zatu]	“six hundred”

From the data in Tables 6.10-6.12, we can tell what happens to each non-continuant when a preceding nasal triggers fortition, summarized in Table 6.13.

⁹⁴ Also ['eni] or ['enĩ]

Table 6.13: Summary of fortition of non-continuant segments in Betsimisaraka

Alternation	Non-fortition example	Fortition Example
f→p	f ulu (“ten”)	enim' p ulu (“sixty”)
v→b	v ari (“rice”)	kilom' b ari (“kilogram of rice’)
s→ts	s ira (“salt”)	homan' ts ira (“eat-salt”)
z→dʒ	z ariko (“bean”)	kilon dʒ ari'ko (“kg of beans”)
l→d	l iku (“fish”)	homan' d iku (“eat-fish”)
r→dʒ	r anu (“water”)	homan' dʒ anu (“drink-water”)
h→k	h ena (“meat”)	homan' k ena (“eat-meat”)

Because one never sees such clusters as *[nl] or *[ŋh], alternations like those seen in Table 6.13 serve to repair what would otherwise constitute an ill-formed structure. This repair maintains voicing—and place of articulation to the extent possible—replacing the liquid or fricative with a representative stop or affricate. This still allows for considerable variety in the alternations: knowing that a given continuant is a liquid does not determine whether its alternant will be a stop ([d] for /l/) or an affricate ([dʒ] for /r/). Fricatives are similarly unpredictable: consider that /v/ surfaces as the stop [b] while /z/ becomes the affricate [dʒ]. Certain minimal changes in place occur as well, such as between glottal /h/ and velar [k], alveolar /l/ and dental [d], or labiodental /f/ and bilabial [p]. All of these irregularities, though, can be accounted for by fortition producing the closest (in a sense that will be formalized at the end of this subsection) stop or affricate, given the phonemic inventory.

By identifying this process as one of post-nasal fortition, it can be viewed as phonetically motivated, easing the transition between nasal stop and continuant. In fact, as pointed out in Ohala (1981), a related process has occurred diachronically in English, wherein some words have developed an excrescent oral stop between nasal stop and oral fricative. To wit, the municipality of Hampshire was originally associated with “ham,” not the meaningless *“hamp.” Even in modern English, a common mistake is to spell “hamster” as *“hampster.” Unlike Malagasy dialects, the English pattern lacks nasal assimilation in these cases (*“hantshire,” *“hantster”), meaning the stop inserted matches the non-assimilated nasal [m] for place.

Suppose, at some point in its development, Betsimisaraka too inserted a stop to ease this transition. As we saw in Section 6.1.1, nasals *do* assimilate in compounds, meaning the excrescent stop will match both surrounding segments for place. In the same way that English speakers require careful articulation to pronounce “hamster” as ['hæmstə̯] rather than ['hæmpstə̯], it is difficult to articulate *[homan**n'sira**] in place of the grammatical [homan**n'tsira**] (“eat salt”), or *[kilonzari'ko] instead of [kilondzari'ko] (“kilogram of beans”). In these cases, the stop-fricative clusters /ts/ and /dz/ are realized as phonemic affricates [ts] and [dz]. The oral stop portion of this affricate, then, acts as a compromise segment between the preceding nasal stop and the oral fricative.

Before moving on to cases where the result of fortition is a stop, let us examine the alternation between [r]/[dʒ]. Recall from section 4.2 that [dʒ] is written <dr>, and that the pronunciation of this segment varies from [dʒ] to [dʒr]. With the latter pronunciation, then, [r]/[dʒr] alternation is not entirely surprising. Turning again to the familiar facts of English phonology, underlying /dʒ/ sequences are pronounced [dʒr], such that “drive” is pronounced [dʒraɪv], not *[dʒraɪv]. Although there is no universal prohibition of stop-rhotic sequences—French, for example, allows both [tʃ] and [dʒ], e.g. [tʃε̃] (“train”) and [dʒε̃] (“drain”)—this English example evinces that it is not unusual for a stop to affricate before a rhotic.

Consider a Betsimisaraka UR like /homan-ranu/ (“drink water”), which will exhibit fortition of /r/. Phonetically speaking, in transitioning from the articulation of [n], where no airflow exits the mouth, to the oral approximant [r], where air flows more freely out of the mouth but where airflow to the nasal cavity is shut off, one is not surprised to find a period of time wherein the velum has shut the nasal cavity but the articulators have not yet allowed airflow past the alveolar place of articulation. Since both segments are voiced, what is heard is the voiced coronal stop [d], which yields the complex cluster [ndr]. Betsimisaraka's nearest phoneme to the cluster [dr] is the one pronounced [dʒr] or [dʒ], and written [dʒ] in this grammar, so the surface form is [homan-dʒanu].

Let us turn now to the fricative/stop alternations, [f]/[p], [v]/[b], and [h]/[k]. From a purely phonetic standpoint, one might have expected, based on the reasoning above, in place of [kilom-bary] (“kilo of rice”) or [kilon]-kena] (“kilo of meat”), to hear instead *[kilombvary] and *[kilonkhena], with an oral stop inserted between nasal stop and underlying oral fricative. These configurations, though, violate Betsimisaraka syllable structure: the inter-syllable clusters [nts], [ndʒ], [ntʃ] and [ndʒ] are allowed as they consist of a nasal (the only acceptable coda) and an affricate, the most complex acceptable native onset. Betsimisaraka, though, does not contain the unusual affricates [k̪h] and [b̪v], and they must thus be simplified before pronunciation.

From this intermediary /...VmbvV.../ string, then, *[...Vm̪b̪vV...] is unacceptable for containing an unattested affricate, and returning to *[...Vm̪vV...] would defeat the purpose of excrescence: to ease the transition between nasal stop and oral fricative. Certainly, syllable structure could be maintained, and excrescence avoided, by converting /...Vm̪vV.../ directly to [...Vm̪vV...]. This result, though, does not contain as many lexical signifiers: by changing the fricatives to stops, but maintaining voicing and (to the extent possible) place⁹⁵, there is more surface evidence of the features of the underlying lexical item. This is

⁹⁵ Some place of articulation *could* be represented on the nasal due to opaque nasal assimilation; after all, this [m] is underlyingly the genitive /n/. In this scenario, though, not only is voicing not represented, but the distinction between /h/- and vowel-initial roots is lost, as both of which will be preceded by [ŋ]. Still, the next subsection concerns analysis of a process that follows exactly this procedure.

not a universal motivation in Betsimisaraka (see 6.1.4 for a case where lexical signifiers are dropped), but one possible inspiration in this case. The only choice, then, is to delete the fricative and leave only the stop.

This leaves only the case of [l]/[d] alternation, which is not an uncommon fortition process cross-linguistically⁹⁶. For the purposes of phonological analysis, it is actually fortunate to have such a non-transparent case occurring in the language. Consider that there is little articulatory difference between the attested [kilon-tsira] (“kilo of salt”) and pronouncing the same compound without fortition: *[kilon-sira]. In both pronunciations, the vocal folds will cease vibrating, the velum will shut off the nasal cavity, and the tongue will shift slightly down from its coronal position; before this last action occurs, though, these articulations are tantamount to pronouncing the voiceless oral stop [t].

Compare this to the fortition alternation between [l] and [d]. In this case, it is easy for a phonetician to distinguish the correct [homan-dɔku] (“eat fish”) from the unattested *[homan-lɔku]. Including /l/ to [d] alternations as part of the same fortition process for other liquids and fricatives defines the alternation more unambiguously as phonological, rather than in the domain of articulatory phonetics: the output from fortition is structure preserving in that it is restricted by the phonemic inventory. Also important for a phonological analysis is that [l] and [d] clearly share articulatory features for voicing and place. The stop

⁹⁶Mielke (2008) describes post-nasal fortition from /l/ to [d] in both Swahili and Bukusu Masaba.

features of the nasal [n] pervade into the pronunciation of the oral liquid [l], resulting in a voiced oral stop, [d]. As in the labial and dorsal alternations, there is no attested affricate *[dʒ̪] to present both the stop features of the [d] and the approximant features of [l] (as there was for [d] and [r] in [dʒ̪]), so the surface pronunciation of the cluster is [d].

Given the preceding phonetic motivation for all aspects of fortition, one possible phonological analysis of these alternations is to follow this motivation step-for-step. What follows is an articulatorily motivated treatment of what we have been calling fortition, although since the excrescent stop is inserted and the original segment joined with it or deleted, it is not functionally fortition of the original segment. After this analysis, a more purely phonological analysis will be offered, including the necessary distinctive features.

The three-step, articulatorily motivated process begins with Excrescent Stop Insertion, which inserts the compromise segment between the two. For instance, excrescent stops are inserted at the morpheme boundary in /homan-ranu/ (“drink water”) and /homam-vari/ (“eat rice”), resulting in the intermediate forms /homan-dranu/ and /homam-bvari/. For the former to surface grammatically, the only change is for the /dr/ cluster to coalesce into the affricate [dʒ̪], whereas for the latter, with no [*bv] affricate, the [v] must be deleted before pronunciation. Excrescent Stop Insertion thus feeds two more rules that clean up the ungrammatical sequences. Affricate Creation renders

complex segments from consonant clusters that are attested affricates; note that it will also create the affricate [dʒ], which will then phonetically be rendered as either [dʒ] or [dʒr] in narrow transcription—essentially, it looks for CCs that can be coalesced into complex single segments, and does so. Finally, CCC simplification deletes the remaining fricative and liquid segments /l/, /f/, and /v/. The rules are lettered, not numbered, as they will not be part of the final rule ordering for this grammar.

A) Excessive Stop Insertion

$\emptyset \rightarrow C_{[-\text{continuant } \alpha\text{place}]} / C_C C_{[\alpha\text{place}]}$

“A stop⁹⁷ is inserted between two consonants that matches the place of articulation of the second.”

B) Affricate Creation

$CC_{[+\text{cor. -lat}]} \rightarrow \widehat{CC}_{[+\text{cor. -lat}]}$

“If a consonant cluster's second segment is a non-lateral coronal⁹⁸, the cluster is simplified into a single affricate.”

C) Complex Consonant Cluster Simplification

$C \rightarrow \emptyset / CC_{-}$

“Delete the final segment of complex consonant clusters..”

⁹⁷The feature used here for “stop” is [-continuant], a class that usually includes affricates. Stops and affricates *can* be further distinguished with the feature [\pm delayed release], but in this articulatorily motivated analysis, Betsimisaraka affricates are the result of excrescence between stops and fricatives, so do not require underlying features. Alternatively, if affricates were interpreted instead as [+strident -continuant] segments—as originally proposed by Kager et al. (1952)—and the Betsimisaraka rhotic were analyzed as [+strident], then all three strident continuants /r, s, z/ become strident stops [dʒ, ts, dz] and the non-strident continuants /f, v, l, h/ become non-strident stops [p, b, d, k]. This would account for fortition, but, in the Sections 6.1.6, strident segments undergo an alternation that /r/ does not, but that /f, v/ do. To employ structure preservation to get around these issues (labial affricates surface as stops, post-alveolar stops surface as affricates) would be tantamount to the analysis proposed here.

⁹⁸[+coronal] segments are alveolars, dentals, post-alveolars, and retroflexes, the last of which do not occur in Malagasy. [\pm lateral] is the feature used in this rule to distinguish the two liquids, but will not be needed in the final analysis in Rule (3).

Table 6.14: Sample derivations of four compounds undergoing the suite of rules that together follow a phonetic derivation of Betsimisaraka fortition. The row labeled “UR” actually contains input from other rules, such as consonant place assimilation from Section 6.1.1. Exceptionally for this work, the rules are referred to by their letter instead of their name.

Gloss	“drink water”	“eat rice”	“eat fish”	“eat salt”
UR	/homan-ranu/	/homam-vari/	/homan-l̪ku/	/homan-sira/
A	/homan-dranu/	/homam-bvari/	/homan-dl̪ku/	/homan-tsira/
B	/homan- <u>d</u> ranu/	---	---	/homan- <u>t</u> sira/
C	---	/homam-bari/	/homan-d̪ku/	---
SR	[homandʒanu]	[homambari]	[homandɔku]	[homantsira]

The table above shows these rules interacting in order, with Ex crescent Stop Insertion (A) feeding Affricate Creation (B) and CCC Simplification (C), and Affricate Creation (B) bleeding CCC Simplification (C). In following the alternations above, note that /h/ and /f/ would behave like /v/, and /z/ would follow the derivation of /s/.

Keep in mind that, while this set of rules represents the phonetic process outlined in the section up until this point, a simpler, though more abstract, analysis is available. All of the interaction in rules A, B, and C can be compressed into the single rule 3.

3) Post-Consonantal Fortition

[-syllabic] → [-cont.] / [-syllabic] _

“Following a consonant, a second consonant will become the corresponding stop or affricate.”

Given that we have simplified the ruleset from three to one, the derivation chart will be simpler, although we are now depending on the grammar to determine that the non-continuant equivalent of /l/ is

[d] :"]and /r/ is [dʒ]. This will be accounted for in the distinctive feature analysis that follows the derivation.

Table 6.15: Sample derivations of four compounds undergoing fortition. As in Table 6.14, the row labeled “UR” actually contains input from select other rules.

Gloss	“eat” “water”	“eat” “rice”	“eat” “fish”	“eat” “salt”
UR	/homan-ranu/	/homam-vari/	/homan-lɔku/	/homan-sira/
Fortition	/homan-d̪ranu/	/homam-bari/	/homan-dɔku/	/homan-t̪sira/
SR	[homandʒanu]	[homambari]	[homandɔku]	[homantsira]

The derivation in 6.14 can be taken as an analysis of the phonetic origin of this process. This grammar, though, opts for derivation as shown in Table 6.15, wherein a single rule accomplishes the fortition, although as noted above, our features will have to show, for example, that [d] is actually the [-continuant] equivalent of [+continuant] [l]. Implementing this, though, will be suspended until the end of this subsection, so that this rule can be observed interacting with previously formulated rules.

Up until now, we have been considering cases where the consonant that has undergone fortition is the second consonant in a cluster. This subsection began by presenting Betsimisaraka's inventory of possible CCs (Table 6.9), and explored the processes by which ungrammatical clusters were made to fit that inventory. To this end, we employed antepenultimately stressed words wherein the underlying coda was nasal, which was preserved in the SRs of compounds. This

leaves the question of what happens to a word ending in one of the grammatical non-nasal codas, either /tʃ/ or /k/, is the first root of a compound.

This is to ask, when a compound has, at the intersection of the roots, underlying /VC₁-C₂V/, where C₁ is oral and C₂ is continuant: will C₂ undergo fortition? As we are employing an ordered-rule paradigm, there are two possible scenarios. In the first, CC reduction precedes Fortition, and thus bleeds it of an environment: the second consonant will remain a continuant. In the second, Fortition precedes CC reduction, and the continuant opaquely undergoes fortition. The latter turns out to be the case for Betsimisaraka.

Table 6.16: Compounds wherein the first root has antepenultimate stress and the second is continuant-consonant initial. Compound glosses are somewhat self-explanatory, e.g. a [tapa'damba] is a smaller piece of cloth, and [saru'dʒanu] is hard-to-find water.

root1	gloss	root2	gloss	compound
'tapaka	“cut”	'lamba	“cloth”	tapa'damba
		'zinga	“water dipper”	tapa'dzinga
		'fe	“leg”	tapa'pe
'sarutʃa	“difficult”	'ranu	“water”	saru'dʒanu
		'heŋatʃa	“shame”	saru'keŋatʃa
		'vadi	“spouse”	saru'badi

As concluded in 6.1.2, if a compound is constructed of an antepenultimately stressed word and a consonant-initial one, the weak final syllable of the former will be dropped in the compound. This was

analyzed as structure preserving: the first root's underlying coda elides because it cannot attach to a syllable. From Table 6.16, though, we can now tell that the CC survives long enough to trigger fortition in the following segment. Some sample derivations are offered below.

Table 6.17: Sample derivations of compounds undergoing both fortition and CC Reduction.

Gloss	"cut" "person"	"cut" "mat"	"cut" "cloth"
UR	/tapak-oluŋu/	/tapak-tsihi/	/tapak-lamba/
Fortition	---	/tapak-tsihi/	/tapak-damba/
CC Reduction	---	/tapa-tsihi/	/tapa-damba/
SR	[tapak'oluŋu]	[tapa'tsihi]	[tapa'damba]

Note in Table 6.17 that, if CC Reduction preceded Fortition, /tapak-lamba/ would surface not as [tapa'damba] but *[tapa'lamba]. In the correct ordering shown in Table 6.17, then, CC Reduction counterbleeds Fortition.

Because these interactions are opaque, a synchronic, phonetically motivated account is no longer open to us. With NC clusters, fortition was hypothesized to be, at some level of representation, due to easing transition from a nasal non-continuant to an oral continuant with a compromise oral non-continuant segment. Nonetheless, this rule could still have an intuitive articulatory component. The three possible weak-syllable onsets, [ŋ], [k], and [tʃ], are all non-continuant, so fortition can still be seen as a remnant of

coarticulation of this feature on the following consonant. With the segment itself deleted, triggering fortition of the following consonant at least preserves the elided consonant's [-continuant] feature.

Given the data from fortition, we are now ready to calibrate the necessary distinctive features for Betsimisaraka. To begin, let us first review the Table 4.7 (renumbered and adjusted below), the consonant chart from Section 4.2.

Table 6.18 Phonemic consonants of Betsimisaraka and their articulations.
The abstract bilabial fricatives are not included below, with discussion of them postponed to Sections 6.1.5 and 6.1.7

	Bilab.	Labiodent.	Dental	Alveolar	Post-Alv.	Velar	Glottal
Stops	p b		t d			k g	
Affricates				ts dz	tʃ dʒ		
Fricatives		f v		s z			h
Nasals	m		n			ŋ	
Lateral				l			
Rhotic				r			

To narrow the many acoustic features above down to the necessary distinctive features for Betsimisaraka, all distinctions unused by phonological processes will be eliminated. The first distinction to eliminate is that between fricatives and liquids (lateral and rhotic), creating the class of continuants with the distinguishing characteristic inability to follow other consonants. Consequently, stops and affricates can be collapsed into the class of non-continuants, since there is no process that distinguishes them. Finally, certain places of articulation

can be linked as well: labiodental fricatives alternate with bilabial stops, so labials need not be distinguished by exact place of articulation⁹⁹. Similarly, glottal [h] undergoes fortition to become [k], so a unified dorsal category is required. Finally, [\pm voice] will be needed so continuants can alter with the proper non-continuant, as this feature is always maintained in fortition. Integrating the [\pm nasal] features proposed in Section 6.1, we are left with the chart below.

Table 6.19 [+syllabic] segments of Betsimisaraka and their distinctive features. Segments that are [+voice] are emboldened while unemboldened segments are [-voice]; nasals are phonetically voiced, so they are emboldened below as well, although there is no phonological function of this polarity.

		+labial	-labial			
			-dorsal			+dorsal
+continuant	-nasal	f v	I	s z	r	h
-continuant		p b	t d	ts dz	tʃ dʒ	k g
+nasal		m	n			ŋ

Almost all necessary distinctions are made above, but the [-labial -dorsal] class needs further distinction. These dentals, alveolars, and post-alveolars form the class of coronals, although, since they are already defined as [-labial -dorsal], a [\pm coronal] feature will not be necessary for Betsimisaraka. This grammar distinguishes them by the usual method, employing the [\pm anterior] and [\pm distributed] features. Anterior segments, pronounced toward the front of the coronal area,

99 As mentioned above, alternations in future subsections *will* require features to distinguish labials from bilabials, but as far as fortition is concerned, no distinction is necessary.

are dentals and alveolars, leaving post-alveolars and the alveolar rhotic [r] as [-anterior]. Distributed segments are pronounced with the point of contact (or near-contact) of the tongue distributed over a larger area, and include dentals and post-alveolars, but not alveolars¹⁰⁰. Knowing this, the coronals can be entirely distinguished as in the table below, with labial and dorsal segments undefined for these two features. /n/ is not included in Table 6.20 because it need not be defined for these features. In discussing nasal dissimilation, Table 6.1 offered a close transcription showing both a dental and a post-alveolar nasal, but these are simply two allophones of /n/ with no evidence for potential phonemic status. [-dorsal -labial] /n/, then, has two allophones that are not instances of neutralization, [+distributed -anterior] (post-alveolar) [ɳ] and [+distributed +anterior] (dental) [ɳ̥].

Table 6.20 Coronals ([-dorsal -labial] segments) of Betsimisaraka and their distinctive features. As above, segments that are [+voice] are emboldened while unemboldened segments are [-voice]. All segments not included are taken as undefined for the features [±anterior] and [±distributed].

	+distributed	-distributed	+distributed
	+anterior		-anterior
+continuant	l	s z	r
-continuant	t d	ts dz	tʃ dʒ

Necessary consonant features are now almost entirely defined,

¹⁰⁰There is no alternation that links [tʃ] and [t] to the exclusion of [ts], even though the first two are part of the natural class [+distributed], from which the last is excluded. This feature, then, is used just to distinguish two separate places of articulation. The same cannot be said of the feature [+anterior], which will prove in 6.1.7 to describe a natural class in Betsimisaraka.

although subsections 6.1.5 and 6.1.6 will show that one more feature will be required.

6.1.4. Lexically Triggered Voiceless Consonant Deletion

So far, we have seen how Betsimisaraka consonant clusters are limited to those wherein the first consonant is a nasal and the second is a non-continuant. This subsection covers cases where, following a single widely used prefix, the distribution is even more limited. This is formalized as a morphophonological process particular to the prefix.

In Section 6.1.1, nasal assimilation was observed through the addition of the present and active verbal prefixes /m-aŋ/ to the beginning of verb roots. To some extent, that technique can also be used to demonstrate fortition, as shown below.

