

A GRAMMAR OF IRARUTU,
A LANGUAGE OF WEST PAPUA, INDONESIA,
WITH HISTORICAL ANALYSIS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAI‘I AT MĀNOA IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

LINGUISTICS

AUGUST 2014

By

Jason Alexander Johann Jackson

Dissertation Committee:

Robert A. Blust, Chairperson
Lyle Campbell
Patricia Donegan
William O’Grady
Ron Bontekoe

Keywords: Irarutu, Phonology, Syntax, Historical linguistics, Austronesian, Papua

This work is dedicated to the speakers of Irarutu
—without you it would not exist—
in hopes that the spoken language perseveres well into the future.

ACKNOWLEDGEMENTS

Thank you, Allah. Nearly a decade ago I started working with South Halmahera-West New Guinea languages, focused in particular on Irarutu. Numerous notes of gratitude are in order for the advice, support, and constructive challenges that were provided by the following people.

First, thank you to the language consultants: Oktavianus Barry, Loruhamas Esuru, Baba Damianus Fenentruma, Yustus Murmana, Bapak Willy Nafurbenan, Otis, Abraham Pomsaru, Adam Refideso, Taher Riroma (Roromana), Sakeus Sasepa, Ibu Beti Wamburie, Clemens Warfete, and Yohanis Werfete. Thank you also, to faculty at Universitas Negri Indonesia di Papua, Cendrawasih (UNIPA): in particular, Andreas Deda and his family, Hengky Mofu, Fersid Mora, Yusuf Sawaki, Yafed Sufi, Hugo Warami, and many others who offered their kindness and friendship. At the Bintuni Office of Statistics, Ibu Suti deserves special recognition.

In addition to numerous faculty at the University of Hawai'i at Manoa (UHM), two special notes of gratitude go to my dissertation advisor, Dr. Robert A. Blust, and previous advisor, Dr. Michael L. Forman, who, through their complementary approaches, helped me develop as a linguist and scholar. I would also like to thank specifically Dr. William O'Grady, Dr. Patricia Donegan, Dr. Ronald Bontekoe, Dr. Lyle Campbell, Dr. Gregory Lee, Dr. David Stampe, Dr. Yuko Otsuka, Dr. Amy Schafer, Dr. Ann Peters, and Dr. Uli Kozok, as well as staff at UHM, former lab supervisor Kurt Brunner, for expert technical knowledge, and former Linguistics department secretary Wendy Onishi. To our department's current, wonderful secretaries, Ms. Jen Kanda and Ms. Nora Lum, thank you for your patience and encouragement.

I am compelled to express my utmost gratitude to my family: my mother Linnea Jackson, my father Dr. Robert Jackson, and my step-mother Eva Mock (whose editing was beyond value); siblings, Vittorio and Liisa, and to my wife, Dr. Kyuseek Hwang Jackson, and our children, Darius and Matilda, for their support, patience, and existence. I am also very grateful for my extended 'ohana. It is likely that there are a number of other people who I should have mentioned here specifically, but have not. I thank you nonetheless. Peace be upon you.

Various stages of this work were supported in part by the UHM Linguistics Department Endowment fund, the Graduate Student Organization travel fund, and an Arts and Sciences award.

ABSTRACT

Irarutu is an Austronesian language that has been classified in the literature as a member of the South Halmahera-West New Guinea subgroup. Some differences with other languages of that subgroup can be attributed to extensive historical contact with speakers of one or more neighboring Trans-New Guinea phylum languages. The Irarutu language is considered endangered. It is spoken fluently by fewer than 6000 speakers on the Bomberai Peninsula in West Papua, Eastern Indonesia, but the total ethnic population is closer to 10,000 people. Nearly all Irarutu speakers know Indonesian, the national language, due to the education system, mass media, and economic factors. This has caused language use to decline, particularly among younger people. This dissertation provides background information (Ch. 1) and describes Irarutu phonology and morphophonology (Ch. 2), morphology and syntax (Ch. 3), and historical phonology as well as diagnostics for classifying the language (Ch. 4). Supplementary materials are provided in several appendices.

The phoneme inventory has fifteen consonants and seven vowels, including a labiopalatal high vowel. In the dialect described in this dissertation, voiced stops are phonetically prenasalized. Consonant clusters abound. Deletion is the strong form of a process called ‘vowel reduction’ that contributes to the complexity of Irarutu consonant clusters. Despite the relatively complex vowel system, native speakers feel that their language has a very consonantal character.

Irarutu has SVO word-order, prepositions, and most modifiers follow their heads, but possessors precede possessed nouns. Furthermore, a contrast between alienable and inalienable possession is expressed morphologically. There are no case markers, but several verbalizing morphemes, including subject markers, an active verb marker *fi-*, and an infinitival/habitual marker *na-* are used in the language. Serial verb constructions are frequent in naturalistic language data. Topicalization, negation, and politeness are achieved through the use of clitics, *=ro* ‘topic’, *=ti* ‘negative’, *=o* ‘polite’.

It is hoped that the present grammar helps this particular language continue to be spoken in the future, by boosting awareness of the language outside of its traditional location and providing resources for its maintenance. Documentation materials of Irarutu can be accessed at the University of Hawai‘i’s digital language archive, Kaipuleohone.

Email inquiries can be addressed to: jjackson@kuboa.com

TABLE OF CONTENTS

Acknowledgments	iv
Abstract	v
List of Maps	xiii
List of Tables	xiv
List of Figures	xvi
List of Abbreviations and Symbols	xviii
Glossing Conventions	xix
 Chapter 1. Introduction.	 1
1.1 Background information	1
1.1.1 Names for the language.	2
1.1.2 Location.	3
1.1.3 Population	7
1.1.4 Dialects.	7
1.1.5 Writing system	9
1.2 The Irarutu community	10
1.2.1 Social structure, economy, and history	10
1.2.2 The Linguistic situation and contexts for use	11
1.2.3 Language endangerment and language documentation	12
1.2.3.1 The Problem of language endangerment	13
1.2.3.1.1 Why Irarutu?	13
1.2.3.1.2 State of endangerment	14
1.2.3.2 Language documentation and language description.	14
1.2.3.2.1 Research in Papua	15
1.2.3.2.2 Metadata	16
1.2.3.2.3 Equipment and software	18
1.2.3.2.3.1 Digital media	18
1.2.3.2.3.2 Equipment	18
1.2.3.2.3.2.1 Audio recorder and microphones.	19
1.2.3.2.3.2.2 Video camera	19
1.2.3.2.3.2.3 Laptop computer	19
1.2.3.2.3.3 Software	20
1.2.3.2.4 Archiving	20
1.3 Previous publications	20
1.3.1 Descriptive publications	21
1.3.1.1 Cowan 1953	21
1.3.1.2 Anceaux 1958, 1961, and Smits and Voorhoeve 1992	22
1.3.1.2.1 Anceaux 1958	23
1.3.1.2.2 Anceaux 1961	25
1.3.1.2.3 Smits and Voorhoeve 1992	26
1.3.1.3 Voorhoeve 1989, 1995	26
1.3.1.3.1 Voorhoeve 1989	26
1.3.1.3.2 Voorhoeve 1995	30
1.3.1.4 Matsumura 1991 and Matsumura and Matsumura 1991	31

1.3.1.4.1 Matsumura 1991	31
1.3.1.4.2 Matsumura and Matsumura 1991	33
1.3.1.5 van den Berg and Matsumura 2008	36
1.3.2 Work on nearby AN languages	37
1.3.3 Comments on neighboring Papuan languages	38
1.3.4 Vernacular and other publications	39
1.4 Overview	40
1.4.1 Methodology	40
1.4.2 Organization	42
Chapter 2. Phonology and Morphophonology	43
2.1 Phoneme inventory	43
2.1.1 Consonants	44
2.1.1.1 Stops	45
2.1.1.2 Fricatives	49
2.1.1.3 Nasals	50
2.1.1.4 Rhotic	50
2.1.1.5 Glides	51
2.1.2 Vowels	52
2.1.2.1 Monophthongs	52
2.1.2.2 Diphthongs	54
2.2 Irarutu feature set	56
2.2.1 Manner features	56
2.2.2 Place features	56
2.2.3 Laryngeal feature	57
2.3 Allophony	57
2.3.1 Consonant allophony	57
2.3.1.1 Release	57
2.3.1.2 Place assimilation: palatalization	58
2.3.1.3 Trilling and rhotic deletion	59
2.3.1.4 Syllabification	59
2.3.2 Vowel allophony	60
2.3.2.1 Careful speech vowel allophony	60
2.3.2.1.1 Schwa epenthesis	60
2.3.2.1.2 [ɪ] epenthesis	61
2.3.2.2 Casual speech vowel allophony	61
2.3.2.2.1 Vowel reduction	61
2.3.2.2.2 Alternation of stressed vowels with [ɪ]	62
2.3.2.2.3 Aberrant palatalization	62
2.3.2.2.4 Palatal glide fortition	62
2.3.2.2.5 Mid-vowel laxing	63
2.4 Suprasegmental phenomena: the syllable and stress	63
2.4.1 Phonotactics, syllabication, and syllable types	63
2.4.1.1 Word initial consonant clusters	64
2.4.1.2 Word medial consonant clusters	68
2.4.1.3 Word final consonant clusters	70

2.4.1.4 Syllabication	71
2.4.1.5 Major and minor syllables	73
2.4.2 Word structure	75
2.4.2.1 Canonical word shape	75
2.4.2.2 Stress and accent	76
2.4.2.3 Resyllabification	78
2.4.2.4 Derived diphthongs	78
2.4.2.5 Nasalization	80
2.5 Loanword phonology	80
2.6 Morphophonemics	82
2.6.1 Allomorphy in verbs	82
2.6.1.1 Subject markers	82
2.6.1.1.1 Nasal place assimilation	83
2.6.1.1.2 Voiceless obstruent fortition	83
2.6.1.2 Stative marker	84
2.6.2 Reduplication	84
2.6.2.1 Full reduplication	85
2.6.2.2 Prefixed partial reduplication	85
2.6.2.3 Suffixed partial reduplication	86
2.6.3 Compounding	86
2.6.4 Cliticization	87
2.7 Summary	87
Chapter 3. Morphology and syntax	88
3.1 Syntactic categories	88
3.1.1 Word formation processes	89
3.1.1.1 Reduplication	90
3.1.1.2 Compounding	91
3.1.1.3 Borrowing and semantic shift	93
3.1.2 Open classes	93
3.1.2.1 Nouns	93
3.1.2.1.1 Function of nouns	94
3.1.2.1.2 Noun morphology	95
3.1.2.1.2.1 Possessive morphology	95
3.1.2.1.2.2 Nominalizer <i>we(n)-</i>	96
3.1.2.1.2.3 Vocative <i>-o/a</i>	97
3.1.2.1.3 Distribution of nouns	97
3.1.2.1.3.1 Clause-level distribution of nouns	97
3.1.2.1.3.2 Phrase-level distribution of nouns	98
3.1.2.1.3.2.1 N precedes adjunct	98
3.1.2.1.3.2.2 N follows adjunct	99
3.1.2.2 Verbs	99
3.1.2.2.1 Function of verbs	100
3.1.2.2.2 Verb morphology	100
3.1.2.2.2.1 Subject marking on the verb	101
3.1.2.2.2.2 Incidental stative <i>ta-</i>	103

3.1.2.2.2.3 Active <i>fi-</i>	104
3.1.2.2.2.4 Stative <i>m-</i>	105
3.1.2.2.2.5 Infinitival/Habitual <i>na-</i>	106
3.1.2.2.2.6 Verbal suffixes <i>-fe</i> , <i>-i</i> , and <i>-nya</i>	107
3.1.2.2.3 Distribution of verbs	108
3.1.2.2.3.1 Clause-level distribution of verbs	108
3.1.2.2.3.2 Phrase-level distribution of verbs	110
3.1.2.3 Adjectives	111
3.1.2.4 Adverbs	114
3.1.2.4.1 Sentential adverbs	115
3.1.2.4.2 Directional/Locational adverbs	115
3.1.2.4.3 Degree adverbs	116
3.1.2.4.4 Manner adverbs	117
3.1.2.4.5 Temporal adverbs	117
3.1.3 Closed syntactic categories	118
3.1.3.1 Pro-forms	119
3.1.3.1.1 Pronouns	119
3.1.3.1.1.1 Personal pronouns	119
3.1.3.1.1.2 Possessive pronouns	120
3.1.3.1.1.2.1 Alienable possessive pronouns	120
3.1.3.1.1.2.2 Inalienable possessive pronouns	122
3.1.3.1.1.3 Reflexive and reciprocal pronouns	127
3.1.3.1.1.4 Demonstrative pronouns	128
3.1.3.1.1.5 Indefinite pronouns	129
3.1.3.1.1.6 Relative pronouns	129
3.1.3.1.2 Other pro-forms: question words	129
3.1.3.1.3 Other pro-forms: pro-sentences, pro-clauses, and pro-verbs	131
3.1.3.2 Noun adjuncts	132
3.1.3.2.1 Role markers	132
3.1.3.2.2 Determiners	135
3.1.3.2.2.1 Definiteness and specificity	136
3.1.3.2.2.2 Demonstratives	138
3.1.3.2.2.3 Deictic demonstratives	139
3.1.3.2.2.4 Discourse demonstratives	139
3.1.3.2.3 Quantifiers and numerals	139
3.1.3.2.3.1 Quantifiers, mass nouns, and plurality	140
3.1.3.2.3.2 Numerals and count nouns	141
3.1.3.2.4 Classifiers	143
3.1.3.3 Verb adjuncts	145
3.1.3.3.1 Tense, aspect, and modality auxiliaries	145
3.1.3.3.1.1 Tense	147
3.1.3.3.1.2 Aspect	147
3.1.3.3.1.2.1 Non-progressive <i>du</i>	147
3.1.3.3.1.2.2 Progressive <i>bu</i>	148
3.1.3.3.1.2.3 Completive <i>su</i>	149
3.1.3.3.1.3 Modality	150

3.1.3.3.1.3.1 Desiderative <i>g-</i>	150
3.1.3.3.1.3.2 Irrealis <i>mo</i>	150
3.1.3.3.2 Other auxiliaries and verbal particles	150
3.1.3.3.2.1 Other auxiliaries	151
3.1.3.3.2.2 Verbal particles	151
3.1.3.4 Conjunctions	154
3.1.3.4.1 Coordinators	154
3.1.3.4.2 Subordinators	156
3.1.3.5 Miscellaneous parts-of-speech	158
3.1.3.5.1 Clitics	158
3.1.3.5.1.1 Topic marker <i>=ro</i>	159
3.1.3.5.1.2 Negative marker <i>=ti</i>	161
3.1.3.5.1.3 Politeness marker <i>=e/o</i>	162
3.1.3.5.2 Copula <i>rau</i> and existential <i>o</i>	163
3.1.3.5.3 Interjections	164
3.2 Grammatical system	165
3.2.1 Grammatical relations	165
3.2.1.1 Grammatical relations in main clauses	165
3.2.1.2 Grammatical relations in relative clauses	167
3.2.2 Word order	167
3.2.3 Transitivity	169
3.2.3.1 Intransitives	169
3.2.3.2 Transitives	170
3.2.3.2.1 Nominative-accusative alignment	171
3.2.3.2.2 Split-S alignment	172
3.2.3.3 Double object pattern	173
3.2.4 Agreement	175
3.2.4.1 Agreement in regular verbs	175
3.2.4.2 Agreement in irregular verbs	176
3.2.4.3 Agreement in serial verb constructions	176
3.2.4.4 Agreement in auxiliaries and other categories	177
3.3 Complex clauses	177
3.3.1 Complement clauses	177
3.3.2 Relative clauses	178
3.3.2.1 Function of RCs	178
3.3.2.2 The relativization hierarchy	179
3.3.2.3 Differences between the relativizers <i>uf</i> and <i>fi</i>	182
3.3.2.4 On relativizers, determiners, and resumption	183
3.4 Speech act distinctions and syntax	184
3.4.1 Declarative	185
3.4.2 Imperative and Vetative	185
3.4.3 Interrogative	187
3.4.3.1 Yes/No questions	187
3.4.3.2 Information questions	188
3.5 Summary	190

Chapter 4. Historical Phonology and Genetic Affiliation	192
4.1 Historical Phonology	196
4.1.1 Retentions	197
4.1.1.1 Consonant retentions *t, *m, *n	197
4.1.1.2 Other consonant retentions *s, *w, *b, *d	198
4.1.1.3 Vowel retentions *a, *u, *i	199
4.1.1.4 Vowel retention *e	199
4.1.2 Innovations	200
4.1.2.1 Recurrent consonant innovations	200
4.1.2.1.1 Fricativization of labial stops *p, *b > f	201
4.1.2.1.2 ‘Unpacking’	201
4.1.2.1.2.1 Anticipatory voicing assimilation	201
4.1.2.1.2.2 Prenasalization of voiced stops	202
4.1.2.1.3 Dorsal stop deletion *k, *q > Ø	203
4.1.2.1.4 Rhoticization *l, *z, *d, *R > r	205
4.1.2.1.5 Other deletions	205
4.1.2.1.5.1 Final consonant deletion	205
4.1.2.1.5.2 Alveolar continuant deletion	206
4.1.2.1.6 Semi-recurrent innovation: glide fortition *y > j	206
4.1.2.2 Vowel innovations	209
4.1.2.2.1 Recurrent vowel shifts	209
4.1.2.2.2 Semi-recurrent and sporadic vowel shifts	210
4.1.2.2.3 Semi-recurrent vowel mergers	212
4.1.2.2.3.1 Prepenultimate vowel deletion	212
4.1.2.2.3.2 Apocope	212
4.1.2.2.3.3 Syncope *V _{penult} > Ø	213
4.1.2.2.4 Sub-recurrent and sporadic vowel mergers	214
4.1.3 Pseudo-cognates and faux cognates	216
4.1.3.1 Pseudo-cognates	216
4.1.3.2 Faux cognates	217
4.1.3.3 Other cognates	218
4.1.4 Other historical innovations	219
4.1.5 Evidence for lexical strata	220
4.2 Genetic affiliation	222
4.2.1 The Austronesian component	222
4.2.1.1 Central-Eastern Malayo-Polynesian (CEMP)	224
4.2.1.2 Eastern Malayo-Polynesian (EMP)	228
4.2.1.3 South Halmahera-West New Guinea (SHWNG)	232
4.2.1.3.1 SHWNG literature review	234
4.2.1.3.2 SHWNG innovations	234
4.2.1.3.3 A note on SHWNG as a dialect chain	241
4.2.1.3.4 Summary: the position of Irarutu	241
4.2.2 Words of unknown provenance and notes on borrowings	244
4.2.2.1 Words of unknown provenance	244
4.2.2.2 Specific borrowings	247
4.3 Discussion	248

4.4 Summary	251
Appendix 1a - Spectrograms	252
Appendix 1b - Irarutu Feature Chart	253
Appendix 2 - Examples of discourse	254
Appendix 3a - Comparative wordlists	258
Appendix 3b - List of sound correspondences	264
Appendix 4 - List of associated data Items	269
Works Cited	270

LIST OF MAPS

MAP 1.1 - Map of the Southwest Pacific	4
MAP 1.2 - Municipalities in the area of Bintuni Bay.	5
MAP 1.3 - The traditional Irarutu speech territory	6

LIST OF TABLES

TABLE 1.1 - Matsumura and Matsumura's 1991 analysis of TAM markers	35
TABLE 1.2 - SHWNG language research	38
TABLE 2.1 - Consonant phonemes	43
TABLE 2.2 - Vowel phonemes	43
TABLE 2.3 - Segments that contrast by position in monosyllabic words	44
TABLE 2.4 - Initial two-member consonant clusters	66
TABLE 2.5 - Initial three- or more member CCs	68
TABLE 2.6 - Medial consonant clusters	69
TABLE 2.7 - Final consonant clusters	71
TABLE 2.8 - Length of words according to syllable count	75
TABLE 2.9 - Examples of derived diphthongs	79
TABLE 2.10 - Verbal paradigm for 'walk'	84
TABLE 3.1 - Syntactic categories	89
TABLE 3.2 - Varieties of Irarutu bound morphology	90
TABLE 3.3 - Examples of compound nouns	92
TABLE 3.4 - Inflectional paradigms for several example verbs	102
TABLE 3.5 - Examples of adverbs	114
TABLE 3.6 - Irarutu free pronouns	120
TABLE 3.7 - Alienable possessive morphemes	120
TABLE 3.8 - Paradigms for two alienable possessives	121
TABLE 3.9 - Two emphatic possessive constructions	121
TABLE 3.10 - Inalienable possessive circumfixes	122
TABLE 3.11 - Example inalienable possession paradigm	123
TABLE 3.12 - Twenty-five inalienable-relationship nouns	123
TABLE 3.13 - Paradigm for inalienable possession involving a compound noun	124
TABLE 3.14 - Thirty-seven inalienable compound relationships	125
TABLE 3.15 - Irarutu numerals	143
TABLE 3.16 - Fifteen Irarutu classifiers	144
TABLE 3.17 - Aspect and modality markers	146
TABLE 3.18 - Verb conjugation for unergatives and unaccusatives	169
TABLE 3.19 - Distribution of pronouns and inflectional prefixes in SVCs	173
TABLE 4.1 - Sixty-seven proposed cognates in Irarutu	194
TABLE 4.2 - Recurrent reflexes of specific PCEMP segments	195
TABLE 4.3 - Six PCEMP segment retentions in Irarutu	197
TABLE 4.4 - Ten recurrent consonant innovations	200
TABLE 4.5 - Six recurrent vowel innovations	209
TABLE 4.6 - Pseudo-cognates and secondary cognates	216
TABLE 4.7 - 'Faux cognates'	218
TABLE 4.8 - Paradigms related to reconstructed personal pronouns	220
TABLE 4.9 - CEMP innovations	224
TABLE 4.10 - Forty 'other PCEMP lexical innovations'	226

TABLE 4.11 - EMP innovations	230
TABLE 4.12 - SHWNG innovations	235
TABLE 4.13 - Two generalized social settings and linguistic correlates.	243
TABLE 4.14 - Ninety-seven Irarutu words without known etymology	246
TABLE 4.15 - Distribution of words with Austronesian vs. unknown etymologies .	246

LIST OF FIGURES

FIGURE 1.1 - Voorhoeve's 1989 analysis of Irarutu consonants	28
FIGURE 1.2 - Voorhoeve's 1989 analysis of Irarutu vowels	28
FIGURE 1.3 - Voorhoeve's 1989 example of contrastive stress	29
FIGURE 1.4 - Voorhoeve's 1989 analysis of subject marking allomorphy.	29
FIGURE 1.5 - Voorhoeve's 1995 refined analysis of Irarutu consonants	31
FIGURE 1.6 - Matsumura's 1991 examples of subject marker as possessive marker	33
FIGURE 1.7 - Matsumura and Matsumura's 1991 schematic for noun phrases. . . .	36
FIGURE 2.1 - Sonority hierarchy for Irarutu vowels.	55
FIGURE 2.2 - Sonority hierarchy for all Irarutu phonemes.	64
FIGURE 2.3 - Examples of two member initial CCs.	65
FIGURE 2.4 - Examples of three- or more member initial CCs	67
FIGURE 2.5 - Examples of medial CCs	69
FIGURE 2.6 - Examples of final CCs.	70
FIGURE 2.7 - Example syllabication of a word with one SSP violation	72
FIGURE 2.8 - Example syllabication of a word with two SSP violations	72
FIGURE 2.9 - Irarutu syllable template	73
FIGURE 2.10 - Sonority profiles of four Irarutu words.	74
FIGURE 2.11 - Acoustic analysis of the word /naji-ro/	76
FIGURE 2.12 - Two examples of minimal pairs showing phonemic stress	77
FIGURE 2.13 - Example of resyllabication.	78
FIGURE 3.1 - Template of the Irarutu noun phrase	98
FIGURE 3.2 - Template for Irarutu verbs	100
FIGURE 3.3 - Irarutu subject markers	101
FIGURE 3.4 - Irregular verb subject marking	103
FIGURE 3.5 - Examples of <i>m</i> -initial verbs	105
FIGURE 3.6 - Examples of detransitivizing suffix <i>-fe</i>	107
FIGURE 3.7 - Some adjective categories	112
FIGURE 3.8 - Examples of degree words	116
FIGURE 3.9 - Irarutu prepositions	133
FIGURE 3.10 - Examples of prepositional phrases.	134
FIGURE 3.11 - Template for demonstrative determiners.	136
FIGURE 3.12 - Irarutu quantifiers	141
FIGURE 3.13 - Irarutu auxiliaries	151
FIGURE 3.14 - Examples of <i>nar</i> 'to give/bring/take'	152
FIGURE 3.15 - Some Irarutu directional verbs.	153
FIGURE 3.16 - Prepositional Noun Modifier Hierarchy	168
FIGURE 3.17 - Examples of various word order patterns.	169
FIGURE 3.18 - The nominative-accusative pattern	172
FIGURE 3.19 - The unaccusative S pattern	172
FIGURE 3.20 - Examples of relativized modifiers	178
FIGURE 3.21 - Implicational hierarchy for raising from relative clauses	179
FIGURE 3.22 - Irarutu question words	189

FIGURE 4.1 - Summary of sound correspondences in proposed cognates	194
FIGURE 4.2 - Summary of multisegmental changes	196
FIGURE 4.3 - Aberrant sound correspondences in ‘pseudo-cognates’	217
FIGURE 4.4 - “Standard” tree model of the Austronesian language family	223
FIGURE 4.5 - SHWNG tree model	233

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviations

AN	Austronesian
IND	Indonesian
IRH	Irarutu
NBN	Nabi, a.k.a. Kuri
PMP	Proto-Malayo-Polynesian
PCEMP	Proto-Central Eastern Malayo-Polynesian
PCMP	Proto-Central Malayo-Polynesian
PEMP	Proto-Eastern Malayo-Polynesian
POc	Proto-Oceanic
SHWNG	South Halmahera-West New Guinea
WMP	Western Malayo-Polynesian
s.o.	‘someone’
s.t.	‘something’
k.o.	‘kind of’
s.k.o.	‘some kind of’

Symbols

C	Consonant
V	Vowel (for phonological description, Chapter 2)
Ṿ	Half-long vowel
σ	Syllable, major (e.g., <i>nar</i> ‘give/cause’)
σ̣	Syllable, minor (e.g., <i>m</i> in <i>mtie</i>)
*	Synchronically ungrammatical (Ch.s 2–3); Diachronic proto-form (Ch. 4)
**	Unattested historical form
O/Obj	Object
S/Subj	Subject
V	Verb (for syntactic description, Chapter 3)
NP	Noun Phrase
VP	Verb Phrase
PP	Prepositional Phrase
⊃	Conditional
∧	Conjunction

GLOSSING CONVENTIONS

(modified Leipzig glossing conventions)

//	Phonemic form
[]	Phonetic form
—	Brief pause in phonetic transcription
()	Non-cognate form in tables/figures that list historical words
.	Syllable boundary (Ch. 2); separates members of multi-member gloss
(Ch. 3)	
;	Separates inflectional information
,	'Sonable' boundary
-	Affix boundary
+	Compound boundary
~	Reduplication boundary
=	Clitic boundary
1	First person
2	Second person
3	Third person
S	Singular
P	Plural
Pn	Plural Inclusive
Px	Plural Exclusive
A	Agent-like argument of canonical transitive verb
ABS	Absolutive
ACC	Accusative
ACT	Active
ADJ	Adjective
ADV	Adverb
AGR	Agreement
ART	Article
AUX	Auxiliary
BEN	Benefactive
CAUS	Causative
CLS	Classifier
CPL	Completive
COMP	Complementizer
COND	Conditional
CONJ	Conjunction
COP	Copula
DAT	Dative
DEF	Definite
DEM	Demonstrative
DET	Determiner
DTR	Detransitivizer
DISJ	Disjunction
DIST	Distal

ERG	Ergative
EXCL	Exclamation
EXPL	Explicative
EXST	Existential
EXP	Experiencer
FOC	Focus
HAB	Habitual
INCHO	Inchoative
INCI	Incidental Stative
INF	Infinitive
INS	Instrumental
INTR	Intransitive
IRR	Irrealis
NEG	Negative
NMLZ	Nominalizer
NOM	Nominative
NPROG	Nonprogressive
NSAP	Non-Speech Act Participant
OBL	Oblique
P	Patient-like argument of canonical transitive verb
POL	Polite
POSS	Possessive
PRED	Predicate
PREP	Preposition
PROG	Progressive
PROH	Prohibitive/Vetative
PROX	Proximal
PPR	Prefixal Partial Reduplication
QUANT	Quantifier
RECP	Reciprocal
REDUP	Full Reduplication
REFL	Reflexive
REL	Relative
RES	Resultative
S	Single argument of canonical intransitive verb
SAP	Speech Act Participant
SPR	Suffixal Partial Reduplication
STAT	Stative
TAM	Tense, Aspect, Mood marker
TOP	Topic
TRN	Transitive
VL	Verbalizer



The Nafurbenan extended family. Standing (left to right): Maksi Nafurbenan, Damianus Fenentruma, Desi Nafurbenan, Moses Fenentruma. Seated (left to right): Willy Nafurbenan, Johanis Nafurbenan (on Willy's lap), Jason Jackson, Beti Wamburie.

Chapter 1. Introduction

Irarutu, an endangered Austronesian language, is spoken in the Indonesian province of West Papua, in and around the Bomberai Peninsula of the island of New Guinea.

Because Irarutu is in danger of extinction, this work is motivated by the need for an adequate description of the language. The introduction covers the following topics that provide a backdrop for the other chapters: (1) background information about the Irarutu language, (2) introduction to the Irarutu community, including a brief discussion of the documentation methodology, (3) synopsis of previous literature with a summary of the genetic classification of the language, and (4) an overview of the remainder of the dissertation, Chapters 2–4.

This description of Irarutu grammar is supplemented as much as possible by digitally archived language documentation materials that were collected in the field. It is also written with as little theoretical bias as possible, while still being contemporaneous to grammar writing, so that it will be accessible to the widest possible audience. Indonesian (which will be abbreviated IND), and data from other languages, are underlined where they play a crucial role in the description.

Irarutu synchronic phonology favors reduction of non-stressed vowels to schwa or zero, especially in rapid speech, which results in frequent consonant clusters. In the grammar of noun phrases, inalienable possessive constructions involving compound nouns sometimes occur with iterated suffixes that repeat part of the possessive circumfix and agree in person and number with speech act participant possessors. Determiners and demonstratives are modified by discourse-sensitive suffixes. Relativizers, *uf* and *fi*, are also sensitive to tracking referents in discourse. Heavy NPs, such as those with more than two adjectives, can occur, but they are rare in naturalistic language data. The verb is inflected by agreement prefixes that reflect the person and number of an agentive subject, whereas pronominal objects can contract to the verb. There are several agglutinative verbalizing morphemes, including *fi*- and *na*-, as well as a detransivizer *-fe*. The verbalizing morphemes frequently occur together. Tense, aspect, and modality markers precede the verb, whereas adverbs can precede or follow the verb and are located in the periphery. Some TAM markers, light verbs, and serial verb complexes are also inflected for the subject of a clause. Argument doubling occurs frequently in a main clause with a topic phrase, which is marked by the clitic *=ro*. Subject doubling in conjoined and subordinated clauses is optional. In naturalistic language data, serial verbs abound. There are remnants of a politeness system with two registers: one elevated and one coarse. Several content words, such as nouns and verbs, have both elevated and coarse counterparts. This area of the language is nearly all but forgotten by most speakers under sixty years old. In raising constructions, the preferred strategy for resumption is to use a demonstrative pronoun *in-situ*, although gapping does occur. Finally, as is typical of the languages of Melanesia, Irarutu historical phonology has been a challenge, due in part to the existence of various lexical strata exhibiting converse, and often conflicting, changes.