Table 6.21: Imperative and active forms of verbs that undergo fortition when the active verbal prefix /maŋ-/ is attached. (Two forms exhibit stress-conditioned high/mid vowel alternation that will be presented in Section 6.2.1. The stress “shift” is caused by the addition of prefixes, as previously described in Table 5.5.)

Imperative	Active	Gloss
vu'lea	mam' b uli	“grow”
ra'honju	man' dʒ ahuju	“tell off”
Iani'a	mand a 'nia	“spend”

Above, the onsets of the imperative forms are the underlying onsets, and their alternants in the active forms have undergone fortition. Of the seven continuants, though, only those that are voiced

alternate following /manj-/. A different pattern emerges when /manj-/ is prefixed to roots whose imperative form—and thus the root itself—has a voiceless onset. Examples of these alternations as they occur in Betsimisaraka are offered in Table 6.23.

Table 6.22: Imperative and active forms of verbs wherein the onset is absent in the active form, although the nasal does preserve its place of articulation. No Betsimisaraka verbs with a [tʃ]-initial imperative were found that also employed /manj-/ to form the active, though OM [tʃan'ga] / [ma'nanga] ("make appear") was familiar to some speakers. Several unrelated vowel and consonant alternations occur among the data below as well, but the relevant deletions and nasal assimilations are emboldened.

imperative	active	gloss
fa'fa	ma' m afa	"wipe"
pu'puha	ma' m upuku	"hit"
sa'sa	ma' n asa	"wash"
tu'hiza	ma' n uhi	"tie together"
tsi'tsemo	ma' n etsina	"plug" (as in a drain)
ki'kera	ma' ŋ ekitʃa	"bite"
hu'rua	ma' ŋ uru	"burn"

The examples above show a pattern analogous to the one in OM (analyzed in 3.4.1) and to Indonesian and other Austronesian languages. The imperative (and thus the root) begins with a voiceless obstruent, which is absent following the nasal in the active. The place features of the voiceless obstruent, however, do not entirely disappear, being preserved in the place features of the prefix-final nasal. This means that [f] and [p] both alternate with [m], [k] and [h] with [n], and all the coronals ([s], [t], [ts], and [tʃ]) with [n].

A rule that deletes voiceless consonants following all nasals or all will not suffice. The second row of data in Table 6.9, which shows all possible CCs for Betsimisaraka, contains five different NCs where the second consonant is voiceless. Nor can the process be restricted to the morpheme boundary: Section 6.1.3. shows how a voiceless continuant undergoes fortition to produce a voiceless non-continuant, for example h→k in [hena] (“meat”) vs. [homanjkena] (“eat meat”).

It is not a general phonological rule that causes this change, then, but a unique characteristic of the prefix. This was identified in Section 3.4.1 as due to a peculiarity of the OM prefix /man-/ , and the only difference in Betsimisaraka is that the underlying form of the prefix is instead /maŋ-/; when this process occurs in Indonesian and other Austronesian languages, it is also restricted to a specific suffix.

To formalize the rule, recall that, though we have been referring to the active present verbal prefix as unanalyzed /maŋ-/ , the /aŋ-/ is the SFP marking the active while /m-/ indicates present tense. The environment for the rule, then, is merely “following the prefix /aŋ-/.”

4) Lexically-Limited Voiceless Consonant Deletion

[-voice] → Ø/ /aŋ-/ ¹⁰¹

“Following the SFP /aŋ-/ , voiceless consonants are deleted.”

¹⁰¹Although some students may be introduced to phonological rules containing segments, a written segment is actually a bundle of features, all of which are many more than is necessary for a rule. The rules in this grammar, then, always employ the absolute minimum number of features; the presence of IPA symbols in this formulation, then, exceptionally indicate a specific morpheme.

“Austronesian nasal substitution” constitutes literally a textbook¹⁰² example of opacity: in Betsimisaraka, the prefix-final nasal /ŋ/, is adopts the place features of segments that do not surface. In an ordered-rule paradigm, this can only be accounted for by place assimilation of the nasal counterbleeding deletion, and thus preceding it. If deletion were allowed to precede, thus bleed, assimilation, we would see surface forms such as *[maŋafa] for /maŋ-fafa/ (“wipe”) and *[maŋasa] for /maŋ-sasa/ (“sweep”).

Because the consonant is deleted, it will not affect the grammar if the consonant undergoes fortition first, producing a medial form for /maŋ-fafa/ (“wipe” active present) of ?/mam-pafa/ before the /p/ is deleted. This grammar opts for the ordering below to prevent this sort of unnecessary change, but only for aesthetic reasons.

Table 6.23: Sample derivations of active present verb forms undergoing several of the rules discussed in preceding sections. The ordering of the first two rules is essential, although the third rule could be placed at any point in the ordering.

gloss	“wipe”	“spend”	“cut”
UR	/maŋ-fafa/	/maŋ-lania/	/maŋ-didi/
CC Assimilation	/mam-fafa/	/man-lania/	/man-didi/
Lexically Limited Voiceless C Deletion	/mam-afa/	---	---
Post-C Fortition	---	/man-dania/	---
SR	[mamafa]	[mandania]	[mandidi]

This rule only makes use of the feature [±voice], which was

102For example as a problem in Halle and Clements (1983, p 125).

already introduced in the previous subsection to ensure fortition maintained this feature when altering [±continuant]. It is thus not necessary to add any features to the analysis in this subsection.

6.1.5. Variant Distribution of Root-Initial [v]

There are some token exceptions to Lexically Limited Voiceless Consonant Deletion, both in OM and in Betsimisaraka. The latter has one where a complex voiceless onset /tʃ/ is allowed to follow /maŋ-/, and another where a [k] that is the onset of the imperative appears as [g] following /maŋ-/. These will be treated simply as irregular verbs.

Research for this study found, like OM, the behavior of root-initial [v] was less consistent than that of other voiced sounds. For example, although the alternation between [vu'lea] (“plant” imperative) and [mam'buli] (“plant” active present) was used as an example of fortition in Table 6.21, it is not difficult to find cases of [v] alternating with Ø as if the former were a voiceless obstruent, as shown in Table 6.25 below.

Table 6.24: Imperative and active forms of verbs wherein the [v] onset of the imperative does not appear in the active.

imperative	active present	gloss
vun'dzea	ma'mundzi	“save”
van'gia	ma'mangi	“visit”
va'kia	ma'maki	“split” or “read”
vu'nua	ma'munu	“kill” or “murder”

It would not be unusual if the conjugation of certain lexical items were not restricted by phonological rules: like every language, Betsimisaraka has irregular verbs. The forms above, though, cannot be put aside so easily: of the only five [v]-initial imperative verb forms elicited with a corresponding active form employing /maŋ-/, only **one** underwent fortition following it.

A two-phoneme solution is presented in Albro (2005)'s analysis of OM: phoneme #1 with [v] and [b] as its allophones, acts just as we would expect /v/ to: like other voiced segments, its deletion is not triggered by the /maŋ-/ prefix, and the nasal coda on the prefix triggers its fortition, so Phoneme #1 will be analyzed as /v/.

As for phoneme #2, with [v] and Ø as its allophones, in 6.1.4, we observed that voiced segments are preserved following /maŋ-/ while voiceless ones are deleted, so underlying features of phoneme #2 might be said to include a specification for [-voice]. It would then be difficult to account for, though, why it surfaces as voiced [v] word-intially as well as following other verb prefixes. Betsimisaraka, like OM, has a prefix /maha-/ which, when attached to the root, yields the causative form of the verb. The causative form of the imperative [van'gia] ("visit") surfaces not as *[mahafan̩gi], but as [mahavan̩gi] (approximately "make (someone) visit"). Betsimisaraka has no intervocalic voicing, and /maha-/ does not cause /f/ or any other consonant to voice, so the onset of the roots in Table 6.25 phoneme

cannot be analyzed as voiceless.

Following Albro, this second phoneme can be analyzed as abstractly a bilabial fricative, so either /ɸ/ or /β/, which will eventually surface as labiodental. Although the former is tempting, since it is deleted after /aŋ-/, we have just shown using the prefix /maha-/ that this will not work. If it is /β/, though, it will already be voiced after /maha-/, and though it is exceptional in being a voiced segment deleted after /aŋ-/, at least segments deleting in that position are not unusual¹⁰³. One feature that can distinguish bilabial and labiodental fricatives is [±labiodental], but, as Albro points out, [±strident] also works, with [f, v] as [+strident] and [ɸ, β] [-strident]. The [+strident] class will prove important in the following subsection, so further discussion of features is suspended until then.

Although the data in 6.24 are not instances of /ɸ/, we cannot preclude that this segment (which will be shown to occur word-finally, and thus to be a phoneme of Betsimisaraka, in Section 6.1.7) from occurring word-initially: it would be deleted after /aŋ-/ (either for being voiceless or being a non-strident labial), undergo fortition to the only

¹⁰³Given the data in this subsection, a lexical class analysis could also account for the alternations without calling upon any non-surface phonemes. In this system, [mambuli] is still the active form of the root /vole/, but [mamaki]'s active is also /v/-initial: /vaki/, not /βaki/. The different behaviors of root-initial /v/ in the same environment are not because the former was actually /β/, but rather because the morphemes belong to separate lexical classes, one in which /v/ is deleted after /aŋ-/, and one in which it undergoes fortition as expected. In 6.1.7, though, we will again see seemingly a single labial fricative exhibiting two distinct behaviors, this time word-finally; while this again could be hypothesized as due to another set of lexical classes, this work opts to follow Albro (2005) in establishing a single pair of abstract phonemes to explain both discrepancies.

available voiceless labial, [p], and elsewhere surface as [f] as bilabial fricatives are not included in the surface inventory. This is to say that there is no way to tell whether a given root-initial [f] is /f/ or /ɸ/.

In the rule below, it is the specific features of bilabial fricatives that target them for deletion. Bilabial fricatives can be defined as [-strident] for Betsimisaraka, because, although they are [+continuant] as well, only continuants are defined for stridency and thus by seeing that a segment has a defined [\pm strident] value, one can tell it is a continuant. The only other defined [-strident] segment is /h/, and while we could target to [-strident +labial], /h/ is deleted after /aŋ-/. We have previously assumed that this was due to its voicelessness, but feature that triggers deletion of /h/ in this position cannot be said certainty. /h/ can be included in the target for the rule defined below with no affect on the language and one fewer feature defining the target.

5) Lexically-Limited Non-Strident Fricative Deletion “Bilabial Fricative Deletion”

[-strident] → Ø/ /aŋ-/ ¹⁰⁴

“Following the SFP /aŋ-/, non-strident fricatives are deleted.”

Adding abstract segments to the inventory is not a strategy here adopted lightly. The only crucial function of Rule 5 is to delete a segment with the feature matrix [+labial -strident] that does not appear on the surface. This adds two complications to the grammar: a new rule and a new feature. The rule, though, is necessary to account

¹⁰⁴As previously, the use of IPA symbols here indicates a specific lexical entry.

for a distribution like in Table 6.24, one which does not follow from the rules already posited, and adding [\pm strident] to our feature set not only gives this rule something to target, but also will be necessary in analyzing the alternation in the next subsection.

To be clear, bilabial fricatives are never heard in Betsimisaraka. This section has shown that they are deleted following /məŋ-/ and Section 6.1.7 will demonstrate how they neutralize to [tʃ] word-finally. All that is left for this subsection, then, is a rule that yields more salient, labiodental fricatives from all other cases of underlying bilabial ones. They become labiodental, phonetically speaking, due to the greater perceptibility of that type of fricative, though phonologically, because the surface inventory lacks [β].

6) **Bilabial Fricative Clean-Up**

[+labial +continuant] \rightarrow [+strident]
“All labial fricatives surface as strident.”

Bilabial Fricative Deletion must precede Bilabial Fricative Clean-Up: the latter counterbleeds the former. There is no need to order the rules that delete two different types of consonants (with some overlap) following /məŋ-/. Bilabial Fricative Deletion, though, must precede Fortition, because the latter counterbleeds the former. If not, Fortition would output a bilabial stop from /β/, which, as a non-continuant, would not be a target for Clean-Up.

Table 6.25: Sample derivation of [v]-initial roots, with one that surfaces faithfully and one that is underlyingly /β/. “UR” contains output from non-interacting rules

gloss	“grow”		“break”	
	imp.	act.	imp.	act.
UR	/vule-a/	/mam-vuli/	/βaki-a/	/mam-βaki/
Bilabial Fricative Deletion	---	---	---	/mam-aki/
Bilabial Clean-Up	---	---	/vaki-a/	---
Post-C Fortition	---	/mam-buli/	---	---
SR	[vu'lea]	[mam'buli]	[va'kia]	[ma'maki]

6.1.6. Coda Deletion and Vowel Epenthesis

In Section 6.1.2, we saw how deletion of the first consonant in a CC maintains Betsimisaraka phonotactics, specifically the requirement that only nasals can serve as codas. Though morphological concatenations produce CCs with non-nasal first segments, CC Reduction deletes those segments. Thus we have seen what happens when C occurs in the environment _C: if it is nasal, it assimilates, but if not, it deletes. In this and the following section, we will examine what happens to a C in the environment _#. In Betsimisaraka, certain consonants are deleted in this position (as will be described in this subsection), while others neutralize to a restricted set of consonants and undergo vowel epenthesis (the following subsection, 6.1.7). The analyses in these subsections build upon analyses of OM reviewed in Chapter 3, with references to the relevant authors.

The distribution wherein a morpheme-final coda consonant appears before a suffix, but is entirely absent when unsuffixed, is the same as we saw for OM in section 3.4.1

Table 6.26: Active and Passive forms of verbs wherein the fricative that occurs just before the suffix in the passive form is entirely absent in the active. URs (with potential codas justified by suffixed passive forms) are provided to show deletion in the active forms. Forms for “continue” and “wrap” also undergo voiceless consonant deletion after /aŋ/, “wrap” exhibits stress-conditioned [u]/[o] alternation.

Active Present (Present+SFP+Root)		Passive (unspecified) (Root+Passive+Unspec.)		Gloss
SR	UR	SR	UR	
mi'tsindzu	m-i-tsindzuv	tsin'dzu <u>v</u> inj <u>a</u>	tsindzuv-i-ŋ	“perceive from afar”
mi'duka	m-i-dukaf	du'kaf <u>i</u> j <u>a</u>	du'kaf-i-ŋ	“flatter”
ma'nuhi	m-aŋ-tuhiz	tu'hiz <u>a</u> j <u>a</u>	tu'hiz-a-ŋ	“continue”
ma'moju	m-aŋ-foŋos	fu'ŋos <u>i</u> j <u>a</u>	foŋos-i-ŋ	“wrap”

Given just the SRs above, one may assume, with Rakotofiringa (1981) and Keenan & Polinsky (1998), that the fricatives are added in the suffixed forms, somewhat unpredictably, and the URs are incorrect in including the fricative. Section 3.4.1 presented this evidence against the model provided by Dziwirek (1989) and Albro (2005) using URs like those in the table above. The latter model will be chosen here for the same reasons as in OM: because it is possible to predict the active based on the passive, making for a regular system.

Given just the data above, one might conclude that consonants delete at the end of the word just as they do at the end of a syllable,

and might set about formulating a rule where $_C$ and $_C$ are the same environment. Throughout this chapter, though, we have seen that other consonants, namely [tʃ], [ŋ], and [k], *can* occur underlyingly word-finally, and rather than being deleted, they undergo vowel-epenthesis, resulting in the “weak syllables” at the end of penultimately stressed words. In the following subsection, we will see that these three consonants are the result of neutralization from a larger set, so the only consonants that delete are those highlighted in Table 6.27.

Given our original consonant inventory, it seems as if all fricatives but /h/ delete; in addition, in the following subsection bilabial fricatives will be shown to neutralize, not elide. A feature is thus required that selects /f, v, s, z/ to the exclusion of /ɸ, β, h/. This requirement is why this grammar follows Albro (2005) in selecting [\pm strident] over [\pm labiodental] as the feature to mark labiodentals: the class of fricatives that delete are [+strident] and the class of those that do not are [-strident].

7) **Strident Coda Deletion (“Coda Deletion”)**

[+strident] $\rightarrow \emptyset / _C$

“When strident segments occur in coda position, they are deleted.”

Before presenting some sample derivations for this rule, it will be best to formalize the alternative method of eliminating potential word-final codas. In Section 5.2, we saw that word-final consonants (now

revealed to be just non-stridents) undergo vowel-epenthesis, resulting in antepenultimate stress. With an ordered rule paradigm, consonants that do not undergo epenthesis do not have to be specified: it will be enough to have Coda Deletion precede, and thus bleed, vowel epenthesis¹⁰⁵. The environment that triggers epenthesis, then, can consist of simply a consonant at the right word boundary. For now, it will be enough to say that the vowel epenthesized is [+syllabic], adding in its place features later, in a process analyzed in 6.2.3.

8) Coda-Eliminating Vowel Epenthesis (“Vowel Epenthesis”)
 $\emptyset \rightarrow [+syllabic] / [-syllabic] _ \#$
“If a consonant occurs at the right word boundary, a vowel is epenthesized following it.”

With both the deletion and epenthesis rules formalized, we can now observe their interactions. In Table 6.27 below, the unspecified-person suffix /-ŋ/ undergoes epenthesis, allowing us to easily see the difference between two coda resolution strategies by comparing derivations of the active and passive forms of verbs. If the order of these rules reversed, epenthesis would bleed coda deletion, producing ungrammatical *[mi'dukafa] and *[ma'monŋusa].

¹⁰⁵Intuitively, a rule wherein non-stridents trigger epenthesis and all remaining consonants are deleted is also possible. The analysis of neutralization offered in the next subsection, though, will be shown to operate more efficiently if the current order is used.

Table 6.27: Sample derivations of active and passive forms of select verbs, all of which resolve codas with epenthesis in the active, and with deletion in the passive. Although previous derivation tables occasionally employed “URs” to which some rules had been applied (which was always noted), URs below are true underlying forms, which facilitates comparison of forms using the same root. A separate row (“other rules”) is provided for the application of nasal assimilation (6.1.1), voiceless consonant deletion (6.1.4), and [o]/[u] alternation (6.2.1).

Gloss	“flatter”		“wrap”	
	active	passive	active	passive
UR	/m-i-dukaf/	/dukaf-i-ŋ/	/m-aŋ-fonɔs/	/fonɔs-i-ŋ/
“other rules”	---	---	/m-am-oŋus/	/fuŋos-i-ŋ/
Coda Deletion	/m-i-duka/	---	/m-am-oŋu/	---
V Epenthesis	---	/dukaf-i-ŋa/	---	/fuŋos-i-ŋa/
SR	[miduka]	[dukafiŋa]	[mamоŋu]	[fuŋosiŋa]

Now that stridency has been demonstrated as a distinctive feature of Betsimisaraka, it should be integrated into our feature matrix. In this analysis, [\pm strident] is a feature with polarity only for fricatives, so only [-syllabic +continuant] segments are included below. Furthermore, the liquids /l, r/ will also be undefined for [\pm strident] here, so will not be included.

Table 6.28: Distinctive feature chart for [-syllabic +continuant] segments that are defined for [\pm strident]. /l/ and /r/ are continuant consonants, but are not defined for this feature.

+labial		-labial	
-dorsal			+dorsal
-strident	+strident	-strident	
-voice	∅	f	s
+voice	β	v	z

6.1.7. Coda Neutralization

In the previous chapter, Section 5.2 presented examples of words that appeared to be underlyingly /k/-, /ŋ/-, or /tʃ/-final, but which, were they allowed to surface faithfully, would violate a prohibition on codas. Betsimisaraka offers two solutions: either a vowel is epenthesized so that the consonant fit the language's syllables structure, or, if it were nasal, the consonant could optionally be deleted, in a process that will be analyzed in Section 6.2.3. For the purposes of the previous chapter, it was enough to assume that only those three consonants can be found in coda position underlyingly. This was based on only those consonants¹⁰⁶ being found as the onset of the final syllable of an antepenultimately stressed word, which was the telltale sign of underlying potential codas.

In this section, as for other dialects in 3.6.4, we will see that segments [tʃ, k, ŋ] are actually the surface forms of a larger group of consonants appearing in underlying word-coda position. As in Section 6.1.6, these non-surface codas will be revealed by comparing active and passive verb forms, because the latter are suffixed and the former are not. This will allow us to compare the form of an underlying coda that undergoes coda-triggered epenthesis with one that is not word-final because it is suffixed. We begin by observing the alternations of the various nasals that neutralize to [ŋ]

¹⁰⁶ Along with [n], which occurs due to nasal dissimilation, described in the following subsection

Table 6.29: Active and Present forms of verbs wherein the active ends with the weak syllable [ŋa]. URs, with codas justified by the pre-suffix consonant of the passive forms, are provided to show neutralization in the active forms. Due to the specificity of examples needed to show these alternations, the data above evince more alternations than just final-consonant neutralization. This includes not only the deletion of voiceless consonants after /mŋ/, as seen in Section 6.4, but also stress-triggered [u]/[o] and [i]/[e] alternations, to be explored in 6.2.1. Additionally, the passive of “approach,” [ha'tonjna] has undergone nasal dissimilation, which will be described in the following subsection.

Active Present (Present+SFP+Root)		Passive (unspecified) (Root+Passive+Unspec.)		gloss
SR	UR	SR	UR	
ma'ŋatunja	m-aŋ-hatoŋ	ha'tonjina	hatonj-i-ŋ	“approach”
mi'tanja	m-i-tan	ta'naja	tan-a-ŋ	“hold”
mi'tandʒinja	m-i-tandrem	tan'dʒemaja	tandrem-a-ŋ	“pay attention”

Notice that the active form for “hold,” [mi'tanja], is not actually stressed penultimately: it evinces [n]/[ŋ] alternation with the passive [ta'naja] because it has a one-syllable root, and because, as noted in 3.2.1, SFPs do not host stress¹⁰⁷. This example is crucial, as well, in showing that these neutralizations are not caused, for example, by the consonants' position between two unstressed vowels, but by the underlying word-coda position.

Coda-preventing vowel epenthesis occurs in all the active forms above, following the rule formalized in the previous subsection. For all of the active present forms, the last syllable is one [ŋa], of the three possible weak syllables described in section 5.2 as indicative of vowel-

¹⁰⁷ Either by fiat or because they are appended, like epenthetic vowels, after stress is assigned

epenthesis. In the passive forms above, stress has shifted rightward vis-à-vis its position on the root. The distribution is such that (except for cases of nasal dissimilation) only one nasal, [ŋ] occurs before epenthetic [a], while all three nasals occur before suffixes.

Aside from the pre-epenthetic-vowel nasal being velar instead of dental, then, the distribution in Betsimisaraka follows the pattern seen in seen for OM in section 3.6.3. Rajaonarimanana (2004) and Keenan & Polinsky (1998) accounted for the OM distribution as a nasal (in that case [n]) alternating unpredictably with [n] or [m] in the suffixed forms. Albro (2005), looking at the same data, interprets instead underlying /m/ and /n/ neutralizing to [n] before epenthetic vowels. As in the previous subsection, then, this grammar follows Albro, as his model offers both accuracy and predictability. Like deletion of stridents (as shown in Table 6.27), neutralization counterbleeds, thus must precede, epenthesis: if the nasal's position at the right word boundary final is what triggers this neutralization, this process could never occur post-epenthesis.

Let us wait, though, to formalize this rule until we have seen the alternations that result in Betsimisaraka's other possible weak syllables: [ka] and [tʃa]. The onset consonants of these “weak syllables” were treated in section 5.2 as the other possible codas, though just as was seen for other dialects in 3.6.3, that limited set of consonants is neutralized from a larger set. The table below compares

the full complement of suffix-preserved underlying codas with the limited set in the final syllable of the active. When examining the table, note in the form for “lick” that a segment that surfaces as [f] before a suffix neutralizes before an epenthetic vowel, rather than being deleted, as in the previous subsection. Fortunately, thanks to Albro (2005) we have already introduced a feature to distinguish instances of surface-identical labials that behave like different phonemes. Since we have seen that strident segments delete, the labial that neutralizes must then be the abstract non-strident labial /ɸ/.

Table 6.30: Active and Present forms of verbs wherein the active ends with one of the weak syllables [ka] and [tʃa]. URs, with codas justified by the passive's pre-suffix consonant, are provided to show neutralization in the active forms.

Active Present (Present+SFP+Root)		Passive (unspecified) (Root+Passive+Unspec.)		Gloss
SR	UR	SR	UR	
man'daka	m-aŋ-dak	da'kaŋa	dak-a-ŋ	“kick”
ma'ŋaraka	m-aŋ-arah	a'rahiŋa	arah-i-ŋ	“follow”
mi'lela tʃ a	m-i-lelaɸ	li'lafin̩a	lelaɸ -i-ŋ	“lick”
man'dʒakut tʃ a	m-aŋ-rakoɸ	ra'kofaŋa	rakof-a-ŋ	“cover”
ma'fantat tʃ a	m-a-fantar	fan'tariŋa	fantar-i-ŋ	“know”
ma'ŋavut tʃ a	m-aŋ-avot	a'votan̩a	avot-a-ŋ	“redeem”

The distributions here show a wider distribution of consonants before the suffix and a much more limited set before the epenthetic vowel. It should be noted here that *only* suffixes preserve the underlying vowel: when compounding before a vowel-initial second

root, the neutralized version of the consonant will always be heard. Although we will continue to refer to the word boundary, it should be remembered that this rule is triggered by the phonological word boundary, lower on the prosodic hierarchy than the morphological word boundary. This should not be *too* surprising: although earlier rules (assimilation, fortition, reduction) operated on CCs spanned that phonological word boundary, we have also seen that stress is assigned leftward from the right phonological—not morphological—word boundary, meaning that if the second root has an odd number of syllables, stress will not alternate throughout the morphological word.

This alternations in Table 6.31 have some aspects in common with the fortition examples we saw in 6.1.3. More specifically, both [h] and [k] undergo fortition and word-final neutralization to [k], and [r] alternates a post-alveolar affricate in both processes, although for fortition it was [dʒ] whereas here it is [tʃ]. If the distribution above followed the fortition pattern, though, we would expect [f]/[p] and [t]/[t] alternations, yielding ungrammatical weak syllables *[*-ta*] and *[*-pa*]. While the fricative [f] does alternate with the less sonorant¹⁰⁸ affricate [tʃ], it also undergoes a place change. So while there are certain similar alternations between fortition and coda-position neutralization, they are not the same process.

¹⁰⁸The stop [t]'s alternation with [tʃ] means its alternant is higher on the sonority scale, suggesting phonetic lenition, rather than fortition, although since the feature system here does not distinguish affricates and stops, this will not present any difficulties

This does not mean the featural commonalities in the alternations in Table 6.30 cannot be qualified. Above, we see both dorsal obstruents [k] and [h] before the passive suffix, but only the stop [k] before the epenthetic vowel. The distribution of consonants that alternates with [tʃ] is slightly more difficult to qualify, consisting of two coronals (the continuant [r] and non-continuant [t]) and the labial continuant [f].

Even with these complications, though, the distribution in Betsimisaraka is simpler than the one Albro (2005) analyzed for OM. Albro presented the form [mi'lelaka] for “lick” (a translation that was verified by OM-speaking informants for this study), cognate to, but distinct from, the Betsimisaraka [mi'lela~~tʃ~~a]. Albro also had a form, [man'drakutra]/[ra'kufana], more closely cognate to the forms for “cover” above. So when, in OM, a consonant surfaced as [f] before a suffix, it was less predictable, whether it would surface as [k] or [tʃ] before an epenthetic vowel. Albro, finding underlying abstract segments preferable to lexical exceptions, proposed that [f]/[k] alternations were underlyingly bilabial fricatives¹⁰⁹ while [f]/[tʃ]

¹⁰⁹In Albro's system, [f]/[tʃ] alternations are due to underlyingly abstract /ɸ/ and [f]/[k] alternations were due to /f/. While no [f]/[k] alternations were found for this project, it does have [f]/Ø alternations in common with OM, as shown in the previous subsection. Albro analyzed these last as being underlyingly abstract /β/, the voiced bilabial fricative, so the underlying form of “follow” is not /dukaf/ but /dukaβ/. For Albro, then, /β/ is deleted whereas /f/ neutralizes to [tʃ], while (as replicated here in 6.1.5) deleting after /an/ rather than undergoing fortition. Albro thus needed a three-way UR divide, to describe why [f] that can be heard before suffixes unpredictably either neutralizes to [k], neutralizes to [tʃ], or deletes. This three-way distribution suggested a three-way UR split, so the two-way distribution in Betsimisaraka suggests a two-way split.

alternations were underlyingly labiodental. This distinction will be employed in Betsimisaraka, as previously noted, to distinguish the phonemes /ɸ/ with allophones [f] and [tʃ] from /f/, with allophones [f] and Ø.

Before extrapolating rules from the alternations, let us take the preliminary step of summarizing the relevant features.

Table 6.31: Summary of the changes that occur when consonants are in morpheme-final position, as shown in the active forms, although these forms have undergone epenthesis. The URs of these consonants are revealed in the suffixed, passive forms, although noted that the ultimate UR in “lick” is not /f/, but /ɸ/. Still, the rule that converts bilabial fricatives to labiodentals occurs early enough to precede neutralization.

gloss	passive	active	change	feature change	
				target	output
“approach”	ha'toŋina	ma'ŋatunja	ŋ → ɳ	[+nasal]	[+dorsal]
“hold”	ta'nanya	mi'tanya	n → ɳ		
“pay attention”	tan'dʒemanya	mi'tandʒinja	m → ɳ		
“kick”	da'kaŋa	man'daka	k → k	(+dorsal) -nasal	[-continuant]
“follow”	a'rahiŋa	ma'ŋaraka	h → k		
“lick”	li'lafiŋa	mi'lelatʃa	ɸ → tʃ	(-dorsal) -nasal	[-continuant] -voice -anterior -labial
“know”	fan'tariŋa	ma'fantatʃa	r → tʃ		
“redeem”	a'votanya	ma'ŋavutʃa	t → tʃ		

While, as noted above, the commonality here is not merely of fortition, it is the case that the only feature that all coda-position consonants have in common is [-continuant]. To generalize, then, one can assume that the fortition exhibited clearly among dorsal segments

is common to all segments found in coda position, while nasals and non-dorsals are also changing place and voicing. This will result in a suite of three rules to provide the limited distribution of coda consonants, but unlike the preliminary three-rule analysis of fortition, the three rules below will remain in the final grammar. As mentioned previously, the exact word boundary referred to in the following three rules is that of the phonological word, as fortition occurs mid-compound.

9) Coda Fortition

$[-\text{syllabic}] \rightarrow [-\text{continuant}] / _ \#$

“Consonants at the right word boundary surface as non-continuant.”

For this rule, /k/, /t/, and all nasals will have the same output as input, although in the cases of /k/, this is the desired surface result. It will also produce the attested SR for one segment that undergoes fortition: /h/ to [k]. For /r/ and /ɸ/, it will output /dʒ/ and /p/, neither of which will surface.

10) Nasal Coda Backing (“Nasal Backing”)

$[+\text{nasal}] \rightarrow [+ \text{dorsal}] / _ \#$

“Nasal consonants at the right word boundary surface as dorsal.”

Only nasal codas recede to dorsal position, so this rule targets them. It will correctly yield the velar nasal [ŋ] from all nasals, ignoring all other codas.

11) Non-dorsal Coda Neutralization (“Coda Neutralization”)

[-dorsal] → [-labial -anterior -voice] / __ #

“Non-dorsal segments at the right word boundary surface as voiceless non-anterior non-labials.”

As long as this rule follows Nasal Backing, all nasals will be dorsal and thus unaffected by this rule. If not, /m/ and /n/ would become a voiceless version of their post-alveolar [ŋ] allophone. This is not an irreconcilable result: nasal backing would then move its place of articulation the rest of the way back, and a clean-up rule might revoice unattested voiceless nasals. Still, this requires nasals to change their place of articulation twice, and another rule, so Nasal Coda Backing will be said to precede Coda Neutralization so it can bleed it.

The [-labial -anterior] place feature specifies post-alveolar pronunciation, directing the /p/ and /t/ output from Coda Fortition to be pronounced [tʃ]. [-voice] is included to convert medial /dʒ/, which is not a valid word-final consonant, to its final form [tʃ].

With these three rules formalized, we are ready to present derivations that show how a large inventory of consonants in coda position result in the three found as the onset of weak syllables. The URs and (to some extent) features used below replicate Albro's (2005) for OM, but with two differences. In OM, nasals neutralize to coronal, just like non-dorsals, resulting in different rules¹¹⁰. And while the

¹¹⁰To be clear, Albro (2005)'s analysis was entirely within the framework of Optimality Theory. In discussing his work in this section, I mean an ordered-rule approach to his analysis. A less complete, but fully rule-based, analysis of this OM data is presented in Dziwirek (1989)'s work on Malagasy Morphophonology.

derivations for “follow” and “redeem” would follow the same steps in OM, as mentioned above, the underlying coda for “lick” is /k/, not /f/, in OM. In elicitation for this study, [f]/[tʃ] alternations were unattested.

Table 6.32: Sample derivations for the suite of rules that neutralize, then eliminate, codas through vowel epenthesis. As previously, the “UR” row actually contains the output of certain rules, such as those regulating [o]/[u] and [e]/[i] alternations to be described in Section 6.2.1

gloss	“pay attention”	“follow”	“lick”	“redeem”
UR	/m-i-tandʒim/	/m-aŋ-arah/	/m-i-lelaɸ/	/m-aŋ-avut/
Coda Fortition	/mitandʒim/	/maŋarak/	/milelap/	/maŋavut/
Nasal Backing	/mitandʒinj/	---	---	---
Coda Neutralization	---	---	/milelatʃ/	/maŋavutʃ/
Vowel Epenthesis	/mitandʒinjɑ/	/maŋaraka/	/milelatʃɑ/	/maŋavutʃɑ/
SR	[mitandʒinjɑ]	[maŋaraka]	[milelatʃɑ]	[maŋavutʃɑ]

The pattern we have seen above can be summarized as: in each of three groups of consonants: nasals, dorsals, and oral non-dorsals, only one segment is allowed word-finally: the furthest back, least voiced, least continuant one, which is [ŋ], [k], and [tʃ] for each group, respectively—even if all nasals are equally non-continuant and (de)voiced. With no suffixes added, these segments undergo epenthesis, and may thus surface as antepenultimately stressed, although monosyllabic roots will still be penultimately stressed. In corresponding suffixed forms, that same segment will alternate with a wider range of consonants that share certain key features with one of

[k, tʃ, ɳ]. This is to say that, by knowing a verb's suffixed (passive and other) form, you can predict its unaffixed (active and others) form, but not vice versa. Some language change (in OM, from Ferrand 1909) following from this fact was analyzed in Table 3.40 in Section 3.6.3.

Recall that, in 6.1.5, eliminating /β/ from surfacing was accomplished through context-free destridentization that produces [v]s from all /β/s not already eliminated. Now that we know there are root-final bilabial fricatives, this Clean-Up rule will have to follow Strident Coda Deletion (proposed in the previous subsection), or else it would feed it. /mi-lelaɸ/ would yield *[mi'lela] rather than the attested [mi'lelatʃa]. The counterfeeding between Strident Coda Deletion and Bilabial Fricatives Clean-Up accounts for the distinction between the [f]/∅ and [f]/[tʃ] alternations, as shown in the derivation table below.

Table 6.33: Parallel derivation of two labial-fricative-final roots (as determined by the pre-suffix segment in the passive) in the active and passive verb form. This derivation shows the crucial ordering of Strident Coda Deletion, which targets /f/ but not /ɸ/, and Bilabial Clean-Up, which destridentizes /ɸ/ to /f/.

gloss	“flatter”		“lick”	
	psv.	act.	psv.	act.
UR	/dukaf-i-ɳ/	/mi-dukaf/	/lilaɸ-i-ɳ/	/m-i-lelaɸ/
Strident Coda Deletion	---	/mi-duka/	---	---
Bilabial Clean-Up	---	---	/lilaf-i-ɳ/	/m-i-lelaf/
Coda Fortition + Neutralization	---	---	---	/m-i-lelatʃ/
Vowel Epenthesis	/dukaf-i-ɳa/	---	/lilaf-i-ɳa/	/m-i-lelatʃa/
SR	[du'kafiɳa]	[mi'duka]	[li'lafiɳa]	[mi'lelatʃa]

The previous two subsections have shown how rules inspired by Malagasy coda prohibition operate on all underlying codas, including the abstract segment /∅/, to either delete them or replace them with one of three possible syllables. These rules have accounted for 13 of the 22 consonants of Betsimisaraka. The remaining eight do not appear to be allowed to occur as codas, even underlyingly. This is to say that, adding suffixes will never reveal the presence of certain consonants: /p, b, β¹¹¹, d, dz, ts, l, tʃ, dʒ/. These might be lexical gaps, gaps in the data, or result from morpheme structure constraints not allowing, for example, labial stops or affricates as word codas.

We now know the reason behind three of the weak syllable onsets of Betsimisaraka, [tʃ, k, η]: The majority of the language's consonants can appear underlyingly in coda position, even though the language does not allow codas on the surface. Strident fricatives are deleted, and the rest are neutralized to one of the weak-syllable onsets. Crucially, stress is then assigned to the penultimate syllable of the word before the coda triggers vowel epenthesis, resulting in the less usual antepenultimate stress pattern. The following subsection accounts for the one, rare weak syllable onset not yet covered: [n].

¹¹¹Of course we would not expect to see a bilabial at the surface. By saying that /β/ is not allowed as the final consonant of the root, we mean we see no alternation between suffixed [v] and unsuffixed (but epenthesized) [tʃ]

6.1.8. Nasal Dissimilation

4.2.1 examined the distribution of nasals and determined that, in Betsimisaraka, /ŋ/ must be a phoneme, despite the seemingly stress-conditioned behavior in the table repeated below.

Table 6.34: Near-minimal pairs wherein the only aspects to change are the stress assignment and the place of the nasals.

Penultimate Stress	Gloss	Antepenultimate Stress	Gloss
ta'naŋa	"town"	'taŋana	"hand"
ma'rini	"near"	'mariŋi	"correct"

By the antepenultimate stress in "correct," we infer an underlying form of /marin/, where the underlying place of the coda nasal is unspecified, since its surface place is determined by coda neutralization. (The epenthetic vowel being [i] will be accounted for through vowel harmony in Section 6.2.5.) It is impossible to get an SR of *['marini], with an non-neutralized underlying coda. There is no reason to assume that the UR for "near" is not /marini/. What is missing from this paradigm, unfortunately, is an SR such as ?[ma'rini], but, as stated in Chapter 4, elicitations for this study produced few minimal pairs for [n] and [ŋ]. Still, despite this notable gap, nothing about these alternations is unexpected.

Following this pattern, The UR for "town" is /tanaŋa/ and for "hand," is /taŋana/. What is unexpected, though, is that ['taŋana] does not show evidence of having undergone neutralization of its underlying

coda. If it had, we would expect a medial form of /taŋan/ and an SR of *[taŋanja]. Not only is this form ungrammatical, but nowhere in the data collected for this study is there a string [...ŋVŋV...]. This gap suggests that such a sequence is not allowed in the language.

Before this is formalized, note that strings of the form [...ŋVŋ...] are not ungrammatical. As long as the next segment in the string is a consonant (which will have to be velar, thus [k] or [g], by CC assimilation), these sequences can be found, as shown below:

Table 6.35: Compounds containing [...ŋVŋ...] strings, produced when the UR of the first root ends in a string [...ŋVN], and the second begins with a velar consonant.

root1	gloss	root2	gloss	compound	gloss
'aŋana	“greens”	'gisa	“goose”	aŋaŋ'gisa	“type of greens”
vi'ləŋina	“cooking pot”	'hena	“meat”	vilaŋŋ'kena	“pot for meat”

As in Section 6.1.1, both of the first roots selected are underlyingly nasal-final, as suggested by the antepenultimate stress assignment. These forms all end with a string [...ŋVnV] because, as we have proposed, [...ŋVŋV] is not allowed, so the second nasal dissimilates. In the compound, though, the similar string [...ŋVŋ...] is found, due to assimilation. Note that this means a specific place of articulation cannot be proposed for the first root's underlying coda: with a vowel following it, it dissimilates from the previous nasal, and it assimilates to the place of a following consonant.

[...ŋVŋ...] strings can be produced by alternation, as shown above, but one apparently monomorphemic form was also found that contains [...ŋVŋ...]: [maŋaŋgi], meaning “fussy” or “difficult to please.” This, along with the forms in Table 6.28, suggests that the prohibition on velar nasals is not against them appearing as successive consonants, but forbids them being pronounced as onsets to successive syllables¹¹².

To succinctly describe the observed trend, we will require a rule of the form below, that operates on nasals only if each one heads a syllable. The rule requires no reference to labiality: although no evidence was found of interaction with /m/, the rule below makes nasals non-dorsal, which does not affect /m/, and is triggered only by velar nasals, again avoiding [m].

14) Nasal Dissimilation

[+nasal] → [-dorsal] / [+nasal +dorsal] V _ V

“Nasals that precede vowels lose dorsality when the previous consonant is a dorsal nasal.”

Remember, though, that we can propose no UR for N in [...ŋVN...] strings: its place is determined either by the preceding nasal, if the next segment is a vowel, or the following consonant. So if CC assimilation is allowed to follow this rule, there is no need for a second

¹¹²Although this work opts to interpret nasal-consonant pronunciations as NC clusters rather than prenasalized stops, this is one area where a pre-nasalization analysis could prove useful. We could posit that it is ŋVŋ that are not allowed, because “fussy” would instead be transcribed [maŋaŋgi]. This analysis, not selected here for consistency, would avoid the upcoming Duke-of-York-style derivation.

“V” in the structure of the rule. This means that nasals can be rendered non-dorsal, and then become dorsal again¹¹³ due to CC assimilation.

The rule, then, can be slightly simplified to the form below.

14') Nasal Dissimilation

$[+nasal] \rightarrow [-dorsal] / [+nasal +dorsal] V$

“Nasals lose dorsality when the previous nasal is also dorsal.”

Crucially, Nasal Dissimilation must also follow the coda neutralization process that changes the place of all coda nasals to [+dorsal]. If the order were switched, the effect of dissimilation would never be seen at the surface. And while Vowel Epenthesis must follow Coda Neutralization, lest it bleed it, the ordering given for the sample derivations in Table 6.37 with Nasal Dissimilation and CC Assimilation is undetermined.

This nasal dissimilation process is the only consonantal process discovered that is entirely unique to Betsimisaraka vis-à-vis OM, although the neutralization process in Betsimisaraka was shown to be somewhat more straightforward than that in the official language. It is not surprising that this should be so, given that the velar nasal only exists in OM as a result of CC Assimilation.

¹¹³The slightly more complex rule in 14 is, of course, still available to any who object to this classic Duke-of-York style derivation, but this grammar will opt for the simpler 14'

Table 6.36: Derivation of monomorphemic “greens” and the compound translated here as “goose greens.” The derivation shows the necessary ordering of Nasal Dissimilation after Coda Neutralization but before CC Assimilation. Due to these rules always determining the place of the underlying coda nasal, it is marked as unspecified.

Gloss	“greens”	“goose-greens”
UR	/aŋaN/	/aŋaN-gisa/
Coda Neutralization	/aŋar/	---
Nasal Dissimilation	/aŋan/	/aŋan-gisa/
CC Assimilation	---	/aŋan-gisa/
Vowel Epenthesis	/aŋana/	---
SR	[aŋana]	[aŋan-gisa]

6.1.9: Distinctive Features for Consonants: Review

Seven features are necessary to distinguish the consonants of Betsimisaraka, given the alternations we have seen in the preceding subsections.