1.1 Background information

This section discusses the names for the language, where it is spoken, by how many people, what the dialects are, and how the language is written. There were two reasons

for choosing Irarutu. First of all, it needs better documentation before it loses its speakers. Because it is a language with a small number of speakers, located in a remote part of the world, this kind of attention has been slow to arrive, and there are several other languages that need similar support. Second, it is a member of the South Halmahera-West New Guinea (SHWNG) subgroup, which is important for Austronesian studies in Eastern Indonesia because it holds the key to unraveling the difficult historical picture of the Austronesian dispersal through West Melanesia. Furthermore, the place of Irarutu within the South Halmahera – West New Guinea subgroup needs justification. Ethnologue (Lewis et al 2013), apparently following Ross (1995), groups Irarutu with the South Halmahera languages, whereas Anceaux (1961), Blust (1974), and Voorhoeve (1989) treat Irarutu as an isolate within SHWNG.

1.1.1 Names for the language

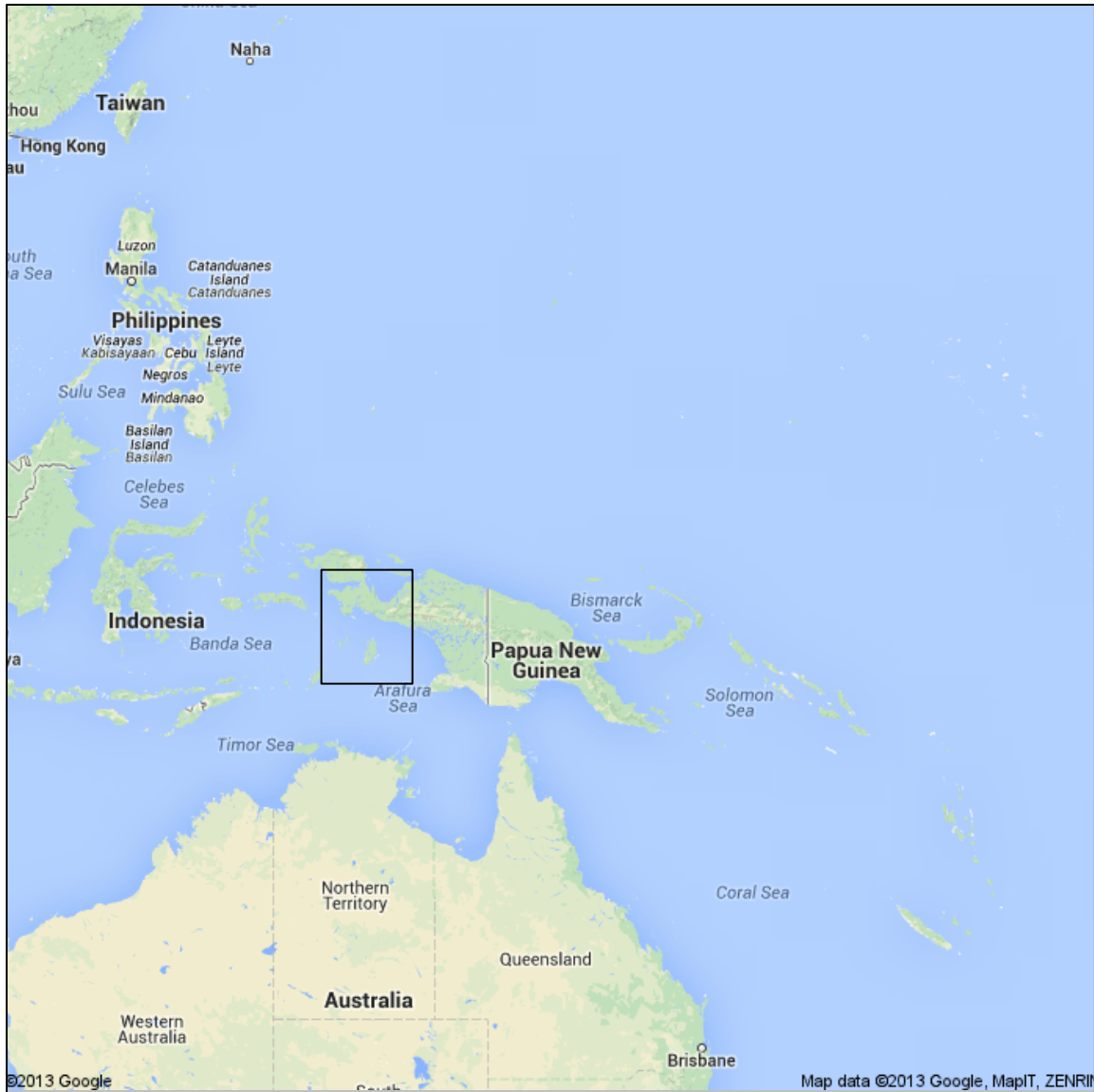
The international language code abbreviation for Irarutu is IRH (ISO 639-3). A closely related linguistic variety that enters the discussion at various points, is Nabi (ISO 639-3 language code abbreviation: **NBN**), mostly referred to as Kuri by native speakers, but Modan in the literature. There has been some uncertainty about how closely related Irarutu and Kuri are. Data shows that they are highly similar in lexicon and have predictable vowel alternations, but they reflect important historical changes differently. Kuri retains Proto-Malayo-Polynesian *-k in some words, whereas it was lost in other languages, and Kuri also retains several lexical items that have been replaced in all other SHWNG languages. For present purposes, Kuri is treated as a highly divergent dialect of Irarutu because a speaker of Irarutu could learn Kuri in a matter of weeks and vice versa. However, it is important to point out that Kuri has also been associated with Wandamen and several surrounding SHWNG languages. Irarutu speakers report that the language called Tandia, from Tandia village, is actually Kuri, but concrete evidence for this claim is not available.

In Western Melanesia it was not traditionally important for speakers to give names to their language. As recently as sixty ago, languages were referred to by the names of the villages where they were spoken. Cowan (1953a, b) cites data from ‘Kaitero’ and ‘Arguni Bay’ (place names), which are now both considered dialects of Irarutu. Variants that have been cited are ‘Irutu’ and ‘Irahutu’ (Anceaux 1958). ‘Kasira’ is another place name that has been used to refer to Irarutu. It is generally held by speakers of the various Irarutu dialects that Fruata (formerly Mabriema), which is located more-or-less in the center of the Bomberai peninsula, is the homeland of the Irarutu community of practice (D. Fenentruma, W. Nafurbenan, Z. Sasepa, and others p.c. 2010, 2013). This is supported by several pieces of linguistic evidence that suggest the Fruata dialect has undergone the fewest number of innovations when compared to other varieties of the language, including loss of prenasalization (which is observed in Kaimana and other southern areas), shift of trilled *r* to *l* (which is observed in East Arguni), and certain instances of lexical syncope (which is observed in the Babo and other northern areas). ‘Irarutu’ has gained general acceptance as a way to refer to the language by speakers of all dialects except Kuri.

The term, *Irarutu*, is an Indonesian spelling of the word [ir³ry'ty], with final stress and substitutions of 'a' for epenthetic schwa and 'u' for [y], which is a high labiopalatal (front round) vowel. In Irarutu, this word means 'true language', and is composed of two parts, iraru [ir³ry] 'language' and tu [ty] 'true'. Furthermore, iraru 'language' is itself composed of two morphemes: /ir/ 'third person plural', which functions as a possessive, and /ry/ 'voice', but in this context, /irry/ means 'language', not 'their voice'. The surface form suggests a structure where *iraru* 'language' is modified by *ty* 'true', and sublexically, *ira* 'their' modifies *ru* 'voice'. In an historic sense, *irarutu* means something along the lines of 'people's true language', but to an Irarutu speaker, it means 'our language'. The existence of a term like this suggests that *Irarutu* contrasts with other languages, such as the national language, Indonesian, or nearby languages, such as Sumuri, Kamrau, and Mairasi, in which Irarutus who live in close proximity to them tend to be fluent. This supports the general notion of a multilingual model of language use on the island of New Guinea (cf. Foley 1998). On a final note, two beliefs are widespread among Irarutus: one is that there is an 'original' language, i.e., Irarutu, and second, Irarutus are the 'original' people.

1.1.2 Location

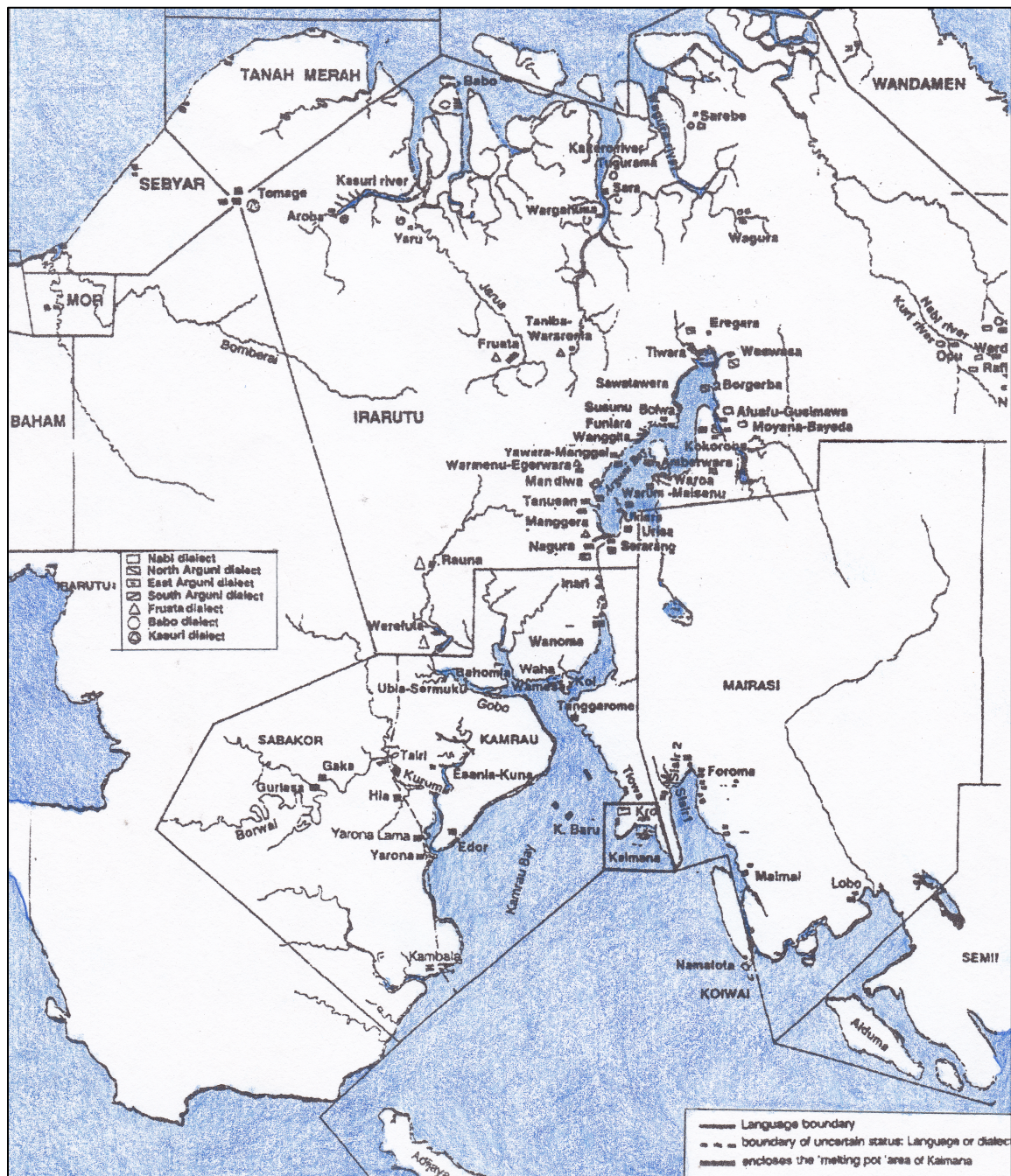
Geographically, Irarutu is located just below the equator at 03°0'-03°41' South Latitude and 132°1'-133°15' East Longitude. It is spoken in West Papua (IND Papua Barat), which is a part of the Indonesian side of the island of New Guinea, see **MAP 1.1**. To clarify some terminology, the island of New Guinea is known as 'Papua' to Indonesians, especially those who live in Papua. So far as the political boundary is concerned, Indonesia includes the western half of the island known – outside the area – as New Guinea. Papua New Guinea is the name of the nation located in the Eastern half of the island. The Indonesian territory was first called West Irian in 1963. After that, it was renamed Irian Jaya, then Papua. Papua is itself divided into two provinces (propinsi): Papua and West Papua. Jayapura is the capital city (ibu kota) of Papua. Manokwari is the capital city of West Papua. Slightly smaller-size population centers, such as Bintuni, are called districts (distrik). They are larger than villages (desa).



MAP 1.1 – The Southwest Pacific showing the island of New Guinea in relation to Australia, Indonesia, the Philippines, and Taiwan. The ‘Bird’s Head’ is indicated by the inset rectangle.

Within West Papua, Irarutu/Kuri occupies a fairly large area (**MAP 1.2–3**) from Cenderawasih Bay, west to about the middle of the Bomberai Peninsula, and from the North coast of the Bomberai peninsula down to, and encompassing, Arguni Bay. The territory is fragmented among several regencies (*kabupaten*): Bintuni, Kaimana, Fakfak, and West Cenderawasih Bay.

North of Bintuni Bay there is a sizable Irarutu-speaking community in Bintuni district (located at the end of the red line that runs down the East coast of the Bird’s Head, in the middle of **MAP 1.2**), the administrative center of Bintuni regency. Kaimana (located towards the bottom of **MAP 1.3**) is the other large population center where Irarutu is spoken. These geographic entities are all located in what has been called the ‘Bird’s Neck’ region of the island of New Guinea.



An undocumented language, Nanggoa, is also reported to be spoken in a few Kuri villages. 3) Mairasi, a non-Austronesian language, is spoken in the Southeast. There are some bilingual villages, so that close interaction between Irarutu and Mairasi speakers exists in those areas. 4) Kamrau and 5) Sabakor are to the South and Southwest. These are non-Austronesian languages but there is interaction between Kamrau, Sabakor, and Irarutus in the Kaimana area. There is a term, 'Iramaika' that is used in the Bomberai

area to show solidarity among three groups: Irarutu (Ira-), Mairasi (-mai-), and Kamrau (-ka). To the west of the Irarutu speech area is a large uninhabited plain that is communal to the peoples of the western portion of the peninsula and those in the East. Northwest of Irarutu there are three additional non-Austronesian languages 6) Mor (different from Austronesian Moor), 7) Sebyar, and 8) Tanah Merah (known today as Sumuri).

On the Bomberai Peninsula, the village Tomage is multilingual, having speakers of Sebyar, Tanah Merah, and Irarutu. Iha and Baham also can be found there. Across Bintuni Bay, in Bintuni district, there are seven groups that are recognized: Sebyar, Sumuri, Moskona, Sougb, Irarutu, Kuri, and Wandamen. For discussion of some of these other languages, the reader is referred to Anceaux 1958. The three largest diasporas of Irarutu speakers are found in Kaimana, located at the Southeast tip of Kamrau Bay, Manokwari, located on the back of the 'Bird's Head', and Bintuni. The first language consultant for the present work was born in Kaimana. Both of his parents were from the Arguni Bay area. He was recruited in Manokwari while he was a student at Universitas Negeri Papua di Manokwari 'Papua State University at Manokwari' (UNIPA). At UNIPA, dorms are designated by region of origin. There are two dorms, Kaimana dorm and Bintuni dorm, where generally more than twenty Irarutu and about a half dozen Kuri speakers live while they attend university.

1.1.3 Population

On account of the two factors mentioned above, that is, multilingualism in Irarutu, Indonesian, and sometimes Kuri, and their location in more than one political district, a precise count for the number of speakers is not entirely realistic. Census data from 2011 and 2012 show that the Irarutu ethnic group in Farfurwar, Babo, Aroba, Kaitero, Kuri, and Bintuni districts comprises approximately 10,000 people. However, most speakers below 60 years of age also speak Indonesian. It may safely be said that there are fewer than seven thousand speakers who are fluent in Irarutu, although this number may be as low as five thousand or lower. As pointed out above, diasporas of Irarutu speakers are spreading throughout West Papua, notably in the towns of Kaimana, Manokwari and Bintuni, and to a lesser extent in Jayapura, Sorong, and Fakfak, which are three other population centers in Papua. In these urbanized areas, Indonesian is dominant, with the result that Irarutu is marginalized. What is most important to note is that the ethnic group is increasing, whereas the number of fluent speakers is not. Anceaux (1958) estimates about 4000 speakers, while Voorhoeve (1989) estimates 5-6000. Clearly, external pressure to learn Indonesian and other global languages, such as English, endangers many minority languages of Eastern Indonesia, including Irarutu (Rusman 1998, Walker 1993, Wurm 2003).

1.1.4 Dialects

It is important to address the relationship between Irarutu and Kuri. As mentioned above, there has been some discussion regarding whether the closely related linguistic variety, Kuri, which has approximately 700 speakers, is a dialect of Irarutu or a separate language (Lewis 2013). From a lexical point of view, Ethnologue (Lewis 2013) states that Kuri and Irarutu are ninety percent (90%) lexically similar, but the basis for the number is not clear. Voorhoeve (1989) classifies Kuri as a dialect of Irarutu, but no language data is

cited to support the claim. Data collected by the author shows that there is recurrent correspondence of Irarutu ‘i’ and ‘u’ with Kuri ‘e’ and ‘o’ in a large number of lexical items, but based a 200-word list, these two varieties share only about 60% lexical similarity. Folk belief holds that there is also a connection between Kuri and Tandia. Tandia is a place, but Irarutu speakers assert that the language called Tandia is really Kuri. Deda (2012) cites approximately forty words and short phrases from Tandia, but it is difficult to appraise the status of the claim that Tandia and Kuri are one language. Data on Tandia is more limited than data on Kuri.

Apart from the differences between Irarutu and Kuri, the amount of phonological variation within Irarutu alone is not clear. There is a high (IND ringen), or smooth (halus), register, and a low, or rough (kasar), register. The Fruata dialect uses more halus forms, e.g., *mbir* ‘boat/canoe’ while Kaimana dialect uses more kasar forms, e.g., *jy* ‘canoe’. From a linguistic standpoint, there has only been enough evidence reported to identify three dialects: Babo, Northeast Arguni Bay, and South Arguni Bay (Voorhoeve 1989; Matsumura and Matsumura 1991). The qualifying grounds are prenasalized voiced stops, which are observed in Babo (and Fruata and Kasuri) where other varieties have plain stops, and a rhotic [r], which is found in most varieties, versus a liquid [l], which is found in Northeast Arguni, e.g., *rre* vs. *lle* ‘day’. Lexical variation also supports the distinction of North Arguni Bay (and Fruata), e.g., *fide* ‘house’, from South Arguni Bay, *san* ‘house’. In addition, Babo differs from Fruata by syncopating lexical items such as [waⁿga’ri] ‘frog’ in Fruata, compared with [wa.ⁿgri] in Babo, as well as segmental shifts such as *mradien* [mraⁿdwen] ‘play’ in Fruata and other dialects versus *mrajuen* [mraⁿjwen] (-d- > -j-) in Babo. By contrast, Voorhoeve (1989) reports that Irarutus recognize six dialects apart from Kuri: Babo, Fruata, Kasuri, North Arguni, East Arguni, and South Arguni. Irarutu dialectology very much deserves further research.

It is important to recognize three factors that play a role in the matter of dialect versus language. Formal similarity, mutual intelligibility (which is closely related to formal similarity), and the opinions of the speakers themselves, that is, whether any political or individual identity considerations complicate the picture, must all be taken into account. Many speakers of Irarutu consider Kuri to be “just a dialect” (personal communication); however, Kuri speakers are in contact with Wandamen speakers in the North, due to their shared border. Both Kuri and Wandamen have been called Modan, but this label only clouds the picture. The distinction between language and dialect deserves some comment. Between the two terms, language and dialect, dialect is always subordinate (Hymes 1968). A language holds more status, is used in a less confined area, or comprises a group of dialects. Particularly for underdocumented and undocumented languages where a standard variety does not exist, a loose association may be observed, such as dialect chains or groups (Florey 2004). Furthermore, a dialectal relationship is one where differences set off a local group’s vernacular in relation to those of other groups within the same broader culture (Gumperz 1968). The Fruata dialect is held by many Irarutus to be the original dialect. Also, according to these criteria, Kuri is best characterized as a dialect of Irarutu.

Although this is a grammar, it is crucial to acknowledge the data's value for the community of speakers who use varieties of the language, especially as an outsider making statements about the status of the varieties as dialects or separate languages (see for example, Malkiel 1968 and Gumperz 1968). In addition, by improving the documentation of a particular variety, one may inadvertently elevate that variety's status, and a standard can emerge based on the variety that was documented (Florey 2004). This information is beneficial because it is crucial to minimize conflict for native speakers by determining if there is a standard or if there are peripheral varieties in the opinions of the speakers themselves. Previous work has focused on Fruata and Gusimawa dialects, which means that these two varieties already enjoy greater recognition than the others. Enhancing the previous work, the Fruata dialect is described in this grammar.

1.1.5 Writing system

In this description, IPA symbols are used for analytic purposes in Chapter 2, on contemporary phonology, and Chapter 4, on historical phonology and genetic affiliation. By contrast, in Chapters 3, a rough phonemic system is implemented, including 'ɪ' for /ɪ/ and 'y' for /y/.

In the appendices, and for writing the language casually, a simpler – and to native speakers more familiar – orthography that uses a greater proportion of conventions from Indonesian is implemented. This is because each speaker of Irarutu uses slightly different writing conventions, but the influence of Indonesian orthography is widespread. Irarutu speakers write phonetic schwa sporadically. When they do, it is generally written 'e', following an Indonesian practice. In this description, schwa is written only where it is pertinent to the discussion. The occasional occurrence of double vowels in native speaker orthography is considered an *ad hoc* means to represent any of three things: i) vowel length, ii) an attempt to distinguish the three-way contrast in non-low labial and palatal vowels, or iii) separate syllables. Data does not support a phonemic length distinction (see §2.1.2), but native speakers may be sensitive to phonetic length and signify it through orthographic vowel doubling. Singleton 'i, u' are regularly used to represent the lax high vowels (/ɪ, ʏ/). Consequently, 'ii' and 'uu' may be considered by native speakers to be a suitable way to represent (phonetically long) tense vowels /i, u/, but this depends heavily on the speaker. In all other cases, double vowels represent separate syllables. For the consonants, native speakers usually write phonetically prenasalized voiced stops as a sequence of a nasal followed by a stop. In this description, prenasalization is only written in phonetic transcription.

For the casual orthography, it is proposed that the phonemes are represented by the letters 'p t c k b d g j f s h m n n g l r w y' and 'i i e a o u ü', where the digraph 'ng' represents the velar nasal [ŋ], 'j' represents a palatal stop [tʃ], 'y' is a palatal glide [j], and 'ü' is a labio-palatal high vowel [ɥ] – some of which are influenced by literacy in Indonesian. The velar nasal is fully predictable. It only occurs as the onset to a prenasalized voiced velar stop [ʎg]. To maintain the phonemic contrast between high tense and lax, the high front vowels /i, ɪ/ are consistently written 'i, ɪ' respectively. It is proposed that the high labio-palatal vowel /ɥ/ should be written as 'ü' to distinguish it from /u/, the high labial (back) vowel 'u', in the casual orthography. This is a more desirable solution than

writing ‘h’ after non-low vowels to distinguish tense from lax segments, which was a convention attempted in 2010. These digraphs cause words to appear both unnatural and more complex than need be (e.g., *fradfihd* ‘five’). It also misrepresents the phoneme /h/.

Determining word boundaries for vernacular literature is sometimes problematic, and the consultants’ intuitions are accommodated where possible. This accounts for some morphemes being variously written as independent elements or simple clitics, although there are instances where the distinction between cliticization and affixation is also fuzzy.

1.2 The Irarutu community

This section introduces some aspects of social structure, economy, and history; the linguistic situation and contexts for using Irarutu; and presents a brief discussion of both language endangerment and documentation.

1.2.1 Social structure, economy, and history

Identity is an important concept for Irarutus, and is based largely upon the village from which one originates. The concept of village needs clarification. In the past, the concept of a village as a social hub did not exist (see Voorhoeve 1989). Today’s village comprises several clusters of living quarters inhabited by a few families each, often related through kinship. Kinship terms help Irarutus embrace the concept of an extended family. In coastal areas, traditional Papuan settlements are erected on stilts over the mangrove to help avoid mosquitos. Modern houses in places such as Bintuni and Kaimana have contemporary Indonesian amenities. Nuclear families are rare but becoming more commonplace in urban areas. In Irarutu, terms for elder sibling and younger sibling are generally used for kinsmen of the same generation, whereas intergenerational terms such as mother, father, and child are not strictly based on biology. Even though the researcher is not ethnically Irarutu, his primary language teacher is his *baba* ‘elder sibling’ who calls the researcher *nefut* ‘younger sibling’.

Irarutu people are traditionally farmers. They tend gardens and cultivate their own food, such as sago palms, cassava, and taro. The average temperature in the region is around 28.5° C (83.3° F); while the average rainfall is about 223 cm (88 inches) per year (Badan Koordinasi Penanaman Modal Daerah 1995). Prized natural resources are not overabundant (Badan Koordinasi Penanaman Modal Daerah 1995, Ratman 1998). The most popular food in Papua is *papeda*, a starch prepared by removing the edible inner substance of the trunk of the sago palm and cooking it, often served with fish and broth. These are important dimensions of anthropological research in Papua (Walker 1990, Foster 1973). Malcolm Walker (1990) reports there are three basic ecological zones that the Irarutu-Kuri community inhabits (supported by Ratman 1998): in Bomberai these are either 1) swampy areas mixed with coastal and riverine features where sago and fishing are practiced, or 2) foothills with small valleys where gardening, hunting, and pig raising are possible, and 3), on Cenderawasih bay, there are coastal lowland areas that are suitable for gardening, fishing and cultivating tree crops. West of the watershed, on both sides of Bintuni Bay, the panorama is characterized by dense jungle. The coast and many rivers are marshy and shrouded by vast stretches of crocodile infested mangrove swamp whereas inland areas are often muddy carstified hills with numerous valleys and

waterways. East of the watershed, the shore of Cenderawasih Bay is adorned with the occasional sandy beach and calm waters that lead out to the Pacific Ocean.

Michiko Matsumura (1992), reports that a young man is ready to start a family when he can build a house, make a fish trap, and cultivate sago. This is in the teenage years. In more recent times, Irarutus from urban areas such as Kaimana have also begun to trade goods, for which they travel to Manokwari, Fakfak, Sorong, Jayapura, Wamena, and the like. One type of highly sought-after traditional commodity in Papua is *bigni* ‘k.o. large ceramic plate’, which is used as dowry, as are such things as golden earrings, money, and a variety of cloths, including *batik*. The plates originated through an extensive trade network that reached to China. However, due to their inherent fragility and diminishing supply, plates are increasingly difficult to acquire. Marriage within the ethnic group is preferred, but exceptions are becoming more commonplace.

The history of the Irarutu people is not well documented. Every clan has its own creation story. Family histories are generally tied to local fauna, such as snakes, cassowaries, pigs, and even mermaids, which are said to be well-documented in the Bomberai area. Many Papuans are adamant that mermaids exist. These sorts of fauna are highly respected. Notably, there are no creation myths tied to crocodiles, a local predator.

Gunung Nabi is an important geographical landmark for Papuans and is considered a holy place (N. Peckham 1981a). *Nabi* is in fact the Indonesian word for ‘prophet’ or ‘holy person’. In 2010, language consultants relayed a creation myth in which ‘Nabi mountain’ (*Gunung Nabi*) is revered as Biblical Mount Sinai (‘Zion’), the place of ascension for both Jesus Christ and the prophet Mohammed, and the origin point of all mankind as well as human language. These stories were affirmed in 2013 by several consultants. It was also mentioned that fragments of Noah’s ark are still able to be viewed in Papua. As pointed out above, the word *irarutu* ‘true language’ has these connotations for Irarutu speakers and connects language to history. Consequently, Papuans generally avoid travelling near *Gunung Nabi* because it is the location of traditional religious ceremonies conducted by the Nabi people. Outsiders who happen to travel there are exempted due to their ignorance.

1.2.2 The linguistic situation and contexts for use

Irarutu has one of the largest populations of speakers in the Bird’s Neck, and there are ample opportunities to use Irarutu with other Irarutu speakers in the home territory, but the Bird’s Neck is located in Eastern Indonesia. This means that for younger speakers of Irarutu to excel in mercantilism or raise their level of education with the aim of gaining employment in a modernizing nation, knowledge of Indonesian is indispensable. Attending university is highly desirable but impossible without proficiency in the national language. Furthermore, many fields of study at universities throughout Indonesia require Bachelor’s theses to be written in English, even though they are often riddled with idiosyncrasies. In this context, English is viewed as *the* language of science and academia.

English also plays a major role in the Internet; however, in Indonesia, the Internet is displayed, and can be navigated, in Indonesian. Text messaging in Indonesia is heavily utilized and fairly cryptic, since it incorporates many acronyms and alternative spellings derived from slang popularized on metropolitan Java, such as ‘q’ for *gue/aku* ‘I (first person singular)’ alongside ‘z’ *saya* ‘I’, and ‘x’ for the determiner *-nya* ‘the’. Alternating capital and lower case letters can be used to include a secondary message. Most members of the younger generation throughout Indonesia have a command of the system. In addition, there are regional dialects of Indonesian that affect how the national language is used. For example, a posting on Facebook by a Papuan youth may incorporate Papuan Malay, abbreviations from metropolitan slang terms, secondary meaning through typography, and perhaps terms from a regional language. A 2013 text message written by a twenty-one year old encoded *terima kasih* ‘thank you’ as *ma kch*, where *ma* comes from the last syllable of *terima* ‘send’ and *kch* derives from *kasih* ‘receive’ (‘c’ = [si]).

Due to the major role of Islam in Indonesia, a third language, classical Arabic, is also important for Irarutus who are Muslim, as well as for Indonesians in general, because it is the language of the Qur‘an, the holy text of Islam. Most Irarutus from Kaimana regency are Muslim, while those from Bintuni regency are Christian: Catholic or Protestant. Kuri speakers are Christian. Christianity and Islam peacefully co-exist in the Irarutu ethnic group. Both religions arrived within the past century. Islam was introduced from the South by way of Kaimana. Christianity took hold in northern areas due to evangelical work by Dutch missionaries around 1960. Each group tolerates the other to a certain degree. Outsiders, especially westerners, who are assumed to be Christian, are avoided more by Muslim Irarutu speakers than by Christian Irarutus, although both groups are initially cautious about outsiders. As far as can be determined to date, there is only a translation of the ‘Gospel of Luke’ from the *New Testament* into Irarutu: *Lukas nfier frgrgir snan frfier dir weni* (Matsumura and Matsumura 1999). Church service is conducted primarily in Indonesian, with perhaps a few English phrases, such as ‘God bless’, and some borrowed English vocabulary. Prayer at the mosque, on the other hand, involves a significant amount of Arabic language, which can be heard over loudspeakers five times per day. The role of Irarutu in religious contexts in urbanized areas is negligible.

1.2.3 Language endangerment and language documentation

In the last decade, the issue of language endangerment has received considerable attention. As a result, scholarship within linguistics has gravitated towards the documentation of endangered languages. Over and above traditional description, documentation provides real-time media of exactly what the language was like when it was documented (see Appendix 4). Recording appropriate metadata, such as where, when, and from whom the data was collected, has a more prominent role than it did in the past, as does analysis based on naturalistic, documented language data. Archiving these materials is a top priority. Language documentation requires specialized training because it is heavily oriented towards producing digital-based media.

1.2.3.1 The Problem of language endangerment

In an age of conservationism, it is assumed that the threat of losing diversity, that is, losing languages, is an important issue for all people, especially for those whose heritage is rooted in a dwindling or extinct language (for an alternative view see Ladefoged 1992). In the 1990s, several publications alerted the linguistics community, as well as the general public, to the urgency of responding to the imminent threat of language endangerment (Hale et al 1992; Himmelmann 1998).

The most striking feature about the distribution of linguistic diversity is that the vast majority of humanity knows how to speak comparatively few languages, in the area of one-twentieth of the total set (Lewis 2013). A conservative estimate of the total number of languages on Earth today is seven thousand. What this means is that most of the linguistic diversity is found in small communities, that is, fewer than 100,000 speakers. These groups of speakers are increasingly being faced with pressure to adapt to one of the larger languages. But, languages exist in a variety of states. Some have lots of speakers and plentiful written literature, while other languages have only a few speakers and sparse documentation. A full spectrum of situations exists between these two categories; and, number of speakers and documentation are not the only factors that need to be taken into consideration. From a language documentationist's point of view, languages with few speakers and little written literature are in danger of extinction. To prevent this, underdocumented languages need immediate attention. Grammatical description (such as the present work) provides rudimentary documentation. In the most dire contexts, it may provide the only record of a dying language. The challenge for the documentation enterprise is to bolster as many endangered languages as possible before they vanish.

1.2.3.1.1 Why Irarutu?

As pointed out above, there are a number of reasons why Irarutu was chosen. Apart from being endangered, it has been classified as an isolate within its subgroup in the Austronesian language family. Although Austronesian is a large family, exceeding 1200 'genetically' related varieties spoken throughout island southeast Asia and the Pacific, including Taiwan, the Philippines, Indonesia, all of Micronesia, all of Polynesia, and coastlands of Melanesia, some areas—especially Eastern Indonesia—still need extensive research. Irarutu is a Melanesian language of Eastern Indonesia. Coincidentally, Melanesia is a term that has a variety of meanings, but for present purposes, the term is useful for specifying a culture area to the North of Australia, stretching from outposts as far West as the island of Timor, and as far East as Vanuatu.

Papua is the largest Melanesian land mass and has been inhabited for approximately 50,000 years or more (Pawley et al 2005). Much of this equatorial island is marshland or jungle; however, there are highlands surrounding a snow-capped, 14,000 foot mountain in the middle of the island. The considerable time-depth of human habitation and often inhospitable geographical characteristics has allowed the languages of Papua to diversify tremendously. The number of language families in Melanesia is even a contentious topic (Wurm, ed. 1975, Wurm 2007). Papua is home to roughly 1000 languages (Lewis et al 2013), most with small numbers of speakers, little to no documentation, and extremely restricted access to outsiders. Considering the number of languages in the area, it is not

surprising that multilingualism is the norm in Melanesia. For these reasons, Papua is a crucial place to study language in its social and cultural context.

The more well-studied Austronesian languages of Melanesia are located East of Cenderawasih Bay and belong to a single, massive subgroup called Oceanic. There are many Austronesian languages in Melanesia that need to be documented, but there is a special need from an Austronesianist's perspective to work with Austronesian languages of Western Melanesia, such as the SHWNG languages in Cenderawasih Bay, in order to improve our understanding of how the Austronesian languages of western Melanesia relate to those in the East (Donohue and Grimes 2008; Blust 2009a; and see §1.3.3).

1.2.3.1.2 State of endangerment

Irarutu is an endangered language although the ethnic group numbers more than 10,000 people and is growing. This is because intergenerational language transmission is waning and pressure from the national language, Bahasa Indonesia (IND), is causing many speakers to compartmentalize their use of Irarutu. It remains to be seen if the Papuan context, where small communities of multilingual speakers have thrived for millennia, assisted by descriptive resources, can help Irarutu to survive. The pervasive role of multilingualism in Papua may provide a model whereby Irarutus incorporate knowledge of Indonesian alongside other local languages, and designate them for use in specific domains, for example, Indonesian for school and Irarutu for identity. Outside the Bomberai Peninsula, code-switching to Irarutu is only appropriate in specific contexts. Code-mixing of Irarutu with Indonesian is prevalent, but this could be argued as an impingement on the vitality of Irarutu. The discussion of loan word phonology (Chapter 2) shows that there are numerous Indonesian loan words in Irarutu. In fact, all Irarutu language use witnessed in this research invariably includes some code-mixing or code-switching.

1.2.3.2 Language documentation and language description

Language documentation provides a means for endangered languages to attain greater permanence than they may otherwise have. The history of language documentation can be traced back through anthropology to anthropological philology, centered on Herder's (1772) thesis concerning the individuality as well as the scientific and humanistic value of the language and literature of every people, regardless of the stage of development. Language description, lexicography, and text analysis play a key role in language documentation, but for some, documentation is considered more nebulous (Himmelman 2012). Where description tends to reinforce a componential understanding of the formal qualities of language, documentation could theoretically include an infinite array of topics. For example, if a native speaker of an endangered language carried around recording equipment, documenting all of the language around her, she could quickly produce a corpus that would far outweigh a single analyst's ability to manage. Due to the resurgence of language documentation in the context of language endangerment, recent work (Himmelman 1998) suggests documenting as much of a language as possible. This is why training native speakers to document their own language is important. But, there is certainly a role for language documentation in ongoing language use. Descriptive documentation can be used to develop pedagogical materials, language maintenance

resources (cf. Hadi 1999), and ethnology (cf. Granadillo and Orcutt-Gachiri 2011). Unbounded documentation is essentially ill-guided, as is the overuse of language maps and reduction of languages to number of speakers. To facilitate contemporary work on endangered languages, there are two specialized resources that help describe and document human language: standardized metadata and digitally formatted media.

1.2.3.2.1 Research in Papua

One aspect of research in Papua that is not evident in the metadata (see Appendix 4) is the research setting and the unique challenges that it presents. Several centuries of research into the languages of the island of New Guinea have yielded but a fraction of the work needed to understand the linguistic situation in Papua (Voorhoeve 1975b). This is also the general case for the larger Melanesian context. There are several reasons why Melanesia is behind the curve insofar as genetic classification and language documentation are concerned. The foremost of these is physical access, although multilingualism, linguistic diversity, and politico-economic obstacles also need consideration.

For an outsider, Melanesia is neither easily accessible nor is it luxurious, however there is an astounding amount of biological diversity, shown for example by the staggering amount of reef marine diversity in the Raja Ampat Islands. Much of the food at the market and sold by vendors is fresh and unprocessed. However, malaria is a common health concern in many areas. Streams are often home to crocodiles. Several other animals are to be taken with caution, such as cassowary and snakes. The first contact with westerners in West Papua is traced back to the Dutch (Kamma 1947, 1994, n.d.), followed by a long colonial history under the Netherlands, briefly Japan, then absorption into Indonesia after its independence in 1949. The field site for the 2013 research, Bintuni, is also known by a Dutch name, *Steenkool*, ‘Coal (lit. stonecoal)’. A common misperception by outsiders of Papuans is that they are a neolithic people including cannibals and headhunters. These may still exist but are the exception rather than the norm. Simultaneously, certain urban hubs, such as Port Moresby, in Papua New Guinea, have frequently appeared among the homicide capitals of the world over the past decade; but statistics like these do not accurately portray the region as a whole. In the experience obtained while gathering these data, the people of West Papua have proven to be both hospitable and genuine.

Travel to the Bomberai Peninsula is time-consuming and presents many unique challenges. For the international segment, there are presently no international airports in Papua, so that international travel is indirect and arduous. It generally takes about four days of continuous vessel hopping and waiting in airports and taxi stations to reach the Bomberai Peninsula from the United States. Travel within Indonesia presents a different hazard. Due to economic factors, several airline companies in Indonesia use older and relatively small aircraft. There is an alarmingly high incidence of aviation tragedy. Within West Papua, there are basically three means of transportation to reach Bintuni and the Bomberai Peninsula beyond: single-propeller airplane, boat, or four-wheel-drive truck. Air travel is the least safe, but it is the fastest (and most expensive). There is a Catholic airstrip in Fruata that was built for clergy members, but it is not generally

accessible to the public. Fakfak, Kaimana, and Babo are more common destinations for single propeller aircraft. Travel by boat is inexpensive but time-consuming and may also require subsequent ground travel. The most common means of travel to the traditional Irarutu speaking area is to access Bintuni Bay, via Bintuni, from Manokwari by four-by-four truck-taxi, followed by ferry from Bintuni to Babo. Strip-mining to prepare for construction of buildings and thoroughfares is commonplace, but road maintenance is patch-work and noticeably imperfect. Semi-improved roads between Manokwari and Bintuni, and Manokwari and Sorong opened in 2005 and 2010, respectively; however, the road system is still developing, which means that travel is possible, but requires certain types of specialized vehicles.

There are no roads into the interior of the Irarutu speaking area. Voorhoeve (1989) reports that it is only a day's walk from the upper reaches of Arguni Bay on the south side of the Bomberai Peninsula to the Kaitero River, which flows into Bintuni Bay in the North. This is the narrowest point on the peninsula. However, my informants state that because there are no roads, it is several days' journey from Babo into the interior by foot. These factors indicate that consideration must be made for travel with documentation equipment, which must be light-weight, securable, and safe from the elements. It is also beneficial if luggage is humble-looking.

1.2.3.2.2 Metadata

It is no longer appropriate for a fieldworker, whether a formally trained linguist, or a native speaker of a language, to cite anonymous data or refer to an un-analyzed data bank. In employing any digital media, it is a top priority to acknowledge the who, what, when, where, why, and how of working with native speakers. This information constitutes metadata. Metadata is the first, and arguably the most important step in developing and managing digital media. Some scholars argue to include the language code, date, location, speaker's name and gender, and type of material contained in a computer file's title, but this leads to a cumbersome label. File names of electronic data become permanently associated with the data itself and must be standardized from the outset. The solution in this description is to use a simple label indicating the researcher and date. Further metadata information can be accessed within each file.

Indonesian is the ideal research tool language for working in Indonesia. It is the national language and used in media, education, law, and government. Elicitation was initially carried out using Indonesian because all of the consultants were bilingual in Irarutu and the national language. Papuan Malay, a dialect of Indonesian, is the *lingua franca* in Papua. Although the researcher studied Indonesian for several years and is competent with it, it was necessary to learn some Papuan Malay to communicate effectively, such as using a periphrastic possessive rather than affixal possession. Some knowledge of Dutch is helpful, too, due to its historical presence. Most older publications are written in Dutch (Kamma 1947,1993; Cowan 1953).

Speakers who contributed to this research, ordered according to year then quantity of data, were, in 2006, Taher Riroma (college student, born in Kaimana); in 2010, Yohanis Werfete (taxi driver, b. 1983 in Kokoroba, also fluent Nabi speaker), Adam Refideso

(college student, Nabi speaker), Sakeus Sasepa (retired, b. 1949 in Bayeda), Abraham Pomsaru (college student), Oktavianus Barry (b. 1983 in Kokoroba village), Yustus Murmana (college student, Nabi speaker), Otis (college student), and Clemens Warfete (college student); and in 2013, Damianus Fenentruma (farmer, b. 1968 in Mabriema), Moses Fenentruma (village leader, b. 1970 in Mabriema), Engelbert Kufiaga (teacher, b. 1965 in Mabriema), Beti Wamburie (housewife, married to W. Nafurbenan, born in Mabriema), Titus Nafurbenan (farmer, b. 1945 in Mabriema), Willy Nafurbenan (government worker, married to B. Wamburie, b. 1963 in Mabriema), Bernard Nafurbenan (retired, b. 1953 in Mabriema), Ansel Kufiaga (b. in Mabriema), Markus Sefire (middle school teacher, married to A. Nimbafu, b. 1965 in Fruata), Aligonda Nimbafu (housewife, married to M. Sefire, b. 1968 in Mabriema), and Siska Wamburie (elementary school teacher, b. 1966 in Mabriema).

In 2006, a visit was made to Manokwari for a two-week pilot study. At the old Faculty of Letters Building at UNIPA, Taher Riroma, who was at the time a college student, was recorded reading a 200-item wordlist that he had prepared based on an Indonesian wordlist, as well as about a dozen simple sentences that he had written as a personal narrative. Similar material was documented for the Austronesian languages Serewen, Biak, Ambai, and Irires (Non-AN; word-list only).

In 2010, a two-week follow-up visit to Manokwari and nearby Aipiri took place. In Manokwari, five Irarutu-speaking college students were recorded as they read parts of a word-list translated into Irarutu by a group of 5-10 Irarutu speaking college students from various parts of the Bomberai area. In addition, two Kuri speaking college students read a word-list prepared by 3-5 Kuri speaking college students. In Aipiri and Manokwari, three speakers of Irarutu and Kuri were recorded. These consultants read words from corrected word-lists, corrected several published sentences from Matsumura and Matsumura 1991, as well as provided their spontaneous commentary. This fieldwork resulted in a 1500-item wordlist for Irarutu and a 1300-item wordlist for Kuri based on 3000 items from Fox 1978 and supplemented by the Intercontinental Dictionary Series (IDS) list developed by Gil and Bowden (n.d.). The Kuri data is important because, apart from fewer than 130 words scattered throughout the literature (Smits and Voorhoeve 1992; Voorhoeve 1975a), there is no other existing documentation.

In 2013, the Fruata dialect of Irarutu was subjected to intense study and participant-observation with Damianus Fenentruma and several other secondary consultants for three weeks. Willy Nafurbenan's family hosted this activity for most of that time. During this visit, approximately five hundred more lexical items were elicited and cross-checked by a number of speakers. Items from the 2010 wordlist were also checked by the primary consultant, more than one hundred sentences were elicited and recorded, a half dozen examples of discourse were recorded by a variety of speakers, and a rudimentary speaking ability was developed by the researcher. Most of the elicitation was performed at the Nafurbenan residence, where over the course of a week nearly two dozen Irarutu speakers would visit the house to discuss various matters and socialize with kinspeople, often late into the night. This is part of the framework for the social setting of an Irarutu home. Experienced speakers are generous in sharing their knowledge of Irarutu, and use

the language in a range of contexts, such as singing, discussion, greetings, and joking. Irarutu speakers are eager to answer questions about their language. The language skills obtained in this manner made it possible to socialize at roadside produce stands and several additional residences, including one of the Irarutu settlements, but – more importantly – to elicit language data using Irarutu. Recording generally took place indoors, or within the perimeter of the Nafurbenan residence, frequently at night, when activity around the house had decreased. However, the whirring of beetles and occasional rain noise can be heard in the background.

1.2.3.2.3 Equipment and software

A non-trivial aspect of language documentation involves understanding the characteristics of digital media, evaluating the equipment used to capture the language data, and discussing the software used to manipulate it for analysis and presentation. It requires trial-and-error to become proficient at handling the suite of digital recording and analysis devices in the documentationist's toolkit. Today there are few standards, but options *ad nauseum*, in selecting which equipment and software to use. One important standard is the '.wav' file type for audio data because it is considered archival quality. Due to an overlap with several other domains of recording purposes, industry standards come into existence frequently and constantly replace older ones. This section briefly reviews the general types of audio recorders, video cameras, microphones, and computers used in this research. Additional information is associated with Appendix 4.

1.2.3.2.3.1 Digital media

In addition to metadata, language documentation is concerned with providing lasting resources in the form of achievable digital media. Working with digital media entails procedures that are widely varied. Beginning with recording and file naming conventions, one takes notes about metadata and captures the language data digitally. Next, the digital files can be transcribed, analyzed, and archived at a digital repository. These procedures allow various means for grooming and presenting language data with specific kinds of hardware to develop multimedia resources and manage data with specialized computer software. This allows a researcher to generate digital resources that will persist well into the future; however it is important to balance practicality (equipment used, research methods, and purpose) with descriptive thickness (detailed analysis, quantity of recording, and various means to access the data), in contributing to a language's documentation (Nathan 2006).

1.2.3.2.3.2 Equipment

A digital format audio recording device and several quality microphones are indispensable. An elaborate setup, such as the one used for the first two field visits for this documentation, includes one audio recorder with a designated channel for a lapel microphone, and a camcorder for capturing visual data through a single built-in, high-quality lens and optionally for gathering audio data through a lapel microphone or its own built-in microphone. Digital media needs to be migrated for storage and to clear space for more recording, so a laptop computer is essential for digital work.

For the formal recording sessions in 2006 and 2010, each of the primary devices was mounted on a tripod, possessed a battery power supply, and had an external power supply with adaptors from American-style 110volt to Indonesian 220volt. There was need for headphones and various accessories, such as a USB connector to migrate the data. An ideal solution to transport the equipment was to use a medium size (approx. 5" high x 9" deep x 13" wide) waterproof hardcase, which provides security, protection from the elements, a portable chair or table, and is not too inconvenient or conspicuous for travel because it fit in a large carry-on sized backpack for air travel.

1.2.3.2.3.2.1 Audio recorder and microphones

The audio recorder and the microphone are the most essential pieces of equipment for language documentation. To accommodate size restrictions, a high quality solid-state handheld digital recording device was preferred. In addition to a high-quality audio recorder, a selection of various microphones is helpful. This is an enormous subject on its own. A combination of lapel and hand-held microphones was utilized in this project. Lapel microphones prevent obstruction between speaker and microphone, whereas a hand-held interview microphone is a good tool because of its robust tactile nature and ready usability.

1.2.3.2.3.2.2 Video camera

This device is the least standardized tool in the documentationist's toolkit, partially due to a divide between consumer and professional grade equipment. Portability restrictions favor consumer grade products. The trend is to avoid moving parts, and so, solid-state devices are considered optimal. SDHC is one type of storage media used for videography. As is true for most consumer-grade camcorders, battery life is short (usually limited to a couple of hours).

The use of a video camera is helpful for the initial stages of the documentation, such as evaluation of the articulatory properties of various segments, and the ability to produce images of the speakers. However, a dedicated, high-quality video camera is non-essential for language description, unless it is used to target a specific paralinguistic feature.

Another difficulty is that camcorders have unavoidable stigma. Subjects are aware of the fact that they are on camera, which causes some people to become overly self-conscious. Occasionally, this impedes a speaker's ability to produce naturalistic language data.

1.2.3.2.3.2.3 Laptop computer

A laptop is also indispensable because it allows data to be managed and manipulated in the field. It is also quicker for taking notes for a skilled typist than hand-writing notes. However, it can be bulky, heavy, and have the highest starting price of the necessary devices. Power supply options and peripherals for the laptop also need consideration. The preferred style of laptop weighs under five pounds, has a full-sized keyboard, and possesses a 13" to 15" display.

1.2.3.2.3.3 Software

Non-proprietary software is the best for language documentation, because it is more accessible than proprietary software, such as Microsoft (MS). However, for this research, MSWord, and MExcel were utilized. They are current standards for word-processing and spreadsheets. Other important software includes Audacity, a good program for viewing waveforms, and Praat, which is suitable for analyzing segments of audio material less than a few minutes long. Transcriber and Elan work well for producing time-aligned transcripts of audio data. Elan can also handle video data. Digital versions of the data in the Appendices are available through a digital archive, which will be discussed next.

1.2.3.2.4 Archiving

In contrast to the constant renewal of equipment and software, digital language archives exist for the specific purpose of standardization and stability. Archives help data persist long into the future with little or no maintenance, to keep the archived material accessible by whomever is granted permission to do so. Researchers who wish to access Irarutu language documentation can request permission from the author (email address provided in the forematerial). All Irarutu people are granted access to the digital files as a token of gratitude.

Irarutu language data is archived at Kaipuleohone, the University of Hawai'i Digital Ethnographic Archive, under the direction of the Department of Linguistics at the University of Hawai'i at Mānoa (UHM). Permission to work with human subjects was granted by the UHM Human Subjects Research Committee (HSRC), in cooperation with the Institutional Review Board (IRB). Permission to record the language consultants was obtained in each case where recording occurred, prior to recording, and in several instances, request for permission to record was repeated during the first few minutes of the recording session.

1.3 Previous publications

This section summarizes previous literature and identifies areas where the literature is different from this analysis. There are just five publications dedicated to Irarutu. A short history of relevant publications summarizes the state of the present understanding of Irarutu. Voorhoeve (1989) states that before the 1950s, there was only one Dutch missionary (Kÿne n.d.) who reportedly made a wordlist for the Irarutu language. Access to this document was not possible. Galis (1955) worked in a relevant geographical area, and Cowan (1953) wrote a monograph about the languages of the western part of New Guinea (present day West Papua). There are two works by Anceaux (1958, 1961) that specifically discuss the relationships of the languages in the Bomberai and Cenderawasih Bay areas. Beginning in 1974, Blust published several seminal works on subgrouping that pertain to the languages of Eastern Indonesia, including Irarutu (1974, 1978, 1982, 1983-4, 1993, 1995, 2009a, 2009b). In Chapter 4, the subgrouping evidence will be reviewed.

Voorhoeve (1989) published a short description of the language. Then, in the early 1990s, Japanese SIL workers, Takashi Matsumura (Matsumura 1991), and his wife, Michiko Matsumura (Matsumura and Matsumura 1991), published two descriptive

works, presumably in conjunction with their translation of the New Testament (Matsumura 1999). Around the same time, Voorhoeve and Smits (1992) published Anceaux's field notes, including 215 words in up to thirteen dialects of Irarutu and two dialects of Nabi. A few years later, Darrell Tryon (1995) included 1200 words and some short phrases that were collected by Voorhoeve along with his (1995) précis-like sketch of the language in the voluminous *Comparative Austronesian Dictionary (CAD)*. The only additional contribution to describing Irarutu since then was a squib about possession in Irarutu (van den Berg and Matsumura 2008). An important debate about the status of the Central Eastern Malayo-Polynesian subgroup was published around the same time (Donohue and Grimes 2008; Blust 2009a). One publication that was not made available, but which supposedly has information on Irarutu is Barr and Barr 1978a. Grace (n.d.) has not been consulted directly, though some data from this source can be found, both in the on-line Austronesian Basic Vocabulary Database (Greenhill, Blust, and Grey 2008), and in the on-line Austronesian Comparative Dictionary (Blust 1995, on-going).

1.3.1 Descriptive publications

Important information from the literature that was outside the scope of this research is included in order to supplement the global understanding of the language and its context. Key points in the description of Irarutu follow. Voorhoeve (1989, 1995) was the first to suggest a specific classification for Irarutu within the SHWNG subgroup. Based on previous work, thirteen or more consonants are expected, along with an unusually large seven-vowel system. Historical Proto-Malayo-Polynesian (PMP) *p became Irarutu *f*, whereas, all instances of PMP *q became zero. Several proto-segments (PMP *d, Z, l, and r) merged as *r. Van den Berg and Matsumura (2008) include /ɪ/ as a phoneme that is not an allophone of /i/. Three non-low back vowels are also reported. Pronouns appear to be of Austronesian origin, some of which attach to verbs. There is a complex system of possession, involving a distinction between inalienable noun phrases, sometimes with complex morphology, and alienable noun phrases. Finally, there is extensive use of a topic marker. Areas for research that project from the literature include: historical phonology because there has only been discussion of select phenomena, synchronic phonology because the sketches miss important details of the language pertaining especially to vowels, and syntax because it remains largely undescribed. There is also need for examples of discourse and a substantial vocabulary.

1.3.1.1 Cowan 1953

As mentioned above, this is an early publication about languages in West New-Guinea. It appears to be the basis for classifying Irarutu as an Austronesian language. It is written in Dutch, but the present understanding of it is summarized below. In the summary section it states, in English, that the publication is an attempt to analyze manuscript vocabularies collected from civil servants, including native-speaker school teachers, from 1947-1949, for the purpose of classifying various languages as Austronesian or non-Austronesian (i.e., 'Papuan'). It was noted that the data was incomplete and often unreliable, partially due to highly variable orthographic representations. This aspect of the orthography is still observed today.

Various sections discuss areas on the North coast of New Guinea, Geelvink Bay (Cenderawasih Bay), the Bird's Head, the Southwest Coast, and the Raja Ampat Islands. Pronominal prefixes on verbs and personal pronouns were used as a primary diagnostic for Austronesian affiliation. Its author suggests that 'Kaitero' (an early name for Irarutu) is an Austronesian language on the north coast of the Bomberai Peninsula, and 'Arguni Bay' (another name for Irarutu), which flanks Arguni Bay, is also Austronesian. A map of West New Guinea was provided, but the interpretation of territories where the Austronesian speech communities (cf. Gumperz 1968 for use of this term) are located is overgenerous. However, it is true that language varieties are more fluid entities than a boundary line on a map is able to express. Cowan concluded that the Austronesian languages are, more or less, closely related.

Seven villages for 'Kaitero' were identified: Babo, Tugerama, Sara, Kaitero, Mabriema, Kasira, and Suga. Forms for seven pronouns in Kaitero (Kt.) and Arguni Bay (Ar.B) were provided, along with the numbers one through ten. Some dialect alternations between these two varieties that are of value to note are Kt. *i-ro* ~ Ar.B *i-fu* '3s Emphatic', and Kt. *matunemaije* ~ Ar.B *îrê* '3P'. *Matunemaije* is described as parallel to Indonesian *mereka itu* '3P DET', where Kt. *matu* means 'person'. The present research illustrates it is composed of *matu-ne-mai-j* 'person-to-there-specific'. The publication proposes that Ar.B *dja* '1s' is comparable to Ajamau *djio*, *djo*, and that Ar.B *ifu* '3s' is comparable to Mansibabêr *ofô* '3s'. Ajamau and Mansibabêr are non-Austronesian. Furthermore, an element similar to the emphatic Kt. *-ro* ~ Ar.B *-fu*, is equated with the Sarimi coast language Sobei's morphemes *to* and *be*.

In the number system, Kt. *eso* corresponds to Ar.B *esu wëma* 'one', Kt. *rivu* ~ Ar.B *ruw* 'two', Kt. *toru* ~ Ar.B *töra* 'three', Kt. *frada vida* ~ Ar.B *refide* 'five', Kt. *teregite* ~ Ar.B *tregegete* 'nine'. The words for 'five' and 'ten' contain the element *fra* 'hand'. Other forms that were cited include: Kt. *tani* ~ Ar.B *ataniuge* 'body'; Kt. *matie* ~ Ar.B *metia* 'eye'; Kt. *maana* ~ Ar.B *mäne* 'bird'; Kt. *teru*, Ar.B *aru* 'egg'. Possessive prefixes *a-* '1s', *o-* '2s', *i-* '3s' attach to head nouns, such as Kt. *a-die*, Ar.B *a-dije* 'my father', 'body', 'eye', etc. The form *ataniuge* 'body' also has a possessive suffix *-uge* '1 POSS'. A few imperatives are cited. Both forms Kt. *mah* and Ar. B *maä* 'come!' seem to suggest something special about the rime (<PMP *maRi). The forms for 'eat' and 'drink' are exceptional because they have initial *g-* in the first and second persons rather than standard pronominal marking. Based on a glimpse of personal pronouns, numerals, morphological possession, and imperatives, Cowan concluded that both Kaitero and Arguni Bay have a strong Austronesian character.

1.3.1.2 Anceaux 1958, 1961, and Smits and Voorhoeve 1992

This section reviews work by Anceaux, who was an active researcher in the Bomberai Peninsula in the middle of the twentieth century. The 1958 publication is a survey of languages there, with emphasis on the non-Austronesian languages. On the other hand, the 1961 publication attempts to classify the Austronesian languages in Cenderawasih Bay, including Irarutu, based on fewer than 300 words and several short phrases. Smits and Voorhoeve 1992 publishes Anceaux's previously unpublicized fieldnotes on a variety

of Papuan linguistic topics. It is included here because it is one of only two sources that contains data on Irarutu.

1.3.1.2.1 Anceaux 1958

This early, eleven page survey of the linguistic situation on the Bomberai Peninsula is focused on non-Austronesian, Papuan languages. It presents data on eighteen Austronesian and non-Austronesian linguistic varieties. There is a map, although it has inaccuracies, which indicates where the languages are found, and a short discussion that suggests which languages are related, with approximate number of speakers and broad genetic affiliation. Anceaux reports that there are several bilingual and multilingual villages where Irarutu is spoken, such as Tomage, Kokoroba, and Barari-Urisa. There is a short discussion of metadata. Furthermore, Anceaux comments on the role of sociolinguistic factors as well as first and second language acquisition in regard to facilitating mutual intelligibility in the Bomberai area.

Data was collected in Fakfak in January of 1956. Anceaux, who worked in close collaboration with George Grace, stated that his conclusions were based on data from native speakers, but secondary sources such as Cowan 1953 and Galis 1955 also played a role. Information from secondary sources was confirmed in nearly all cases by a native speaker. The length of the wordlists varied. Grammar notes were said to be focused on verb conjugation, but this data was not published.

Basic information about Irarutu and its immediate neighbors is relayed below. According to Anceaux, names for some languages are somewhat arbitrary because several did not traditionally have their own names, whereas others had more than one name, depending on location. The following conventions are introduced here to signify the genetic affiliation of the languages. Square brackets indicate that a language is [Papuan], i.e., non-Austronesian, whereas angled brackets indicate that a language is <Austronesian>.

- <Onin>, an important trade language in coastal areas of the Bomberai Peninsula, closely related to <Uruangnirin>; approximately 600 speakers.
- [Mor], not <Moor>, no tense, but an intricate verb system; ~60 speakers.
- [Barau], spoken in the linguistically diverse Tomage village located at the junction of <Irahutu> and ‘Tomage language’ possibly [Tanah Merah]; [Barau] has intricate verb morphology, but no tense; there is a strong bilabial fricative; accent is achieved by high pitch; final vowels have no accent and are deleted in casual speech; a vowel preceding an accented syllable is prolonged, but some words have no accent; accent in polysyllabic words is lexicalized; ~150 speakers.
- [Tanah Merah], referred to as ‘Sumeri’ or ‘Sumerine’ (Galis 1955) – speakers now call it ‘Sumuri’; long words with open syllables and infrequent consonant clusters, but many vowels are deleted in rapid speech, and, as a result, shorter words with frequent consonant clusters arise; there is intricate morphology and verbs have tense; ~400 speakers.
- <Irahutu>, alternately spelled ‘Irarutu’ and ‘Irutu’, called <Kasira> in the North and <Arguni Bay language> in the South; there are “lots of differences” between dialects; many villages around Arguni Bay and further North to Bintuni Bay; [Mairasi] is spoken in Kokoroba and Barari-Urisa; ~3850 total speakers. A report by Van Beek that was sent to the Bureau of Native Affairs classified four dialects:

- 1) Hills between Arguni Bay and upper course of Kaitero river (Voorhoeve's 'Fruata'), spoken in: Manggera, Kupriai, Warmenu, Egerwara, Warafuta, Temia, Rauna, Ewaragegra, Mambriema, Marjedi, and Taniba; pop. ~950.
 - 2) Central Arguni Bay (Voorhoeve's 'South Arguni'), spoken in: Mandiwa, Tanusan, Manggai, Djawera, Sawatawera, Susunu, Wanggita, Funiara, Mafua, Warua, Tugumawa, Amberwara, Barari, Urisa, Waromi, Maisenu, Seraran, Nagura, and Matua (inhabitants of this village moved to an area northwest of Kaimana); pop. ~1300.
 - 3) North and East Arguni Bay, spoken in: Kokoroba, Mojana, Bajeda, Gusimawa, Afuafu, Gusi, Tantura, Borogerba, Weswasa, Tiwara, and Eregara; pop. ~800.
 - 4) Babo, similar to 'Fruata', spoken in: Wergunusa, Sara, Tugrama, Suga, Jaru, Aroba, Kasira, and Babo (where 'Modan' is also spoken); pop. ~800.
- <Nabi>, located northeast of <Irahutu>, and often held to be a dialect of it; Nabi is spoken in: Wagura, Sarebe, Naramasa, Taramanate, and Nabi; ~550 speakers.
 - [Iria] or [Kamrau], spelled 'Kambrau', 'Kumrau', 'Kumbrau'; closely related to [Asienara]; three dialects; ~900 speakers.
 - [Mairasi], also called 'Kaniran' by Galis 1955, spoken East of Kamrau Bay and in the interior to the northeast; Mairasi speakers cohabit villages with Irahutu speakers; thought to be losing ground to Irahutu, which is "easier to learn"; linguistically very close to "Etna Bay language"; ~1000 speakers.
 - <Kaiwai>, a trade language, known also as 'Kuiwai', 'Kowiai', or 'Namatota'; spoken alongside <Uruangnirin> in the small, 50-person village Nusaulan; also spoken in Kaimana and on Kajumerah (West of Etna Bay); it is the Easternmost Austronesian language on the South coast of New Guinea; ~600 speakers.