Table 6.37: Distinctive features for [-syllabic] segments of Betsimisaraka. All segments have a value for [\pm labial], [\pm dorsal], [\pm continuant], [\pm nasal], and [\pm voice] this last indicated by boldness rather than marked polarity. [\pm distributed] and [\pm anterior] are not defined for labials and dorsals, and only the fricatives indicated are have a polarity for [\pm strident].

		+labial		-labial						
				-dorsal			+distr.	+dors.		
				+distr.	-distr.	+distr.				
				+anterior		-ant.				
-cont.	+nasal	m		n				ŋ		
	-nasal	p b		t d	ts dz	tʃ dʒ	k g			
+cont.		ɸ β	f v	l	s z	r	h			
		+strident	+strident		+strid.			-strid.		

6.2 Alternations Affecting Vowels

Of processes primarily affecting vowels, Betsimisaraka has two that are related to, but distinct from, processes observed in Official Malagasy in Section 3.4.2, and two that distinguish it entirely. Firstly, OM was seen to have stress-conditioned vowel heightening in certain environments, although the Betsimisaraka system affects more vowels in more positions (6.2.1). Additionally, both varieties of the language have systematic reduction of hiatus through deletion of vowels and creation of diphthongs (6.2.4), but (as already mentioned in 4.1.2) only Betsimisaraka has coalescence to mid-lax vowels. Although there was no evidence for vowel harmony in OM, Betsimisaraka seems to have two different such processes: a harmony-like alternation affecting only forms of the 1S suffix (6.2.2), and a productive harmony of the epenthetic vowel, also called copy epenthesis (6.2.3). Section 6.2.3 will also cover the coarticulation alternations in Betsimisaraka, occurring in fewer environments than in OM, and one last, optional consonantal alternation, involving deletion of final nasals.

6.2.1 Stress-conditioned Alternations of High and Low vowels

Recall from Chapter 3 that Betsimisaraka has five phonemic vowels, consisting of the five cardinal vowels /a, e, i, u, o/, the last of which distinguishes it from OM. Stress-conditioned [e]/[i] and [o]/[u] alternations were used in 4.1.1 to motivate the phonemic status of /o/.

due to the lack of minimal pairs found. In this section, the process behind those alternations will be formalized, starting by reviewing the data repeated from Tables 4.2 and 4.3.

Table 6.38: Stress-conditioned height alternation of non-low vowels as evidenced by active and imperative forms of representative verbs. For “cut” and “assemble,” the same vowel is high no matter stress assignment, whereas in “look” and “drop by,” the vowel is mid-height when stressed, and high when unstressed.

gloss	active	imperative
“cut”	man'didi	di'diə
“look”	mi'dzeri	dzi'revə
“assemble (for a meeting)”	mi'vuri	mivu'riə
“pass”/“drop by”	man'dalu	mandə'lɔvə

This distribution above, with four non-low vowels in stressed syllables and just two when unstressed (not counting [ə], which does not alternate with these vowels), suggests neutralization, and the vowel that occurs in stressed positions is taken to be the UR. The alternative, where some high vowels lower in stressed syllables while others stay high, is less systematic.

The rule proposed below governs this alternation, finding unstressed vowels and raising their height. Because stress is a target feature for this alternation, the rule must follow stress assignment.

- 15) Stress-conditioned Vowel Heightening
 (“Vowel Neutralization”)**
 [+syllabic -stress -low] → [+high]
 “Non-low, unstressed vowels are pronounced as high.”

An example derivation using this rule, along with those analyzed previously, is offered below.

Table 6.39: The root /dalov/¹¹⁴ (“pass/drop by”) derived in the active and imperative. Although other rules are shown in the derivation, it is not exhaustive. The only necessary ordering shown is that of stress assignment feeding vowel neutralization.

“pass/drop by”	Active	Imperative
UR	/maŋ-dalov/	/maŋ-dalov-a/
CC Assimilation	/mandalov/	/mandalova/
Strident Coda Deletion	/mandalo/	---
Stress Assignment	/man'dalo/	/manda'lava/
Vowel Neutralization	/man'dalu/	/mandə'lovə/
SR	[man'dalu]	[mandə'lovə]

Recall from section 4.1.1 that these alternations represent a generalized version of a process in OM, analyzed here in section 3.4.2. There, post-tonic /e/ reduced to [i], surfacing faithfully in other unstressed syllables. No similar process could occur for /o/ in OM, though, due to the lack (or, at most, extremely limited distribution) of that phoneme in the language. In any case, certain participants' knowledge of the more restricted OM process is one possible account for why they could pronounce underlying /dzerev-a/ (“look” imp.)¹¹⁵ also as [dze'revə], with the unstressed vowel unneutralized.

¹¹⁴ For most forms in previous examples, we chose an imperative without the prefix /maŋ/ (actually the present /m-/ and the SFP /aŋ-/), but this lexical item did not sound grammatical in its unprefixed form. This means that, by fortition, it could conceivably be underlyingly ?/lalov/.

¹¹⁵ Additionally, the more usual translation of “look” in Betsimisaraka is [mi'zaha] rather than [midžeri], suggesting the pronunciation could be influenced by lexical knowledge of OM.

Place features necessary for vowels at this point are [\pm low], which identifies the potential targets as [-low], and [+high], which is the end result of the alternation. At this point, there is no need to assign [a] a value for [\pm back], its [+back] polarity will be justified in the following subsection. Similarly, since it is ignored by stress-conditioned height alternation, it does not yet require a [\pm high] value, although its [-high] feature will be critical in 6.2.4.

Table 6.40: Distinctive features necessary so far for [+syllabic] segments.

		-back	+back
+high	-low	i	u
-high		e	o
	+low		a

6.2.2: Backness Harmony of the 1S Suffixes

Although lexically limited, one feature distinguishing Betsimisaraka from OM is backness harmony on the 1S suffix. This suffix is used for both possession of nouns and for agreement on passive verbs, with related but distinct harmony in these two cases: both involve the suffix surfacing as [ki] following /i/, though the elsewhere allomorph of the suffix differs.

In OM, there is a single, fairly immutable 1S suffix used both as a possessive marker and for agreement on passive verbs. It is written as <ko> in the standard orthography and pronounced [ku]—or [k'u]

following /i/ (about which more below the table): examples are shown below with /-ku/ following each of OM's four vowels and for each of the two functions of the suffix.

Table 6.41: The behavior of 1S suffix <-ko> in both uses, **in Official Malagasy**; data for Betsimisaraka will be shown below. The OM data were found in a Malagasy dictionary, and the forms checked with an OM speaker.

possessives	gloss	passives¹¹⁶	gloss
'ra-ku	"my blood"	isa'ura-ku	"thanked by me"
'trundru-ku	"my fish"	anu'lu-ku	"changed by me"
'vadi-k̥u	"my spouse"	va'ki-k̥u	"read by me"
laki'le-ku	"my key"	dze're-ku	"observed by me"

The phenomenon seen above is allophony, not allomorphy: as seen in section 3.1.1, palatalized allophones of velar stops invariably follow /i/. Although [k̥] does occur in certain contexts in Betsimisaraka, as will be shown in the following subsection, it never occurs in the 1S suffix. The rest of this section, then, is dedicated to the facts of Betsimisaraka, beginning with the straightforward distribution of the 1S suffix used on passive, and moving on to the more diverse distribution of the 1S possessive suffix.

Table 6.42 provides examples of high-vowel-final passives roots in Betsimisaraka, all with the 1S suffix added. In all these examples, the vowel in the 1S suffix mimics the final vowel of the root. This is

¹¹⁶The translations are literal; in idiomatic use the "passive" is the most likely form of the verb to be used, and thus closer in pragmatic function to "I am reading" or "I am changing." This generalization is true for Betsimisaraka as well for OM.

distinct from the coarticulation effect seen in OM, where a form like [i'kehiki] above is [i'kehiku]: a *[ku] allomorph of the 1S was never heard in Betsimisaraka,

Table 6.42: The 1S passive suffix attached to high-vowel-final roots in Betsimisaraka. Column headings refer to the final vowels. To save space, glosses are of roots, not passives

[u] passives (1S)	gloss of root	[i] passives (1S)	gloss of root
a'lavu-ku	“drop”	i'kehi-ki	“accept”
'afu-ku	“be able to do”	hin'dʒami-ki	“borrow”
aru'su-ku	“enter”	'vidi-ki	“buy”

In Vavatenina, where the bulk of the data for this project was elicited, the pattern seen above was consistent, with the [-ki] allomorph following /i/-final roots and [-ku] following /u/. The pattern was less straightforward for a small sampling of speakers who were natives of the Northern city of Maroantsetra, such as the project's principle language assistant. For them, certain exceptions to this pattern could be found at the surface.

Table 6.43: Data from two Betsimisaraka dialects for 1S passive forms of two lexical items that employ the [-ki] allomorph of /-ku/, even though the harmonic environment for use of this allomorph is not evident in the Northern variety.

Maroantsetra (North)	Vavatenina (South)	gloss
lu'ku-ki	lu'kui-ki	“colored by me”
su'tʃu-ki	su'tʃui-ki	“drunk by me”

In Northern Betsimisaraka, with the [u]-final root, one might have

expected instead *[luku-ku] and *[sutʃu-ku], but these were rejected as sounding like OM forms. To clarify: OM uses [ku] in all cases, whereas the Betsimisaraka pattern seen above had [ki] following front vowels. If the Maroantsetra variety is systematic, it is likely due to the UR of the root in both varieties being /ui/-final. In Vavatenina, the [ui] is allowed to surface, whereas in Maroantsetra, hiatus resolution (explained more fully in Section 6.2.4) deletes the unstressed vowel. If the data above is representative, then, in Maroantsetra the SRs are opaque, due to the suffix harmonizing with a non-surface vowel. Still, a full on-site investigation, with multiple speakers in Maroantsetra, is necessary for a full description of this phenomenon.

Vavatenina Betsmisaraka's 1S suffix shows frontness harmony following high vowels, and as the data below shows, following low vowels, with [a]-final stems taking [-ku]. This generalization motivated placing /a/ in the [+back] column in the above distinctive feature chart (Table 6.40). Fieldwork only produced two instances of the rare /e/ phoneme in final position, one taking [-ki] as expected given its final front vowel, and one unexpectedly taking the back-vowel suffix [-ku].

Table 6.44: The behavior of the 1S passive suffix on some non-high-vowel-final forms in Betsimisaraka.

[a] passives (1S)	gloss of root	[e] passives (1S)	gloss of root
lu'kua-ku	"color"	ampi'te-ki	"cross"
za'ha-ku	"look"	dzi're-ku	"look"

This may be, again, because /džere/ is the usual OM root for “look,” not as commonly heard by Betsimisaraka speakers as the native root /zaha/, shown in the passive form on the left column. If it is a recent borrowing from OM, or just a foreign word, it may follow that, like OM, it does not trigger harmony in the 1S suffix¹¹⁷.

Elicitation for this study produced only a single /o/-final root. The 1S passive for “brush” is [bru'soki], not *[brusoku], as one would expect in a fully harmonic system. Perhaps complicating matters, this word is a straightforward borrowing from the French “brosser” ([bro'se]), although we do not here propose that the [e] in the original French form somehow triggers the front-vowel allomorph). At this time, then, we can conclude nothing about harmony of the mid vowel forms, since the data provide only three: a form where harmony occurs ([ampi'te-ki]), a borrowing from OM where it unexpectedly does not ([dži're-ku]), and a borrowing from French where harmony does occur, but without a trigger ([bru'so-ki]).

Having seen that harmony of the 1S suffix on passives is a reliable system, if only for high and low vowel triggers, we turn our attention to the distribution of the [-ki] allomorph when used as a possessive. It will be shown that, in Southern Betsimisaraka, it occurs as expected, following nouns ending in /i/, but that in Northern

¹¹⁷If a word's etymology is truly determining the type of suffix allowed, this would be similar to some English speakers dislike of pluralizing Latin-derived “-us” words with the native “-es,” such as “cactus,” “alumnus,” “nucleus,” etc.

Betsimisaraka, it only occurs following /i/-final kinship terms.

Furthermore, in both dialects, when [-ki] cannot be used, the alternate form is not [-ku], but an alternate lexical item, either [-naha] in the North or [-(a)nahi] in the South.

In Northern Betsimisaraka, the harmonic 1S allomorph [-ki] is used for possession similarly to how it is used for passives, although as can be seen below, it only attaches to kinship nouns. For non kinship nouns, and for non [i]-final nouns, it has an entirely different allomorph.

Table 6.45: Behavior of 1S possessive suffixes in Betsimisaraka, including both non-kinship terms and kinship terms, as well as both [a]- and [i]-final forms.

[-naha] possessives		[-ki] possessives	
Possessive form	gloss	Possessive form	gloss
sira naha	“my salt”	vadi-ki	“my spouse”
mama naha	“my mom”	rahavavi-ki	“my sister”
aumbi naha	“my cow”	rahalahi-ki	“my brother”
tadi naha	“my rope”	reni-ki	“my mother”

The dependence on family relation is so strong that the same morpheme will take a different suffix depending on its semantic role: consider that [lalahi] (“man”)—containing the same male-indicating morpheme /lahi/ as [rahalahi] (“brother”)¹¹⁸—has the 1S possessive form [lalahi-naha], never *[lalahi-ki]. Other family members that are

¹¹⁸As in OM, Betsimisaraka actually has two words for brother, one for a brother to a male and one for a brother to a female. To put this another way, different nouns are employed in translating “his brother” and “her brother” into Betsimisaraka.

not [-i]-final, though, are never possessed with [-ki], including such basic terms as [mama-naha] for “my mom.”

The semantic element does not appear to be a factor in Southern Betsimisaraka. In Vavatenina, all nouns *can*¹¹⁹ take the non-[-ki] form, which is pronounced [anahi] in most circumstances, but [-nahi] following /a/ due to hiatus prohibition. Unlike subjects from Maroantsetra who insisted on [-ki] only following kinship words, though, Vavatenina residents accepted or even preferred forms such as [aumbi-ki] (“my cow”) or [tadi-ki] (“my rope”). Furthermore, [i]-final forms taking [-ki] were heard in conversation, and even applied to borrowings from other languages such as English ([gisi-ki] “my goose”) and French ([foriseti-ki] “my fork”).

In both Northern and Southern Betsimisaraka, the pattern must be viewed as morphophonological, as no phonological rule will account for alternation between [-ki] and [-anahi] or [-naha]. Still, the data present a clear picture for the grammar of Southern Betsimisaraka wherein all [i]-final nouns can (and often are) possessed using [-ki], [a]-final nouns with [-nahi], and others (consisting almost entirely of [u], due to the limited distribution of [o] and [e]) taking [-anahi]. For Northern Betsimisaraka, the picture is both clearer (because the consonant-initial suffix requires no hiatus-resolution following [a]) and

¹¹⁹The original methodology to elicit this data involved reading the subjects noun after noun, to which they invariably attached [anahi] or [-nahi]. It was the secondary language assistant’s suggestion that the investigators try again to ask participants if given forms were acceptable.

more nuanced (due to the intrusion of the semantic element of kinship).

Phonologically conditioned allomorphy of the type examined above has been observed in other languages; Paster (2009) presents the case of the Hungarian 2S present-tense indefinite verb suffix, which takes the disparate forms [-ol] if the root ends in a sibilant, [-sz] elsewhere. Sun's (2008) analyzes the Korean nominative suffix, which has allomorphs [-i] and [-ka], depending if the root is consonant- or vowel-final. In both cases, as in Betsimisaraka, it is difficult to see the phonological relationship between the allomorphs¹²⁰. As to rules of possession changing depending on whether the noun is one of kinship, Italian's requirement of a definite article preceding possessive pronouns does not apply¹²¹ in the case of kinship terms.

To offer a possible historical explanation, at some stage of development Betsimisaraka may have possessed a more clearly phonologically determined [-ki]/[-ku] alternation of the 1S possessive suffix at the surface, just like the one it still has for the 1S *passive* suffix, as seen above in Tables 6.31-6.33. Over time, though, the [-ku] form came to be supplanted by a form that was more closely related to

120English "a"/ "an" alternation, historically, derives from "an" losing its consonant due to cluster reduction before a consonant, but no rule deleting all nasals before other consonants existed at any stage of development. "A" and "an," then, are more similar than [-ki] and [-anahi], but are still phonetically independent.

121In Italian, then, when speaking of family members, the definite article that is usually required before possessive pronoun, is not used. For example, Italians say "il mio abito" ("my suit") and "la mia mela" ("my apple") but simply "mio nipote" ('my nephew") and "mia zia" ("my aunt.")

the form the 1S pronoun takes as the subject of a sentence: compare [zaha] (1S subject pronoun) and [naha] (1S possessive pronoun).

Whether or not the historical hypothesis holds, the following is the pattern for Betsimisaraka 1S possessives. [k]-initial suffix can be used if it is the uniquely Betsimisaraka (or at least non-Merina) [-ki], but for phonological/semantic environments where [-ku], rather than [-ki], would be triggered, Betsimisaraka speakers prefer to use their native innovation, either [-naha] in the North or [-anahi] in the South. The pseudo-political consideration need not occur each time a speaker needs a 1S possessive pronoun, but rather may account for the tendencies that determined what is now more clearly a case of phonologically conditioned allomorphy.

Section 6.2.3: Harmony of Epenthetic Vowels in Betsimisaraka

In 6.1.6 and 6.1.7, the process of vowel epenthesis for Betsimisaraka was formalized, wherein an /a/ would be added to consonant-final URs. The examples given in those sections, with [a] as the quality of the epenthetic vowel, were all attested in Betsimisaraka. Still, Betsimisaraka is completely distinct from OM, and similar to other Northern dialects (as described in Section 3.6.2). in also allowing copy epenthesis. This is a more complete harmony process than even the backness harmony of the 1S passive suffix examined in the previous subsection: as seen in Table 6.46, harmony on the epenthetic vowel

includes harmony by height as well as backness.

Table 6.46: Optional vowel harmony of epenthetic vowels in Betsimisaraka. Consonant-final URs are proposed based on antepenultimate stress. As the possible qualities for the epenthetic vowel are [a], [i], and [u], depending on the previous vowel, ungrammatical surface forms are provided as well.

gloss	UR	SR	ungrammatical SR(s)
“many”	/betsak/	['bets a k a]	*['betsaku], *['betsaki]
“four”	/efatʃ/	['efatʃ a i]	*['efatʃu], *['efatʃi]
“sand” (N)	/fasik/	['fas i k a, 'fasiki]	*['fasiku]
“skin” (N)	/huditʃr/	['huditʃ a i , 'huditʃi]	*['huditʃu]
“dust” (N)	/vuvuk/	['vuv u k a, 'vuvuku]	*['vuvuki]
“leg”/“foot”	/tungutʃ/	['tungutʃ a i , 'tungutʃu]	*['tungutʃi]

The resulting distribution is systematic: the epenthetic vowel is either [a] or else matches the final vowel of the root¹²². Among Betsimisaraka speakers, there is even a third possibility for epenthetic vowel forms that was not included in Table 6.47: forms wherein the vowel is [a] but the quality of the previous vowel is expressed only as coarticulation on the intervening consonant. If the final vowel of the root is /i/, then, the underlyingly morpheme final consonant can be palatalized on the surface, whereas if the vowel is /u/, then the coda can be labialized. This coarticulation may be related to OM palatalization, presented in Chapter 3, wherein the UR /fasik/ (“sand”) was pronounced [fasik^ja], not *['fasika], because all instances of /k/

¹²²The environment where this occurs, between the two final, unstressed vowels in an antepenultimately stressed word, limits the relevant vowel inventory to /i, a, u/. Harmony to *underlying* mid vowels is discussed later in the chapter.

show this effect following /i/. Although *labialization* in OM has only been briefly noted (Rajaoranimanana 2004), in Betsimisaraka, dorsal-final URs such as /vuvuk/ (“dust”) can be pronounced with rounding of the transition between [k] and [a], thus [vuvuk^wa]. The full distribution for Betsimisaraka, then, is that a root ending in a high vowel followed by an underlying potential coda can be pronounced with an epenthetic matching high vowel or with epenthetic [a]; if the latter vowel, coarticulation is optional on an intervening dorsal segment. The same speaker might offer [fasika], [fasik^ja], and [fasiki] in the same elicitation session.

Table 6.47: Possible pronunciations of a root that underlyingly has a dorsal stop as its final consonant.

gloss	proposed UR	Possible SRs		
		epenthesis ([a])	coarticulation	epenthesis (harmony)
“sand” (N)	/fasik/	'fasika	'fasik ^j a	'fasiki
“dust” (N)	/vuvuk/	'vuvuka	'vuvuk ^w a	'vuvuku

Even if OM does have labialization, though, the systems are not the same. Betsimisaraka coarticulation is limited, like full vowel harmony, to forms with epenthetic vowels. While the dialects overlap at least on allowing the pronunciation [fasik^ja], OM requires this palatalization in all possible cases. The following coarticulations were offered by an OM speaker but rejected as “Merina” by a Betsimisaraka speaker, even on cognate forms: [a'tsik^ja] (2P excl.), [di'k^jani]

(“reason”), and [mik^{ja}'ruana]¹²³ (“research”). Conversely, unpalatalized forms were deemed incorrect by a native Merina speaker. This coarticulation, then, is just a less far-reaching form of the harmony seen in Table 6.46, and will thus be treated along with it.