Northeast of <Nabi>, the language <Wamesa>, also called 'Modan', is spoken; however, 'Modan' has also been used to refer to Nabi (Lewis 2013). Wamesa is found in a large area that extends East to the coast of Geelvink (Cenderawasih) Bay, where it is called 'Wandamen' or 'Windesi', and Northwest to Steenkool, where it is called <Bintuni>. Bintuni is an Irarutu word that roughly means 'swallowed land' (*bin-tun-i* 'land-swallow-it'), in the sense of, 'possessed or inhabited territory'.

Whereas all Austronesian languages were seen as related, mutual relationships between Papuan languages were not made clear. Some pairs of languages, such as [Mairasi] with [Etna Bay] and, despite dissimilar vocabulary, [Iria] with [Asienara], were thought to be closer than others. Anceaux comments that some non-Austronesian languages are surely related to languages from the Vogelkopf (Bird's Head); e.g., [Barau] appears related to [Arandai] on the opposite side of the McCluer Gulf, which includes Bintuni Bay. Some of the non-Austronesian languages seem to have borrowed terms from Irarutu or some other nearby Austronesian language. Finally, although Austronesian languages occupy a considerable area in the Bomberai Peninsula, they are not the dominant group. At the time of publication, Austronesian languages had about 7000 speakers, whereas non-Austronesian languages had approximately 9300 speakers. On the west side of the peninsula, Austronesian languages are confined to the coast and isles, as is typical for Austronesian groups; however, in East Bomberai, Austronesian speakers inhabit the interior.

1.3.1.2.2 Anceaux 1961

This monograph, called ‘the Linguistic Situation on the Islands of Yapen, Kurudu, Nau and Miosnum, New Guinea’, is a valuable resource for SHWNG studies. It attempts to analyze the genetic affiliation of twenty West New Guinea (WNG) languages located in Geelvink (Cenderawasih) Bay. There are two maps. A notable comment is that certain sound changes, as well as the division between language and dialect, do not correspond well with proposed language boundaries. Conclusions are insightful, but impressionistic, because they are based on examples of agreement versus disagreement in word forms, rather than sound correspondences. There is some metadata, which describes the nature and origins of the data he uses, around 260 lexical items from Irarutu in “simplified spelling”. Possessive constructions for the body part ‘hand’ and the kin terms ‘mother’ and ‘father’, as well as verbal morphology for ‘to walk’ are included. Several hypotheses concerning migration are suggested, but these have become obsolete. Anceaux’s main conclusion about Irarutu is that it belongs in a subgroup of its own, but resembles Austronesian Mor (‘Moor’) closest, with 32% cognacy.

The data in this publication originate from a variety of sources. There were “standard” 360 item wordlists and short sentences, which were circulated to officials in 1947. Lists of 100 words for basic vocabulary were circulated in 1955. For Irarutu, Anceaux used his own materials in addition to nine lists of the 100-word basic vocabularies representing Funiara, Aroba, Yaru, Tugrama, Tomage, Werafuta (Warafuta), and Warmenu villages. He also used two “standard” lists from Kaitero and Arguni Bay. The lists were completed by a variety of people, but detailed metadata are not indicated. The Irarutu data is limited to 261 items, with a further qualification that some items have multiple entries. This is problematic because sources for the data are not specified. Furthermore, no attempt was made to distinguish borrowings from neighboring Austronesian versus non-Austronesian languages.

Similarly to Cowan, Anceaux observes that body parts are necessarily possessed. The general word *-mtie* ‘eye’ is dispreferred to *a-mtie* ‘my eyes’, *o-mtie* ‘your eyes’, etc. In the word *-fra-* ‘hand’, there is a possessive prefix that agrees in person and number with the possessor. For this word, the first and second person possessive forms also have suffixes but these do not distinguish number. *Oframe* ‘2S hand’ contrasts with *eframe* ‘2P hand’, and *ifra* ‘3S hand’ contrasts with *irfra* ‘3P hand’. Kin terms such as *a-jie* ‘my father’ and *a-den* ‘my mother’ do not have possessive suffixes. Irarutu shares some grammatical devices with WNG languages—in particular, prefixes on verb forms. The verb ‘to walk’ is conjugated for person and number, *oba* ‘2S walk’, *ba* ‘2P walk’, *ifa* ‘3S walks’, and *nfa* ‘3P walk’; however, the use of grammar for the purpose of classification shows that instances of disagreement, especially in pronominal prefixes, outweigh correspondence.

One notable hypothesis about the settlement of Austronesian speakers in Cenderawasih Bay that Anceaux’s data seems to suggest involves two stages: first, Austronesians navigated around the Bird’s Head then down along the coast towards the area where Ron is presently spoken, and moved across the south coast, eastwards to the Moor islands, then up to Yapen, where Kurudu is spoken. Stage two was a backmigration towards the

current Wandamen area. Irarutu probably moved west from the Ron area during stage one. This hypothesis is corroborated by accounts from the language consultants, who report that Tandia, which is fairly close to Ron, is the original homeland of Austronesians in Cenderawasih Bay. Anceaux also mentions the belief among some Austronesians of Yapen island, as well as Wandamen-Windesi speakers, that their people originated from the Wandamen area.

1.3.1.2.3 Smits and Voorhoeve 1992

Relevant data in this publication partially overlaps with Anceaux 1961; however, it also includes a small amount of material on Nabi that is not available elsewhere.

Unfortunately, as noted for Anceaux's 1961 publication, the orthography is highly variable, so it cannot be analyzed with any degree of accuracy.

1.3.1.3 Voorhoeve 1989 and 1995

Voorhoeve 1989 is longer than Voorhoeve 1995, but the contents are essentially identical, except that the 1989 paper includes a brief discussion of how to classify Irarutu. On the other hand, Voorhoeve 1995 is supported by a 1200-item vocabulary.

1.3.1.3.1 Voorhoeve 1989

Voorhoeve 1989 is the first publication dedicated to describing Irarutu. Although only twelve pages, it is packed full of information. There is some metadata that clarifies data collection methods. The publication also contains some traditional and mythological beliefs that do not appear elsewhere. There are two detailed maps, one of Babo to Arguni Bay, and one of Kamrau Bay to Kaimana (adapted as **MAP 1.3**, above). The plates illustrate language boundaries of neighboring languages, and include village names and locations, providing unsurpassed information about where Irarutu is spoken. Voorhoeve listed phonemes in an accountable manner and described many morphemes with reasons why particular morphemes were problematic, such as *-ro*, *-ri* and a ligature, *ne*, which is not mentioned in subsequent publications. Very basic syntax was also described. The author cited four changes that support classifying Irarutu as a SHWNG language. These will be discussed in Chapter 4.

Voorhoeve's data was collected in 1982 from one speaker of Fruata dialect who was living in Manokwari, and one speaker of Kasuri dialect in Bintuni. Follow-up materials were collected in Werefuta in 1986 from two consultants who spoke the Fruata dialect. Voorhoeve also worked in Kaimana with one speaker from Mariedi village (Fruata dialect) and consulted a casual informant from Susunu village (South Arguni Dialect). The following neighbor languages were identified: to the East is Mer, to the South are Mairasi and Kamrau, to the West are Baham, Mor, and Sebyar, and to the Northwest is Tanah Merah.

Voorhoeve analyzes 'Irarutu' as an impressionistic rendering of *irərotu* 'their language true', or 'their own language'. A sequence of two flapped *r*'s with intervening schwa often becomes a trilled *r*, which explains Anceaux's spelling 'Irutu' (but see Matsumura 1991). By 1989, the name 'Irarutu' was commonly accepted, but Anceaux's spelling

‘Irahutu’ was not confirmed by Voorhoeve or the present research. In fact, *h* is only observed in Arabic loans, so the *h* in ‘Irahutu’ remains unexplained.

Voorhoeve does not repeat Anceaux’s village names because many had changed due to “gradual condensation of scattered hamlets into fewer but larger villages” (see also M. Matsumura 1992). Based on an amalgam of 1971 census data from two subdistricts, Arguni with 3000 speakers, and Babo with 2000 speakers, Irarutu was thought to have 5000-6000 speakers in 1989. Voorhoeve’s informants identified seven variants based on lexical data and *rimia* (IND), an Indonesian term that refers to intonation, stress pattern, and speed of speech; however, only two differences are found in phonemes: prenasalized versus plain consonants, and *l* or no *l*.

- 1) ‘Nabi’ is spoken in seven villages, five are in valleys of the Nabi and Kuri rivers: Nabi, Rafideso, Werdadne, Owa, Opu; two are on the Wagura river: Wagura and Sarebe. Nabi river villages are bilingual in ‘Nanggwa’, a language that is not mentioned elsewhere (such as Voorhoeve 1975a; Barr and Barr 1978; Lewis 2013), but it could be an alternate name for <Tandia> to the East, or [Mer] to the South.
- 2) ‘Babo’, the name of the administrative center of the Babo sub-district, is spoken in six villages, two on the Wagura river: Sarebe and Wagura, and four in the lower reaches of the Kaitero river: Tugurama, Sara, Warganusa (Wergunusa), and Babo.
- 3) ‘Kasuri’ has prenasalized consonants and is spoken in three villages, two on the Kasuri river: Aroba and Yaru, and Tomage, which has a mixed demographic of Sebyar, Tanah Merah, and Irarutu speakers.
- 4) ‘Fruata’ has prenasalized consonants and is spoken in six villages: the relatively new settlement, Fruata (from combined Mambriema and Mariedi), Taniba-Wararoma, Warmenu-Egerwara, Manggera, Rauna, and Werefuta-Temia. Warmenu-Egerwara and Manggera relocated closer to Arguni Bay. Werefuta and Rauna used to be located further up the Gono River. Tradition says they used to live in the hills to the West.
- 5) ‘South (Coastal) Arguni’ is spoken in fifteen villages: Susunu, Sawatawera, Funiara, Wanggita, Yawera-Manggai, Mandiwa, Tanusan, Nagura, Serarang (Seraran), Ukiara, Warum-Maisenu, Waroa, Amberwara, Mahua, and Bofwer, which is a newer settlement and the administrative center of the Arguni sub-district. Bofwer has a mixed population, but Voorhoeve does not identify the other group.
- 6) ‘East (Coastal) Arguni’ is spoken in four villages: Borgerba (Borogerba), Afuafu-Gusimawa, Moyana (Mojana)-Bayeda, and Kokoroba.
- 7) ‘North (Coastal) Arguni’ distinguishes *l* and is spoken in three villages: Tiwara, Weswasa, and Eregara.

In addition, significant populations of Irarutu speakers from the Arguni area relocated to Kaimana, Tiowa, and Kroi. Prenasalized voiced stops are found in the inland dialects of Fruata and Kasuri. The phoneme *l* is only found in the North Arguni dialect. Formal grounds to distinguish all seven dialects are lacking. Voorhoeve has no data from ‘East Arguni’ or ‘Babo’, and for ‘North Arguni’ he reports on an old 200-word wordlist compiled by the Dutch missionary, Kÿne. The data for ‘South Arguni’ is based on Anceaux’s 1961 wordlist and a short wordlist recorded by Voorhoeve himself, whereas ‘Kasuri’ data was limited to a short wordlist and some grammar notes.

Voorhoeve identified thirteen consonants: /p, t, k, b, d, g, m, n, f, s, w, y, r/, and comments on some of these segments’ phonetic qualities. Distinctive claims are

summarized below in **Figure 1.1**. All segments can occur in initial, medial, or final position, except where noted.

/p/ was introduced by way of Indonesian loan words.
 /t/ is dental-to-interdental.
 /b/ in initial position may have slight prenasalization. Elsewhere it is prenasalized and may devoice.
 /d/ is dental-to-interdental. Initially, it is often prenasalized. Elsewhere it is always prenasalized.
 /g/ in initial position is often prenasalized. Elsewhere it may devoice.
 /f/ is described as a bilabial fricative. In initial position, it may sometimes affricate; between two vowels, it sometimes voices.
 /r/ is described as a flapped vibrant with either weak vocalic onset or devoicing in initial position. Two adjacent flaps merge as a trill.
 /w/ is characterized as a bilabial voiced semivowel in initial position; other instances are analyzed as nonsyllabic [ɰ].
 /y/ is analyzed as a palatal fricative with strong friction. It alternates with a voiced palatal affricate [dʒ]. In Kasuri /y/ is a high front semivowel [j]; but, in Fruata, /y/ only occurs initially, elsewhere it is analyzed as nonsyllabic [ɿ].

FIGURE 1.1 – Voorhoeve’s (1989) analysis of Irarutu consonants.

Voorhoeve proposed a seven vowel system /i, e, ɛ, u, o, ɔ, a/. Five of these segments were qualified specifically, as in **FIGURE 1.2**.

/e/ [ɪ], [ɛ̞] (‘raised e’)
 /ɛ/ [ɛ̞] (‘raised e’), [ɛ] (‘raised e’ seems like it should be ɛ̞ ‘lowered e’ in the original)
 /u/ [u], [ɯ] (‘fronted u’), [ɯ̞] (‘raised u’)
 /o/ [o], [ɔ̞] (‘raised o’), [ɯ̞] (‘lowered u’)
 /a/ [a, ɑ, æ]

FIGURE 1.2 – Voorhoeve’s (1989) analysis of Irarutu vowels.

In addition, a mid-central vowel, schwa [ə], was also included, but it is analyzed as non-phonemic because it is predictable. It occurs after a final consonant before a pause; but this vocalic release can also echo a preceding vowel. For example, [‘su:s^ə] alternates with [‘su:s^u] ‘breast’. However, schwa is also said to be able to bear stress when a word is uttered in isolation, or for extra emphasis, e.g., in elicitation, [‘no:ndə] ‘he holds it’ can alternate with [no’ndə] ‘he holds **it**’. In another context, [ə] automatically separates two adjacent consonants, as in /ɔ msbegt fanɛ/ [ɔ məsə’mbengətə fa’nɛ] ‘you shot the pig’. Schwa does not occur between a voiced stop consonant and *r*, or *f* and *r*. In “pre-tonic” position, a vowel can reduce to a mid-central vowel [ɛ̞], which is non-distinct from [ə].

[ɔɔ'e ŋgáta] (ɔ '2P', -rɔ 'TOP', e 'Q', gáta 'who') 'who are you?'
 vs.
 [ɔɔ' ŋga tá] (ɔ '2P', -rɔ 'TOP' ga, 'eat' tá 'feces') 'you eat shit'.

FIGURE 1.3 – Voorhoeve's 1989 example of contrastive stress in Irirutu.

Stress usually falls on the last vowel *fa'ne* 'pig'. Voorhoeve indicates stress when it is non-final, suggesting that stress is phonemic. He provides one example. However, the targets are not comparable: *gáta* 'who' is one word whereas *ga* 'eat' *tá* 'feces' are two. There are corrected typos in the example (originally published as [ɔɔ' ŋgatá] 'you eat shit') in **FIGURE 1.3**.

Voorhoeve chooses not to use strictly phonemic spelling. This makes the language appear different from what it actually is. Voiced stops are written with prenasalization medially and finally, but not initially; and schwa is included between two consonants as well as finally.

For morphology, Voorhoeve suggests that transitive and intransitive action verbs are marked to agree with second person (*m-* 'you/you all') and third person (*n-* 'he, she, it/they') agents, but a first person subject has no overt marking (\emptyset 'I/we'). These agreement morphemes occur with vowel-initial stems, stressed syllable t-initial stems, or stems that start with *k*, *d*, *g*, *n*, *r*, or *y*. Other combinations fall into four categories of morphophonemic change, adapted below as **FIGURE 1.4**.

-
- 1) *n-* + unstressed syllable with stem-initial /t/ → [d]; *n-təfe* [dfe] 'he spits'
 - 2) *m-* + stem initial /f/ → [b]; *m-fa* [ba] 'you go'
 - 3) *m-* + stem initial /m/ → [m]; *m-mərir* [mrir] 'you stand'
 - m-* + stem initial /b/ → [b]; *m-bəɾəfun* [brfun] 'you hide'
 - 4) *n-* + stem initial /s/ → [s]; *n-si* [si] 'he sees'
-

FIGURE 1.4 – Voorhoeve's 1989 analysis of subject marking allomorphy.

Similar to Cowan, Voorhoeve also identifies words that use *g-* for 1st and 2nd persons, *-en* 'sleep', *-in* 'drink', *-a* 'eat'. The verb follows particles indicating aspect (*bu* 'Compleitive' or *do* 'Progressive') and mode (*ga* 'intention, wish'); however, see Matsumura and Matsumura (1991) for a slightly more elaborate system.

The personal and possessive pronoun systems are elaborate, divided into four series. However, there is clearly a basic series: (*y*)*a* '1S'; *ɔ* '2S'; *i* '3S'; *it* '1Pn'; *am* '1Px', (no form cited) '2P', *ir* '3P', to which various morphemes attach to derive the other series.

Series I can serve as grammatical subject or object; furthermore, *a* '1S' can be a first person subject agreement marker on the verb, or it can be used for possession. The general pattern for possession is: the possessor precedes the possessed. Head nouns are

not usually marked for possession, except for body parts (*fra* ‘arm’, *rit* ‘skin’, *nəmu* ‘thigh’), one kin term (*ta fand* ‘elder same sex sibling’ as in *a tangg fand* ‘my elder brother’), and the word *no* ‘name’. These nouns take possessive suffixes that agree with first (*-ngg*, *-ungg*, *-ongg*) and second (*-m*, *-um*, *-om*) person possessors. If the possessor is a full noun, an optional resumptive possessive pronoun occurs between possessor and possessed, e.g., *a di’ē san* ‘my father’s house’ alternates with *a di’ē i san* ‘my father’s **house**’ (emphasis in original).

Series II comprises emphatic pronouns used for focus or foregrounding. These are composed of a basic pronoun and *-rɔ*. Nouns and noun phrases can also occur with *-rɔ*.

Series III is made up of possessive forms that occur predicatively, and occur with *-ri(r)*. These pronouns follow a possessed noun to emphasize the possessive relationship: *a ‘mɔ* ‘my child’ contrasts with *‘a mɔ yari* ‘definitely my child’.

Series IV pronouns are emphatic versions of series III, that is, series III pronouns with the element *-rɔ* from series II. However, only first and second person singular data were reported. An example is *ya’rirɔ* ‘(it’s) definitely mine’.

One particularly interesting aspect of Voorhoeve’s analysis that is not found elsewhere is the assertion that Irarutu previously had a possessive particle *ne*, which could take possessive suffixes, but is now only found in petrified constructions. His clearest example is *ɔ nəmətamən* < **ɔ nem* tamən ‘your son-in-law’.

Voorhoeve noticed that verbs and nouns can be reduplicated. Reduplication affects the 1st or 2nd syllable. All reduplicated forms that Voorhoeve cites are nouns: *sut* ‘fathom’ and *sasut* ‘measuring tool, yardstick’; *mərir* ‘to get up’ : *mərarir* ‘tempat pencarian’ (no English gloss was provided by Voorhoeve, but the phrase means roughly ‘place of searching’ or ‘prospecting place’).

Canonical word order is reported as SVO. The head of a noun phrase precedes its modifiers: *mɔ kəsi* (child small) ‘a small child’. Some nouns require a classifier before the numeral modifier such as *o’ mangge fu eso* (coconut CLS one) ‘one coconut’.

1.3.1.3.2 Voorhoeve 1995

Voorhoeve 1995 is a condensed version of Voorhoeve 1989, minus the discussion of classification. It appears in Tryon’s *CAD*. However, the significant addition is an extensive 1200 word vocabulary. These words, including several dozen short phrases and clauses, are from the Fruata dialect. However, each of the 1200 items in the publication is presented with seventy-nine other Austronesian languages from various subgroups, so that looking at data from any one language from a global perspective is rather difficult. The publication includes a basic map and some notes on dialectography. As in the previous publication, presentation of the segmental phonemes discusses allophony. Voorhoeve 1995 reduces the discussion of pronouns to one set, which can be modified by one of the following morphemes: *-rɔ*, *-ri*, or a combination *-rirɔ~ririo*.

/b/ has the allophones [b^m b^m p]
 /d t/ have dental, and prenasalized, allophones [d^ɲ d^ɲ t^ɲ t^ɲ]
 /g k/ alternate with [g^ɲ g^ɲ k]
 /ɸ/ is [ɸ] in final position, in initial position it can be [p^hɸ] a weak bilabial affricate, and in medial position, it is sometimes voiced [β].
 /r/ can surface as [r̃ R] word finally
 /y/ can be either [y] or [d͡ʒ]

FIGURE 1.5 – Voorhoeve’s 1995 refined analysis of Irarutu consonants.

The consonant inventory has one major alteration—the use of ‘ɸ’ instead of ‘f’, representing the descriptive label, voiceless bilabial fricative more accurately; however, there are certain contexts where ‘f’ is used instead of ‘ɸ’. The description of consonants is slightly simplified, and is recast using the term allophony instead of range of phonetic variation. The characters used to represent the phonemes, and some conditioning environments, are improved, see **FIGURE 1.5**. Typographical errors in the examples used to substantiate phonemic stress were corrected, but the examples were not improved.

The paradigm for basic pronouns was completed by inclusion of *ε* ‘2P’. Usage of the first person singular pronoun was altered by stipulating that *a* can be a possessive pronoun or subject, but not a grammatical object. Departing from his earlier analysis, the emphatic variant of *-ri* is described as optionally containing a second *-i-*, that is, *-ririo*; yet, the variation needs explanation. The first person possessive suffix, used with body parts, and so forth, is written without prenasalization [-g -ug -og]. The gloss of *mɔ kəsi* (child **small**) from ‘a small child’ was changed to ‘baby’.

1.3.1.4 Matsumura 1991 and Matsumura and Matsumura 1991

These two works were published together, with only one set of references, which was located at the end of Matsumura and Matsumura 1991. Matsumura 1991 describes basic Irarutu phonology. Matsumura and Matsumura 1991 sketches Irarutu grammar. The Matsumuras have published at least seventy-five pages of linguistic description of Irarutu. Takashi Matsumura is affiliated with SIL. His work is connected through the *Ethnologue* (Lewis 2013) website listing for the Irarutu language.

1.3.1.4.1 Matsumura 1991

Matsumura 1991 is the first paper to focus specifically on Irarutu phonology. It provides an overview of the phonemes and phonotactics of Irarutu based upon the Gusimawa dialect, including consonant and vowel phonemes alongside phonetic representations. There is some discussion of allophony and allomorphy, in particular the agent marker on the verb (compare Voorhoeve 1989). As in Voorhoeve 1989, the consonant ‘p’ is argued to be a loan phoneme found in Indonesian borrowed vocabulary. Consequently, Matsumura suggests that *f* (which he writes as ‘p̥’) is the inherited voiceless counterpart of *b*. A significant amount of the publication is devoted to analysis of Irarutu phonotactics, which depends on the analysis of schwa. The appendix is about as long as the body of the paper and contains approximately 369 words, grouped by syllable count and shape. Very little metadata is included, and no connection to previous literature is

provided. Discussion of the phonological differences across dialects would be appropriate (although these are discussed in Matsumura and Matsumura 1991).

Matsumura discusses the location of Irarutu and mentions that there are three dialects distinguished by their phoneme inventories. For the Gusimawa dialect, Matsumura identifies fourteen consonants: /p, t, k, b, d, g, ɸ, s, j, l, r, m, n, w/, similar to Voorhoeve, but with the addition of *l*, and *j* instead of *y*. Symbols for the segments are provided, but without specific descriptions. The nasal [ŋ] is analyzed as an allophone of /n/ when it occurs before a [+back, +stop]. It stands out that /w/ is included as a phoneme, but /y/ is described as allophonic and optional. Matsumura proposes only six vowel phonemes /i, e, a, o, ʊ, u/, which he presents with descriptions, grouped in three categories: front /i, e/, central /a/, and back /u, ʊ, o/. High front *i* has the optional allophone *y* after alveolars. Evidence for the segmental phonemes could be more robust. Matsumura argued that consonants only occur in initial or medial position because all syllables are open. Voiced continuants form a cluster with the alveolar flap [Cr], but no examples of /l, m, n, w/ followed by /r/ are provided.

Schwa is analyzed as non-phonemic, and an allophone of underlying unstressed /e/. Notably this characterization has at least four exceptions: schwa does not occur word-initially, but surfaces as unaccented [ɛ], schwa also surfaces as unaccented [ɛ] in vowel sequences in conjunction with /w/ and when preceding /j/; however, [ɛ] is not directly associated with /e/. All examples of final syllables have a phonemic vowel or schwa; so, schwa is relevant to Matsumura's analysis of phonotactics. To account for this, Matsumura provides alternative analyses of schwa: it could derive from /a/ in unstressed syllables, it could be phonemic (but does not contrast in stressed syllables), or it could be a 'transitional' vocoid.

The description of vowel sequences is problematic. Matsumura provides fourteen examples of words with adjacent vowels but asserts that "no vowel clusters [...] form a complex nucleus of one syllable". In particular, this presents a problem for description of the allophone [y], which occurs frequently in the language due to obligatory desyllabification of certain vowels in the appropriate conditioning environment, such as in /matie/ → [matye] 'eye'.

Stress is said to be phonemic, but no supporting evidence is presented. From a phonetic standpoint, he states that "vowel length does not occur on the stressed syllable"; however, "vowel length occurs in the stressed syllable when the stressed syllable is followed by the syllable with a schwa" (Matsumura 1991: 48). "Gliding" pitch occurs on stress syllables; however, it is not clear what this means.

Syllable structure is a major topic in the publication. It is generalized that each "syllable has one vowel as syllable nucleus and optional consonant or consonant cluster as onset". Furthermore, a syllabic nasal could function as syllable nucleus. His template is: syllable = (C)(C2)V/N, where (C2) is [r]. The three syllable types he proposes are: V, CV, and CCV, supported by examples of monosyllabic, disyllabic, trisyllabic, and quadrisyllabic words. As expected, no monosyllabic words end in schwa.

Matsumura notices morphophonemic deletion of unstressed, word final /e/ ([ə]) before a vowel-initial suffix or a vowel-initial word, for example, /sóte/ ‘girl’ + /adi/ ‘the’, surfaces as [sótadi] ‘the girl’. Stress prevents elision. Stressed final -e surfaces as [e] in /debé/ ‘hit’ + /o/ ‘2S’, and is realized as [dəbéo] ‘hit you’.

For allomorphy of agent marking *m-* ‘2S’ and *n-* ‘3S’ on verbs, Matsumura states that *n-* appears as [m] before verb stems that start with a bilabial consonant, while the allophone [ŋ] appears before verb stems that begin with “back” consonants (presumably [k] and [g]). Syllabic nasals are analyzed as allophones of plain nasals, but without examples.

Matsumura observed that the subject marker is occasionally used as a possessive marker (see **Figure 1.6**), as in ‘father-in-law’, where, in the third person, the prefix *n-* syllabifies and the free form subject marker *i* ‘3S’ deletes. However, this example needs clarification because several morphemes, including -g- in the first person singular, are not accounted for.

ja_a-	/a-ge-támene/	[agətá:mənə]	agtámn	‘my father-in-law’
o_m-	/o-me-támene/	[omətá:mənə]	omtámn	‘your father-in-law’
i_n-	/Ø-n-támene/	[ntá:mənə]	ntá:mn	‘his father-in-law’

Figure 1.6 – Matsumura’s 1991 examples of subject marker as possessive marker.

The discussion of allophony in liquids is different from that of Voorhoeve. Where two voiced alveolar flaps, /r/, are separated by schwa, Matsumura asserts that the result can become a voiced alveolar lateral *l*, for example, /ire/ ‘they’ + /ró/ ‘FOC’ [ire-ró] alternates with [iló] ‘they (FOC)’. This is best explained as a feature of the Gusimawa dialect.

Concerning orthography, Matsumura points out a strong correlation between Irarutu sounds and Indonesian sounds, which might indicate that Indonesian letters would do a good job representing Irarutu. He selects fifteen characters ‘p t k b d j g f s l r m n ng w’. The digraph ‘ng’ represents [ŋ]. As for vowels, he chooses only five, ‘i e a u o’, where ‘e’ represents /ɛ, ə/, and ‘u’ represents /u, ʊ/. But this collapses a phonemic distinction.

1.3.1.4.2 Matsumura and Matsumura 1991

This is the first specialized publication on Irarutu grammar. The data provided by Matsumura and Matsumura (1991) is from the Fruata dialect, not the Gusimawa dialect, which is expected based on Matsumura 1991. Another contrast with Matsumura 1991 is that Matsumura and Matsumura 1991 contains information on dialect variation and provides the evidence for distinguishing three dialects: 1) North Arguni with two liquids, *l* and *r*, but with plain stops, 2) South Arguni with only *r* and plain stops, and 3) Fruata/Babo with only one liquid, *r*, but with prenasalized stops. For classification, the authors comment there is a high percentage of cognacy with Proto-Oceanic, but otherwise appeal to Anceaux (1958) and Voorhoeve (1989).

The paper contains 172 example sentences and introduces a fair range of phenomena, such as pronouns, noun phrases, verbs, clause types, directionals, case, and lexical

material including question-words, time expressions, and several pronominal paradigms. However, several topics, such as definiteness, disjunction, animacy, and case are mentioned but not described. The authors comment that their focus is on words and phrases rather than clauses and sentences. The morpheme *-ro* is referred to as a topicalization clitic, although there is no specific discussion of cliticization or how to differentiate topicalization from focus. No diagnostics are provided for analyzing particular morphemes as affixes or freestanding words. The Matsumuras also comment that an unusual aspect of Irarutu predication is the pervasiveness of directionals in relation to the verb phrase. A significant portion of their paper is devoted to directionals. The interpretation of ‘Irarutu’ differs slightly from that of Voorhoeve. The Matsumuras believe it to be a combination of two words, *iraru* ‘voice/language’ and *tu* ‘true’. They also include some folk beliefs that provide some foundation for the name.

The article describes statements as having level or descending terminal intonation. Canonical word-order is also described as subject-verb-object (SVO). Imperatives contain a verb stem with implied second person subject, and a series of imperatives can occur in a single sentence.

Yes/no questions are signaled by an interrogative (i.e., rising) intonation. Content questions have high intonation on the *wh*- question word and contain the clitic *-e*. The question words ‘who’, ‘what’, ‘where’, ‘why’, and ‘how’, have direct equivalents, but ‘when’ does not. The morpheme *gata* ‘who’ occurs at the end of a sentence that inquires about people. *Nia* ‘what’ occurs at the end or beginning of a sentence that inquires about things. The combination of *ri* ‘possession’ + *gata* ‘who’ + *uf* ‘which’ inquires about a ‘relative pronoun’, whereas *nen* ‘how’, followed by the optional preposition *ne(ne)*, asks about place or location. There is no word for ‘when’. Instead, a time word followed by *nia* is used to inquire about ‘what day, what time, how many days’, etc. The phrase *nia se* is used for future events. Time words precede the subject of a clause. The term *nfnanuf* ‘why’ always occurs initially and asks the cause of the addressee’s action. This word may be decomposable, but no declension is offered. Three words, *nir* ‘with’, + *nia* ‘what’, + *uf* ‘which’, are combined to question the purpose of something. To form manner questions, *fnaneno* ‘how’ is used, whereas *fnano* ‘how’ occurs in contexts such as ‘how can I help?’ To question the extent of a predication, an ‘extent word’ + *fnaneno* ‘how’ or *nia* ‘what’ occurs. Alternatives can be questioned by using *uf* ‘which’ + *nen* ‘how/where’.

The fourth sentence-level topic is negation. To negate a statement, the morpheme *ti* ‘not’ occurs clause-finally. Negative imperatives do not use *ti*, but conjugated forms of *fnate* ‘do not’: *o-bnate* ‘I do not’, *e-bnate* ‘you do not’. *n-fnate* ‘he does not’ + *of* ‘Obligatory Negative Imperative’ do. *Ti* is not used for expressing negative desire. Rather, a preverbal element with negative semantic, *bu*, is conjugated to agree with the person of the agent: *ngabu* ‘I don’t want’, *ngobu* ‘you don’t want’, *ngibu* ‘He/she doesn’t want’; however, the authors do not discuss the element *ngV-*. Two negative responses are cited: *firati* ‘no’, and *niati* ‘empty, none’.

Agent markers are said to prefix to the verb, and they note that these markers are different from free pronouns that occur in what they see as the subject slot. Furthermore, a verb phrase may contain one or more verb stems and optional preverbal elements that indicate tense, aspect, or mode, listed in order of occurrence in **TABLE 1.1**.

Tense	+	Aspect	+	Mode	+	Verb
<i>e</i> PST		<i>du</i> PROG		<i>g</i> DESIDERATIVE		
		<i>bu</i> NON-PROG				
		<i>su</i> COMPLETIVE				

TABLE 1.1 – Matsumura and Matsumura’s 1991 analysis of Tense, Aspect, Mood markers.

The non-progressive marker *du-* is in complementary distribution with progressive *bu-*. It is reported that in certain combinations, *g-* can be used for future tense, especially with a third person actor-subject. The morphemes *du-*, or *su-*, + *g-* always refer to something in the future, and *e-* + *g-* implies future tense. For past action, *e-* + *bu-*, or *su-*, are used. By combining tense, aspect, and mode morphemes, i.e., *e bu su g*, a meaning of ‘past action that is not yet realized’ can be achieved.

Other elements that the Matsumuras associate with the verb phrase are verbal modifiers for manner, time, location, benefactive, instrumental, comitative, causative, and referential (‘with’). Manner is typically expressed by a postverbal adverb. Adverbs are treated as verbs. For example, the morpheme *fi* is analyzed as a verbalizer that attaches to an adjective to derive an adverb. There is a table containing twenty-two temporal terms, such as ‘tomorrow’, ‘two days hence’, etc. These terms usually precede the verb, but can also precede the subject. Location modifiers are analyzed as prepositional phrases that follow the verb phrase, and are signaled by the generic preposition *nene* ‘at, in, on, for’. Benefactives also appear with *nene*. By contrast, instrumental, comitative, causative, and referential modifiers are indicated by *nir* ‘with, and, because of, about’. Speech-act verbs also occur with conjunctive *nir* ‘because’ to introduce a clause.

The Matsumuras’ analysis of directionals is based on four distinctive features that belong to a verb or combination of verbs. Directional verbs can be the main verb or supplement another verb with directional information. The first directional feature pertains to two major opposite directions in relation to the speaker, *fa* ‘go (away from speaker)’ versus *ma* ‘come (towards speaker)’. The second directional feature is determined by elevation, *iet* ‘go up (anything higher than the present location)’ versus *briet* ‘come up’. Ascent entails descent, *ro* ‘go down’ and *bro* ‘come down’. The third feature is context sensitive to direction across rivers, doorways, and streets, *ri* ‘go across’ versus *bri* ‘come across’. To help explain some of the intricacies of this system, they refer to a map of Arguni Bay. They assert that *fi* ‘make, have’ specifies the direction of a patient. It can be causative and must occur with a directional verb, as in /m-fi-n-iet/ *bi-n-iet* (2S-CAUS-3S go.up) ‘you put it up’. The fourth directional feature relates to *ra* ‘carry (take/bring)’, which they say always needs a directional verb to complete its meaning. One example they provide contains four verbs: *ba* ‘go’, *ud* ‘get’, *ra* ‘take’, and *briet* ‘come up’. The last verb carries the directional information.

Noun phrases do not contain case markers, but rely upon word order. The Matsumuras state that the phrase-final element *-ro* occurs frequently and indicates focus; but they also mention that it could be analyzed as a relativizer or a temporal conjunction ‘when’. In their analysis, focused elements occur clause initially. One constraint that is pointed out is that *-ro* does not usually occur with *ja* ‘first person singular’ (compare Voorhoeve 1989). The subject is the main focus in unmarked situations, so that *-ro* is not necessary, especially with pronominal subjects. Indirect objects appear in bi-transitive clauses, and are marked by *ragi* ‘give to’.

There are seven free pronouns, distinguishing first, second, and third persons, as well as singular and plural number. The first person plural distinguishes inclusive from exclusive referents. Dual pronouns are formed from a free pronoun and a numeral suffix. Free pronouns are required in independent clauses, but are optional in dependent clauses and in sequential text where the subject is “clear and continuously on the main line”.

Supporting previous publications, possessive pronouns are divided into alienable and inalienable. In Irarutu, inalienable possession relates to parts of the human body and is indicated by what appears to be pronominal circumfixes. Alienable possession is indicated by a prefix.

	PreMod	Head	Post Mod
NP =	POSS+	N+ ADJ	:Quant P + demonstrative + partitives

FIGURE 1.7 – Matsumura and Matsumura’s (1991) schematic for noun phrases. ‘Quant P’ is an abbreviated for ‘Quantifier phrase’.

A template for the noun phrase is presented, adapted as **Figure 1.7**, above. In cases where two nouns are in sequence, the second is attributive to the first, except for possession. Modifiers such as classifiers, numerals, and demonstratives follow the head noun. Two adjectives can accompany a single noun phrase, but a third adjective is relegated to a relative clause. For alternative noun phrases given in series, a disjunctive is observed after every linked element. ‘Post modifiers’ need clarification.

1.3.1.5 van den Berg and Matsumura 2008

This publication is a squib about possession in Irarutu. It is the most recent publication about Irarutu and gives a good overview of the possessive system. It is based on the East Arguni dialect from Gusimawa village. There is a template for possession that is justified by a variety of generalizations. Their discussion of the interrelation of morphology-syntax-semantics is noteworthy. The paper is framed by a phonology summary with discussion of allomorphy in the possessive system. The publication briefly mentions linguistic theory relating to possessives (Himmelmann 2005, Klamer et al 2007, Lichtenberk 1985); surprisingly, the authors declare that the paper is oriented towards data, not theoretical or historical inquiry. The most sensational feature they note is apparent infixation of possessive morphology.

These authors recapitulate the order of constituents presented in Matsumura and Matsumura 1991, highlighting the fact that possessors precede possessed heads. Major generalizations follow: there is a difference between alienable and inalienable possession, alienable possession is the default class and uses a possessive prefix (or proclitic), whereas inalienable possession has an apparent circumfix but could be a combination of a prefix and a suffix. Van den Berg and Matsumura analyze the apparent infix as being derived from frozen compounds. An alternative possessive strategy that they cite has the clitic *=ri* after the possessor but before the possessed noun. In total, six varieties of inalienable possession are identified. The distinction between inalienable and alienable possession is lost for third person referents.

1.3.2 Work on nearby AN languages

Progress on the internal classification of the SHWNG subgroup is slow to unfold. Lewis (2013) cites forty-one SHWNG languages. Only fifteen of these have some kind of published linguistic description (see TABLE 1.2). A full analysis of the relevant publications is outside the scope of the present work. However, for illustrative purposes the SHWNG languages are listed alphabetically with references to published linguistic descriptions or present research. Several undergraduate theses at UNIPA, such as Karubaba (2000), focus on some aspect of the languages of West Papua; however, due to their inaccessibility, it has not been possible to look closely at all of the work that has been done. It is likely that several undergraduate research projects could be added to the list.

There are some general surveys of interest. Van der Crab 1862 and De Clerq 1890 are early surveys of the Moluccas and Halmahera, respectively, Teljeur 1982 is a more recent survey of the same area, Silzer and Ajamiseba 1981 is a fairly recent survey of Yapen island, Walker 1983a is a survey of the western end of the Bomberai peninsula, and Sawaki 2006 is a recent survey of the Raja Ampat islands. Silzer and Heikkinen (1984) and Silzer and Clouse (1991) present several useful maps of 251 languages in Indonesian Papua and many references to older linguistics-related literature.

Three Austronesian languages spoken in the Northwest tip and West Coast of the Bomberai Peninsula, Onin, Sekar, and Uruangnirin, subgroup with CMP languages of Southeast Maluku under Yamdena-Bomberai (Blust 1993). In the south, Koiwai is putatively a CEMP isolate (R. Walker 1990). Data on two languages are unavailable and are probably names for areas, not languages: Bedoanas (Wurm 2000 cites 180 speakers) and Erokwanas (Wurm 2000 cites 200 speakers). For inhabitants of West Papua, these two names refer to locations rather than languages. It is important to note that interest in the languages of Indonesian Papua is increasing. Nikolaus Himmelmann and several of his associates have been working at the Center for Documenting Endangered Languages, which is supported in part by UNIPA. Also, UNIPA hopes to offer a post-baccalaureate degree program in Linguistics starting in 2015. The combined efforts of these organizations will certainly provide more language data in the years to come.

<u>Ambai</u> (Karubaba 2000; P. Silzer 1982a, b, 1983; S. Silzer 1979; Silzer and Silzer 1985)	Irarutu (Fields n.d.; Mandopma 1992; M. Matsumura 1992, 1997; T. Matsumura 1984, 1985a,b, 1991, 1999; Matsumura and Matsumura 1991; van den Berg and Matsumura 2008; Voorhoeve 1989, 1995)
<u>Serui-Laut</u> (Slump 1924–38)	Iresim/Yiresiem (Kamholz p.c.)
<u>Woi</u> (Sawaki forthcoming)	Kawe ———
<u>Pom</u> ———	Kurudu ———
<u>Ansus</u> (Price and Donohue 2009; Saragih 2006)	Legenyem ———
<u>Munggui</u> ———	Maden ———
<u>Papuma</u> ———	Marau ———
As ———	Matbat (Remijsen 2002, 2010, n.d.(a, b))
Biak (Fautngil et al. 1994; Kern 1885; Mandowen 1999; Meyer 1874a, b; Mofu 2005, 2008; Ottow 1862; Patz 1978; Soeparno 1975, 1976, 1977a, b; Steinhauer 1985, 1986; van den Heuvel 2006; F. van Hasselt 1902a, b, 1905, 1936, 1947; J. van Hasselt 1868, 1876, 1876/1893; van Hasselt and van Hasselt 1947; Wanma n.d.)	<u>Maya</u> (Remijsen 2002; van der Leeden 1980, 1993, 1995)
Biga (Remijsen 2002)	<u>Banlol</u> , <u>Tepin</u>
<u>Buli</u> (Maan 1940, 1951)	Mor (Laycock 1974; Kamholz p.c.)
<u>Maba</u> ———	Taba/East Makian (Bowden 1997, 2001)
<u>Patani</u> ———	Tandia (Deda 2012, Kijne n.d.(a))
<u>Weda</u> /Sawai (J. Whisler n.d.; R. Whisler 1992, n.d.; Whisler and Whisler 1995)	Wabo ———
Busami ———	Waigeo/Amber ———
<u>Dusner</u> (Dalrymple and Mofu 2011, 2012; Deda et al 2011a, b; Kijne n.d.(b))	Wandamen (Bink 1891; Cowan 1955; Flaming 1981, 1983a,b; Henning and Henning 1991; Ongkodharma and Nelwan 1983; Ongkodharma et al n.d.; Warami n.d.)
Ron/Roon (Kijne n.d. (b))	Waropen (Held 1942a, b, 1956, 1957; Kijne n.d.(b); Ramar et al 1983; Saggars 1979, Walker 1983b)
<u>Meoswar</u> ———	Wauyai ———
Gane/Gimán (Teljeur 1983, 1990, 1994)	Yaur (Kamholz p.c.)
	Yeretuar (Kamholz p.c.)

TABLE 1.2 – South Halmahera West New Guinea languages and (linguistic and anthropological) research that is related to their documentation. Alternate names are given after a backslash ‘/’. Underline indicates names of dialects; whereas double underline indicates names of languages that are thought to belong to, or constitute a dialect chain, as per Voorhoeve (1989).

1.3.3 Comments on neighboring Papuan languages

Nearby Papuan languages, listed clockwise starting in the Southeast are: Mairasi, Kamrau (Kamrau), Buruwai, Baham, Mor, Kembrano (Kemberano), and Tanah Merah (Lewis 2013). Mairasi had approximately 3300 speakers in 1996. There is a description of Mairasi phonology (Peckham 1991), and Mairasi Grammar (Peckham 1978) in which it is postulated that the language has two contrastive tones and prenasalized stops. Southern Mairasi speakers are bilingual in Koiwai, whereas Northern Mairasi speakers are bilingual in Irarutu. Mairasi people are said to be semi-nomadic, and have experienced slavery by nearby groups. Mairasi is also the name of a language family that contains Mairasi, Semimi (1000 speakers in 1991), and Mer (85 speakers in 2000), which are located to the East and South of Mairasi. Kamrau (approx. 1500 speakers in 1993)

and Buruwai (Anceaux's Sabakor, 1000 speakers in 1990) belong to the Asmat-Kamoro subgroup, Sabakor branch, of the large Trans-New Guinea phylum. Baham, which is nearby, but not adjacent to Irarutu, had 1100 speakers in 1993, and belongs to Trans-New Guinea, under West Bomberai. Mor is also a Trans-New Guinea language, belonging to its own subgroup, and is endangered. In the year 2000 it had merely 25 speakers (Lewis 2013). Kembrano (Sebyar) is classified as a South Bird's Head language within Trans-New-Guinea. It had 1500 speakers in 1987. Tanah-Merah, a Trans-New Guinea language in its own subgroup, had only 500 speakers in 1978. Finally, it was suggested by Reesink (1998, 2005) that there may be a connection between Irarutu and languages of the Bird's Head, such as Sougb. Concrete evidence has yet to be presented.

Several features in Irarutu, can be seen as areal. Focusing on the phonological material in Anceaux 1958, Voorhoeve 1989, and Matsumura 1991, it can be seen that Irarutu's tendency to have long words with open syllable types in careful speech but shorter words with consonant clusters, which is suggested by Voorhoeve's orthography, or Matsumura's analysis of syllable structure, is similar to Anceaux's generalization about word shape in Tanah Merah. Also, the 'strong bilabial fricative' and phonetic lengthening of vowels in Kamberano (Anceaux's Marau) appear to correlate with Matsumura's description of Irarutu. Structural similarity strongly suggests that transfer in other areas, particularly in the case of lexical items, may have also occurred. The task of identifying the source of Irarutu's non-Austronesian lexical items has not progressed very far, but judging from the fact that all of the neighboring non-Austronesian languages belong to the Trans-New Guinea family, it seems likely some Trans-New Guinea etyma will arise. Nanggoa (Voorhoeve's 'Nanggwa') could also be a candidate, but data is not available.

1.3.4 Vernacular and other publications

Two additional kinds of relevant literature are vernacular literature, such as a translation of the Gospel of Luke (Matsumura and Matsumura 1999), and non-linguistic descriptions of the Irarutu people, such as M. Matsumura's (1992) description of kinship and marriage.

The only vernacular publication that is widely known is *Lukas nfier frgrgir snan fifier dir weni* 'the Gospel of Luke' (Matsumura and Matsumura 1999). However, there are two Irarutu music videos that can be found on YouTube: 'Mumri Nuf' and 'Mo Irana'. Both are sung by the artist Suryani Paurada. Lyrics accompany the music in the style of karaoke, for sing-along. They are well-known to most Irarutu people.

There are two anthropologically oriented publications regarding Irarutu. M. Matsumura published a concise paper that discusses kinship and marriage in the Irarutu community (1992, revised and reprinted 1997). Mandopma (1992) described the system of traditional culture knowledge in Fruata village, but that research is not included in this description.

1.4 Overview

Because this description of Irarutu contributes to documentation of the language, it was important to remain theory-neutral as much as possible. The framework for any given description of a language is based on a variety of approaches, depending upon the researcher's understanding of general linguistic theory, personal interests, and the target language. In this regard, dissidents are prone to fault present conceptions of documentation for lack of standardization, both in approaching language data and in the ability to tie it in with a general conception of language. For documentationists, it is hoped that success within the discipline will bring about reform in the field of general linguistic theory by way of standardization, that is, the development of an adequate theory of language that can be used universally in the effort to document endangered languages, such as Comrie and Smith's extensive 'Lingua Descriptive Studies: Questionnaire' (1977) and Dixon's Basic Linguistic Theory (2011–12).

On account of the larger language documentation context, this description of Irarutu aims for a theory-neutral descriptive framework that will be easy for future researchers to compare with other languages, especially SHWNG languages. It is merely a picture of what is known about the language as seen through the lens of contemporary linguistics. Doubtless, even for the Fruata dialect, there will be other researchers whose interests are different, or whose conception of language is different. There will arise alternative interpretations and representations of the data, such as found in the publications that were reviewed above, but it is improbable that any such analysis, or any future research effort, will be able to answer all questions that the data raise.

One assumption worth articulating here is that in describing any human activity, it is not entirely possible to separate what is happening at a specific moment from the conditions that brought it about. That is why synchrony (Voorhoeve 1989, 1995; Matsumura 1991; Matsumura and Matsumura 1991) and diachrony (Blust 1978, Ross 1995) both play important roles in describing a living language. Furthermore, it is acknowledged that language is frequently viewed as being componential. Various aspects of language can be isolated from one another such as phonology, morphology, syntax, and so forth. At the same time, any single component of language is also interconnected with many others (Pike 1947). For example, languages have certain sounds that are used to differentiate meaning when combined to form morphemes and words; but to discover the meaningful sounds of a language, one must consider not just words but all venues that use sounds distinctively, such as clauses.

This description does not explore in any great detail the social or psycholinguistic parameters that play a role in the language, nor is it an ethnography of communication (Hymes 1968). Such projects can occur later.

1.4.1 Methodology

Each visit to Papua was filled with intensive elicitation, documentation, and participant-observation. Between field visits, the data was analyzed and internalized, research was done within secondary sources, and questions were formulated about the language. This process allowed for less time in the field than would be necessary otherwise to achieve

the same results, but it has taken longer overall. At the same time, the length of study signifies a continuing commitment to Irarutu.

In 2006, speakers of several SHWNG languages were recruited in order to collect basic language data. A 200 word Indonesian language list adapted from the Swadesh wordlist, kindly supplied by Hugo Warami, was utilized. The wordlist was given to the language consultants and they were asked to translate Indonesian words into Irarutu. They were also asked to prepare short personal narratives, including information such as place of birth, occupation, religion, parents' work, and number of siblings. The speakers were subsequently asked to read items from the wordlist and their personal narrative in their native language. These events were recorded on mini-disc and mini-DV. The setting was fairly formal. The speaker sat against a wall facing the microphones and camcorder. The researcher sat off camera but beside the consultant in order to look at the list together.

In 2010, a longer, 3000 item Indonesian language wordlist was prepared, based on the English language index in Fox 1978. The list was printed twice, once for Irarutu and once for Kuri. These wordlists then went to students at Kaimana/Bintuni dormitories where younger native speakers provided Irarutu words and phrases to describe the Indonesian concepts in the wordlist. Flaws in this method were redundancy and excessive variation. For the Irarutu data, there were several different kinds of handwriting on the lists, so it is assumed that several different speakers contributed. Four speakers of Irarutu, and three speakers of Kuri were recorded. Kuri language data has been included at several points, but not discussed at length. Each new participant in a recording session was encouraged to make appropriate changes as he saw fit. However, none of the college student consultants felt comfortable speaking at length in Irarutu. An older male speaker of Irarutu was asked to review the wordlist and make additions or corrections as well as produce several sentences in Irarutu. A significant number of the sentences he provided are reviews of data from Matsumura and Matsumura 1991, or translations of Indonesian sentences. However, there are several examples of spontaneous speech. Again, the recording format was fairly formal. The recordings were corrected for redundancy after the visit, and several hundred words from the Intercontinental Dictionary Series (kindly supplied by David Kamholz) were added.

In 2013, a refined wordlist was brought to the field, the researcher had a better mental vocabulary, and numerous questions about noun phrases, verb phrases, prepositional phrases, clauses and complex sentences were prepared. Several new approaches based on previous visits were utilized, to improve elicitation of Irarutu discourse, such as the short conversation between husband and wife (see Appendix 2). After finding the primary consultant, a highly-qualified language teacher named Mr. Damianus Fenentruma, and a host family, the project focused on four activities: to correct and expand the wordlist, to record naturalistic language data, to elicit examples of specific grammatical structures, and to speak Irarutu. Several additional consultants helped expand the wordlist, provide syntax data, and produce examples of discourse. Language data was written down in a notebook and on a laptop computer. Several times in public, and without access to the notebook, I was prompted to use Irarutu with Irarutus known to the host family. In a relatively short time, it became possible to understand the spoken language, speak the

language, and elicit data in it. Obviously there is much more to learn and even more to do in the process of documenting Irarutu, but being immersed in an Irarutu language context provided some intuition about the basic components of the language and supplied a representative sample of relevant data. The recording sessions were considerably more relaxed. No camcorder or lapel microphones were used during recording sessions.

1.4.2 Organization

Several grammars of Austronesian languages have influenced the organization and contents of this grammar, especially Bowden (2001), Guérin (2007), Robinson (2008), and Thieberger (2004). For the sake of clarity, a simple structure is used in organizing the description. The first chapter provided background information on Irarutu, including a literature review. Chapter 2 describes a range of synchronic phonology and morphophonology topics. Chapter 3 describes key points in Irarutu morphosyntax and syntax, with a significant portion devoted to parts-of-speech. Chapter 4 explores the historical phonology and genetic affiliation of the language. Appendix 1a gives acoustic data that demonstrate particular characteristics that are relevant to the description of Irarutu phonology. Appendix 1b provides a feature chart. Appendix 2 contains two short examples of discourse. Appendix 3a provides a list of comparative material. Appendix 3b lists specific sound correspondences identified in proposed cognates in appendix 3a. Appendix 4 lists media files and other metadata associated with the research. Works cited are located after the appendices.

Most Irarutus above 30 years old are aware that the Irarutu language is declining. Many want to find a way to protect their community of practice. Some priorities for future work are therefore: lexicography, dialectology, collecting more examples of discourse, ethnography, psycholinguistic experimentation, and, most importantly, developing pedagogical resources. These documentation tasks are necessary for the effort to conserve and maintain the language and traditional knowledge. In this regard, one of the main, albeit inadvertent, contributions of the present description is a set of orthographic conventions that meets the needs of various uses of the language, from casual representation to detailed analysis.

Chapter 2. Phonology and Morphophonology

This chapter describes contemporary Irarutu phonology and morphophonology. The first section (§2.1) identifies phonemes using minimal and near-minimal pairs and introduces basic phonotactic distributions of those phonemes. The second section (§2.2) discusses the subphonemic features of Irarutu segments, and the third section (§2.3) describes in more detail several instances of allophony. The fourth section (§2.4) describes, in more detail, Irarutu phonotactics and the need to differentiate two types of syllables. The section continues by describing stress placement and suprasegmental properties of the phonological word. The fifth section (§2.5) briefly discusses loanword phonology, while the sixth section (§2.6) describes several instances of allomorphy that are relevant to the description of Irarutu morphosyntax. A summary is provided at the end (§2.7).

2.1 Phoneme inventory

The Irarutu phoneme inventory is introduced using phonetic terminology. The consonant phonemes are depicted in **TABLE 2.1**, vowels in **TABLE 2.2**.

	Labial	Alveolar	Palatal	Velar	Laryngeal
Stop, voiceless	/p/	/t/	/c/	/k/	
Stop, voiced	/b/	/d/	/j/	/g/	
Fricative	/f/	/s/			/h/
Nasal	/m/	/n/			
Tap/trill		/r/			
Glide	/w/				

TABLE 2.1 – Consonant phonemes. /c/ occurs most frequently in ‘loan’ words. /h/ occurs exclusively in loanwords. /r/ can be a non-lateral approximant or a tap. /w/ is a gliding labial approximant.

This analysis differs from previous literature (see Chapter 1) by including two segments found in recently ‘borrowed’ words: /c/ and /h/. The segments /c/ and /j/ are classified in **TABLE 2.1** as stops, but they are close in articulation to the complex segments [tʃ] and [dʒ] because they have mildly fricated release (see Appendix 1a); however, based on phonological evidence presented in §2.2 below, these segments pattern with stops.

	Labial	Labiopalatal	Palatal	Neutral
High (+tense)	/u/		/i/	
High (-tense)		/y/	/ɪ/	
Mid	/o/		/e/	
Low				/a/

TABLE 2.2 – Vowel Phonemes. ‘Labial’, which is predictable for the back vowels /u/ and /o/, can be interpreted as Ladefoged’s (2001) ‘round’, whereas ‘palatal’ can be interpreted as ‘front’. ‘Neutral’ is a property of achromatic vowels, i.e., not palatal or labial. /y/ is a labiopalatal lax high vowel [Y].

This portrait of vowels is similar to the treatment by van den Berg and Matsumura (2008) in having a distinction between tense and lax high vowels, but differs from the analysis of Voorhoeve (1989), who proposed that the tense/lax distinction was associated with the mid vowels, that is, /e, ε, o, ɔ/. Although it has not been mentioned in previous publications, in casual speech there is a generous amount of overlap, among the three palatal vowels, i.e., /fi/ [fi ~ fi ~ fe] ‘active verb marker’, and to a lesser extent, the labial vowels, e.g., /ny/ [nY ~ nu] name and /nu/ [nu ~ nU] ‘island’. Further details are provided in §2.1.2.

2.1.1 Consonants

Each phoneme of Irarutu is described in this section using articulatory terminology, and supplemented by comments about relative frequency and distribution according to position in words, initial (C-), medial (-C-), and final (-C). The fifteen consonant phonemes /p, t, c, k, b, d, ʃ, g, f, s, h, m, n, r, w/ are identified according to voicing, manner of articulation, and the articulators used to produce them.

The consonant inventory is fairly symmetrical. Voicing is distinctive for stops. Voiced and voiceless stops are distinct from one another /p : b/, /t : d/, /c : ʃ/, /k : g/. The difference is further emphasized by the fact that voiced stops are phonetically prenasalized [ᵐb], [ᵑd], [ᵑʃ], [ᵑg]. Unlike the stops, which occur at four points of articulation, the fricatives and nasals only occur at two points of articulation. The bilabial stop is distinct from, but aligns with, the labiodental fricative /p : f/, and the alveolar stop is distinct from the corresponding fricative /t : s/. The glottal fricative /h/ has not been mentioned in previous analyses of Irarutu (Matsumura 1991, Voorhoeve 1989). Bilabial and alveolar stops are distinct from nasals /b : m/, /d : n/. The rhotic /r/ is a non-lateral approximant [ɾ], or optionally a tap [ɾ]. It can also become a full trill [r] in certain environments. It is distinct from the voiced stop /r : d/. The labial glide /w/ aligns with one of the major points of articulation for the stops.

<u>/C-/</u>	<u>/-C/</u>
[pa] ‘nail’	[ᵑdap] ‘can/able’
[ti] ‘negative marker’	[it] ‘we (inclusive)’
[kas] ‘kit’	[ᵐbwek] ‘betel nut’
[ᵐba] ‘you go’	[ᵐbaᵐb] ‘mush’
[ᵑdir] ‘good’	[refiᵑd] ‘side’
[je] ‘meat/flesh’	[aᵑdifeʃ] ‘determiner’
[ᵑgin] ‘I drink/you drink’	[fiᵑg] ‘fly’
[fi] ‘active verb marker’	[nif] ‘recount’
[si] ‘to see’	[mis] ‘lure’
[ma] ‘come’	[am] ‘we (exclusive)’
[ne] ‘to, at’	[sen] ‘money/gold’
[ri] ‘have’	[ir] ‘they’

TABLE 2.3 – Segments that contrast by position in monosyllabic words.

Because Irarutu possesses a significant number of monosyllabic lexical items, a chart (cf. Hockett 1955, 1958) can be constructed that pairs words containing each phoneme in contrastive positions, i.e., initial versus final (see **TABLE 2.3**). For example, [pa] ‘nail’ starts with /p/, whereas [dap] ‘can/able’ ends with /p/.

Twelve of the fifteen Irarutu consonants /p, t, k, b, d, ʝ, g, f, s, m, n, r/ are observed in initial and final position. The remaining three segments /c, h, w/ have positional constraints, which are described below, where each consonant phoneme is reviewed and justified in relation to segments that contrast in single feature values. Redundancy of presenting minimum pairs in illustrating the segments is kept to a minimum. The presentation is ordered according to manner of articulation, for which general comments are given before describing the individual segments.

2.1.1.1 Stops

There are eight stops /p, t, c, k, b, d, ʝ, g/, which, as mentioned above, contrast according to four points of articulation (labial, alveolar, palatal, and velar), as well as voicing (voiced or voiceless). This is the largest natural class according to manner of articulation. All stops can occur in initial position. /c/ is the only stop not attested in final position. A voiceless stop can differ in place from a preceding nasal, as in /fim^hta/ [fim^hta] ‘weapon’, but is often separated from it by slight vocalic release in careful speech. When the sequence is homorganic, the release feature is absent, even in careful speech. That is, /ntamn/ is realized as [n^hta^hmn] ‘her mother-in-law’, with no release after the nasal onset. In contrast, voiceless stops are optionally aspirated when they form consonant clusters with certain other consonants in careful speech, /tftfrie/ [tftfrie ~ t^hft^hfrie] ‘dragonfly’. They are often aspirated before vowels in careful speech /ntu/ [ntu ~ nt^hu] ‘young, new’ and can be aspirated at the end of a word /fit/ [fit ~ fit^h].