Unlike the data seen above for root-final [k] and [tʃ], the quality of the epenthetic vowel following nasals is restricted to the harmonic vowel possibility. For example, given the distribution above one might expect /masinj/ (“salty”) to have the alternate forms *[‘masinjə] and [‘masinjɪ], but the former form is in fact unattested. One *does* hear the OM cognate form [‘masinə] (with an alveolar nasal), due to Betsimisaraka speakers' ability to codeswitch to the official variety. Still, if the Betsimisaraka UR of /masinj/ is used, the only epenthetic vowel allowed is the harmonic one. This means that, when the final consonant of the root is nasal, the only type of epenthetic vowel allowed in Betsimisaraka is the harmonic one. In an extreme, but still evidence-grounded, interpretation, this evidence of unwillingness to use a compromise form (evidence for which was seen in Velonandro's 1983 survey of Northern dialects) may indicate that *all* uses of [a]-epenthesis are code-switching to OM.

Additionally, nasal-final roots in Betsimisaraka have an alternate SR, with deletion of the potential word-final coda, as shown below.

¹²³Rejected even with a velar nasal onset to the final syllable, although, without the palatalized [k], [mika'ruaña] is the usual Betsimisaraka word for “research.”

Table 6.48: Optional Vowel harmony of Epenthetic Vowels in Betsimisaraka. Although exact numbers were not produced, the two possible pronunciations seemed to be approximately equally used, between and among speakers.

gloss	UR	SR	ungrammatical SR(s)
“rain”	/majoran/	[ma'ŋorəŋa, ma'ŋora]	*[ma'ŋoranju], *[ma'ŋoranji]
“salty”	/masinj/	['masinji, 'masi]	*['masinju], *[masinja]
“person”	/olunj/	['olunju ¹²⁴ , 'olu]	*['olunji], *['olunja]

The rule that removes these nasals can be formalized as below:

16) (Optional) Nasal Coda Deletion

[+nasal] → Ø/_#

“Nasals are deleted word-finally”

The harmonic vowel examples seen in the data tables in this section show yet more evidence for the epenthetic vowel argument for antepenultimate stress in Malagasy, and represent a process of vowel harmony that is unknown in the fairly well described phonology of OM. While OM explanations for antepenultimate stress offer a thorough justification based on the phonotactics of Malagasy in general, this unique process of complete vowel harmony affecting only final vowels of antepenultimately stressed words further distinguishes epenthetic vowels from those occurring in roots.

Although antepenultimately stressed words represent most of the available examples of words in which harmony is possible, it is a

¹²⁴This lexical item also has an OM cognate ['uluna], which may be known to Betsimisaraka speakers, but, like *[masinja] above, the hybrid form (*['oluŋa], with [a] epenthesized to a Betsimisaraka UR) is unattested.

mistake to think that harmony operates, for example, on the vowels of two back-to-back unstressed syllables. Such an analysis would be supported by the vowel harmony literature, appealing to Kaun's (2004) and Walker's (2009) interpretation of vowel harmony as a process whereby vowels spread their features onto segments where those features will be more salient. Walker presented evidence of short vowels spreading features to long vowels and even unstressed vowels spreading features to stressed vowels. One could imagine a system, then, in which an unstressed vowel might spread its features to another unstressed vowel so, with those features applying to two segments rather than one, they might be more salient.

Such an analysis, while appropriate elsewhere, is unworkable in Betsimisaraka because antepenultimately stressed words are not, in fact, the only ones on which harmony can be found. Consider the adjective ['*ritʃi*] ("dry") with associated verb forms in the active ([*man'dʒitʃi*]) and passive ([*ri'tinjɪ*]). The [t]/[tʃ] alternation is indicative of word-final neutralization, suggesting a root /rit/ that undergoes both neutralization and epenthesis¹²⁵. The epenthetic vowel is harmonic with the previous vowel in both the adjectival and active verbal forms, yet the harmony is between vowels of different stress levels. Because two-

¹²⁵CVC strings where the V was not [a] were difficult to come by, and /o/ and /e/ were infrequent phonemes, so it is not surprising that no data was found of CoC or CeC roots. We thus cannot observe harmony from surface mid vowels to high epenthetic vowels. This is surely what would occur, though, as we do see cases of underlying mid vowels triggering harmony. Compare, for example, the root /velom/ ("live")—justified by the imperative [*ve'loma*]—for which the present is ['*veluŋu*].

syllable, penultimately stressed words do not routinely exhibit vowel harmony, a stress-based explanation is thus untenable in Betsimisaraka. Instead, an underspecified vowel is inserted, and it gets features from the nearest vowel. This follows from Nevins's (2013) work on vowel harmony, wherein all vowel harmony can be motivated by vowels' need for features. A word like /rit/ is assigned stress to its only syllable, as usual, and then a vowel is attached, in need of features: it gets these features from the nearest vowel, or else gets default low, back place features.

Copy epenthesis patterns were previously stated to be found in other Northern dialects, and they are not unknown elsewhere in the Austronesian language group. Both the Formosan language Maga Rukai (de Lacy 2002, p. 150) and Malayo-Polynesian Selayar (Broselow 2008) exhibit it. The latter language even presents a case where words are typically penultimately stressed, unless they are underlyingly consonant-final, in which case a copy of the previous vowel is epenthesized to maintain syllable structure!

This is, of course, the pattern we have seen in Betsimisaraka, although it does not appear to be optional in Selayar. As we have seen, in the specific Betsimisaraka pattern, word-final codas are forbidden, so some segments delete word-finally (6.1.6) and others epenthesize vowels (6.1.7), with a rule that must apply after penultimate stress assignment. This rule is repeated below:

8) Coda-Eliminating Vowel Epenthesis (“Vowel Epenthesis”)

$\emptyset \rightarrow [+syllabic] / [-syllabic] _ \#$

“If a consonant occurs at the right word boundary, the minimal vowel is epenthesized between them.”

As mentioned when formulating this rule, the inserted vowel will be defined only for one feature, [+syllabic]. An alternative approach would have two separate epenthesis rules: copy-epenthesis across nasals and either copy- or [a]-epenthesis across consonants in general. The problem with this approach is it ignores the generalization that epenthesis happens either way, whereas inserting an underspecified vowel that will get its features later does capture this observation.

Furthermore, the temporary situation wherein a segment, while specified [+syllabic], has no specification for place can be exploited to account for both [a]-epenthesis and the copy-epenthesis seen in Betsimisaraka. To wit, a rule can target the unique $[\emptyset \text{place}]$ specification¹²⁶ and provide the associated segment with defined place features, either as the minimal vowel [a] or as matching the preceding vowel. First, though, recall that harmony across nasals is obligatory:

17) Trans-Nasal Place Harmony (“Trans-N Harmony”)

$[+syl. \emptyset \text{place}] \rightarrow [\alpha \text{place}] / [+syl. \alpha \text{place}] N _$

“Vowels following a nasal and undefined for place features are realized as a copy of the preceding vowel.”

¹²⁶Typically, $[\emptyset \text{feature}]$ values cannot be exploited in rules, so this is a nonstandard approach. In the context of this grammar, rules that target $[\emptyset \text{place}]$ can be interpreted as formalizations using the language of rewrite rules to achieve the sort of feature-filling that is usually displayed autosegmentally, for example in Goldsmith (1989). In the case of harmony, features spread from a vowel with place features to one without, whereas in the [a]-epenthesis case, [a] is defined as having the minimal features for place, so that a [+syllabic] element is automatically interpreted as [+low].

Whether an epenthetic vowel following an oral consonant was harmonic or not was more varied, between and among speakers, with a different version of the place-feature assignment rule applying in either case. 18a produces the same result as OM for non-nasal underlying final consonants, while 18b, like 17 and 16, produces a distinct Betsimisaraka form.

18a) Default Place Feature Filling (“Default Feature-Filling”)

[+syl. Øplace] → [+low +back]

“Vowels undefined for place features are realized as [a]”

18b) Trans-Oral Place Feature Harmony (“Trans-C Harmony”)

[+syl. Øplace] → [əplace] / [+syl. əplace] (C) C _

“Vowels undefined for place features are realized as a copy of the previous vowel”

With regards to other associated rules, as in the OM ordered-rule analysis, stress assignment must precede epenthesis in Betsimisaraka. Assignment of features to the epenthetic vowel (in either version) must itself follow epenthesis, as the vowel must be present to harmonize. Betsimisaraka's optional deletion of nasal codas must precede epenthesis, lest it never have the chance to occur

All that is left for analysis of harmony, then, is to see representative derivations demonstrating the rule interactions. Because there are two optional rules (root-final nasal deletion or not, copy-epenthesis or [a]-epenthesis), four distinct derivations produce all attested forms, depending on which set of rules is applied.

Table 6.49: Four derivations that show the four grammars available to speakers of Betsimisaraka Malagasy. Stress assignment is not included below, as its ordering has already been covered. Note also that, in grammars featuring nasal-coda deletion, there will be no need for trans-nasal assignment of place features.

Word-final nasal deletion and [a]-epenthesis

gloss	“big”	“skin”	“person”
UR	/betsak/	/huditʃ/	/oluŋ/
Nasal Coda Deletion	---	---	/olu/
Vowel Epenthesis	/betsakV/	/huditʃV/	---
Default Feature Filling	/betsaka/	/huditʃa/	---
SR	[betsaka]	[huditʃa]	[olu]

Word-final nasal deletion and copy-epenthesis

gloss	“big”	“skin”	“person”
UR	/betsak/	/huditʃ/	/oluŋ/
Nasal Coda Deletion	---	---	/olu/
Vowel Epenthesis	/betsakV/	/huditʃV/	---
Default Feature Filling	/betsaka/	/huditʃi/	---
SR	[betsaka]	[huditʃi]	[olu]

NO word-final nasal deletion and [a]-epenthesis

gloss	“big”	“skin”	“person”
UR	/betsak/	/huditʃ/	/oluŋ/
Vowel Epenthesis	/betsakV/	/huditʃV/	/oluŋV/
Trans-N Harmony	---	---	/oluŋu/
Trans-C Harmony	/betsaka/	/huditʃa/	---
SR	[betsaka]	[huditʃa]	[oluŋu]

NO word-final nasal deletion and copy-epenthesis

gloss	“big”	“skin”	“person”
UR	/betsak/	/huditʃ/	/oluŋ/
Vowel Epenthesis	/betsakV/	/huditʃV/	/oluŋV/
Trans-N Harmony	---	---	/oluŋu/
Trans-C Harmony	/betsaka/	/huditʃi/	---
SR	[betsaka]	[huditʃi]	[oluŋu]

Let us close this section by returning to the optional coarticulation strategy, which could follow Default Feature Filling in the two top sub-tables of Table 6.49. At first, coarticulation seems to run counter to the idea that it is the epenthetic vowel's need for features that drive the copying of place features. After all, how does it help the vowel get its features if they only get as far as the intervening consonant?

There are two reasons underspecification still seems like an appropriate motivation. Firstly, recall that in OM, coarticulation *always* manifests on dorsal consonants (/g, k, h/) following high front vowels, but that Betsimisaraka limits the spreading of the final underlying vowel's place features to cases of epenthesis: cases where they alternatively spread all the way to the next vowel. Since it is the existence of an added vowel that triggers the spread, whether the features manifest on a consonant or vowel, the reason for the spread must have its root in that vowel's underspecification.

Secondly, recall that, although the phenomenon will not be covered in detail here, Section 5.1 briefly mentioned that unstressed vowels can be devoiced, *especially* high vowels and *especially* final vowels of antepenultimately stressed words. The most likely vowels to be devoiced, then, are the harmonic epenthetic vowels. The place features that spread to them, then, are very likely not to be noticed, which may well explain why they should be articulated elsewhere.

6.2.4: Hiatus-Resolution Strategies

Betsimisaraka, like Official Malagasy, does not allow full vowels to be pronounced in sequence, at least in the discourse register. Both varieties allow hiatus underlyingly; this can be due to morphological concatenation, but evidence exists that it is also allowed root-internally. Like in OM, Betsimisaraka strategies to resolve hiatus typically give priority to the stress-assigned vowel, with the other vowel minimized in different ways depending on the combination. Two of the hiatus resolution processes are those discussed in section 3.4.2: glide formation (of high vowels) and vowel elision (of low). Although the facts were disputed, OM may have coalescence of certain VVs to [e] and/or [o], the latter of which is not a phoneme in that language variety. Unlike in OM though, and as presented in 4.1.2, certain high-low vowel clusters coalesce into mid-lax monophthongs, rather than forming diphthongs. An analysis where certain monophthongs and all diphthongs are the result of underlying hiatus will be necessary to account for cases where the two vowels in question originate from separate morphemes, and will prove more parsimonious as pertains to the vowel inventory.

Except for cases of coalescence, stress is the deciding factor in Betsimisaraka for determining which hiatus resolution strategy is chosen. Crucially then, outside of compounds, which act like full words when it comes to stress assignment, all underlying vowel pairs will

have one and only one that is assigned stress, either primary or secondary: recall from section 5.2 that primary stress is assigned to the penultimate vowel of the UR, then secondary stress to alternating vowels back from there. Outside of compounds, then, there is no instance where two syllables in a row are stressed, and, the only case where two in a row can be unstressed is due to vowel-epenthesis. Epenthesis, though, will never cause a VV string, as it is triggered by word-final consonants: the two unstressed vowels will always be separated by a consonant. When encountering underlying VVs, then, it can safely be assumed that stress will be assigned just to one vowel, which will aid in choosing the surface representation for this pair.

We can now turn to examining the hiatus-resolution strategies themselves. One way to determine the existence of underlying hiatus is looking at cases where vowels from separate morphemes meet under concatenation. Although Betsimisaraka uses different passive suffixes with different verbs, they reliably consist of a single vowel. When the root is vowel-final, then, passives will exhibit underlying hiatus. Furthermore, in cases where single-syllable person suffixes are attached, penultimate stress will fall on the passive suffix, meaning the vowel of the suffix will be prioritized over the final vowel of the root.

In the cases shown in the table below, the hiatus resolution strategy is to desyllabify the unstressed vowel, creating a semi-vowel glide that does not violate the prohibition against hiatus. To keep this

glide from violating restrictions on CCs, though, the glide attaches to the stressed vowel, creating a diphthong. Below, the final vowels of the root are high, and thus can desyllabify to become glides, resulting in uniformly rising diphthongs. Although no evidence was found of [o] used as a passive suffix, the second table below offers evidence of rising diphthongs where the main vowel is [o], this evidence coming from /o/-initial roots following the high-vowel final prefix /mi-/.

Table 6.50: Desyllabification of unstressed high vowels, in service of hiatus resolution. Vowel-final roots are justified by active forms. Unstressed /u/ and /i/ become the corresponding glides /w/ and /j/, respectively, which surface as onglides to the stressed vowel, creating a rising-sonority diphthong. To be clear, then, the string [k^wi] represents a stop followed by a diphthong, *not* a labialized stop followed by a monophthong. “UR” below refers to a stress-assigned form.

active	root	suffix		UR (passive)	SR (passive)	gloss (root)
		passive	3S			
[man'duku]	/luku/	/i/		/luku'ini/	[lu'k ^w ini]	“color”
[ma'lazu]	/lazu/	/e/	/ni/	/lazu'eni/	[la'z ^w eni]	“wilt”
[ma'ŋuru]	/iru/		/a/	/iru'ani/	[i'r ^w ani]	“burn”
[mi'vuri]	/vuri/			/vuri'ani/	[vu'r ^w ani]	“assemble”

Table 6.51: Desyllabification of unstressed high vowels, in service of hiatus resolution. Vowel-initial roots are justified by passive forms. Unstressed /u/ and /i/ become the corresponding glides /w/ and /j/, respectively, surfacing as onglides. Again, the string [m^wo] is a nasal followed by a diphthong, *not* a labialized nasal followed by a monophthong.

imperative	active				
	prefix	root	UR	SR	gloss (root)
ova	mi-	ova	/miova/	[m ^w o va]	“change”
odi	hi-	odi	/hiodi/	[h ^w o di]	“come home”

Although the evidence above is helpful in showing that diphthongs heard at the surface must begin as underlying vowel clusters, they are limited. Because in passive and active morphology, the passive suffix vowel and the initial vowel of a two-syllable root will each bear stress in their respective cases, they will always surface faithfully. This means that, with only the evidence above, one could equally well hypothesize that it is the *second* of two vowels that surfaces faithfully, and not the stressed one.

It is for this reason that we take recourse to orthographic representations for further examples. This allows for a greater inventory of vowel-vowel interactions, as well as a case where the stress falls on the first vowel, showing that falling-sonority diphthongs are possible as well. These forms were chosen from cases in which the lead language assistant's chosen spelling for the word included VV sequences. The sequences were furthermore typically found to have alternate, more careful pronunciations featuring hiatus.

Table 6.52: Desyllabification of unstressed high vowels, in service of hiatus resolution. Unstressed /u/ and /i/ become the corresponding glides /w/ and /j/, respectively.

orthography	stress-assigned UR	SR	Gloss
⟨igniani⟩	/iŋi'ani/	[i'ŋjani]	"today"
⟨voemba⟩	/vu'emba/	['v ^w emba]	"bean"
⟨antoandro⟩	/antu'andru/	[an't ^w andru]	"mid-day"
⟨akeo⟩	/a'keu/	[a'ke ^w]	"now"

Given the data in the three tables above, we can begin to formulate a process wherein unstressed high vowels become glides adjacent to stressed vowels. The formalization of the rule below states that the environment for gliding is “before or after” a stressed vowel, as there is no agreed-upon notation to state “next to,” but this should be considered a disjunction only as it pertains to formalization. Because of this, there is no reason to throw out the “after” portion just because it is difficult to formalize and because there is only one example ([a'ke^w] “now”: the generalization that adjacency to stress vowels always triggers this type of hiatus resolution is true and should not be sacrificed due to formalization difficulties.

19) Hiatus-Triggered Unstressed Vowel Desyllabification (“Desyllabification”)

$[-\text{consonantal}] \rightarrow [-\text{syllabic}] / \{ \begin{array}{l} [+ \text{syllabic} + \text{stress}] \\ [+ \text{syllabic} + \text{stress}]_ \end{array} \}$

“Vowels become corresponding glides when they occur next to a stressed vowel.”

Not every vowel in Betsimisaraka has a corresponding glide. The mid-glides [e] and [o] are not heard in the language, not surprisingly considering that only unstressed vowels desyllabify and [e] and [o] are not heard in unstressed syllables, as demonstrated in Section 6.2.1. There is no reason to assume that a mid vowel in an unstressed position adjacent to a stress vowel would maintain its place features while becoming a glide, considering that even their full vowel versions

are raised when not stressed. In any case, examples would be almost impossible to find: the orthography-based search method will never turn them up because they are written as the corresponding [+high] segment when unstressed.

The obvious question raised by desyllabification is what happens to low vowels, since low glides are phonetically impossible. As can be seen below, if the unstressed vowel is /a/, it is simply not pronounced when adjacent to another vowel:

Table 6.53: When an underlying form contains a VV sequence featuring an unstressed /a/, that vowel does not surface. The mid-derivation form below is the UR, having had stress assigned and then epenthesis.

orthography	mid-derivation form	SR	gloss
<miboaka>	/mi' b uaka/	[mi' b uka]	“go out” (active)
<fandriagna>	/fan' d rianya/	[fan' d rinya]	“sleep” (gerund)

The deletion of /a/ can be motivated by its lack of corresponding glide. Analytically, this also follows from the way Desyllabification was constructed above, such that all vowels become [-syllabic], so that we have been left with instances of impossible [+low -syllabic +consonantal] segments. These impossible segments are simply deleted; the feature [\pm consonantal] is actually not required for the target of this rule, as only vowels are specified as [+low] anyway.

20) **Deletion of Aberrant Glides**

[+low -syllabic] $\rightarrow \emptyset$

“All low glides are deleted”

The two examples offered in Table 6.44 show the surface realizations of /ua/ and /ia/ as [u] and [i], but this subsection has not yet presented the phonetic output for /ai/ and /au/. That is because those underlying vowel sequences have a different SR that, as previously mentioned, distinguishes Betsimisaraka from OM.

Previous examples in this subsection followed a process wherein unstressed vowels next to stressed ones desyllabified, and showed what exactly this meant for high and low vowels, while ignoring mid vowels which never occur unstressed. In most cases we have seen, the first vowel desyllabifies and the resulting diphthong is of rising sonority, although note [ake^w] (“now”), which has falling sonority. The latter is rare because the most common vowel in the language is [a], and (unlike in OM) the diphthongs [a^w] and [a^j] are almost never heard¹²⁷. Instead, these two vowels coalesce into one with height in between the two, but which is only heard in cases that can be analyzed as hiatus-triggered.

The table below employs, like in Table 6.50, formation of passives to demonstrate hiatus resolution. Included are examples where a root ends with /a/ and where the passive suffixes /u/ or /i/ are used. (The URs of these suffixes are justified in cases where roots are consonant-

¹²⁷One prevalent exception is in the word spelled *aomby* and pronounced universally ['a^wmbi] or even [a^w'umbi]. This word, which describes the hump-backed zebu cattle that represent an intrinsic part of Malagasy culture, can be heard daily, but never as *['ɔmbi]. Looking at /aiza/ (“where”) above, we know that coalescence can occur word-initially. Although it is possible that the exception is lexical, it cannot be cultural, as the OM word, spelled *omby* and pronounced ['umbi] does not have the initial /a/.

final, as seen above in Section 6.1.7)

The data below shows alternation evidence for the mid lax vowels appearing only as a result of underlying hiatus. Without the evidence below, we could not know for certain whether the mid lax vowels [ɔ] and [ɛ] are themselves phonemes. They are heard root-internally as well, but given the alternation examples above, the mid-lax vowels below can also be posited to be the result of coalescence. This allows for a smaller vowel inventory, while adding no more rules than the one necessary to account for the alternations below.