The voiceless bilabial stop /p/ is relatively infrequent, but it does occur in a number of loanwords from Indonesian. In initial position, examples are [p^hintr] ‘smart’ (cf. IND pintar), [pa] ‘nail’ (IND paku), [pwar] ‘knock’ (IND menampar), [pjar] ‘nourish’ (IND pelihara), [putr] ‘turn’ (IND putar), in medial position [‘ka.p^har] ‘head/chief’ (IND kepala), [‘kap^h.ri] ‘ship’ (IND kapal), and in final position [mse^hp] ‘kick’ (IND sepak), and [‘ⁿdap] ‘can’ (IND dapat). However, there are /p-/ initial words of presently unknown provenance, such as [‘pu.a] ‘big, much’, [pi^h.pi.si] ‘money’ (Kaimana dialect), and [‘pa.nim] ‘bottle’. In final position, /-p/ is rare but it does occur occasionally. Examples of this segment in contrast with the voiceless stops /t, c, k/, the voiceless fricative /f/, and the voiced bilabial stop /b/, are given below.

p : t

[pof] ‘k.o. stairs’ : [tof] ‘cliff’

[‘ⁿdap] ‘can/be able’ : [‘mmat] ‘green/unripe’

p : c

[pa] ‘nail’ : [‘car] ‘instruct’

p : k

[pa] ‘nail’ : [‘ka.ku.ri] ‘tree kangaroo’

[a.‘pa.pr] ‘butterfly’ : [ka.‘kan] ‘flower’

p : f

[pa] ‘nail’ : [fa] ‘go’
[dap] ‘can’ : [taf] ‘yellow’

p : b

[pa] ‘nail’ : [ᵐba] ‘you go’

The voiceless alveolar stop /t/ occurs frequently, and is attested in initial position [te] ‘feces’, medial position [a.tif.ro] ‘top/thatch’, and final position [ne.fut] ‘younger sibling’. It ranges from post-dental [t̪] to alveolar [t]. Examples of /t/ in contrast with the voiceless stops /c, k/, the fricative /s/, and the voiced stop /d/ are given below. For a minimal pair of /t/ versus /p/, see the description of /p/.

t : c

[tar] ‘split’ : [car] ‘instruct’

t : k

[tar] ‘split’ : [kar] ‘bracelet’
[taᵐb] ‘add’ : [kaᵐ.bja] ‘bury’

t : s

[tar] ‘split’ : [sar] ‘false’
[ɪt] ‘we (inclusive)’ : [ɪs] ‘earthquake’
[ᵐbit] ‘child’s play’ : [ᵐbis] ‘can’

t : d

[tᵐbe] ‘hit’ : [ᵐdbe] ‘he hits’
[nut] ‘stab’ : [nuᵐd] ‘he brings’
[taᵐd] ‘line’ : [tat] ‘shell’

The voiceless palatal stop /c/ is relatively rare, but only appears in words that are thought to be borrowed from Indonesian, such as [coᵐb] ‘try’ (cf. IND coba ‘try’), and [car] ‘instruct’ (IND ajar ‘teach/instruct’). Notably, [c] is also an optional allophone of /t/ in certain contexts, e.g., /utie/ ‘sago starch’ is realized as [utie] (careful speech) or [uce] (casual speech). For more details, see §2.3. /c/ occurs in initial position [car] ‘instruct’, but is not attested, phonemically, in medial or final position. Examples of /c/ in contrast with the voiceless stop /k/, and the voiced palatal stop /j/ are given below. For examples of /c/ versus /p/ and /c/ versus /t/, see the descriptions of those segments, above.

c : k

[car] ‘instruct’ : [kar] ‘bracelet’

c : j

[car] ‘instruct’ : [jar] ‘I give (casual speech tempo)’

The voiceless velar stop /k/ is fairly common in consonant clusters with another /k/, [kke] ‘jaw’ and [kkor] ‘chicken’ (for more details see the description of reduplication in §2.6), and when it precedes /r/, as in [m.kr.kur] ‘fresh, young’, [krafr] ‘choker (k.o. necklace)’, [fukr] ‘net (for catching things)’, and [krikr] ‘smooth’, but it is somewhat infrequent on its own in initial and final positions, [kar] ‘bracelet’, [kakuri] ‘bat’, and [ᵐbwek] ‘betel nut’. It shows greater frequency of occurrence in medial position: [wakuri] ‘tree

kangaroo’ and [ᵐba.ri.ku] ‘lighter’. An example of /k/ in contrast with the voiced velar stop /g/ is provided below. Examples of this segment in contrast with the voiceless stops /p, t, c/ are provided above.

k : g

[kaᵐd] ‘pants’ : [ᵑga] ‘eat’

[kar] ‘bracelet’ : [ᵑgar] ‘dig’

Phonemic voiced stops in the Fruata dialect are phonetically prenasalized (^NC): [ᵐb], [ᵐd], [ᵐj], [ᵑg]. This property is thought to increase distinctiveness between voiced and voiceless stops because the lowered velum for the initial portion of these segments allows greater duration of voicing. One important exception is /j/, which is – more frequently than not – realized without prenasalization, but there is a considerable amount of interspeaker variation. Some speakers prenasalize this segment with a high degree of regularity. Otherwise, exceptions to prenasalization in voiced stops are extremely rare.

Prenasalized stops are different from nasal-plus-stop sequences. A nasal is separated from a following voiced stop by vocalic release, i.e., /nd/ becomes [nᵐd]. For example, /ndridn/ becomes [nᵐdriᵐdn] ‘sick’, which contrasts with /dridn/ [ᵐdriᵐdn] ‘cold’. Another example is the near-minimal pair /timbe/ [tᵐᵐbe] ‘all’, in contrast with /tbe/ [tᵐᵐbe] ‘hit’ (see §2.3 for discussion of *r*-epenthesis).

The voiced bilabial stop /b/ occurs frequently, partly because it results from allomorphy of *f*-initial verbs with second person agents. It occurs in initial and medial position [ᵐbaᵐba] ‘elder sibling’, as well as final position [coᵐb] ‘try’. Voorhoeve (1989) reports that /b/ may devoice in medial or final position. However, for the Fruata dialect speakers consulted for this description, /b/ is phonetically prenasalized, and never devoices. Examples of /b/ in contrast with the voiceless stop /p/, the voiced stops /d, j, g/, the nasal /m/, and the glide /w/ are presented below. The contrast of /b/ versus /p/, is given above.

b : d

[ᵐbir] ‘canoe’ : [ᵐdir] ‘good’

[ᵐboᵐb] ‘smoke’ : [ᵐdoᵐb] ‘satchel’

b : j

[ᵐbu] ‘not yet’ : [ᵐju] ‘canoe’

[ᵐbir] ‘separate’ : [ᵐjir] ‘gum’

b : g

[ᵐbis] ‘great’ : [ᵐgis] ‘gums’

b : m

[ᵐba] ‘you go’ : [ᵐma] ‘come’

[ᵐbri] ‘come in’ : [ᵐmri] ‘go in’

b : w

[ᵐbar] ‘lung’ (underlying root in /brbar/) : [ᵐwar] ‘root’

The voiced alveolar stop /d/ is articulated at a post-dental [ɖ] or alveolar [d] place and is attested in initial [ᵐdir] ‘good’, medial [oᵐdeno] ‘your mother’, and final positions

[na^mbaⁿd] ‘big’. Interestingly, there are a few recorded instances in the literature within the past hundred years where /d/ varies with the rhotic /r/, for example duguin ~ runguin ‘head’ (Anceaux 1961, Smits and Voorhoeve 1992). Unfortunately, the specifics of this variation are not clear. Examples of this segment in contrast with the voiced stops /j, g/, the nasal /n/, and the /r/ are provided below. For the contrast of /d/ and /b/, see the description of /b/ above. Palatalization of /d/ causes the contrast between /d/ and /j/ to be problematic in some instances.

d : j

[ⁿdu] ‘not.yet’ : [ju] ‘canoe’
 but, *d ~ j*
 [aⁿdje] ~ [aⁿje] ‘my father’

d : g

[ⁿdap] ‘can/able’ : [ⁿga] ‘speech act participant eats’
 [muⁿd] ‘you hold/take’ : [muⁿg] ‘much, big’

d : n

[ⁿdu] ‘not.yet’ : [nu] ‘name’
 [ⁿdir] ‘good’ : [nir] ‘with’
 [tafaⁿd] ‘elder sibling’ : [fran] ‘widow (randa)’

d : r

[ⁿdir] ‘good’ : [rir] ‘seed’

The voiced palatal stop /j/ is analyzed as a stop rather than an affricate because some speakers prenasalize it in a range of lexical items, especially in an intervocalic context, but some speakers do so inconsistently, while others rarely – if ever – do. In addition, there are no known instances of prenasalized /j/ in initial position in the free pronoun /ja/ [ja] ‘I’, *[ⁿja]. This consonant is phonetically close to the affricate [dʒ], articulated with both the tip and blade of the tongue, [j ~ dʒ ~ (ⁿ)dʒ]. In several non-nativized loanwords, the stop [j] varies with the glide [j] (see §2.2.3). There is also a historical connection between these two segments. In instances where PCEMP *y (which represents a palatal glide, /j/, in the Austronesian diachronic literature) is reflected as the onset of a syllable, it is reflected as /j/ [⁽ⁿ⁾j] in present day Irarutu. See Chapter 4 for more discussion of Irarutu historical phonology. Furthermore, in the closely related linguistic variety, Kuri, many lexical items still reflect *y as [j]. /j/ contrasts in place of articulation with /b, d, g/, see above descriptions of /b/ and /d/ for examples of contrast with these segments and /j/. No known examples show contrast in manner between [j] and the palatal glide [j].

j : d

[⁽ⁿ⁾ju] ‘canoe’ : [ⁿdu] ‘not.yet’
 [aⁿje ~ aje] ‘my father’ : [aⁿdena] ‘my mother’

j : g

[ja] ‘I’ : [ⁿga] ‘eat’

j ~ j

[ja] ‘I’ ~ [ja] ‘I’ (Cowan 1953)
 [s^mbajan] ~ [s^mbajan] ‘prayer/mass’

The voiced velar stop /g/ occurs is frequent. It is found in initial and medial position [ʔgiʔgi], as well as final position [mʊʔg] ‘much’. One fairly clear example of a voiced stop that is not prenasalized in the Fruata dialect, at least in careful speech, is initial /g/ in /gmon/ [gmon] ‘reflection’, but it does become prenasalized in casual speech [ʔgmon]. See description above for examples of contrast.

2.1.1.2 Fricatives

There are three phonemic fricatives in Irarutu, /f, s, h/. All of them voiceless.

The voiceless labiodental fricative /f/ in Irarutu has been analyzed in the literature as a bilabial fricative (Matsumura 1991). This could be true for the East Arguni dialect, but in material relating to the same segment in other dialects (e.g., 2010 video footage of a language consultant smiling while articulating the segment, with the upper teeth and lower lip as the articulators), there are strong indications that it is better analyzed as a labio-dental fricative [f] despite strong lip protrusion, which is a common anatomical feature of Irarutu speakers. In addition, all else being equal, it is common for speakers to articulate /f/ as [v], with voicing. The likelihood of this phoneme surfacing as a voiced segment can be predicted by environment, such as in intervocalic position /refefa/ [refefa ~ reveva] ‘afternoon’, but even so, the two realizations occur in free variation. Since /f/ is frequently voiced, it is helpful to note that /f/ contrasts with /w/: [far ~ var] ‘keep’ versus [war] ‘root’. As mentioned in the description of /b/, above, there is an allomorphic connection between /f/ and /b/ (see §2.6 for more details). /f/ occurs in initial position [fa] ‘go’, medial position [refefa] ‘afternoon’, and final position [nof] ‘wind’. Examples of this segment in contrast with the voiceless fricatives /s, h/ are provided below. For an example of the contrast between /f/ and /p/, see above (§2.1.1.1).

f : s

- [fu] ‘bony protrusion’ : [su] ‘completive’
- [ʔfi.ti] ‘forest rat’ : [ʔsi.ti] ‘betel (nut)’
- [fwir] ‘center, navel’ : [swir] ‘k.o. wooden utensil’
- [ʔgaf] ‘to write’ : [fas] ‘rice’
- [friʔr] ‘wing, wide’ : [friʔs] ‘k.o. ethnic dance’

f : h

- [fa] ‘go’ : [hak] ‘rights’

The voiceless fricative /s/ occurs frequently. It ranges from alveolar [s] to slightly post-alveolar [sʰ], and is attested in initial position [simbua] ‘chili’, medial position [esu] ‘one’, and final position [fas] ‘rice’. /s/ contrasts in place of articulation with /f/ (see above) and /h/, and in manner of articulation with /t/ ([is] ‘earthquake’ : [it] ‘we (inclusive)’ and see description of /t/, above). Unlike /f/, voicing has never been observed in any contexts in the realization of /s/, i.e., [z] cannot be substituted for [s].

s : h

- [su] ‘completive’ : [hukum] ‘justice’
- [ʔta.her] ‘man’s name’ : [ser] ‘k.o. leaf’

The voiceless glottal fricative /h/ is perhaps the newest addition to the Irarutu phoneme inventory. It is only found in recent loanwords from Arabic via Bahasa Indonesia, or perhaps some neighboring language, in proper nouns and legal terms, and it is still limited to this lexical stratum. Two of the consultants for the present research, both from the Kaimana area, whose inhabitants are predominantly Muslim, have /h/ in their names: [ʔa.hɛr] ‘man’s name’ and [loru.hama] ‘woman’s name’ (this word appears to be a non-nativized borrowing, i.e., it contains a segment [l] that is otherwise not found in the Fruata dialect of Irarutu). The segment /h/ is found in initial position [hak] ‘rights’ (IND *hak*), [hukum] ‘justice’ (IND *hukum*), and medial position [ʔa.hɛr] (man’s name). It is not attested in coda position. Examples of contrasts are given above.

2.1.1.3 Nasals

Irarutu nasal phonemes are inherently voiced. They are articulated with oral closure at either of two points of articulation, bilabial or alveolar. It is worth noting that the voiced plosives are prenasalized /^mb, ⁿd, ^ɲj, ^ŋg/, which causes many people to write a velar nasal ‘ng’ preceding the voiced velar stop /g/, but this represents non-phonemic material. Nasal segments are frequent in Irarutu.

The bilabial nasal /m/ is involved in two productive areas of the grammar system (see §2.6). It occurs in initial position [ma] ‘come’, medial position [rimun] ‘citrus’, and final position [am] ‘we (exclusive)’. Examples of contrast between this segment and the alveolar /n/, the stop /b/, and the approximant /w/ are given below.

m : n

[ma] ‘come’ : [na] ‘he eats’
 [mif] ‘polite’ : [nif] ‘count’
 [mʔ.gi^ŋgr] ‘you ask’ : [nʔ.gi^ŋgr] ‘he asks’

m : b

[mis] ‘lure’ : [^mbis] ‘able to’
 [^mdam.ri] ‘lamp’ : [ⁿam^mb.riⁿd] ‘old (for people)’

m : w

[mar] ‘you give’ : [war] ‘root’

The alveolar nasal /n/ occurs in initial [nir] ‘and’, medial [nene] ‘to’, and final position [^mbin] ‘land’. /n/ contrasts in place with /m/, see above. Examples of contrast between /d, r/ are provided below. No clear examples are known to show contrast of [n] and [j].

n : r

[nut] ‘pierce/stab’ : [rut] ‘sound’
 [^mbin] ‘land’ : [^mbir] ‘canoe’
 [mʔ.gan] ‘hollow’ : [mʔ.gar] ‘dig’

2.1.1.4 Rhotic

The Irarutu consonant /r/ contrasts with the phonemic glide /w/. Speakers of Irarutu report that certain dialects have a lateral /l/ instead of /r/; and, Matsumura and Matsumura (1991) report that in Northeast Arguni dialect, /l/ is an allophone of /r/ in certain contexts.

The segment occurs in initial [ru] ‘two’, medial [irary] ‘language’, and final [^mbir] ‘canoe’ position. It contrasts in manner with /d, n/. Examples of contrast between /r/ and /d/ are provided in the description of /d/, and examples of /r/ versus /n/ are given in the description of /n/. There is one example where [r] appears to contrast with [j] (‘male headdress’ versus ‘month’), but this is not exactly an appropriate comparison.

r : j
[sre^m.bro] ‘male headdress’ : [sje^m.ba] ‘month’

2.1.1.5 Glides

One glide partakes in the Iraputu phoneme inventory, /w/ [w]. The palatal [j] can be considered an allophone of /i/ conditioned by obligatory desyllabification (/i/ → [j]) because all examples of [j] can be syllabified as [i] in careful speech. Neither [w] nor [j] serves as the nucleus of a syllable, nor does either play a role in consonant clusters. In a previous analysis, Voorhoeve (1989) chose only to write glides in word-initial position before a vowel, but used non-syllabic vowels elsewhere.

In careful speech, native speakers can syllabify glides, [‘ni.a ~ nja] ‘what’ and [‘sⁿgwif.r ~ ‘sⁿgu.i.fɾ] ‘move’, [f^hwer ~ f^hf.u.er] ‘sugar ant’, which suggests the phonemic form of these words are /nia/, /sguifr/, and ‘ffuer’. However, in forms such as [wi.tu] ‘forest/jungle’, [wo] ‘line’, [wer] ‘water’ and [war] ‘root’, native speakers do not syllabify word-initial /w/, *[u.i.tu], *[u.o], *[u.er], and *[u.ar]. There are no examples that show syllabified /i/ in word-initial position and no known examples combine [w] and [u], or [j] and [i]. These facts suggest that, in addition to /u/, which has a desyllabified allophone in some environments, there is also an underlying glide /w/ (or that some of the relevant examples contain diphthongs): /witu/ (or /ɰitu/), /wo/ (/ɰo/), /wer/ (/ɰer/), and /war/ (/ɰar/). On the other hand, desyllabification of /i/ → [j] is obligatory in certain environments, but there is no /j/ (or that its allophones are no longer distinctive). This is supported from a historical perspective. Reflexes of PCEMP *y- ([j-]) merged with the Fruata dialect segment /j/ [j] (see Chapter 4 for further details). The closely related linguistic variety Kuri lacks initial [j], and reflects [j] instead. Medially, [j] is observed in loanwords, but alternates with nativized [j], as in IND sembahyang ‘prayer/church service’, which is borrowed as [‘s^mba.jan], but occasionally alternates with [‘s^mba.jan] (which native speakers comment sounds ‘Indonesian’).

The phoneme /w/ is attested in word initial position [wifu] ‘knee’, and medial position [‘sa.wat] ‘snake’, but not in final position. In the environment of occurring between a consonant and a vowel, at least one example shows that [w] contrasts in place of articulation with [j]. Free variation of [w ~ j] in the counter example [fwir ~ fjir] can be attributed to desyllabified /y/ in /fyir/ ‘navel’, compare careful speech [‘fy.ir].

w : j
[fi pwar] ‘slap/knock’ : [fi pjar] ‘nourish’
[f.fwer] ‘sugar ants’ : [f.fjer] ‘say (active)’
but, *w ~ j*
[fwir ~ fjir] ‘navel’

2.1.2 Vowels

There are two types of vowels, monophthongs (V), and diphthongs (VV or VY). Following Donegan (1986), vowels are described according to tongue height, coloring (palatality and labiality), and a tense versus lax distinction. A vowel without color (palatality or labiality) is said to be achromatic, as with /a/ [ə ~ ɑ], and non-phonemic schwa [ə]. Labiality is predictable for two back vowels in Irarutu, /u/ and /o/, but one labial vowel in Irarutu is also palatal /y/ [Y], that is, not predictably back. Importantly, it appears that three vowels play a role in Irarutu derived diphthongs: /i/, /u/, and /y/, but the role of /y/ in this capacity is marginal.

2.1.2.1 Monophthongs

There are seven monophthongs in Irarutu /a, e, ɪ, i, o, u, y/, plus phonetic schwa [ə], which does not contrast with any phonemic vowels. Schwa occurs frequently, due its role in [ə] epenthesis.

The best way to show the phonemic status of the vowels is with three sets of words. The first set of words illustrates the vowels /a, e, i, o, u/ are single segment words (V).

- a* [a] ‘I (in relative clauses)’
- e* [e] ‘you all’; ‘tree’; ‘yes’
- i* [i] ‘he/she/it’
- o* [o] ‘you’
- u* [u] ‘rattan’ (Matsumura 1991)

The second set establishes another contrast the remaining contrast missing from the first set of words.establishes six of the seven contrasts /a, ɪ, i, o, u, y/:

- a* [m^ɰgar] ‘dig’
- i* [m^ɰgir] ‘wring’
- ɪ* [m^ɰgɪr] ‘k.o. low lying hills similar to a plateau’
- y* [m^ɰgyr] ‘buy’
- u* [m^ɰgur] ‘lazy’
- o* [m^ɰgor] ‘bite’

The third set shows the palatal vowels /i, ɪ, e/ contrast

- i* [ti] ‘penis’, [fi] ‘vagina’
- ɪ* [tɪ] ‘no, not’, [fɪ] ‘verbalizer/nominalizer’
- e* [te] ‘tea’, [fe] ‘thing’

The low achromatic lax vowel /a/ represents Donegan’s (1986) vowel [ɑ]. It has the highest sonority and lowest tongue height of all the Irarutu vowels. /a/ is the most common, but not the only possible, nucleus in derived diphthongs. As a lax vowel, it contrasts with /ɪ/ (high palatal) and /y/ (high labiopatal). In addition to the two sets of words cited above, two pairs of words help to show how /a/ is distinct from mid vowels in the chart presented in **TABLE 2.2**.

a : e

[mtmat] ‘death’ : [mtmet] ‘a little bit’

a : o

[ma] ‘come’ : [mo] ‘child’

The mid palatal (tense) vowel /e/ mostly surfaces as [e], but can be realized as any member from a wide range of values, [ɛ ~ e ~ ɪ ~ i], which overlaps with the phonetic values of /ɪ/ and /i/. /e/ does contrast with mid /o/. The best, or most functional, example of this contrast is [e] ‘you (plural)’ versus [o] ‘you (singular)’.

The mid labial (tense) vowel /o/ ranges between [o ~ ɔ]. In contrast with the corresponding palatal vowel, it does not overlap with the phonetic realizations of the high labial vowel /u/. A set of three words that have not been mentioned above nicely shows that the labial vowels contrast with one another.

o [no] ‘rise’

u [nu] ‘island’

y [ny] ‘name’

All else being equal, the four high vowels /i, ɪ, u, y/ are the hardest to distinguish from one another.

The high palatal tense vowel /i/ often functions as the non-syllabic on-glide or off-glide [j] in Irarutu derived diphthongs (see §2.1.2.2). /i/ contrasts in tenseness with /ɪ/, and in tongue height with /e/. Several widely cited examples in the literature (Voorhoeve 1989, Matsumura 1991) show /i/ is distinct from /ɪ/. These are presented below, followed by several sets of words that show /i/ contrasts with both /ɪ/ and /e/.

i : ɪ

[gin] ‘I drink/you drink’ : [ᵑgm] ‘I sleep/you sleep’

[nin] ‘he/she/it drinks’ : [nɪn] ‘he/she/it sleeps’

[rir] ‘industrious’ : [rɪr] ‘strong’ (Voorhoeve 1989)

i : ɪ : e

[ti] ‘penis’ : [tɪ] ‘negative marker’ : [te] ‘disjunctive’; ‘tea’; ‘thing’

[fi] ‘vagina’ : [fɪ] ‘relativizer’ : [fe] ‘thing’

[si] ‘see’ : [sɪ] ‘shoe!’ : [se] ‘then’

The high palatal lax vowel /ɪ/ is somewhat less frequent in lexical items in careful speech in relation to [i] and [e], but, as already noted, these two segments, several very frequent lexical items allow variation across the [-low +pal] vowels, such that [gin] ‘drink’ often surfaces in casual speech as [gm] (homophonous with [gm] ‘sleep’), [tɪ~ti] ‘negative’, [fɪ~fi] ‘relativizer’, [fe~fi] ‘thing’, and often but not always surface as [ɪ], causing the lax high palatal vowel to seem more common in casual speech. Furthermore, its articulatory range overlaps with the articulatory range of /e/, as in free alternation of the first vowel in the word /seba/ [se^mba ~ sɪ^mba] ‘month’. Reliable examples of contrast are given in the first set of words at the beginning of this section, and in the description of /i/, above.

The high labial vowel /u/ ranges over [u ~ ʊ] and is about as frequent as the labiopalatal vowel /y/. It contrasts in tongue height with /o/, and in coloring with /i/ and /ɪ/ and /e/. /u/ contrasts with /y/ in both laxness and coloring.

u : *o*

[uf] ‘one of, which, who’ : [of] ‘prohibitive (don’t)’

u : *i*

[rur] ‘bone’ : [rir ~ rir] ‘seed’

u : *ɪ*

[ftut] ‘form’ : [mtit] ‘fall’

u : *y*

[tur] (~[tor]) ‘three’ : [tyr] ‘star’

[fu] ‘fruit’ : [fy] ‘flower’

[ru] ‘two’ : [ry] ‘voice’, ‘origin/source’

The high labiopalatal lax vowel /y/ is unusual cross-linguistically, especially so in a seven vowel system, such as Irarutu’s. In careful speech /y/ is central to palatal, compared with (back) /u/, and strongly labialized through labial protrusion. Several of the 2013 consultants impressionistically compared it to German *ü* [y], asserting that it is very far forward, and demonstrated by lip protrusion (for example, they were unsatisfied by the quality of my approximations when I articulated the vowel with the tongue in a back position). In casual speech, /y/ ranges over [y ~ ʏ ~ ʊ], but can also merge with [u]. It occurs with equal or greater frequency than /u/. In fact, the name of the language ‘Irarutu’, actually derives from /iraryty/ [ir³ry³ty], with two labiopalatal vowels /ry/ ‘voice’, and /ty/ ‘true’, but no /u/’s. The spelling of the name of the language will not be altered. /y/ contrasts with palatal /ɪ/, and the high vowels /i/ and /u/.

y : *ɪ*

[fy] ‘flower’ : [fi ~ fi] ‘female genitalia’

y : *i*

[jy] ‘canoe’ : [ji] ‘explicative’

y : *u*

[myn] ‘mud’ : [mun] ‘k.o. sound (Voorhoeve 1989)’

2.1.2.2 Diphthongs

The Fruata dialect of Irarutu does not have unit diphthongs, but the Papuan Malay discourse marker [jo ~ ʔjo.i] ‘hey’ is frequently used by Irarutu speakers, and there are signs that at least one dialect of Irarutu (Central Arguni dialect) has phonemic /ɛ̯e/. The evidence that supports the claim that diphthongization is present within Irarutu, as a group of linguistic varieties, is based on sonority, the sonority sequencing principle, syllable count, and stress assignment. These factors distinguish true diphthongs from three alternative realizations of two adjacent vowels: phonetic diphthongs (as in dialect variation), derived diphthongs (a result of speed of speech in Irarutu), and adjacent vowels in separate syllables (cf. Reh 2012).

Sonority provides grounds for analyzing diphthongs and differentiating them from glide and vowel sequences. Sonority can be described according to the sonority hierarchy (Donegan 1986, Zec 1994). Therefore, relevant aspects are presented here. Diphthongs either decrease in sonority and accentability across subconstituents, or increase in sonority and accentability. Everything else being equal, lower vowels are more sonorous than higher vowels, i.e., sonority is inverse to vowel height. Furthermore, lax vowels are more sonorous than tense vowels. According to Donegan (1986), as sonority increases, color (labiality or palatality) decreases, and vice versa. This factor plays a peripheral role in the description of Irarutu diphthongs, however, it is assumed that when two colors are combined, sonority decreases accordingly. This impacts /y/, which is phonetically lax but appears to rank closer to the high tense vowels /i/ and /u/, than to the high lax vowel /ɪ/, especially in the derivation of a word like /fyir/ → [fwir ~ fjir] ‘navel’, which loses either its palatality or its labiality, respectively, when spoken. The use of the sonority sequencing principle (Selkirk 1984) for the purpose of analyzing diphthongs can be construed as: the segment that has the highest sonority will be the syllable nucleus and can take stress. This depends on an explicit ranking of a language’s monophthong vowels, see **FIGURE 2.1**.

Least sonorous	glide:	/w/	[w]	<
	high tense vowels:	/i, u/	[i, u]	<
	high lax vowel:	/ɪ, y/	[ɪ, ʏ]	<
	mid vowels:	/e, o/	[e, o]	<
Most sonorous	low vowel:	/a/	[a]	

FIGURE 2.1 – Sonority hierarchy for Irarutu vowels, presented in order of increasing sonority ranking (decreasing vowel height) to parallel the presentation of the phoneme inventory. The symbol ‘<’ can be read as ‘has less sonority than’.

Cross-linguistically, diphthongs are commonly known to arise from vowel substitutions in different dialects. This is true in Irarutu. For example, /seba/ ‘moon/month’ is realized in the various dialects as [se^mba] (Fruata), [sie^mba ~ ʃje^mba] (Nagura and Kuri), or [sɪba] (Kaimana/Central and West Arguni). Although a comprehensive treatment of Irarutu dialectology is not possible at this point in time. A better understanding of the interrelationships of the various dialects would help improve the description of diphthongs in Irarutu.

On a related point, it is undeniable that Irarutu has ‘phonetic diphthongs’, which are essentially phonetic byproducts of articulating certain vowel sequences, such as /nia/ [ni.a ~ ni.^ɪa] (careful speech tempo) ‘what’, in which the phonetic glide [ɪ] functions to facilitate the transition between the adjacent syllable nucleus /i/ and /a/. This example of a phonetic diphthong also illustrates that Irarutu has sequences of adjacent vowels, that is, [i.a], which are quite prevalent in the language. Depending on speech tempo, and the location of stress when two vowels are adjacent, and if one of the vowels has very low sonority, i.e., /i, u/ and possibly /y/, the sequence may be susceptible to derived diphthongization (see §2.3.2 for further discussion), or the unstressed vowel may be deleted (see §2.4).

2.2 Irarutu Feature Set

The description of Irarutu phonemes relies on the following set of features, taken primarily from Lass (1984): syllabic, approximant, consonantal, sonorant, continuant, nasal, tense, high, low, labial, palatal, coronal, dorsal, and voiced.

Seven features pertain to manner, (syllabic, approximant, consonantal, sonorant, continuant, nasal, and tense). Six concern place (high, low, labial, palatal, coronal, and dorsal). One feature is a laryngeal feature (voicing). Each feature that is discussed below has phonetic (articulatory and acoustic) correlates as described in the literature (cf. Donegan 1986, Carr 1993, Hayes 2009).

2.2.1 Manner Features

There are seven manner features: syllabic, approximant, consonantal, sonorant, continuant, nasal, and tense.

Syllabic [+syllabic] segments are the vowels /y, i, ɪ, e, a, o, u/.

There are two **approximants** in Irarutu, /r/, and /w/. /r/, a non-lateral, is usually articulated as a rhotic /ɹ/ or tap [ɾ], but in some contexts, it can become a trill.

Approximants are the most sonorous type of consonant. [+approximant] segments are /w, r/. [-approx] consonants are nasals, fricatives and stops: /m, n, f, s, h, b, d, ʒ, g, p, t, c, k/. The vowels /y, i, ɪ, e, a, o, u/ are also [-approx].

The **consonant** feature [±consonantal] serves as a cut-off point between [+cons] /r, m, n, f, s, h, b, d, ʒ, g, p, t, c, k/ and the other segments. The glide /w/ is [-syll, -cons].

Another subcategory of segments in Irarutu are **sonorants** /m, n, w, r, y, i, ɪ, e, a, o, u/. The obstruents /f, s, h, b, d, ʒ, g, p, t, c, k/ are [-sonorant]. Obstruents ([-sonorant]) are preferentially voiceless.

Continuant functions to distinguish vowels /y, i, ɪ, e, a, o, u/, approximants /w, r/, and /f, s, h/ from the stops /b, d, ʒ, g, p, t, c, k/.

Nasal segments are /m, n/.