Table 6.54: Formation of mid lax vowels as the surface realization of sequences of /a/ followed by high vowels. Depending whether the /-i/ or /-u/ passive suffix is employed, /a/-final roots will combine with this suffix to form [ɛ] or [ɔ]. Note that the choice of /-ki/ or /-ku/ allomorph of the 1S suffix is due to vowel harmony, as described in Section 6.2.2.

active	root	Suffix		UR (passive)	SR (passive)	gloss (root)
		passive	1S			
[ma'ŋisa]	/isa/	/i/	/ki/	/isa'iki/	[isɛki]	"count"
['mila]	/ila/			/ila'iki/	[ilɛki]	"need"
[mi'asa]	/asa/			/isa'iki/	[isɛki]	"work"
[mi'bata]	/bata/			/bata'iki/	[batɛki]	"lift"
[mi'ova]	/ova/	/u/	/ku/	/ova'uku/	[ovɔku]	"change"

Table 6.55: Examples where the orthography suggests an underlying hiatus and the pronunciation includes a lax mid vowel, which, combined with the morphological data above, suggests a coalescence analysis.

orthography	stress-assigned UR	SR	gloss
<anao>	/a'nau/	[a'nɔ]	(2S)
<atao>	/a'tau/	[a'to]	"do" (root)
<aiza>	/a'iza/	['ɛza]	"where"

The stress assignment to these underlying forms, as shown above, is /a'nau/ and /a'iza/, which in the former case places stress on the low vowel, but in the latter case places stress on the high vowel. Note, as well, that the unusual, final stress in [a'nc] (2S) can be accounted for by stress-assignment preceding vowel coalescence, whereas otherwise this would have to be an exception. Unlike when resolving hiatus by desyllabification or vowel deletion, then, coalescence follows the same pattern regardless of stress assignment. It cannot be simply fed by syllabification where [a^w] and [a^j] are further altered into [ɛ] and [ɔ], and should instead precede the other two hiatus resolution strategies.

21) Vowel Coalescence (“Coalescence”)

[+syl +low][+syl +high αback] → [-high -low -tense αback]

“An underlying hiatus where the first vowel is low and the second high surfaces as a lax mid vowel with the backness value of the high vowel.”

The rule above does not entirely follow in the tradition of rewrite rules, as it has two feature matrices merging into one. There is no standard way to record this in rewrite rules, so in case the rule above is not sufficiently rigorous, the following two rules can take its place. Still, since these rules will happen in immediate sequence, they will continue to be referred to as a single coalescence rule.

21i) Vowel Coalescence I

[+syl +low] → [-low αback -tense] / [+syl +high αback]

21ii) Vowel Coalescence II

[+high] → Ø / [-tense]

"When a low vowel precedes a high vowel, the former becomes a mid-lax vowel with the backness value of the latter. The back vowel is then immediately deleted, motivated here by it following a lax vowel."

We now have rules to account for every possible combination of vowels. To summarize, the first process to occur (because it is not stress-dependent) is the last described: the coalescence process that renders a single lax mid vowel from a low-high vowel sequence. Following that, unstressed vowels found to either side of a stressed vowel are desyllabified. This results in attested glides for high vowels and in one impossible segment for the low vowel. This ordering and the resulting feeding and bleeding processes are shown in the table below.

Table 6.56: Derivation showing the interaction of the three rules proposed in this subsection, with coalescence first, lest it bleed the following rules, and desyllabification feeding deletion. ⊗ is employed in the table for the impossible low glide that is deleted by the final rule.

Gloss	"change" (active)	"change" (passive 1S)	"go out" (active)
UR	/mi-ova/	/ova-u-ku/	/mi-buak/
Coalescence	---	/ovɔku/	---
Stress	/mi'ova/	/o'vɔku/	/mi'buak/
Epenthesis	---	---	/mi'buaka/
Desyllabification	/'mjova/	---	/mi'bu⊗ka/
Deletion	---	—	/mi'buka/
SR	['mjova]	[o'vɔku]	[mi'buka]

Having completed our discussion of hiatus, the last subsection covering vowels, we can list all distinctive features necessary. This chart will not be phonemic, as it must list the surface allophones that result from vowel interactions.

Table 6.57: All distinctive features necessary for [-consonantal] segments. Allophones are shown in parentheses.

		-back	+back		
+high	-low	(j)	(w)	-syl	
		i	u	+syl	
-high		e	o		
		(ɛ)	(ɔ)		
+low			a		

+tense
-tense

Several important changes have been made since the previous iteration of this chart, not the least of which is relabeling all the segments as [-consonantal], since glides share this feature with full vowels, while not being recognized as [+syllabic]. To that end, a column indicating polarities for [\pm syllabic] has been added to the right, as has a column for [\pm tense] to allow for the mid-lax vowels that are the result of coalescence. Certainly /i/ and /u/ could be viewed as [+tense], assigning high vowels a value for that feature, but with no [-lax] high vowels in Betsimisaraka, this cannot be phonologically motivated.

5.3 Final Rule Ordering and Conclusion

The most crucial aspect of rule ordering, as alluded to frequently throughout Chapters 4 and 5, is that Penultimate Stress Assignment must precede Coda-Eliminating Vowel Epenthesis. Out of this order, antepenultimate stress would not exist in the language, but these two processes are also the lynchpins of the phonology. Generally speaking, all consonant alternations at least *can* precede epenthesis, whereas all vowel-based alternations follow stress assignment. Because Epenthesis is the only rule that interacts with both, this section will cover rule ordering for consonant-based alternations in 6.3.1, then vowel-based in 6.3.2. Each section will conclude with a Hasse Diagram, visually displaying the crucial ordering.

6.3.1 Crucial Rule Orderings: Consonants

This subsection presents a prose justification for the final rule ordering for all alternations involving consonants, followed by a visualization of that ordering. In using both approaches, this section can utilize thorough prose analysis for rule orderings together with the diagram's ability to show the big picture all at once. To this end, and to ensure that the two presentation styles work in concert, each justification will be marked with a letter that can be used as reference to the relevant ordering-arrow in the Hasse Diagram in Figure 6.1.

To begin, consider one straightforward aspect of the ordering,

revealed from analysis of Consonant Cluster Place Assimilation, the process whereby the first consonant assumes the place features of the second. This assimilation must precede Lexically-Limited Voiceless-Consonant Deletion (**↓A**), which eliminates voiceless consonants following the SFP /aŋ-/. Section 6.1.4 presented the case wherein nasals from underlying NC clusters preserve the place features of the deleted voiceless consonants, which they would not do if the order were reversed. Following the same counterbleeding reasoning (/aŋ-/ assimilates to /ɸ, β/ before these bilabial fricatives are deleted following it) CC Assimilation must also precede Lexically Limited Bilabial Fricative Deletion (**↓B**).

Returning to crucial orderings, Bilabial Fricative Deletion must also precede Post-Consonantal Fortition (**↓C**), which turns the second consonant in a CC into its [-continuant] equivalent as analyzed in 6.1.3, to bleed it of a chance to operate on /β/. If it did, it would produce [b], an ungrammatical result that these abstract fricatives were introduced to prevent in Section 6.1.5. Fortition must in turn precede CC Reduction (**↓D**) due to counterbleeding: as the cases at the end of section 4.3 showed: consonants can be seen to have undergone fortition even if the consonants that triggered this deletion do not surface.

Returning again to the beginning of this chain, Lexically-Limited Bilabial Fricative Deletion must also precede Bilabial Fricative Clean-Up (**↓E**), which renders labiodental stops from bilabial ones. Clean-Up

would bleed Bilabial Fricative Deletion before the bilabial fricatives could utilize their features distinct from labiodentals, which, again, would be against the point of proposing abstract phonemes. Similarly, Clean-Up must also follow Strident-Coda Deletion ($\downarrow \mathbf{F}$), which deletes all potential strident codas, lest it be allowed to feed it by producing additional /f, v/ for it to delete. If Clean-Up preceded Strident-Coda deletion, bilabial fricatives would merge with labiodental ones and be deleted, rather than exploiting their featural differences to account for the disparate distribution of seemingly root-final /f/s.

Coda Fortition (an aspect of potential coda neutralization, described in 6.1.7) is another process that must not be allowed to apply before Strident Coda Deletion ($\downarrow \mathbf{G}$), since both bilabial and labiodental fricatives would all undergo fortition to bilabial stops, which would not be targeted for elision. Coda Fortition must then follow this rule, again due to counterbleeding.

Similarly, Coda Neutralization *could* also bleed Strident Coda Deletion: if the stridents were neutralized, they would produce aberrant [-anterior -labial -voice +strident] segments. Theoretically, this is not a problem, since Strident Coda Deletion would still delete them, but since there are no strident continuants among the non-anterior coronals—[r] is the resident non-strident continuant—perhaps [\pm strident] cannot be hosted among these place features, meaning these former stridents would not delete? None of these

queries necessitate an ordering between Strident-Coda Deletion and Coda Neutralization, but they are worth considering.

On the other hand, Bilabial-Fricative Clean-Up must precede Coda Neutralization, although in this case to feed it ($\downarrow \mathbf{H}$). Coda neutralization never officially removes the [-strident] feature from / ϕ / to render [tʃ] and though, again, this may not be necessary, it is enough to achieve an ordering.

Almost all other consonant rules crucially must precede Coda-Eliminating Epenthesis. Coda fortition ($\downarrow \mathbf{I}$), Nasal Backing ($\downarrow \mathbf{J}$), and Coda Neutralization ($\downarrow \mathbf{K}$) together, make all segments, within their Dorsal, Nasal, and “Other” domains, the furthest back voiceless stop/affricate; together with the optional nasal-coda elimination ($\downarrow \mathbf{L}$), and Strident-Coda Deletion ($\downarrow \mathbf{M}$), they all make reference to consonants at the right word boundary, and crucially for the analysis in Section 6.1.7, must precede epenthesis.

This only leaves Nasal Dissimilation, which prevents two velar nasals from heading sequential syllables by making the second one an alveolar. The environment for this one is created by Nasal Backing, which creates velar nasal, and Epenthesis, which makes those nasal velars syllable onsets, so Nasal Backing feeds Dissimilation, and must proceed it ($\downarrow \mathbf{N}$). Nasal Dissimilation is also fed by Epenthesis, which provides the second vowel for its / $\eta V \eta V$ / environment, so it is the only

rule that must¹²⁸ follow Epenthesis ($\downarrow O$).

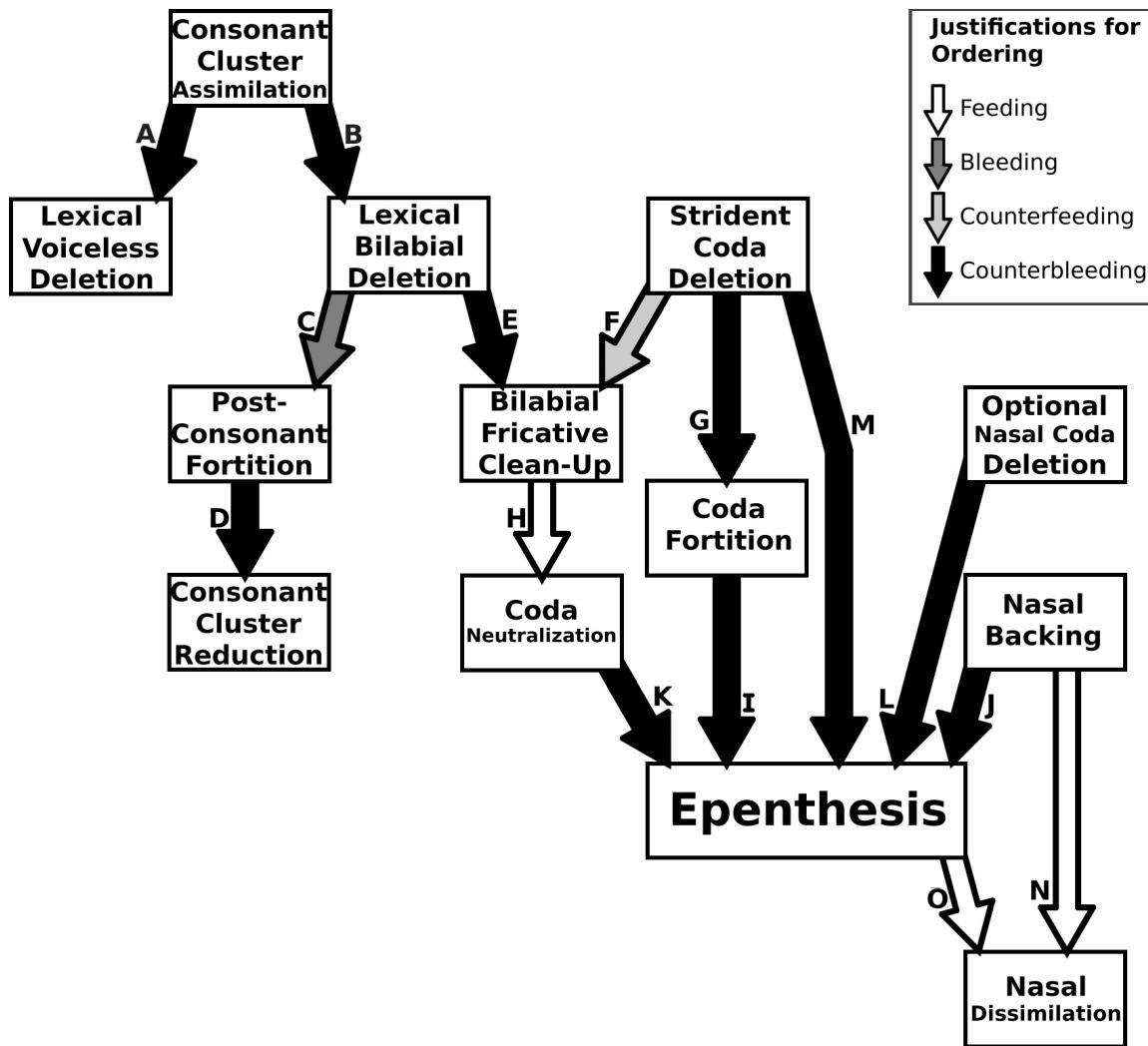


Figure 6.1: Hasse Diagram, representing all crucial orderings in the consonant-affecting rules of Betsimisaraka Malagasy. Arrows represent necessary orderings: rules are placed above, and connected by arrows to, rules they must precede. Letters next to arrows refer to the justification offered for those orderings, although arrow colors summarize which kind of justification is used. The vast majority of orderings, at least considering the formalizations used in this grammar, are opaque: counterbleedings and counterfeedings.

128At the beginning of this subsection, we said that all consonant-affecting rules could precede epenthesis; this was on the understanding that, as mentioned in 6.1.8, a slightly more complex rule could be used in a non-Duke-of-York approach, which did not require the epenthetic vowel.

All crucial orderings in the domain of consonants have been established, almost entirely due to counterbleeding: rules usually must occur later because otherwise they would deprive earlier rules of the environments that would cause them to exist. These cases then represent opacity, along with the single case of counterfeeding. Only four cases involving derivational transparency (three feeding and one bleeding) were used to establish orderings. One important conclusion from the Hasse Diagram above is thus the large amount of opacity in the language.

6.3.1 Crucial Rule Orderings: Vowels

The ordering for vowels is a less complex system, if only because fewer rules need be ordered. The most fundamental step in establishing this order is to begin with Penultimate Stress Assignment: all rules depend directly on the assignment of [\pm stress] in their triggers or environment, would change stress assignment if they preceded it, or else must follow another rule that Stress Assignment must precede.

To begin with, Stress Assignment directly feeds Stress-Conditioned Vowel Heightening, so must precede it ($\downarrow \mathbf{P}$). Non-low vowels are only raised if they are [-stress], a meaningless feature before stress assignment. Two other rules depend on [-stress], so must come after this feature is assigned: Vowel Coalescence ($\downarrow \mathbf{Q}$), in which

sequences of unstressed high vowels adjacent to low vowels will coalesce therewith, producing lax-mid vowels, and the similar process of Vowel Desyllabification ($\downarrow \mathbf{R}$), where unstressed high vowels form diphthongs with adjacent vowels. These, then, are cases of feeding, but these rules are also ordered by *counterfeeding*. Stress is assigned counting backwards based on underlying vowels, but both Coalescence and Desyllabification decrease the number of vowels, which would change where stress is assigned. If syllabification were to come first, examples of final stress, such as those seen in Section 5.3, would never occur. In the Hasse diagram (Figure 6.2), these will be marked as counterfeeding-based orderings, because counterfeeding produces opaque SRs, and one purpose of color-coding the Hasse Diagram is to display the large amount of opacity in the language's alternations.

Vowel Heightening also feeds, and thus must precede, Desyllabification ($\downarrow \mathbf{S}$) to ensure that no mid glide surfaces. Among Coalescence and Desyllabification, the former must precede the latter ($\downarrow \mathbf{T}$) by counterbleeding, lest the low-high vowel clusters surface as low-high diphthongs instead of mix-lax monophthongs. Finally, at the end of this chain, Desyllabification feeds, thus precedes, Aberrant Glide Deletion ($\downarrow \mathbf{U}$). Unstressed, low glides adjacent to stressed vowels are created by Desyllabification and thus could never be deleted without this rule applying first.

Let us now turn to the chain of processes that lead to Feature-Filling Vowel Harmony. This process could not occur without Coda-Eliminating Vowel Epenthesis preceding it ($\downarrow \mathbf{V}$) so it can feed it place-featureless vowels. As motivated in Section 5.2, Epenthesis, in turn, must follow Stress Assignment ($\downarrow \mathbf{W}$), without which antepenultimate stress would never occur. On the surface, epenthetic vowels' place is filled in based on the features of the previous vowel, even if that "vowel" is just the offglide of a diphthongs or even barely remains as merely the [\pm back] feature of a mid-lax vowel¹²⁹. For this reason, Feature-Filling Vowel Harmony must precede both Vowel Coalescence ($\downarrow \mathbf{X}$) and Desyllabification ($\downarrow \mathbf{Y}$).

That ordering will suffice for versions of the grammar where speakers exhibit across-the-board copy epenthesis, with or without having deleted word-final nasals pre-Epenthesis. Other speakers, though, harmonize place-featureless vowels only across nasals, and other epenthetic vowels surface as [a], just like in OM. The Feature-Filling Clean-Up process, which assigns [+low +back] features to all place-featureless vowels, must follow Trans-Nasal Harmony ($\downarrow \mathbf{Z}$): the latter would certainly bleed the former, producing an unattested version of Betsimisaraka much more similar to OM¹³⁰. This last step and

¹²⁹Section 6.2.2 offers no formalizations of the two vowel harmony systems on 1S suffixes, concluding they were due to allomorphy, not allophony. Still, if they were included in the upcoming Hasse Diagram, they would also precede Coalescence and Desyllabification, since [-ki] is heard following [u^j] (/ui/) and [ɛ] (/aj/).

¹³⁰If there is a variety of Malagasy where a UR /olunj/ ("person") has SR *['olunja], with phonemic /o, ɲ/ but [a]-epenthesis, it was not revealed by this study.

this last ordering are unnecessary in a grammar of Betsimisaraka with complete cross-consonantal vowel harmony.

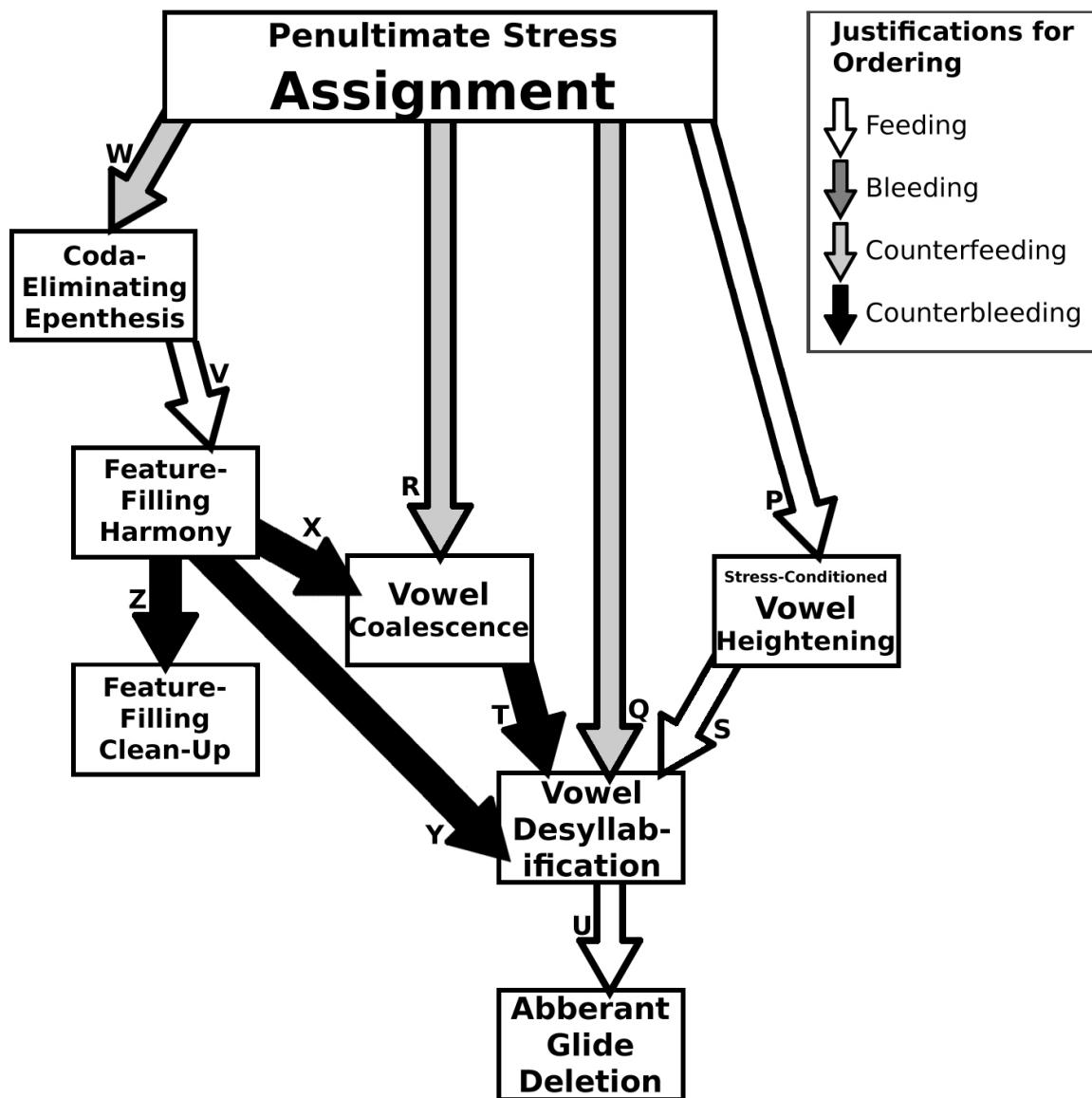


Figure 6.2 Hasse Diagram, representing all crucial orderings in the vowel-affecting rules of Betsimisaraka Malagasy. Arrows represent necessary orderings: rules are placed above, and connected by arrows to, rules they must precede. Letters next to arrows refer to the justification offered for those orderings, although arrow colors summarize which kind of justification is used. Again, as formalized here, these orderings imply opaque generalizations.