Tense differentiates the vowels /i, e, o, u/ from three ‘lax’ (i.e., [-tense]) vowels /ɪ, y, a/. All else being equal, the lax vowels have relatively weak color (labiality or palatality) for their degree of tongue height. Lax vowels may also have no color [-labial, -palatal], as is the case for /a/. This feature captures the contrast between /i/ versus /ɪ/ and, in some sense, /u/ versus /y/ (Chomsky and Halle 1968; Donegan 1993).

2.2.2 Place features

The six place features are high, low, labial, palatal, coronal, and dorsal.

The Irarutu **high** vowels include /i, ɪ, u, y/, whereas the feature **low** defines the vowel /a/. The features [-high] [-low] characterize the mid vowels /e/ and /o/.

Labial segments are articulated in Irarutu with strong lip protrusion in segments where the tongue is usually back from neutral position, as for the vowels /o, u/, but labiality is also found in the non-back vowel /y/. Labial also applies to consonants that are articulated with the lips /p, b, f, m, w/.

Palatal segments are articulated with the tongue forward compared to the neutral position. [+palatal] vowels are /i, ɪ, e, y/. [+palatal] combines with [+labial] for the vowel /y/. [+palatal] consonants are /j, ɟ/. [-palatal] consonants are /p, b, t, d, k, g, f, s, h, r, m, n/. The feature palatal affects consonant-vowel interaction in Irarutu. For example, the consonant /t/ often becomes palatal [c] preceding a high palatal vowel, shown by /mtie/ → [mtie ~ mce] ‘eye’, /mumtie/ → [mumce] ‘morning’, and /utie/ → [uce] ‘papeda (k.o. starch)’, but /mtyn/, with a labio-palatal, also becomes [mcyn] ‘cook’/‘shoot’.

Coronal consonants /t, d, s, r, n/ are articulated with the blade or the tip of the tongue and contrast with **Dorsal** [+dorsal] phonemes /k, g/, which are articulated with the body of the tongue. Dorsal is also a property for the predictably ‘back’ [+labial] vowels /o, u/. The place of one segment, /h/, is defined by all negative place features: [-labial, -coronal, -palatal, -dorsal].

2.2.3 Laryngeal feature

In addition to manner and place features, a laryngeal feature, **voiced**, is necessary. [+voiced] segments are /y, i, ɪ, e, a, o, u, w, r, m, n, b, d, ɟ, g/. [-voiced] segments are /f, s, h, p, t, c, k/. Vowels and sonorants are inherently voiced in Irarutu. [-voiced] implies ‘spread glottis’, which is observed in the segment /h/.

2.3 Allophony

Subphonemic features help to describe allophonic variation in Irarutu consonants (§2.3.1) and vowels (§2.3.2). Speed of speech and speakers’ attention to their own speech play a significant role in the allophony of consonants, but is even more important for vowel allophony. The discussion of vowels is organized according to a basic two-way distinction in speech tempo.

2.3.1 Consonant Allophony

In everyday language use, casual speech is predominant, but careful speech also plays a role to provide emphasis or clarification – especially in elicitation settings. Among the alternations that primarily involve consonants, there are differences in release for voiced and voiceless segments that are determined by environment, especially position in a word. Apical stops are palatalized in certain environments. Two alternations are tied to /r/.

2.3.1.1 Release

Three positions effect the release characteristics of consonants, word final (or more generally, utterance final) position, preceding a vowel, and preceding a consonant in a consonant cluster. Furthermore, voiceless consonants behave somewhat differently than voiced consonants.

In final position, voiceless obstruents can be unreleased (-C[̚]), plain (-C), or aspirated (-C^h), e.g., /fidap/ [fi.[̚]dap[̚] ~ fi.[̚]dap ~ fi.[̚]dap^h] ‘can/able’. Voiceless continuants can be plain or aspirated /nof/ [nof ~ nof^h] ‘wind’. Voiced final sonorant consonants can be plain, but tend to have vocalic release (C^ə) /san/ [san ~ san^ə] ‘house’. Furthermore, in final position, hyperarticulation causes aspiration to become vocalic release [fi.[̚]dap^h ~ fi.[̚]dap^ə] and [nof^h ~ nof^ə]. This range of allophonic variation is especially prominent in careful speech and citation forms, but not in connected speech. The generalization that governs the substitutions is, aspiration and vocalic release are used to emphasize or clarify a final consonant.

In casual speech before a vowel, non-laryngeal voiceless obstruents can be de-aspirated for plain release /pɪntr/ [p^hɪntr ~ pɪntr ~] ‘smart’, /fidap/ [f^hi.[̚]dap ~ fi.[̚]dap] ‘can/able to’, whereas sonorant consonants have plain release /dap/ [n[̚]dap] ‘can’. Again, aspiration is used to emphasize a particular segment and can be exaggerated in careful speech.

Generally in careful speech where a voiceless obstruent precedes another consonant, except for a voiced [-cont], the voiceless consonant has plain release, or can be aspirated, /tʃfrie/ [t^hf^hfrie ~ t^hf^hf^hfrie] ‘dragonfly’, /fsfusr/ [f^hsfusr ~ f^hs^hfusr] ‘a lot’; however, where a voiceless [-cont] precedes another voiceless [-cont], the first one can be unreleased /mtiet/ [mt[̚]cjet] ‘fresh’, which increases the likelihood that a following a voiceless [-cont] will also be unreleased [mt[̚]cjet[̚]]. In sequences of multiple similar segments, aspiration tends to skip at least one segment. Aspiration is more probable for initial clusters than for medial or final clusters. A form like [f^hsfus^hr] is extremely rare.

Syllabification of stranded consonants is an additional factor that determines the release feature of a voiceless segment [tʃfrie], [fsfusr] (see §2.4 for further discussion). In casual speech, voiceless consonants preceding a voiced obstruent generally have plain release [f[̚]gf[̚]gaf], but can be aspirated [f^{h̚}gf^{h̚}gaf], or given vocalic release [f^{ə̚}gf^{ə̚}gaf], in cases where emphasis is crucial, such as to signify a morpheme boundary, e.g., between /fi-/ ‘active verb marker’ and /gfgaf/ ‘scratch’, which in careful speech is [fi[̚]gf[̚]gaf] ‘(active) scratch’. This is a second context where aspirated release can vary with vocalic release (C^h→C^ə), similar to the behavior of -C. An additional complexity of release before a voiced stop is that vocalic release can merge with the nasal onset of the stop, resulting in a [f^{ə̚}gf^{ə̚}gaf].

In careful speech, voiced consonants that precede another consonant in a consonant cluster have vocalic release (C^ə), e.g., /gfgaf/ [g^{ə̚}f[̚]gaf] ‘write’, but, paralleling the release features of voiceless obstruents in consonant clusters, the release feature of a voiced consonant is also governed by syllabification of high ranking candidates as well as syllabification constraints.

2.3.1.2 Place assimilation: Palatalization

In casual speech, /t/, which is [+coronal, -palatal], can become [+palatal] before a high palatal vowel, especially /i/ and /y/, but it can retain its underlying position in careful speech. For example, /mtiet/ [me.tjet ~ met.cjet ~ me.cjet] ‘fresh’, /mumtie/ [mum.tie ~ mum.ce] ‘morning’. Furthermore, palatal vowels tend to delete following a palatal, or

palatalized, consonant [mtie ~ mce] ‘eye’, via hypothesized intermediate stages of derived diphthongization [mteje] and CC reduction [mcje]. This occurs regularly in compounds such as /mtie+wer/ [mcewer] ‘goggle-eyed’. Notably, the high labio-palatal vowel [y] also induces the same substitution, /mtyn/ [mtyn ~ mcyn] ‘cook’, supporting the conclusion that /y/ is indeed palatal, and that its palatality causes a [+cor] to become [+pal]. In contrast, the [-high +pal] vowel in /te/ ‘incidental stative marker’ does not become *[ce], nor does the [-tense +high +pal] vowel /ɪ/ in /mtɪt/ ‘fall’ become *[mɛɪt].

Voiceless coronal continuants can be optionally palatalized before a desyllabified high palatal vowel: /siem/ [sjem ~ jjem] ‘rain’, /sieri/ [‘sje.ri ~ ‘jje.ri] ‘crocodile’. Palatalization generally does not occur before a monophthong /sɪbua/ [‘sɪ^m.bua] *[jɪ^m.bua] ‘chili’, /ssi/ [s^hsi] ‘fast’ *[s^hʃi], and /sy/ [sy ~ sY] ‘tail (classifier)’ *[jy], *[jY].

2.3.1.3 Trilling and rhotic deletion

Sequences of two [+cor +approx] segments optionally merge as a trill [r], represented here as [r̥]. /r/ is generally the first element in clusters only with another rhotic, as in /rre/ [r̥³re] (careful speech) vs. [r̥.e] (casual speech) ‘day’, but can follow a range of consonants, such as /krru/ [kr̥³ru ~ kr̥.³u] ‘thunder’, /frro/ [fr̥³ro ~ fr̥.³o] ‘run’, /mrro/ [mr̥³ro ~ mr̥.³o] ‘long’.

Trilling is also supported by dialect variation. Matsumura (1991) reports that the East Arguni dialect has changed /rr/ to [l] in certain contexts; but notably, the same environments sometimes do not yield [l]. Matsumura does not cite [i.lu.tu], only [i.ra.ru.tu] ‘Irarutu’; however, Anceaux (1961) does cite the form ‘*irutu*’. Native speakers who were consulted for the present description only supported the form [ir̥³.ry.³ty], but [i.r̥.y.³ty] is supported by the trill allophony in the Fruata dialect. Trilling does not appear to arise from morpheme combination.

However, /r/ is always deleted before /m/ in the forms /m-ar-ma/ ‘you-take-come’ [ma.³ma], *[mar.³ma], and /m-ar-mri/ ‘you-take-go.into’ [ma.³mri].

2.3.1.4 Syllabification

Syllabification in [+consonantal] segments is a product of the process of syllabication, discussed in §2.4.1.4, and is closely tied to speed of speech (see also Donegan and Stampe 1978). In contrast with careful speech, where consonants in clusters are somewhat tidily delineated by release features, the majority of casual (i.e., rapid tempo) language data indicates that several types of consonants, especially word-initial sonorants [r̥] [m̥] [n̥], but also continuants and voiced stops, behave differently from instances of these segments in other environments. For example, according to the available phonetic data, when the phoneme /r/ is in a consonant cluster, release of a preceding segment is emphasized in careful speech, as in /tru/ [t̥^hru ~ t̥³ru] ‘egg’ and /frifr/ [f̥^hrifr] ‘wide’.

From a perceptual point of view, vocalic release is more salient in careful over casual speech, and after voiceless segments – due to transition to voicing – than after voiced segments. However, in clusters of three or more consonants where /r/ is a non-initial segment, the vocalic release that precedes /r/ contributes to /r/ becoming [+syllabic]. The

word /br̥bar/ ‘lungs’ appears as [ᵐb̥r̥.ᵐbar], with the rhotic gaining the feature [+syllabic] in the consonant cluster to emphasize it in careful speech, but [ᵐbr̥.ᵐbar] in casual speech.

2.3.2 Vowel Allophony

Allophony in vowels is considerably more dependent on speech tempo than the consonants. Therefore, the description of careful speech phenomena (§2.3.2.1) and casual speech phenomena (§2.3.2.2) can more easily be separated.

2.3.2.1 Careful speech vowel allophony

In normal careful speech, and to some extent, the exaggerated careful speech of elicitation settings, the release characteristics of consonants have varying degrees of influence on syllable structure, resulting in two varieties of epenthesis: insertion of schwa [ə], and insertion of [ɪ].

2.3.2.1.1 Schwa epenthesis

As mentioned in the description of consonant release features above, to emphasize a consonant in Irarutu, whether it is voiced or voiceless, slow speech tempo enhances vocalic release, which can derive from aspiration in voiceless obstruents, attains syllabic status ($C^{\circ} \rightarrow C\text{ə}$), as in /tru/ ‘egg’ [$t^h\text{ru} \sim t^{\circ}\text{ru}$] (casual speech) versus [$t\text{ə}.\text{ru}$] (careful speech) and therefore under the right conditions it can appear to function as the nucleus of its own syllable. The strength of word final vocalic release is salient enough to speakers of Irarutu to cause /nof/ ‘wind’, which is generally realized as [nof] (dialectally [nof]), to be pronounced [$\text{'no}^{\text{f}} \sim \text{'no.fə}$] (dialectally ['no.fə]), and is transcribed as *nufa* in the title and lyrics of the popular song *Mumri Nufa* ‘dry season wind’ (but the representation of Irarutu words using Indonesian spelling conventions, often by non-Irarutus, employs inserting the symbols ‘a’ and ‘e’, where these would not be included by an Irarutu speaker). On a phonological level, schwa [ə] is non-contrastive, but it is used in certain speech acts, such as reciting varbal art, to indicate a break between lines of a poem or composition.

Matsumura (1991) and Voorhoeve (1989) both comment that there are two general environments where a phonetic schwa is inserted. One is after final consonants. However, in the present description, appearance of word-final schwa is attributed to speakers' heightened awareness of final segments and slower speech rate in careful speech. Furthermore, insertion of a final schwa is most likely to occur at the end of a phrase in casual speech, rather than after each word that has a coda. One form in particular, /sus/ 'breast', is cited with epenthetic schwa [sus^ə] or an 'echo vowel' [sus^u] (Matsumura 1991; Voorhoeve 1989). The assertion that there is an 'echo vowel' in Irarutu is possible but seems conjectural, especially considering that the reconstructed form is *susu (see Chapter 4); therefore, the final vowel could equally well be seen as a reduced form of the reflex from the proto-form, or as an echo vowel. An argument against 'echo vowels' is that they are exceedingly scarce in Irarutu.

The other environment where previous researchers noticed schwa is between adjacent consonants except between a voiced obstruent and /r/, or /fr/. For example, in the form

/təmn/ ‘father-in-law’, Voorhoeve (1989) asserts that schwa is inserted between /m/ and /n/ in the coda, and after the final consonant, resulting in trisyllabic form [ˈtə.mə.nə]. However, native speakers pronounce this as a monosyllable [ˈtəmn] in casual speech. Another example is /ssi/ ‘preparation’, which occurs as [s^hsi ~ s^əsi] in casual speech, but [sə.ˈsi] if a speaker wishes to emphasize the initial /s/. In general, schwa epenthesis is a result of over-emphasizing the release feature of a consonant, such as aspiration or vocalic release, in careful speech, and it can occur between voiced obstruents and /r/ or /f/ and /r/.

2.3.2.1.2 [ɪ] epenthesis

Epenthesis of [ɪ] can be justified as a variation of schwa insertion (Ø → [ə]). In comparison with schwa epenthesis, [ɪ]-epenthesis is heavily regulated by context. It only occurs after a [+coronal] consonant, e.g., /t/, in a consonant cluster in careful speech. In other words, it appears to be the case that epenthetic [ə] becomes [ɪ] after a coronal consonant, /tbe/ → [t^{hm}be ~ tə^mbe → tɪ^mbe] ‘hit’. [ɪ] is not inserted in the conditioning environment in casual speech.

2.3.2.2 Casual speech vowel allophony

Five types of vowel allophony are observed in casual speech in Irarutu: vowel reduction, alternation of stressed vowels with [ɪ], an aberrant form of palatalization, glide fortition, and laxing.

2.3.2.2.1 Vowel reduction

In complementary distribution with the careful speech process of schwa epenthesis, there is a common tendency in casual speech for monophthongs in unstressed syllables, most commonly penults, to reduce phonetically to schwa or zero. The weak form of vowel reduction can be represented as V → [ə] (possibly through laxing) and is common when an Irarutu speaker is monitoring his connected speech, but distinct from careful (e.g., clarifying or citation form) speech. The strong form of vowel reduction, V → Ø, is functionally opposite of [ə]-epenthesis, and has an intermediate stage where [ə] reduces to vocalic release [ʔ]. In general it seems that the faster the tempo of speaking, the stronger the form of vowel reduction.

Vowel reduction applies most frequently to disyllables, such as /besu/ [ˈmbe.ˈsu] (careful speech) ‘only’, to derive a monosyllabic form in casual speech. Dialectally, it occurs in the penult of trisyllables, such as the Babo dialect pronunciation of /wagari/ ‘frog’, to derive a disyllable [wa.ˈgri], which is trisyllabic [wa.ˈga.ˈri] in casual speech in the Fruata and other dialects.

/besu/ → [ˈmbə.ˈsu] (weak VREDUC) ~
 [ˈmbʔsʲu] ~ [ˈmbʃju] (strong VREDUC) ‘only’
 /wagari/ → [wa.ˈgə.ˈri] (weak VREDUC) ~
 [wa.ˈgʔri ~ wa.ˈgri] (strong VREDUC) ‘frog’

The multi-faceted speed of speech dimension of vowel reduction is supported by the structure of phonological words (see §2.4, below for more discussion). For example, /te

o m-tit/ ‘you will fall’ is realized as [tom.'tit] in regular-speed casual speech (also indicated by resyllabication of [m] and nasalization, described in §2.4, below), but was repeated as [tə.o.'mtit] when another speaker was asked to repeat what the previous speaker said, but was defined, word-by-word, as [te.o.'mtit] by both speakers. This phenomenon suggests that vowel reduction applies to a larger domain than individual words, because /te/ ‘incidental stative’ and /o/ ‘you (free pronoun)’ contracts according to the weak form of vowel reduction to [tə.o], but undergoes the strong form when it is realized as [to], paralleling vowel reduction at the word level, as described for /besu/.

2.3.2.2.2 Alternation of stressed vowels with [ɪ]

There is also another form of vowel reduction, that allows stressed vowel phonemes to unexplainably alternate with [ɪ]. For example, in careful speech, /warada/ ‘k.o. knife’ usually surfaces as [wa'raⁿda], but in casual speech, the same word optionally appears as [wa'ɾɪⁿda] (*[wa.rə.ⁿda] has not been observed).

2.3.2.2.3 Aberrant palatalization

Palatalization is generally a process that affects [+coronal] consonants that occur before a palatal vowel. Palatalization also enhances the [+palatal] color of a vowel that follows a [+coronal] consonant, such that [-palatal] /s/ becomes [+palatal], [s^j ~ ʃ], preceding a palatal vowel, e.g., /y/.

/besy/ → [besy] (careful speech)
[^mb^əs^jy] ~ [^mb^jy] (casual speech) ‘only/just’

2.3.2.2.4 Palatal glide fortition

Palatalization is also supported by a process of glide fortition that is observed in dialect variation, adaptation of loan words into the Fruata dialect of Irarutu, and the historical phonology of the language. As mentioned in the descriptions of the palatal stop and semi-vowels in §2.1, above, the Nagura dialect, and Kuri, have forms such as [ja] ‘I’, with a semi-vowel onset, whereas other dialects, including the Fruata dialect, have [ja] a [+consonantal] segment in the onset, but without prenasalization.

The Irarutu phoneme [j] is a reflex of historical /i/, phonetic [j]. In the Fruata dialect, synchronic variation is observed in loanwords that contain [j], e.g., IND ‘*sembahyang*’ [səm.'bah.jaŋ] ‘church service’ is realized variously as [s^mba.'jaŋ] (which my consultants think sounds very ‘Indonesian’) or [s^mba.'jaŋ] (which sounds more like a nativized lexical item to an Irarutu speaker’s ear). From a diachronic perspective, instances of historical /j/ (represented in the literature as *y) in the onset position are reflected as present-day Irarutu /j/, e.g., *maya > IRH [m.'ma.je] ‘tongue’ (see Ch. 4 for additional details).

Each of these forms of evidence support ‘glide fortition’ as an allophonic rule in Irarutu whereby desyllabified /i/ ([j]) in onset position becomes [j], through increased constriction in the vocal tract. In other environments, [j] is allowed. Glide fortition is a fairly reliable process that indicates native Irarutu words. For instance, if [i-] does not de-syllabify or undergo glide fortition, it is likely to be a loan form. This is the case for the discourse marker [i.'o] ‘hey/yo’, which is an areal feature of Papuan Malay used by

speakers throughout the area and does not undergo substitution of [ɟ] for [j], or expected [j] for [i], because it is not an Irarutu form.

2.3.2.2.5 Mid-vowel laxing

A considerable number of words in Irarutu contain the sequence of /ie/, such as /fier/ ‘say’, /sieri/ ‘crocodile’, /siem/ ‘rain’, and so forth. In a several of these lexical items, the mid-vowel /e/ becomes lax [ɛ]: [fjɛr], [ʼsjɛ.ri], [sjɛm]. The mid vowel /o/ also appears to become lax [ɔ] following desyllabified /u/ [w], e.g., /wo/ [wo ~ wɔ] ‘line’, /uoffi/ [wo.f^h.fi ~ wɔ.f^h.fi] ‘glow’, but with lexical exceptions such as /niauo/ [ʼnja.wo] * [ʼnja.wɔ] ‘cat’. Following desyllabified /i/ [j], in /sioru/ [sjo.ru] ‘kind of knife’, /o/ remains [o]. On the other hand, /nof/ ‘wind’ does alternate dialectally and individually, among same dialect speakers, between [nof] and [nɔf].

2.4 Suprasegmental phenomena: the syllable and stress

This section describes a range of suprasegmental phonological phenomena. §2.4.1 presents data on phonotactics and the process of syllabification, and it distinguishes two general syllable types. §2.4.2 describes the structure of Irarutu words, stress and accent, resyllabification, derived diphthongs, and nasalization.

On a metalinguistic note, Irarutu speakers feel that their language has a particularly consonantal character. One speaker, who was talking about the sound of various vocabulary items, remarked that there are lots of consonants and not many vowels. Although there are seven monophthongs, compared with neighboring Austronesian languages’ five vowel system, speakers do not consider vowels to be a prominent feature of the language. The allophonic process of vowel reduction, introduced above, is one process that deletes unstressed vowels and helps contribute to this native-speaker perspective. It is shown below that Irarutu also has complex phonotactics, similar to those discussed in Hajek and Bowden (2002). Consonant clusters can occur at the beginning [nfa(n)] ‘walk’, middle [nir^hge] ‘because’, and end of words ‘because’ [fukr] ‘net’, or at multiple sites simultaneously [fsfusr] ‘a lot’. Therefore, despite the relatively modest size of the phoneme inventory and ratio of consonants to vowels (16 Cs : 7 Vs), several phonological processes, phonotactic constraints, and prosodic structure cause spoken Irarutu to give the impression that the language does have a markedly consonantal, e.g. percussive, gurgling, popping, hissing, and crackling character.

2.4.1 Phonotactics, syllabification, nasalization, and syllable types

Irarutu syllable structure is described in the context of the sonority hierarchy (**Figure 2.2**) and the sonority sequencing principle. This section presents phonotactic evidence from attested consonant cluster onsets, medial consonant clusters, and consonant cluster codas (§2.4.1.1), as well as syllabification (§2.4.1.2), in an effort to formulate a template for syllable structure (§2.4.1.3), by proposing that a distinction between major (i.e., able to bear stress) and minor (i.e., not able to bear stress) syllables is strongly correlated with the feature [±syll], not just syllable count.

Least sonorous	1) voiceless stops	/p, t, c, k/	<	
	2) voiceless fricatives	/f, s, h/	<	
	3) voiced stops	/b, d, ʒ, g/	<	
	4) nasals	/m, n/	<	
	5) rhotic	/r/	<	
	6) glide	/w/	<	cannot be stressed
	7) high vowels :	/y, ɪ, i, u /	<	can be stressed
Most sonorous	8) mid vowels:	/e, o/	<	
	9) low vowel:	/a/		

Figure 2.2 – Sonority hierarchy for all Irarutu phonemes. The symbol ‘<’ means ‘has less sonority than’. The cut-off for segments that can be the nucleus of stressed syllables occurs between glides and high vowels. The segments with the least sonority are the voiceless stops 1). The segment with the highest degree of sonority is the low vowel 9).

It has been observed, cross-linguistically, that the most sonorous segments are closest to the nucleus of a syllable, whereas the least sonorous segments will be at the edges (cf. Zec 1994). –An additional dimension of the Fruata dialect of Irarutu syllable structure involves prenasalization in the voiced stops, which have an inherently falling sonority profile. In Irarutu phonology, voiced stops have greater sonority than any voiceless obstruent (i.e., fricatives or stops). The complex nature of these segments, which always presents a violation of the SSP onset positionⁱ, introduces a complication that must be resolved during (re)syllabification.

2.4.1.1 Word initial consonant clusters

Two-member consonant clusters are frequently found in onsets in Irarutu words, but a number of examples with three-member consonant cluster onsets (CCCV), and exceedingly rare instances of four-member (CCCCV) and five-member phonemic consonant clusters (CCCCCV) are attested.

A note on representation must be added at this point. The standard method in contemporary phonology for representing consonant clusters, which uses subscripts of increasing numerical values starting at the beginning of a word, such that a two member initial consonant cluster syllable is represented as C₁C₂V, for the description of Irarutu phonotactics has shortcomings. First, on account of the fact that every syllable will have a C₁, it implies that C₁ is the inherent part of the onset for all syllable types. This representation fails to capture the fact that if a consonant is present, the one that is adjacent to the syllable nucleus behaves differently from consonants that are non-adjacent to the nucleus, which have increased likelihood to become [+syll].

The solution presented here, albeit fairly clumsy, is to use subscript descriptors where the traditional methodology is insufficient, such as, ‘C_{adjacent}’ for a consonant that is adjacent to the syllabic segment, ‘C_{proximal}’ for a consonant that occurs close to – but not adjacent to – the syllabic, ‘C_{distal}’ for the segment that is furthest from the syllabic. This reduces the representational innovation as much as possible, while still indicating explicitly which position a consonant occupies in relation to a syllable nucleus as well as to other

consonants in a consonant cluster, regardless of word boundaries and simultaneously allowing the description to be consistent with the existing literature.

There are thirteen consonants that are likely to occur in either position in a two-member consonant cluster. The phoneme /h/ is a relatively new addition to the phoneme inventory and does not occur in any known clusters. The glide [w-] only occurs in word-initial position. A form such as [nrwi.'en] 'he is aware' is therefore analyzed as /nruien/, with only two consonants in onset position.

Forty-four initial consonant clusters that involve two consonants are illustrated in **Figure 2.3**. It stands out that neither [p] nor [ɟ] occurs as either the initial consonant (C₁) or the following member (C₂) of any consonant cluster. Even with only two segment consonant clusters, there are several instances where the sonority profile violates the sonority hierarchy, such as clusters with a voiceless stop as C₂, e.g., [tt, bt, ft, mt, nt]. About half of the two-member initial CC-s follow the sonority sequencing principle, the other half violate it. Specific sequences of consonants are presented in **TABLE 2.4**.

[ru ttur]	'three o'clock'	[nfun]	'he makes'
[^m btun]	'cook/shoot'	[^m bsi]	'dried'
[fta]	'stomach'	[ssi]	'process of building'
[wer ntur]	'low tide'	[msutr]	'sit with two legs out'
[mtie]~[mce]	'eye'	[nsir]	'dirty'
[kkor]	'chicken'	[gmon]	'reflection'
[t ^m bif]	'bag'	[mmat]	'green/unripe'
[s ^m bajan]	'mass/prayer service'	[nmi]	'live'
[n ^m ba ^m br]	'close'	[tni]	'body'
[r ^m bua]	'squirrel'	[^m bnisr]	'mosquito'
[matu ⁿ d ⁿ dan]	'farmer'	[fne]	'pig/pork'
[f ⁿ defa ⁿ dr]	'floor on stilts'	[nneno]	'at where'
[m ⁿ da ⁿ dn]	'you peel'	[tresu]	'six'
[^m b ⁿ guen]	'look'	[krikr]	'smooth'
[ⁿ g ⁿ ga ⁿ gun]	'nutmeg'	[^m bramu]	'arrow'
[f ⁿ gje]	'back'	[ⁿ dri ⁿ dn]	'sick'
[m ⁿ genir]	'bitter'	[fra]	'hand'
[n ⁿ gor]	'it bites'	[sre ^m bro]	'man's headdress'
[r ⁿ guin]	'head'	[mro]	'exit'
[kfi]	'hat'	[nrwien]	'he is aware'
[ffuer]	'k.o. ant'	[rre]	'day'
[sfar]	'more'		

Figure 2.3 – Examples of initial CCs.

	p	t	k	f	s	b	d	g	m	n	r
p											
t		<u>tt</u>		<u>ft</u>		<u>bt</u>			<u>mt</u>	<u>nt</u>	
c									<u>mc</u>		
k			<u>kk</u>								
f		<u>kf</u>		<u>ff</u>	<u>sf</u>					<u>nf</u>	
s					<u>ss</u>	<u>bs</u>			<u>ms</u>	<u>ns</u>	
b		<u>tb</u>			<u>sb</u>					<u>nb</u>	<u>rb</u>
d				<u>fd</u>			<u>dd</u>		<u>md</u>		
g				<u>fg</u>		<u>bg</u>		<u>gg</u>	<u>mg</u>	<u>ng</u>	<u>rg</u>
m								<u>gm</u>	<u>mm</u>	<u>nm</u>	
n		<u>tn</u>		<u>fn</u>		<u>bn</u>				<u>nn</u>	
r		<u>tr</u>	<u>kr</u>	<u>fr</u>	<u>sr</u>	<u>br</u>	<u>dr</u>		<u>mr</u>	<u>nr</u>	<u>rr</u>

TABLE 2.4 – Initial two-member consonant clusters. C_{distal} elements are listed across the top and C_{adjacent} down the left column. Twenty-seven sequences that violate the sonority sequencing principle are underlined. Grey shading indicates sequences that conform to the SSP on a phonological level. Boxes indicate sequences that violate the SSP on a phonetic level due to prenasalization in voiced stops.

FIGURE 2.4, below, summarizes forty-four examples of three- or more initial CCs. The largest number of consonants that occur in an initial CC- is five: [tftfr-]. Three types of complex initial consonant clusters are identified: I) those that follow the sonority sequencing principle, II) those that violate it once, and III) those that violate it twice. Each violation correlates with a prominence in the sonority contour that can cause a segment to become [+syllabic].

Generalizations about the complex initial consonant clusters in **FIGURE 2.4** are somewhat tricky to capture, but can be summarized as in **TABLE 2.5**. There are two main concerns: illustrate what sequences occur and identify patterns in the consonant clusters that violate the sonority sequencing principle. Sequences of consonants with equal sonority are considered violations of the sonority sequencing principle in Irarutu because they are most often realized in a separate timing unit by native speakers [fɾ.'ro]. One violation of that generalization is [nm.'se] ‘until’, where the two nasals are pronounced together as one syllable [nm].

Only a few examples (3/44) of complex, i.e., three-or-more-member initial consonant clusters, (/kdr/¹³⁵, /tgr/¹³⁵, /sgr/²³⁵), obey the sonority sequencing principle (type-I CC-s) on a phonological level, while more than half (24/44) of the examples violate it once (type-II CC-s), and remaining examples (17/44), violate it twice (type-III CC-s). Including a fourth attested sequence, /tfr/, a total of three three-or-more- member consonant clusters in the onset do not violate the sonority sequencing principle on a phonological level, i.e., /kdr/¹³⁵, /tgr/¹³⁵, /sgr/²³⁵ /, but do violate it on a phonetic level due to prenasalization: [tⁿgr-]¹⁴³⁵, [kⁿdr-]¹⁴³⁵, [sⁿgr]²⁴³⁵. The only attested sequence of three consonants in a CC that does not violate the SSP on a phonetic level is /tfr-/ [tfr]¹²⁵. Each of these violations of the SSP at the phonetic level is caused by prenasalization in voiced stops.