All the rules proposed in this chapter have now been ordered, showcasing the seventeen-to-nineteen segmental rules (depending on whether the optional rules are employed) and single, unified prosodic rule necessary to account for all alternations in phonology in Betsimisaraka Malagasy covered in this chapter.

Chapter 7

MAJOR CONTRIBUTIONS AND CONCLUSIONS

As the first phonological grammar of Betsimisaraka Malagasy, and as a work in the understudied field of Malagasy phonology (of any variety), this grammar demonstrates the distinctness of this language variety. In a country where the government's designated language academy states that Official Malagasy is the form everyone speaks, with only minor differences, this work joins Bouwer (2005, 2007) and other previous dialectology research in confronting that idea. Within Malagasy linguistics, the findings from the project aid in comparing and contrasting Malagasy language varieties. The phonological diversity demonstrated here, though, also demonstrates a need for a phonological approach to be taken in dialectology across the island. The findings from this work and works like it also have the power to affect the way language is viewed in general in Madagascar.

Chapter 3 makes one important contribution: it gathers phonological insights proposed by previous authors and gleaned from previously published data to constitute the most thorough review of what is known about the phonology of Official Malagasy and other dialects. Although there has been no lack of analysis in the articles and

other works cited, there has not been the kind of academic back-and-forth that could lead to a consensus on phonological interpretation. The issues of whether Malagasy has prenasalized stops or not, or whether antepenultimate stress is a reflection of an underlying potential coda have been covered in most deeper analyses of the language's phonology. Nevertheless, though Albro (2005) does consider some previous proposals, this is the first work to incorporate so many relevant findings, with pros and cons of each competing hypothesis.

The primary contribution of this work, though, is its description of a previously undescribed phonology. More general works in dialectology agree that Central languages like OM and Eastern varieties like Betsimisaraka are generally believed to be fairly close. Dez's (1963) traditional divide, between /li, ti/ and /di, tsi/ dialects, for example, places both OM and Betsimisaraka on the same side of the isoglosses. This project, though, found several factors distinguishing the two: the presence of /o/ and /ŋ/ phonemes (4.1.1, 4.2.1), regular use of vowel coalescence to [ɛ] and [ɔ] (4.1.2), generalized stress-conditioned vowel-raising (6.2.1), allomorphies on two different 1S suffixes (6.2.2), related-but-distinct word-final consonant neutralization (6.1.7), vowel harmony on epenthetic vowels (6.2.3), and nasal dissimilation (6.1.8). Additionally, OM's general palatalization of dorsal consonants following /i/ was not present in the Betsimisaraka

described. That two seemingly closely related language varieties can be so different illustrates the great need for describing the phonologies of all the island's dialects.

The sounds of many dialects of Malagasy have already been analyzed, but a phonological approach is still required. Just as one example: recall that, in inventory, dialects from the Southeast, such as Taisaka and Southern Betsimisaraka, were said to have the series /ʃ, ʒ, s, z/ where OM (and Northern Betsimisaraka) have /s, z, ts, dz/. What the authors (Deschamps 1936 and Dez 1963) did not mention was how this affects fortition: in languages of the latter type /s, z,/ → /ts, dz/ following other consonants, so is it true in languages of the former type that /ʃ, ʒ/ → /s, z/ in that environment? If so, this would constitute evidence against the phonetic motivation for fortition proposed in 6.1.3: if an underlying /nʃ/ string becomes [ns], there is no excrescent stop easing the transition. But if an excrescent stop were inserted in Taisaka, this would make sense for /ns/, which could surface as the [nts] found in the majority of dialects, but would /nʃ/ become [ntʃ]? If so, since Taisaka is one dialect that pronounces OM /tr̩/ as /tʃ/, this would represent a merger of /r/ and /ʃ/ neutralizing after fortition. Because modern phonological methods do not seem to have been applied to these Southeastern languages, though, these questions remain to be answered.

Potential avenues of investigation remain within Betsimisaraka

phonology. Although minimal pairs were discussed in cases where Betsimisaraka and OM differed, many previous authors cited here provided very thorough listings of minimal pairs for every pair of featurally similar segments. One place to start would be to look at minimal pairs listed for closely-related varieties and see if they exist in Vavatenina.

Another project that could employ the data collected here would be one focusing on variation. This aspect was touched upon in analyses of some alternations. When variation was mentioned, it was typically motivated by interactions with OM, optional rules, or both. Given the great variety found in spoken language, though, a statistical approach to how often different responses were elicited would provide a fuller picture of Betsimisaraka as it is spoken in Vavatenina. Furthermore, all alternations covered here were conducted based on single words or short phrases, leaving the phonology-syntax interface almost entirely unexplored.

Just comparing the description and analysis of Vavatenina Betsimisaraka here with previous accounts, though, can contribute directly to deciphering the Betsimisaraka branch of the still disputed family tree of Malagasy languages. The copy epenthesis pattern found in Vavatenina is unlike the across-the-board [a]-epenthesis found to the west (in OM) and universal [i]-epenthesis to the southwest (in Taisaka). Evidence for (at least) optional copy-epenthesis was found to the north

and northwest: in Takarana and Northern Sakalava, respectively. If the harmony on epenthetic vowels is a Northern feature, one might expect Vavatenina to not exhibit certain features of the Eastern group. Indeed, although Dez (1963) described Southern Betsimisaraka as having [j] in cognates of OM [z], calling this a trait common to Eastern dialects, Vavatenina had [z] as well. Furthermore, Dez's own finding was that Southern Betsimisaraka does not have /o/, which Vavatenina's form did. Vavatenina Betsimisaraka, then, having /o/, not having */j/, and evincing copy-epenthesis, can be grouped with Northern Betsimisaraka and likely with the Northern division of dialects in general. The difference between Vavatenina and even-further-north Maroantsetra's use of 1S possessive pronouns (in 6.2.2), then, can be taken as a further dialectal distinction within Northern Betsimisaraka.

It was never a goal of this work to single-handedly convince the people of Madagascar of the distinctions in their speech varieties. However, the findings in this phonological grammar can and should be used in conjunction with other works of dialectology as evidence for greater diversity than was previously understood.

What we have, then, is a comprehensive analysis of one variety of one dialect of Northern Betsimisaraka, a language spoken by a million people, but whose syntax had been described only once and whose phonology had never before been described. The results of this analysis help us understand how this language variety relates to those

spoken by others in the Betsimisaraka ethnic group and others in Madagascar. The comprehensive generative phonology approach suggests future studies in Northern Betsimisaraka and in other dialects. Finally, together with the dialectologies that have come before—and those that will follow—the findings here can add to an understanding of the depth of linguistic diversity in Madagascar.

REFERENCES

- Adelaar, Alexander. (2013) "Malagasy Dialect Divisions: Genetic versus Emblematic Criteria." *Oceanic Linguistics*, Volume 52, no. 2. Univ. of Hawai'i Press.
- Albro, Daniel. (2005) *Studies in Computational Optimality Theory, with Special Reference to the Phonological System of Malagasy*. Dissertation, UCLA.
- Andrianasolo, Fidèle-Joseph. (1989) *Interprétations fonctionnelles et combinatoires des problèmes de phonologie malgache*. Doctoral Thesis, University of Paris, VII.
- Bakovic, Eric and Colin Wilson (2000) "Transparency, Strict Locality, and Targeted Constraints" *WCCFL 19 Proceedings*, ed. Billerey and Lillehaugen, pp. 43-56. Somerville, MA: Cascadilla Press.
- Baron, Rev. (1893) "Notes on the Betsimisaraka and Tanaraka Dialects." *Antananarivo Annual*, Issue XVII, 54-67.
- Bouwer, Leoni. (2005) "Towards a sociolinguistic profile of Madagascar : issues of diversity" *English usage in Southern Africa Volume 36 Issue 1*, 98-116.
- Bouwer, Leoni. (2007) "Intercomprehension and Mutual Intelligibility Among Southern Malagasy Languages." *Language Matters; Studies in the Languages of South Africa volume 38 issue 2*.
- Bowern, Claire. (2008) *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.
- Broelow, Ellen. (2008) "Stress-epenthesis interactions. In B. Vaux and A Nevins (eds.) *Rules, Constraints, and Phonological Phenomena*, 121-148 Oxford: Oxford University Press.
- Builles, Jean-Michel. (1987) *Le Mérina : description phonologique et grammaticale de la variété de malgache parlée à Tananarive*. Doctoral Thesis, University of Paris V.

- Chelliah, Shobhana L. and William J. de Reuse. (2010) *Handbook of Descriptive Linguistic Fieldwork*. Springer Science & Business Media.
- Chomsky, Noam and Morris Halle. (1968). *The Sound Pattern of English*. Harper & Row.
- Clements, Nick. (1991) "Place of Articulation in Consonants and Vowels: a Unified Theory." *Working Papers of the Cornell Phonetics Laboratory* 5, 77-123.
- Dahl, Otto Christian. (1951) *Malgache et Maanjan. Une comparaison linguistique*. Oslo: Egede Instituttet
- Dalmond, Pierre. (1842) *Vocabulaire et Grammaire pour les Langues Malgaches Sakalava et Betsimisara*. St. Denis, Imprimerie de Lattupe.
- de Lacy, Paul (2002) *The Formal Expression of Markedness*. Doctoral Dissertation, University of Massachusetts, Amherst.
- Descheemaeker, A. (1969) *Vocabulaire des Betsimisaraka Atsimo Zafindriamanaña*, Fianarantsoa, unpubl. ms.
- Deschamps, Hubert. (1936) *Le dialecte antaisaka (Langue Malgache)*. Doctoral Thesis, Faculté des Lettres de Paris. Tananarive: Imprimerie de l'Emyrne Pitot de la Beaujardière.
- Dez, Jacques. (1960) *Le Dialecte betsimalaraka du Sud, région de Nosy Varika*, unpubl.
- Dez, Jacques. (1963) "Aperçus pour une dialectologie de la langue malgache." *Bulletin de Madagascar*
- Dez, Jacques. (1991) *La Linguistique Malgache: Bref Aperçu Historique*. Archives et Documents de la Société d'Histoire et d'Epistémologie des Sciences du Langage (SHESL): Paris.
- Dez, Jacques. (1992) *Lexique Sommaire du Parler Vorimo, D'après L'étude Grammaticale du Dialecte Betsimalaraka du Sud du Pasteur J. Ruud*, unpubl.
- Dziwirek, Katarzyna. 1989. Malagasy phonology and morphology, *Linguistic Notes from La Jolla* 15: 1-30.

- Erwin, Sean. (1996) "Quantity and Moras: An Amicable Separation." *The Structure of Malagasy*. pp 2-31. Los Angeles.
- Ferrand, Gabriel. (1909) *Essai de phonétique comparée du malais et des dialectes malgaches*. Paul Geuthner, Paris.
- Finley, Sarah. (2008) "Alternatives and Additions to Turbidity Theory: Turbid Spreading" *Formal And Cognitive Restrictions On Vowel Harmony* Dissertation for Johns Hopkins University: 72-137.
- Flacourt, Etienne de. (1658) *Dictionnaire de la langue de Madagascar*. COACM, Paris.
- Garvey, Catherine J. (1964) *A Sketch of Malagasy Grammar*. Center for Applied Linguistics, Washington D.C.
- Goldsmith, John. (1985). "Vowel Harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian" *Phonology Yearbook 2*: 253-275.
- Griffiths, David. (1854) *Grammar of the Malagasy Language in the Ankova Dialect*. Woodbridge, Edward Pite.
- Halle, Morris and Clements George N. (1983) *Problem Book in Phonology*. MIT Press.
- Hayes, Bruce. (2009) *Introductory Phonology*. Wiley-Blackwell.
- Jakobson, R., G. Fant, and M. Halle. (1952) *Preliminaries to Speech Analysis: the Distinctive Features and their Correlates*, MIT Press, Cambridge, Ma.
- Kager, Rene. (1999) *Optimality Theory*. Cambridge University Press.
- Kaun, Abigail. (2004) "The typology of rounding harmony." *Phonetically-Based Phonology*, Bruce Hayes, Robert Kirchner, and Donca Steriade, eds., Cambridge: Cambridge University Press, 87-116.
- Keenan, Edward L. and Polinsky, Maria (1998) "Malagasy (Austronesian)" *The Handbook of Morphology*, A. Spencer and A. Zwicky, eds. Blackwell.
- Keenan, Edward L. and Razafimamony, Jean Poulin. (1995) "Malagasy Morphology: Basic Rules" *UCLA Occasional Papers in Linguistics*, volume 17, pp 31-48.

- Kikusawa, Ritsuko. (2006) "A Malagasy (Northern Betsimisaraka) Text with Grammatical Notes" *Journal of Asian and African Studies*, No. 71.
- Kikusawa, Ritsuko. (2007) "Sentences Expressing Three Arguments in Betsimisaraka Malagasy" Lecture, International Institute for Asian Studies, Leiden, The Netherlands.
- Martin, Andrew. (2005) "Loanwords as pseudo-compounds in Malagasy." *Proceedings of the 12th Annual Conference of the Austronesian Formal Linguistics Association, UCLA Working Papers in Linguistics*, Vol. 12, pp. 287-295.
- McCarthy, John J. (2007) *Hidden Generalizations: Phonological Opacity in Optimality Theory*. Equinox.
- Mielke, Jeff. (2008). *The emergence of distinctive features*. Oxford University Press.
- Nelson, Ryan. 2012. "A Principled Account of Malagasy Deverbal Nouns." *Proceedings of the 29th West Coast Conference on Formal Linguistics*, ed. Jaehoon Choi et al., Somerville, MA , pp. 196-204.
- Newman, Paul and Martha Ratliff, eds. (2001) *Linguistic Fieldwork*. Cambridge: Cambridge University Press.
- Odden, David (2013) *Introducing Phonology*. Second Edition. Cambridge University Press.
- Ohala, John J. (1981) "The listener as a source of sound change." In: C. S. Masek, R. A. Hendrick, & M. F. Miller (eds.), *Papers from the Parasession on Language and Behavior*. Chicago: Chicago Ling. Soc. 178 - 203. .
- Paster, Mary. (2009) "Explaining phonological conditions on affixation: Evidence from suppletive allomorphy and affix ordering" *Word Structure* 2.1: 18-47
- Pater, Joe. (1999) "Austronesian nasal substitution and other NC effects," *The prosody-morphology interface* 79, p. 310.
- Paul, Ileana. (1996) "The Active Marker and Nasals in Malagasy." *UCLA Occasional Papers in Linguistics*. pp. 49-57.

- Prince, Alan S. and Paul Smolensky (1993). "Optimality Theory: Constraint Interaction in Generative Grammar." MS, Rutgers University/University of Colorado, Boulder.
- Rajemisa-Raolison. (1969) *Grammaire Malgache*. Centre de Formation Pédagogique, Fianarantsoa.
- Rajaonarimanana Narivelo (1995) *Grammaire Moderne de la Langue Malgache*. L'Asiathèque
- Rakotofiringa, Hippolyte. (1981) "Les Consonnes dites 'Euphoniques' en Malgache-Merina." *Bulletin de l'Institut de Phonetique de Grenoble*, vol. 10/11, p. 251-164.
- Randriamasimanana, Charles (1999) "The Malayo-Polynesian Origins of Malagasy" In Neile A. Kirk & Paul J. Sidwell, editors, *From Neanderthal to Easter Island. A tribute to, and a celebration of, the work of W. Wilfried Schuhmacher presented on the occasion of his 60th birthday*. Association for the History of Language, 1999: 26-43, Melbourne, Australia.
- Sagey, Elizabeth. (1986) *The Representation of Features and Relations in Non-Linear Phonology*. Doctoral Dissertation, M.I.T.
- Simon, Pierre. (1988) "Ny fiteny fahizahy. reconstitution et periodisation du malgache ancien jusqu'au XIV^e siècle." *Travaux et documents 5 du CEROI*. Paris: Institut des Langues et Civilisations Orientales.
- Suh, Y. (2008) "Korean Suffix Allomorphy in OT." *The Proceedings of SUNY/CUNY/NYU Mini-Conference: Linguistics in the Big Apple*.
- Thomas-Fattier, Domique (1982) *Le Dialecte Sakalava du Nord-Ouest de Madagascar. Phonologie. Grammaire. Lexique*. Paris, SELAF.
- Uhrbach, A.J. (1987) A Formal Analysis of Reduplication and its Interaction with Phonological and Morphological Processes. PhD Dissertation. University of Texas.
- Velonandro, ed. (1983) *Lexique des dialectes du Nord de Madagascar*. Tulear: CEDRATOM, Valbonne: CNRS-CEDRASEMI.
- Walker, Rachel. (2009) "(Non-)Adjacency in Harmony Systems" CLS 45.

Appendix A

SWADESH LIST FOR BETSIMISARAKA MALAGASY

The first steps of this project, even before Vavatenina was chosen as a field site, was to elicit a Betsimisaraka Swadesh list. Answers were elicited through charade, picture-identification, and open question asked in Betsimisaraka. The responses below, transcribed in IPA, are limited to those given by the second language assistant and the eleven subjects who participated in Vavatenina, whose version of Northern Betsimisaraka is the focus of this project.

The list below is not exhaustive: if an alternate version was overheard outside of Swadesh list elicitation, it is not included below. Included are all subjects answers that the language assistants judged had the meaning intended, so may perhaps be somewhat biased by that languages assistants own judgments. Slashes indicate slightly different pronunciations that likely correspond to the same UR, and commas indicate unrelated synonyms.

Most nouns below are monomorphemic, although when they are compounds or reduplications, a separate, more literal gloss is provided. Verbs verb-like adjectives are given in the active present form.

- | | | |
|----|----------|--|
| 1. | I (1S) | 'zaha |
| 2. | You (2S) | a'no |
| 3. | We (1PI) | 'atsiki / 'atsika, |
| 4. | this | i'ti (<i>article</i>), i'tiu (<i>noun</i>) |

5.	that	'zeŋi (<i>others too, depending on distance, visibility, etc.</i>)
6.	who	i'zovi
7.	what	'inu
8.	not	'tsi
9.	all	dzi'abi
10.	many	'fontʃi
11.	one	a'reki / a'reka, 'reki / 'reka
12.	two	a'rua, 'rua
13.	big	ma'venti
14.	long	'lava
15.	small	'heli, ma'dinika / ma'diniki
16.	woman	vi'avi
17.	man	la'lahi
18.	person	'oluŋu / 'olu
19.	fish	'ləku
20.	bird	'voruŋu, 'voru
21.	dog	am'bua, ki'va
22.	louse	'hɔ
23.	tree	ka'kazu
24.	seed	'vua, 'mondʒa
25.	leaf	'ravinji / 'ravi
26.	root	'vahatʃa
27.	bark	hudin-ka'kazu ("skin of tree")
28.	skin	'huditʃa / 'huditʃi
29.	flesh	no'futʃu
30.	blood	'ra
31.	bone	ta'holanja / ta'hola
32.	grease	dza'bora
33.	egg	a'tudi, an'tudi
34.	horn	'tandʒuka, 'tandʒuku
35.	tail	'nuhi, vulum-budi ("hair of bottom")
36.	feather	vulu-'vulu, vuluvulum-'boruŋu ("hair-hair of bird")
37.	hair	'vulu
38.	head	'loha
39.	ear	'tadinji / 'tadi
40.	eye	'masu
41.	nose	'oruŋu / 'oru
42.	mouth	'vava
43.	tooth	'nifi
44.	tongue	'lela
45.	claw	hohum-'bibi ("claw of animal")
46.	foot	'tuŋguʃu / 'tuŋgutʃa
47.	knee	lu'halitʃa / lu'halitʃi
48.	hand	'taŋana / 'taŋa
49.	belly	'votʃaka, 'boku
50.	neck	am'buzuŋu, am'buzu
51.	breasts	'nunu
52.	heart	'fo
53.	liver	'ati

54.	drink (V)	migi'aka ("gulp"), 'homaŋa / homa ("consume")
55.	eat (V)	homaŋa / homa, mihinaja (<i>trans.</i>), misa'kafu (<i>intrans.</i>)
56.	bite (V)	ma'ŋekitʃa / ma'ŋekitʃi
57.	see (V)	ma'hita
58.	hear (V)	ma'reŋi
59.	know (V)	ma'hε
60.	sleep (V)	'mandʒi
61.	die (V)	'mati
62.	kill (V)	ma'munu
63.	swim (V)	miulu'maju
64.	fly (V)	ma'ŋembaŋa
65.	walk (V)	mitsangatsangana ("stand-stand")
66.	come (V)	'tonga, 'avi ("from")
67.	lie down (V)	mandžimandži ("sleep-sleep")
68.	sit (V)	man'tutʃi, mipetʃaka (<i>also</i> "stay, inhabit")
69.	stand (V)	mi'tsangana / mi'tsanya
70.	give (V)	maya'mia
71.	say (V)	mi'volanja / mi'vola
72.	sun	ma'sova
73.	moon:	da'volanja / davola, 'volanja / 'vola
74.	star	la'kintanya / la'kinta, 'kintanya / 'kinta
75.	water	'ranu
76.	rain	ora'ŋandžu (<i>N</i> : "rain of day"), mayoranya / mayora (V)
77.	stone	'vatu
78.	sand	a'lanya / a'lana, fasiki / fasika
79.	earth/soil	'tani
80.	cloud	'rondžuju / 'rondžu
81.	smoke (N)	tsemboko / tsemboka, tsembuk-afu ("smoke of fire")
82.	fire	'afu
83.	ash(es)	'dzofu
84.	burn (V)	mi'rehitʃi / mi'rehitʃa (<i>intrans.</i>)
85.	path	'gidzu ("trail"), 'lalaŋa / 'lala ("road")
86.	mountain	ta'neti
87.	red	'mena
88.	green	'mahitso
89.	yellow	'asaka, 'mavu (<i>also</i> "grey")
90.	white	'fotsi
91.	black	'məntiŋi / 'mənti
92.	night	aliŋi / ali
93.	hot	ma'fana, ma'mε
94.	cold	manga'tsiaka
95.	full	'fenu
96.	new	'vaovao
97.	good	'tsara
98.	round	buri'buri
99.	dry	'mεŋi / 'mε
100.	name	a'ŋaranya / a'ŋara

Appendix B

GLOSSARY OF ELICITED FORMS

This glossary contains every Betsimisaraka word used in Chapters 4, 5, and 6, along with a sampling of others elicited for this project. It represents neither a true Betsimisaraka-English dictionary nor even a record of every word elicited during this project. These would unquestionably be useful resources, but were not within the goals of this project.

Words are all listed in the orthography preferred by the primary language assistant, which is based strongly on OM orthography and not distant from the IPA. -Differences from IPA are those below:

Table B.1 Correspondences between Betsimisaraka spelling and IPA. Other regional alphabets employ <ñ> or <ň> for [ŋ]. Aside from <ai> and <ao>, no VV diagraphs are included, as their pronunciation varied between hiatus, hiatus with a glide inserted, and devoicing or elision of the unstressed vowel (see 6.2.4). Also not included are small differences in place, such as the dental stops or changes due to nasal assimilation

vowels		consonants	
orthography	IPA	orthography	IPA
<ai> or <ay>	ɛ	<gn>	ŋ
<ao>	ɔ	<j>	dʒ
<ô>	o	<ts>	ts
<o>	u	<tr>	tʃ
<y>	i	<dr>	dʒ

⟨ai⟩ and ⟨ao⟩ are not true digraphs, as they represent underlying /ai/ and /au/ that coalesce into mid-lax monophthongs. They are included above because—unlike other /VV/ sequences—they have consistent, monophthongic pronunciations. The digraphs ⟨gn⟩, ⟨ts⟩, ⟨tr⟩, and ⟨dr⟩, though, each represent single phonemes. For this reason, they are treated as single letters for the purposes of alphabetical order. ⟨gn⟩ is ordered directly above ⟨n⟩, and ⟨dr⟩, ⟨ô⟩, ⟨ts⟩, and ⟨tr⟩ immediately follow their corresponding letters.

Malagasy orthography traditionally only marks unexpected stress, i.e. stress that is not on the penultimate vowel, or else the antepenultimate vowel if the final syllable is ⟨ka⟩, ⟨tra⟩, or ⟨na⟩. In Northern Betsimisaraka, due to the copy epenthesis pattern described in 6.2.3, it is a bit more complex: stress will be antepenultimate if the final two vowels are the same quality, and the final consonant is ⟨k⟩, ⟨tr⟩, ⟨gn⟩, or (due to nasal dissimilation) ⟨n⟩. For example, ⟨aligny⟩ (“night”) is pronounced [‘alɪŋjɪ] while ⟨amaregny⟩ (“tomorrow”) is [ama’reŋjɪ]. In cases exceptional to the rule above, stress is marked with a grave accent, as in ⟨marìny⟩ (“near”).

Decisions about parts of speech are not based in syntactic theory, but can here be considered to be an aspect of the gloss. For example, ⟨mirônko⟩ (“close”) is glossed as an adjective not because it doesn't have certain characteristics English speakers would associate with verbs, such as tense-marking or the ability to form a complete

sentence with a noun. Rather, it is merely to distinguish it from “close,” the English verb.

French and English etymologies are provided by tradition of Malagasy dictionaries, and for general interest. They can be considered recent loanwords for the purposes of analyses focusing on the subject. Etymologies from Malay or Bantu languages, or even those with unnoticed European origins, are not provided.

All verbs and adjectives are listed in the active, present form, consisting of the present prefix <m-> followed by an SFP (often <a->, <aha->, <agn->, <agnagn->, or <i->) and finally the root. All passives used in the body of this work are included below in their active form. Many words identified below as prepositions, such as <avy> (“from”) and <ambôny> (“above”), also agree for tense with the prefix <t->. In sentences that take place in the past, then, active verbs would have the <n-> prefix and these words would appear as <tavy> and <tambôny>.

Unlike the Swadesh list in Appendix A, the words below are not entirely limited to the Northern Betsimisaraka found in Vavatenina. Because limited elicitation was also possible with two residents of the Northern town of Maroantsetra, words from there are also included below, but are indicated as such.

A

afaka, v. be able to, can
aketo, loc. here
akôho, n. chicken
akôry, pro. how
angy, loc. there
afara, p. behind
afô, n. fire
alohagna, p. in front of
ambany, p. under
ambadigny, p. next to
angy, n. father
ankanjo, n. clothing
agnana, n. leafy greens
agnaragna, n. name
agniny, n. delta
ahitry, n. grass
aiza, pro. where
akaiza, pro. where
akagny, loc. there (*third person*)
akao, loc. there (*second person*)
akeo, adv. now
aketo, loc. here
akondro, n. banana
akôry, pro. how
alahelo, n. sadness
alanagna, n. sand
alavasagna, n. length
alika, n. dog (*Northern B.*)
aligny, n. night
alôha, p. in front of
alôhaka, p. before (*time*)
amaregny, adv. tomorrow
amboa, n. dog
ambôny, p. above
amy, p. in, to, at, for
aminy, p. about, with
ampy, adj. enough
anabavy, n. sister (*to a female*)
anao, pro. you (*singular*)
andre, pro. you (*plural*)
andro, n. day
angitry, n. smell
angôfy, n. fingernail
ankana, n. worm

ankera, n. leftovers, breakfast
antegna, n. self
antoandro, n. mid-day
antody, n. egg
antony, n. reason
antsasagna, adj. half
aomby, n. cow
arabè, n. street
araiky, adj. one
ary, conj. and (*used to connect clauses*)
aroa, adj. two
asa, n. work
asaka, adj. yellow
aty, n. sliver
atiala, n. forest
atsika, pro. we (*inclusive*)
atsimo, n. east
avake, conj. then
avy, p. from
aza. negation used in commands
azafady, int. sorry, excuse me
azo, v. can, be able to

B

baobaby, n. baobab (tree)
bazary, n. market, bazaar
be, adj. big (*also adv. “very”*)
betsaka, adj. many
biby, n. animal
bibilava, n. snake (*lit. “long animal”*)
bisikileta, n. bicycle¹
boka, n. leper
boribory, adj. round
bosy, n. cat (*Maroantsetra.*)
bôtra, adj. chubby

D

davôlagna, n. moon
dia, pron. 3S
dimy, adj. five
dômy, adj. die, dice

¹ French “bicyclette” (bicycle)

E

elatra, *n.* wing
enigny, *adj.* six
efa, *adv.* already
efatra, *adj.* four
eto, *loc.* here

F

fa, *conj.* but
fady, *n.* taboo
fagnôrogno, *n.* fanorona (*a popular board game*)
faly, *adj.* happy
farany, *adv.* most
fasiky, *n.* sand
fao, *conj.* but
fe, *n.* leg
feno, *adj.* full
fianaragna, *n.* school
firy, *pro.* how much/many
fito, *adj.* seven
foagna, *adj.* empty
folera, *n.* flower
folo, *adj.* ten
fontry, *adj.* many
forsety, *n.* fork²
fô, *n.* heart
fôhy, *adj.* short
fômba, *n.* way
fôtaka, *n.* mud
fôtsy, *adj.* white

G

gidro, *n.* monkey (*or similar creature*)
gijo, *n.* path
gisy, *n.* goose
goelahy, *adj.* big (*of a person*)

H

hafa, *adj.* other
hao, *n.* louse

² French “la fourchette” (*the fork*)

hariva. tonight (side)

havagna, *n.* family
havanagna, *n.* right (side)
havia, *n.* left (side)
hehy, *v.* laugh (*Northern B.*)

helatra, *n.* wing
hely, *adj.* little

hena, *n.* meat
hevity, *n.* idea

herinandro, *n.* week
hasin-drano, *n.* fish

hegnatra, *n.* shame
hevity, *n.* idea

hira, *n.* song
ho. future marker used for non-active verbs

hoatrino, *n.* how much (*cost*)

hodin-kakazo, *n.* bark (*lit. “skin of wood”*)

hoe, *p.* of

hoditra, *n.* skin

hômanga, *v.* eat, drink

I

i, *d.* the

ia, *int.* yes

ihany, *adv.* only

igny, *d.* that

igniany, *adv.* today

ilay, *d.* the (*aforementioned*)

ino, *pro.* what

iregny, *p.* those

isaka, *adj.* every

isany, *adj.* every

ity, *pro.* this

izao, *adv.* now

izegny, *pro.* this/these

izôvy, *pro.* who

izy, *pro.* he, she, it

J

jabôra, *n.* fat

jiaby, *adj.* all

jôfo, *n.* ash

K

- kakazo**, *n.* wood
kamarady, *n.* friend
kamo, *adj.* lazy
karaha, *adj.* like, similar
karazany, *n.* kind, type
kibo, *n.* belly
kilasy, *n.* class³
kintagna, *n.* star
kiso, *n.* knife
kivà, *n.* dog
koa, *adv.* too, also
kobay, *n.* stick (*for striking with*)
kôliny, *n.* collar
kôro, *n.* course, class⁴

L

- lafariny**, *n.* flour⁵
lafo, *adj.* expensive
làkana, *n.* canoe
lakilè, *n.* key⁶
lalagna, *n.* road
lalàna, *n.* law⁷
lalahy, *n.* man (male)
laligny, *adj.* deep
lamba, *n.* cloth, clothing
lambo, *n.* pig, boar
lagny, *n.* stench
lany, *adj.* all gone, diminished
laoko, *n.* fish
lasety, *n.* dishes⁸
latabatra, *n.* table⁹
lava, *adj.* long, tall
lavaka, *n.* hole
lavak-ôrogno, *n.* nostril (*lit.* “hole of nose”)
lavitry, *adj.* far

3 French “classe” (class)

4 French “cours” (course)

5 French “la farine” (the flour)

6 French “la clé” (the key)

7 French “la loi” (the law)

8 French “l’assiette” (the plate)

9 French “la table” (the table)

lela, *n.* tongue (*phys.*)

lera, *n.* hour¹⁰

loatra, *n.* surplus, too much

lohalitry, *n.* knee

loko, *n.* color

lôha, *n.* head

M

- madiniky**, *adj.* little
mafana, *adj.* hot
mahafantatra, *v.* know
mahafaty, *v.* kill
maharegny, *v.* hear
mahety, *adj.* narrow
mahia, *adj.* thin (*of people*)
mahita, *v.* see
mahitigny, *adj.* black
mahitso, *adj.* green
maigny, *adj.* dry
mahafantatra, *v.* know
(something, someone)
mahay, *v.* know, be able to
mahampy, *v.* suffice
maharitry, *v.* last, take (time)
mahatombo, *v.* grow up
maintsy, *v.* (*used as part of expression* tsy maintsy
meaning “must”)
malaka, *v.* take
malalaka, *adj.* wide
malazo, *v.* wilt
malemy, *adj.* soft
mama, *n.* mother (*informal*)
mamafa, *v.* wipe, sweep
mamay, *adj.* hot
mamaky, *v.* split, read
mamangy, *v.* visit
mambaby, *v.* carry, support
mamboly, *v.* grow
mamonjy, *v.* save
mamono, *v.* kill
mamopoko, *v.* hit
mamôgno, *v.* wrap

10 French “l’heure” (the hour)

- magnambara**, v. tell
magnambela, v. allow
magnamboatra, v. fix
magnamia, v. give
magnampy, v. help
magnano, v. make, put, do
magnaraka, v. follow
magnatogno, v. approach
magnatsatoka, v. stab
magnavotro, v. redeem
magnekirity, v. bite
magneno, v. sing
magnerigny, v. repeat
magneriky, adj. bad, unfavorable
magnerigny, v. repeat
magnisa, v. count
magnoro, v. burn
magnôragna, v. rain
magnisa, v. count
manakalo, v. exchange, barter
manakôry, pro. how
manao, v. do
mandàka, v. kick
manohy, v. tie together
manasa, v. wash
mandaniha, v. go
manapaka, v. cut
mandalo, v. pass, drop by
mandeha, v. go
mandidy, v. cut
mandoa, v. vomit
mandòko, v. color
mandrahognoy, v. tell off
mandrakotro, v. cover
mandregnyn, adj. morning
mandry, v. sleep
mandrimandry, v. lie down, recline ("sleep" redup.)
mandritry, v. dry
mandregnyn, v. hear
mandrôra, v. spit
manetsigny, v. plug
mangady, v. dig
mangatsiaka, adj. cold
mangôro, v. hunt (*Maroantsetra*)
manify, adj. thin (of material)
manohy, v. continue
manoloko, v. color
manomboko, v. begin
mantotry, v. sit
marigny, adj. correct
mariny, adj. near (*Maroantsetra*)
marivo, adj. shallow, near
maro, adj. many
masigny, adj. salty
maso, n. eye
masôva, n. sun
matahotro, v. fear
matanjaka, a. strong
maty, v. die (also adj. dead)
matevy, adj. thick
matify, adj. thin (of material)
matomboko, adj. thick
mavesatra, adj. heavy
maventy, adj. big
megntra, adj. shy
mena, adj. red
miadagna, adj. slow
miady, v. fight
miaigny, v. breathe
miampita, v. cross (*intr.*)
mianatra, v. study
miaraka, v. mix
miasa, v. work
mibata, v. lift
miboaka, v. go out
midoka, v. flatter
mihaza, v. hunt
miheritrerity, v. think
mihinanga, v. eat, drink
mijorotro, v. pull
migiaka, v. drink
mihezaka, v. try
mikaroagna, v. search, research
mila, v. need
milelatra, v. lick
mimoehy, v. laugh
mino, v. believe
mipetraka, v. live, sit, stay
miolomagno, v. swim

miompy, *v.* breed
miôdy, *v.* go home
miôva, *v.* change, exchange
mirehitry, *v.* burn
miregny, *v.* sing
mirônko, *a.* close
misaotro, *v.* thank
miseky, *v.* swim
misy, *v.* there is
misotro, *v.* blow, drink
mitagna, *v.* hold
mitandrigny, *v.* pay attention
miteny, *v.* speak
mitokiky, *v.* laugh
mitraotro, *v.* meet
mitsangana, *v.* stand
mitsentsitry, *v.* suck, suckle
mitsonjo, *v.* perceive from afar
mividy, *v.* buy
mivily, *v.* turn
mivory, *v.* assemble for a meeting
mivôlagna, *v.* say, speak
mofo, *n.* bread
morogno, *n.* edge, border

N
ny, *p.* of (*in genitive expressions*)
nify, *n.* tooth
nohigny, *n.* tail
nono, *n.* breast

O
oviagna, *pro.* when
ohiny, *n.* tail

Ô
ôhatra, *n.* example
ôlogno, *n.* person
ôlombelo, *n.* human
ôragna, *n.* shrimp
ôragnandro, *n.* rain (*lit.* “rain of day”)
ôrogno, *n.* nose

P
papa, *n.* father (*informal*)
pati, *n.* pasta¹¹
paomy, *n.* apple¹²
piso, *n.* cat

R
ra, *n.* blood
rahalaly, *n.* brother (*to a male*)
rahavavy, *n.* sister (*to a female*)
rano, *n.* water
ravigny, *n.* leaf
reny, *n.* mother
ritry, *adj.* dry
ronono, *n.* milk
rôndrognو, *n.* cloud

S
sakafo, *n.* food
saosy, *n.* sauce¹³
sarotro, *adj.* difficult
sasany, *adj.* half
satria, *conj.* because
sy, *conj.* and
sira, *n.* salt
siramamy, *n.* salt (*lit.* sweet salt)
sisa, *n.* rest (remaining portion)
sokajiny, *n.* type, kind
sôfy, *n.* ear
sôma, *n.* game
sôrognو, *n.* high tide
sôtro, *n.* spoon

T
tady, *n.* rope
tadigny, *n.* ear
tagnana, *n.* hand
tanàgna, *n.* hand
taheza, *n.* back
tandroko, *n.* horn

11 French “pâtes” (pasta)
 12 French “pomme” (apple)
 13 French “sauce” (“sauce”)

tanety, *n.* mountain, high ground
tany, *n.* soil, earth
taolagna, *n.* bone
taogno, *n.* year
tara, *adj.* late
taranaka, *n.* descendant
tavela, *adj.* remaining, left
tavoangy, *n.* bottle
tehigny, *n.* branch
telo, *adj.* three
tegna, *adv.* very
tenda, *n.* throat
teny, *n.* word, language
titezagna, *n.* bridge
titezagna, *n.* bridge
toetra, *n.* behavior
tongotro, *n.* foot, leg
tôkony, *v.* should
tôgno, *v.* tuna¹⁴
tônga, *v.* come

TR

trandraka, *n.* tenrec
tratra, *n.* chest

TS

tsara, *adj.* good, beautiful
tsary, *adv.* never
tsemboka, *n.* smoke
tsembok-afø, *n.* smoke (of fire)
tsy. not
tsihy, *n.* mat
tsinay, *n.* guts

V

vady, *n.* husband/wife, spouse
vary, *n.* rice, meal
vahaza, *n.* Westerner, white person
vahiny, *n.* guest
vahitra, *n.* root
valilahy, *n.* brother-in-law

14 French “thon” (tuna)

valo, *adj.* eight
vao, *adj.* new (often reduplicated vaovao)
vanio, *n.* coconut
vapaza, *n.* papaya
vato, *n.* stone
vava, *n.* mouth
velogno, *v.* live
viavy, *n.* woman
vilagniny, *n.* cooking pot
vitsivitsy, *adj.* some
voa, *n.* seed, grain
voan-dalagna, *n.* present from a trip (*lit.* “seed of the road”)
voangy, *n.* orange
voan-kazo, *n.* fruit (*lit.* “seed of wood”)
voemba, *n.* bean
volo, *n.* hair
volum-borogno, *n.* feather (*lit.* “hair of bird”)
volovolo, *n.* feather (“hair” redup)
vorogno, *n.* bird
vovoko, *n.* dust
vôla, *n.* precious metal, money
vôlagna, *n.* month
vôrogno, *n.* bird
vôtraka, *n.* belly
vôzogno, *n.* neck

Z

zaha, *pro.* I
zahay, *pro.* we (exclusive)
zaho, *pro.* I
zareo, *pro.* they
zarikô, *n.* bean¹⁵
zato, *adj.* hundred
zavatra, *n.* thing, something
zaza, *n.* child
zazakely, *n.* baby
zegny, *pro.* that
zinga, *n.* pitcher-like water dipper
zôky, *n.* older sibling

15 French “les haricots” (the beans)

Appendix C

IRB APPROVAL LETTERS

As this project involved working with human subjects, Internal Review Board approval was required. On the following pages, approval letters are provided from 2011, 2012, and 2013. Note that these letters are written to Jeffrey Heinz, not the author: according to NSF regulations, the professor in charge of the dissertation is the principal investigator and the writer of the dissertation (the author) is the co-PI. As this project was funded by the NSF, these same roles were used on the IRB applications. All human subject interaction was thus supervised by the co-PI, who was in turn supervised by the PI.



RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: May 31, 2011

TO: Jeffrey Heinz
FROM: University of Delaware IRB

STUDY TITLE: [246417-1] Doctoral Dissertation Research: Phonology of Betsimisaraka Malagasy

SUBMISSION TYPE: New Project

ACTION: APPROVED

APPROVAL DATE: May 31, 2011

EXPIRATION DATE: May 30, 2012

REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # 6

Thank you for your submission of New Project materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding and continue throughout the study via a dialogue between the researcher and research participant. The requirement of documentation of consent has been waived for this study. Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

Figure C.1: IRB approval letter for 2011



RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: June 14, 2012

TO: Jeffrey Heinz
FROM: University of Delaware IRB

STUDY TITLE: [246417-2] Doctoral Dissertation Research: Phonology of Betsimisaraka Malagasy

SUBMISSION TYPE: Continuing Review/Progress Report

ACTION: APPROVED

APPROVAL DATE: June 14, 2012

EXPIRATION DATE: June 13, 2013

REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # 6

Thank you for your submission of Continuing Review/Progress Report materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

Figure C.2: IRB approval letter for 2012



RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: June 13, 2013

TO: Jeffrey Heinz
FROM: University of Delaware IRB

STUDY TITLE: [246417-3] Doctoral Dissertation Research: Phonology of Betsimisaraka Malagasy

SUBMISSION TYPE: Continuing Review/Progress Report

ACTION: APPROVED

APPROVAL DATE: June 13, 2013

EXPIRATION DATE: June 12, 2014

REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # 7

Thank you for your submission of Continuing Review/Progress Report materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

If you have any questions, please contact Jody-Lynn Berg at (302) 831-1119 or jlberg@udel.edu. Please include your study title and reference number in all correspondence with this office.

Figure C.3: IRB approval letter for 2013