Type-I) CC-s with no SSP violation (3)			Type-III) CC-s with two SSP violations		
[k ⁿ dra]	<u>135</u>	‘chair’	(17)		
[t ^ŋ gra]	135	‘ear’	[ṃ.ṛ.'ro]	<u>455</u>	‘long’
[s ^ŋ gritn]	235	‘tingly feel’	[ṃ.ŋgṛ.'far]	<u>43</u> <u>52</u>	‘return’
Type-II) CC-s with One SSP violation			[ṃ.kṛ.'kur]	<u>41</u> <u>51</u>	‘young/green’
(24)			[ṃ.ṃ.'tyn]	<u>441</u>	‘cook’
[^m bṛ.'rar]	<u>355</u>	‘fast’	[ṃ.ŋ.'nis]	<u>444</u>	‘to peel’
[ⁿ dṛ.'fat]	<u>352</u>	‘to fish’	[ṃ.ṭ.'tjet]	<u>411</u>	‘refreshed’
[ⁿ dṛ. ^m bun]	<u>353</u>	‘neck’	[ṃ.ḳ.'ka]	<u>411</u>	‘you bathe’
[^ŋ gṛ.'far]	<u>352</u>	‘return’	[ṃ.ŋḍ. ⁿ dur]	<u>433</u>	‘group’
[tṛ.'ro]	<u>155</u>	‘tall/mountain	[ṃ.tf̣.'taf]	<u>41</u> <u>21</u>	‘to fan’
top’			[ŋ. ^m bṛ.'funi]	<u>43</u> <u>52</u>	‘hide’
[kṛ.'ru]	<u>155</u>	‘thunder’	[ŋ. ^m bṛ. ^ŋ gin]	<u>43</u> <u>53</u>	‘even (just)’
[fṛ.'ro]	<u>255</u>	‘run’	[ŋ.f̣.'tifn]	<u>421</u>	‘love’
[sṛ.'fufn]	<u>252</u>	‘top of head’	[ŋ.f̣.'fe]	<u>422</u>	‘lightning’
[sṛ.'rin]	<u>255</u>	‘drug’	[^m bṛ.ṛ.'fot]	<u>3552</u>	‘heart’
[ṃ.'sru]	<u>425</u>	‘you descend’	[^m ḅ.ŋgṛ. ^ŋ gir]	<u>33</u> <u>53</u>	‘clarify’
[ṃ.'s ^m be ^ŋ gt]	<u>423</u>	‘shoot’	[^m ḅ.ṣ.'si]	<u>4322</u>	‘mud’
[ṃ.'s ^ŋ gre]	<u>4235</u>	‘wait, guard’	[s ^ŋ g̣.f̣.'fo]	<u>2322</u>	‘kiss’
[nṃ.'se]	<u>442</u>	‘until’			
[mṭ.'met]	<u>414</u>	‘little-by-little’			
[mṭ.'not]	<u>414</u>	‘only/just’			
[ŋ. ⁿ dri ⁿ dn]	<u>435</u>	‘smoke’			
[ŋ.'fnano]	<u>424</u>	‘how does it feel’			
[ŋ.'fi sṇ.'nan]	<u>244</u>	‘whisper’			
[ŋ.'frua]	<u>425</u>	‘half’			
[t ^m ḅ.'sufn]	<u>132</u>	‘study later’			
[ⁿ ḍ. ^m bri]	<u>335</u>	‘urinate’			
[^ŋ gf̣. ^ŋ gaf]	<u>323</u>	‘scratch’			
[tf̣.'tfrie]	<u>12125</u>	‘dragonfly’			
[ʃ.'fri]	<u>225</u>	‘above ground’			

FIGURE 2.4 – Examples of three- or more member initial CCs. Data is grouped into three ‘types’ according to number of sonority sequencing principle (SSP) violations. Superscript numbers between the phonetic form and the gloss are used to represent the amount of sonority a segment has, cf. **FIG 2.2**, and therefore helps to show the sonority profile and violations of the sonority sequencing principle, which are underlined here.

Type-I CC-s, i.e., sequences of consonants that do not violate the SSP in the onset, illustrates a trend for [r] (the most sonorous of the consonants) to be adjacent to the nucleus of the syllable. Eighteen different types of initial complex consonant clusters (CC-s with three or more members) have [r] in this position. No instances of [p, c, m, h] are observed in this position in a complex consonant cluster. Some of these gaps may be due to chance.

	Ex.	C _{distal} -	C _{proximal} -	C _{adjacent} -	V/C
Type I-	/tbs-/	[+voiced]	*[+son]	≤ [+approx]	*[-cont -voiced]
Type II-	/tftfr-/	—	—	≤ [-son]	[+cons +syll]
Type III-	/mtft-/	—	—	—	[+syll], *[-voiced]

TABLE 2.5 – Initial three- or more member CCs. The ‘=’ can be read as ‘equal in sonority to’. ‘≤’ means ‘equal or less sonority than’. The template for *Type-I* (CC-s w/o SSP violation), in red face across the various examples, illustrates that the syllable types are build cumulatively and have their own phonotactic constraints. For example, *Type-III* (CC-s with two SSP violations), which can only ever be one segment, only occurs if Type-II (CC-s with one SSP violation), which can be one or two segments, is present if *Type-I*, which can be up to three segments long, is present.

Type-II builds on type-I, and type-III builds on type-II. It is assumed that for every additional consonant present within each of the syllable types (I–III), that the sonority sequencing principle applies for well-formedness, meaning that the consonant that is furthest from the syllabic segment has less sonority than a segment that is closer to the syllabic. For example, [tf.tfrɪe] ‘dragonfly’ contains a type-I CC [tfr-], which is composed of the voiceless obstruent [t], the continuant [f], and the sonorant [r]. Preceding that is a type-II CC, which, in this example, is also composed of the voiceless obstruent [t], and the continuant [f]. Native speakers’ pronunciation of words such as [tf.tfrɪe] ‘dragonfly’ suggest the [+continuant] consonant /f/ in the type-II CC becomes [+syllabic] in order to avoid an SSP violation.

There are two lexical exceptions that violate the generalizations made in TABLE 2.5. Both allow [t] after [m], /mtmet/ [mt.tmet] ‘a little bit’, and /mtnot/ [mt.tnot] ‘only/just’. Pronunciations such as *[m.tmet] and *[m.tnot] are not native-like.

Co-occurrence constraints implied in TABLE 2.6 determine which segments can occur where. For example, even though it could be possible for three [r]s to combine in a complex initial CC, this does not occur, but two /r/s do occur in several examples in type-II syllables, e.g., [frro] ‘run’.

2.4.1.2 Word medial consonant clusters

Medial consonant clusters are less abundant than initial clusters. Hypothetically the widest variety of consonant clusters could occur medially, because codas are adjacent to onsets, and Irarutu has a large number of multi-member complex consonant cluster onsets and, as shown in the next section, several types of consonant cluster codas. **Figure 2.5** illustrates the relatively limited diversity of medial consonant clusters, which are limited to twenty-five different combinations.

-C.C-		[o. ^m b .na.fe.ʝof]	‘taboo’
[ap.pap.ro]	‘butterfly’	[fun.nya]	‘do’
[^m ba ^m b .tun]	‘go prepare’	[sir.ni.ef.ta. ^m br]	‘potato’
[mum.tie]	‘morning’	[rat.rum]	‘hill’
[san.ti]	‘homeless’	[ok.'ro.fa]	‘(family name)’
[ⁿ dar.tu]	‘k.o. bowl for	[sre ^m b .ro]	‘man’s headdress’
stacking’		[^t u ^g g .ro]	‘big bamboo’
[tim. ^m be]	‘all’	[fafr.'war]	‘long skirt’
[nar.ʝa]	‘bring me’	[^m dam.ri]	‘lamp’
[is. ^g gi ⁿ g]	‘house knife’	[sun.'rifr]	‘bitter’
[nir. ^g ge]	‘because’		
[fin.fa]	‘to go to’	-C.Ç.C-	
[fis.sjet]	‘speed’	[fi ⁿ g .f. ⁿ gaf]	‘book’
[san.sun]	‘shirt’	[fi. ⁿ gr. ⁿ gir]	‘tell story/story’
[ʝa ⁿ g .mon]	‘my reflection’	[fr.'ru.f ⁿ .ni]	‘this door’
[ter.mus]	‘thermos’		

Figure 2.5 – Examples of medial consonant clusters. Two and three-consonant clusters occur medially. Three-member consonant clusters surface with the medial consonant belonging to its own syllable.

One example, [fafr.'war] ‘long skirt’, is notable because in native speaker pronunciation, stress falls on the word *war* ‘root’, forcing the consonant cluster [fr] to remain in the penultimate syllable [fafr], rather than being divided across the two syllables, i.e., [-f.r-], or function as onset to the final syllable [fr-]. The second syllable in this word begins with [w].

	p	t	k	f	s	b	d	ʝ	g	m	n	r
p	pp											
t				<u>ft</u>		<u>bt</u>				<u>mt</u>	<u>nt</u>	<u>rt</u>
k												
f											<u>nf</u>	
s					ss						<u>ns</u>	
b										<u>mb</u>		
d												
ʝ												<u>rʝ</u>
g					sg							<u>rg</u>
m									gm			<u>rm</u>
n						bn					nn	<u>rn</u>
r	pr	tr	kr	fr		br			gr	mr	nr	

TABLE 2.6 – Medial consonant clusters -CC-. Codas (-C) are listed across the top. Onsets (C-) are given on the left column. Underline identifies twelve medial clusters in which the coda has greater sonority than the onset. Gray shading indicates fourteen -CC-s where -C is phonologically as sonorous as, or less sonorous than, C-. The box indicates a sequence that violates the SSP on phonetic grounds.

One point of interest is that about half of the data suggests that Irarutu contradicts a cross-linguistic trend (similar to English) to have codas that are more sonorous than the following syllable's onset. Two patterns are observed in the three examples of three-member medial consonant clusters cited in TABLE 2.5. Both involve syllabifying the segment that is non-adjacent to a vowel. One pattern also involves parsing the syllabified consonant into its own syllable and packaging the other consonants into the other syllables, /figfgaf/ [fi^g.f.^gaf] 'book'. Parsings such as *[fi^g.f.^gaf] and *[fi^gf.^gaf] are non-native-like. The other pattern parses the coda of the preceding syllable as the onset of the medial, minor (i.e., headed by a syllabified consonant) syllable, [fi.^ggr.^gir] 'tell story/story', [fr.^gru.fⁿ.ni] 'this door'.

2.4.1.3 Word final consonant clusters

Word final position has the least variety of attested consonant clusters. Only fifteen consonant clusters occur at the end of words, see Figure 2.6 and TABLE 2.7. One known example of a word with a three-member consonant cluster in coda position is found in a loan word [pɪntr] 'smart'.

[fi ^g gt]	<u>31</u>	'prepare garden'	[risr]	<u>25</u>	'satiated'
[^m bruft]	21	'teach'	[msutr]	<u>15</u>	'sit with two legs out'
[jamt]	41	'wound'	[fukr]	<u>15</u>	'net'
[kams]	42	'k.o. grass for skirt'	[n ^m ba ^m br]	<u>35</u>	'close'
[ⁿ dritn]	<u>14</u>	'tend garden'	[f ⁿ d ^a fa ⁿ dr]	<u>35</u>	'floor on stilts'
[rɪ ⁿ dn]	<u>24</u>	'wall'	[ja ^g gr]	<u>35</u>	'rubber tree'
[tifn]	<u>24</u>	'bee'			
[rimn]	<u>44</u>	'heavy'	[pɪntr]	<u>415</u>	'smart'
[mifr]	<u>25</u>	'land'			

Figure 2.6 – Examples of final CCs, organized by increasing sonority in the final segment.

Most final clusters (11 of 16) violate the sonority sequencing principle. If -C_{distal} occurs, nearly half of the time the data suggests it will be [r], 25% of the time it will be [n], although three examples of [-t] and one [-s] are also attested. Word final consonant clusters present a challenge in the description of Irarutu syllable structure because they add a trochaic subsyllabic element (strong-weak) in an otherwise iambic system (weak-strong), compare FIGURES 2.3–4 with FIGURE 2.6. I will consider the amphibrach pattern (weak-strong-weak) in words with complex onsets and codas to be a variant of the iambic canon.

	p	t	k	f	s	b	d	ɟ	g	m	n	r
p												
t				ft					gt	mt		
k												
b												
d												
ɟ												
g												
f												
s										ms		
m												
n		<u>tn</u>		<u>fn</u>			<u>dn</u>			<u>mn</u>		
r		<u>tr</u>	<u>kr</u>	<u>fr</u>	<u>sr</u>	<u>br</u>	<u>dr</u>		<u>gr</u>			

TABLE 2.7 – Final consonant clusters. –C_{adjacent} is listed across the top; –C_{distal} is listed down the left. Clusters that violate the SSP are underlined. Gray shading indicates sequences of consonants that conform to the SSP.

Due to the number of violations of the sonority sequencing principle, the range and variation of consonant cluster phonotactics strongly supports the [+syllabic] feature can apply to consonants. It is applied post-lexical, that is, during (re-)syllabication, suggested by the fact that when a word such as /mtie/ [m̥tje] ‘eye’ is possessed by /a-/ ‘my’, the resulting form is [am.tje] ‘my eye’, never *[a.m̥tje]. In this sense, syllabic is an optional feature for consonants. Furthermore, because additional consonants can be grouped with syllabified consonants, e.g., [fi.ᵐgr.ᵐgir] ‘tell story/story’ the syllable in Irarutu needs special consideration.

2.4.1.4 Syllabication

Syllables are essentially abstract timing units. Syllabication is therefore the process of optimally distributing segments into timing units. It is an operation that assigns each vowel in the phonemic representation of a word to a syllable (σ), and it may vary in careful and casual speech. In general, a consonant that precedes a vowel is associated to the syllable as its onset, C_{adjacent}-. A consonant that follows a vowel, and is not already selected as an onset for a separate syllable, is associated with that vowel as its coda, -C_{adjacent}. An additional consonant that precedes the onset (e.g. C_{distal}-) can join the syllable if it obeys the sonority sequencing principle (SSP), see /mgyr/, below.

This procedure of associating consonants as onsets or codas continues until all segments are associated with syllables. However, if a consonant is stranded due to violating the SSP, the consonant forms a derivationally secondary syllable (°), because it is [+consonantal] and acquires the value [+syllabic] (alternately see O’Grady and Archibald 2011 on ‘extrasyllabic’ segments). Additional consonants are assigned as onsets to secondary syllables, provided they follow the SSP, or they can become [+syllabic] if they violate it. An example of how a word with an SSP violation is syllabified is given in **FIGURE 2.7**, below.

	i)	ii)	iii)	iv)
Phonetic form				[m̥ ^ɰ gyr]
Syllable tier	σ	σ	σ	σ _σ
Associations		/	/ \	/ \
Abstract form	CCVC	CCVC	CCVC	CCVC
Phonological form	/mgyr/	/mgyr/	/mgyr/	/mgyr/
Gloss	‘buy’			

FIGURE 2.7 – Example syllabication of a word with one SSP violation. i) shows the vowel is assigned to a syllable, ii) shows the onset is associated, iii) shows the coda is associated, iv) shows that violation of the SSP causes C_{distal} to be stranded.

The procedure is straightforward in simple cases, but in many words (see **FIGURE 2.4**), three or more consonants can occur in a consonant cluster, often with multiple SSP violations. **FIGURE 2.8** illustrates syllabification of a word with a four-member CC onset and a medial CC.

Voiced obstruents have a range of variation that deserves additional description. For syllabication of nasal-voiced obstruent sequences in careful speech, prenasalization tends to be realized as the coda of a syllable that is headed by the vocalic release of the preceding syllabified segment: /NC_{[+voice][-continuant]}/ → N^ə.^NC → N^ə.^NC, as in /n-/ ‘he’ + /mgur/ [n̥^ə.m̥^ɰ.'gur] ‘he (is) lazy’.

i)	ii)	iii)	iv)	v)
				[n̥ f̥n̥ 'ta ^ɰ g re]
σ σ	σ σ	σ σ	σ σ σ σ	σ σ σ σ
	/ /	/\ /	/\ /	/ /\ /
CCCCVCCV	CCCCVCCV	CCCCVCCV	CCCCVCCV	CCCCVCCV
/nf n t a g r e/				/n f n t a g r e/
‘regarding’				

FIGURE 2.8 – Example syllabication of a word with two SSP violations. Operations i)–iv) are parallel to those in **FIGURE 2.7**. An additional step, v), shows an onset is assigned to C₃.

Due to the implied preference for codas over complex onsets in the procedure of syllabification combined with the ability for vocalic release to become fully syllabic in careful speech, the initial nasal portion of a voiced stop can also approach full segmental status [NəN.C]. This property enhances the careful versus casual distinction between /ndrɪdn/ [n̥ən.'drɪⁿdn ~ n̥^ə.'ⁿdrɪⁿdn] ‘sick’ and the similar word /drɪdn/ [n̥'drɪⁿdn] ‘cold’. In these transcriptions, the word for ‘sick’ is two syllables: syllabified [n̥] and the rest [drɪⁿdn]. This phenomenon also occurs when the nasal segment is nonsyllabic, /timbe/ [tɪ.m̥əm.'be ~ tɪm.'^mbe] ‘all’, which contrasts with /tbe/ [tɪm.'be ~ t^m'be] ‘hit’. Native speakers often write /timbe/ as ‘timembe’, which suggests this is the reason why they also write /dir/ [n̥dir] as ‘ndir’.

2.4.1.6 Major and minor syllables

Evidence from phonotactics, especially word initial CCs, and syllabification suggests that two syllable types are needed to accurately describe Irarutu language data. These are analyzed as major (stress-able) syllables and minor (not stress-able) syllables. A template for Irarutu syllable structure that synthesizes the data presented above is proposed in **FIGURE 2.9**.

(C)(C), (C)(C)(C)V(C)(C)
Minor Syllable, Major Syllable

FIGURE 2.9 – Irarutu syllable template. A ‘major syllable’ has a [+syllabic] nucleus that can be a monophthong or a derived diphthong in casual speech. ‘Minor syllables’ are found in certain words, have a [+consonantal] [+syllabic] nucleus, always co-occur with a major syllable, but are almost never stressed. The boundaries of a minor syllable are predicted by the SSP and indicated by the symbol ‘,’. See discussion in §2.4.1 for additional phonotactic constraints.

Out of a sample set of 114 arbitrary items taken from a passage of a recording session of isolated words that were not patterned with any particular intent, sixty-seven words can be described in reference to the ‘major syllable’ structure. The template for major syllables in **FIGURE 2.9** captures several facts. Irarutu allows vowels to stand alone as syllables, e.g., [e.rit] ‘bark’. Onsets occur frequently and can be singleton, two-, or three- member consonant clusters. Codas (which are optional and slightly less frequent than onsets) can be singleton or two-member consonant clusters (only one example of a three-member –CC is attested). The remaining forty-seven items (41%) need a different template to account for segments incompatible with major syllables due to the sonority sequencing principle. This syllable type is captured by the label ‘minor syllable’ in **FIGURE 2.9**. The only segment type that is not attested as the [+syll] member of the ‘minor syllable’ appears to be voiceless stops, see **FIGURE 2.4** and **TABLE 2.5**.

The contrast in possible types of syllable nucleus has been shown cross-linguistically, but with a strong areal association in Southeast Asian languages, to indicate that there is structure at a subsyllabic level (cf. Sievers 1901, Huffman 1972, Edmondson and Quan 1989, Matisoff 1990, Bennett 1994, Piggott 2004), and which can be defined according to a unit called the ‘sonable’ (Stampe p.c., 2001). This structural unit is one or two consonants that are not parsed into a nearby syllable but function in a similar way to a syllable by having a sonority peak and a role in the prosodic structure of the language. Being able to refer to the sonable helps to describe how numerous lexical violations of the sonority sequencing principle are resolved in spoken Irarutu by providing a way to predict how phonological material is parsed. For example, the sonority peak of a sonable is also a probable location for an associated release feature to be upgraded to syllabic status in careful speech. To supplement the description of minor syllables and how they correlate with sonables in the remainder of this chapter, a sonable boundary will hereafter be indicated by a comma ‘,’ to differentiate relevant material from syllable boundaries, which are indicated with the period ‘.’.

Two minor syllables (prosodically organized groups of consonants in a consonant cluster) can occur in sequence, but never without a nearby major syllable. According to the literature on subsyllabic structure, if minor syllables are present in a language, usually only one can precede a major syllable. Therefore, it is noteworthy that Irarutu allows two minor syllables to occur before a major syllable, such as [m,tʃ,'taf] ‘to fan’.

To illustrate the difference between minor and major syllable types, examples that show the sonority profiles of four monosyllabic words are given in **FIGURE 2.10**, below. It must be pointed out that the treatment of extrasyllabic material impacts the prosodic structure of the language, and, because extrasyllabic material such as the sonable is described as having a [+syll] segment, some interaction with syllable count is expected.

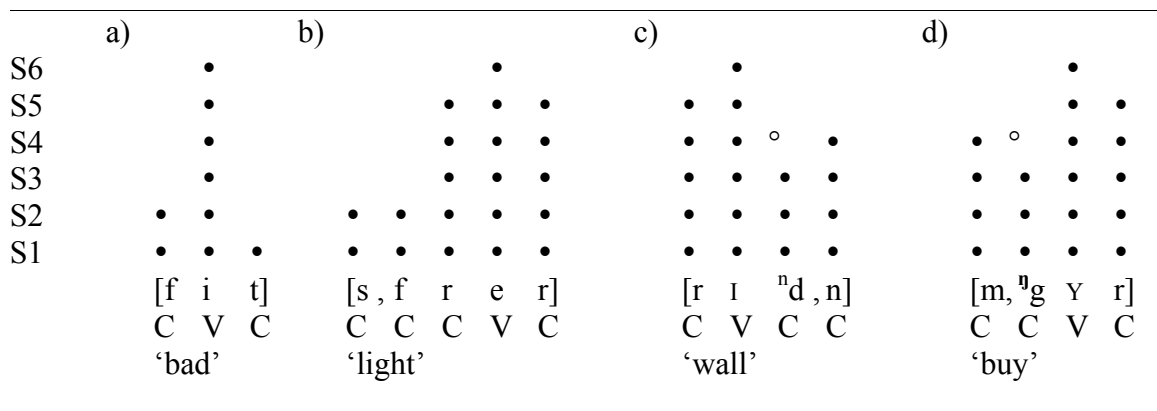


FIGURE 2.10 – Sonority profiles of four Irarutu words. S1-voiceless stops, S2-voiceless continuants, S3-voiced obstruents, S4-nasals, S5-approximant, S6-vowels. The use of hollow dots is an ad-hoc way to represent the prenasalization of voiced obstruents because nasality ranks fairly high on the sonority scale but, in this instance, cannot be given equal status as a full nasal segment.

Although the four words illustrated above generally appear to follow the SSP-prescribed profile, close examination reveals there are two sorts of violations: level sonority, e.g., b) [sf-] ²², is a less severe violation than a sonority sequence where the second consonant is more sonorous than its neighbor when it should be less sonorous, e.g., c) [-ⁿdn] ³⁴, in the coda or d) [m^ʱg] ⁴³ in the onset.

The characterization in Matsumura (1991) is that Irarutu has a (C)(r)V syllable shape. Later, van den Berg and Matsumura (2008) state that Irarutu phonotactics allow a variety of initial and medial consonant clusters, but instances of a nasal plus a stop, e.g., [nt] and [mb], have a syllabic nasal. They distinguish ‘close transition’ consonant clusters [pr, fr, br, dr, kr, gr, sr] along with nasal-plosive sequences from ‘open transition’ consonant clusters such as [CəC]. They also describe word final schwa [-C^ə] as somehow “weaker” than other instances of schwa. On the other hand, as seen in Voorhoeve 1989, to represent phonetic schwa in the orthography causes a word such as /msbegt/ [m.s^mbe^ʱgt] to look as though it has five syllables ‘mə.sə.be.gə.tə’, where in casual speech it has two syllables. Each of these peculiarities is dispensed with by accounting for extrasyllabic material according to the role of the sonable in prosodic structure, which gives rise to

minor and major syllable types identified in **FIGURE 2.9**. When used alongside the SSP, these concepts allow us to predict the behavior of consonants in certain types of Irarutu consonant clusters, such as determine which one will become [+syllabic] in particular contexts.

2.4.2 Word Structure

Four topics are addressed below to enhance the description of Irarutu words structure and suprasegmental features: the canonical shape of words, stress and accent, resyllabication, and nasalization.

2.4.2.1 Canonical word shape

Phonological words are generally no longer than four syllables, as illustrated in **TABLE 2.8**. Monosyllabic and disyllabic words are plentiful. In citation form, or for clarification purposes, disyllables are the most abundant, due to schwa epenthesis in monosyllabic words, e.g., /it/ [itə] ‘we (inclusive)’ (see §2.3, above). However, the number of Irarutu monosyllables, including true monosyllables (/wer/ [wer] ‘water’) and minor-syllable-plus-major-syllable monosyllables (/frfier/ [fɪ,fjɪr] ‘speak’), exceeds the number of phonemic disyllables (/tugro/ [tũ^ɰ.gro] ‘big bamboo’). There are some trisyllabic roots, but trisyllabic stems are more frequent, due to agglutinative (inflectional) morphology. Four-syllable words are attested, but these are generally due to compounding or reduplication of disyllabic roots.

Monosyllabic words (σ)

V	[a] ‘I/my’; [e] ‘you all’; [i] ‘he/she’; [o] ‘you (singular)’
CV	[^ɰ ge] ‘want’; [^ɰ du] ‘already’
VC	[it] ‘we (excl.)’; [am] ‘we (incl.)’; [ar] ‘I give’
CVC	[wer] ‘water’; [nir] ‘and’; [^m bir] ‘canoe’
CCV	[tni] ‘body’

Disyllabic words (σσ)

Ç.CV	[m.'tɛ] ‘eye’
CV .CV	[ʃɛ.ri] ‘crocodile’; [sɪ. ^m bɯa] ‘chili’
CVC.CV	[tu ^ɰ g.ro] ‘big bamboo’
CV .CVC	[we. ^ɰ du ^ɰ d] ‘work’

Trisyllabic words (σσσ)

C.CV .CV	[m.'ma.jɛ] ‘tongue’
CV .CV .CV	[wa. ^ɰ ga.'ri] ‘frog’; [^ɰ gi. ^ɰ gi.'si] ‘chili’
VC.CVC.CVC	[am.ta ^ɰ g.'rum] ‘my face’

Quadrisyllabic words (σσσσ)

CV. V.CV. V	[so.a.so.a] ‘lizard’ (reduplicated CVV root)
CV.CV.CV.CV	[ra. ^m bi.ra. ^m bi] ‘k.o. knife’ (reduplicated CVCV root)
CV.CV.CV.CVCC	[si.nje.fe.ta ^m br] ‘sweet potato’

TABLE 2.8 – Length of words according to syllable count.

2.4.2.2 Stress and accent

Stress in Irarutu is not predictable. Himmelmann and Ladd (2008) identify pitch, duration, and stress as the primary components of lexical prosodic structure in the context of describing the typology of prosodic systems paralleling tone, accent, and intonation. Stress and accent overlap in expressing prominence through acoustic properties including intensity, pitch, volume, and spectral energy (see the spectrogram in **FIGURE 2.12** for an example of how these properties interact in Irarutu). In casual speech, monosyllabic words have only one major syllable that can bear stress, but the locus of stress in minor-plus-major-syllable monosyllables is known to fall on a minor syllable (see **FIGURE 2.11**) in at least one example. Disyllabic words, i.e., words with two major syllables, can be either iambic, that is stressed on the final syllable (weak-strong), or trochaic, stressed on the penult (strong-weak). Minor-plus-single-major-syllable (i.e., monosyllabic) words belong in the iambic category. The stress patterns in disyllabic words strongly suggests that stress is phonemic, i.e., lexical. Van den Berg and Matsumura (2008) mention “stress is possibly phonemic, although there are few convincing minimal pairs.” Indeed it is difficult to find examples of contrastive stress, but stress does appear to be phonemic.

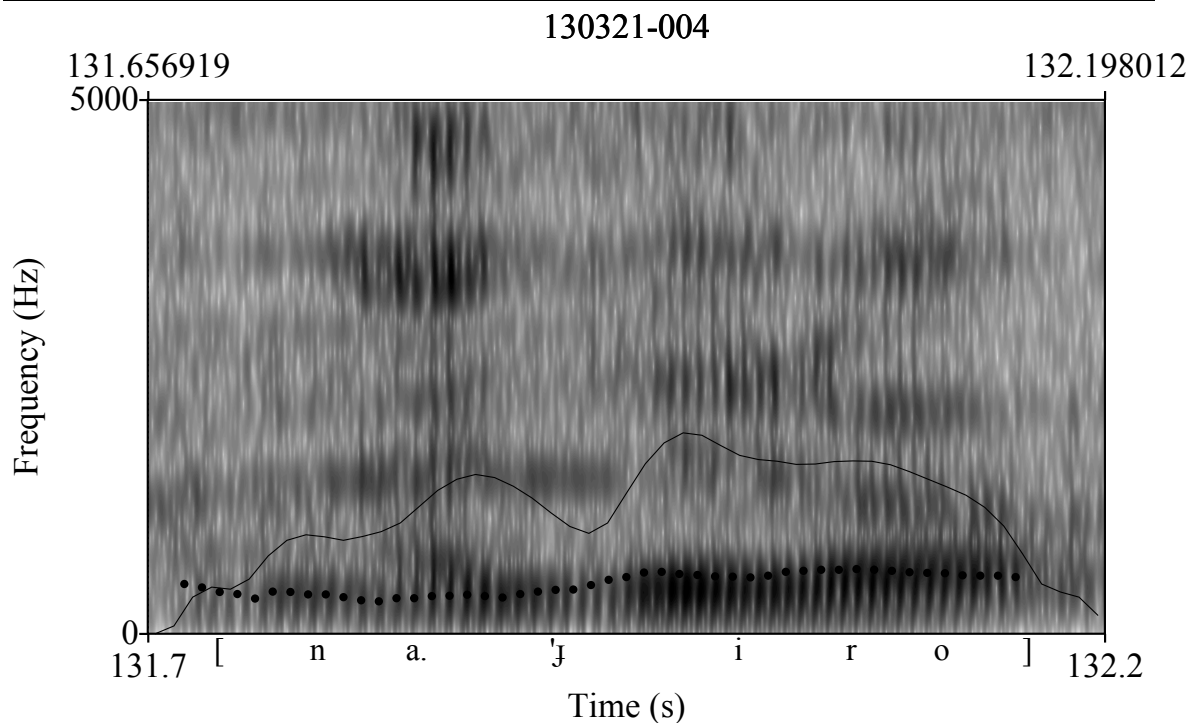


FIGURE 2.11 – Acoustic analysis of the word /naji-ro/ [na.'ji.ro] ‘expletive (if that’s so)’ illustrates how intensity contributes to stress. The spectrogram shows the formants, with added intensity (solid line, measured between 50–100 decibels, peaks at 68 dB), and pitch (as dots). The file name is listed at the top. Frequency (in Hz) is on the y-axis. Time (in seconds) is on the x-axis. The speaker is male.

The number of words with final stress is roughly proportional to the number of words with penultimate stress. For trisyllabic and longer words, stress usually falls on the final

syllable or the penult, but because stress is lexically determined, it can fall on the antepenult or the preantepenult. These generalizations concur with the literature (Voorhoeve 1989, van den Berg and Matsumura 2008), but the best published example of phonemic stress does not isolate it as the only contrastive element, therefore there is room for speculation. Two examples of pairs of words that contrast in location of stress are provided in **FIGURE 2.12**.

Final stress	: Penultimate stress
[fi.'na] 'like/as'	: ['fi.na] 'fire'
[r.'re ~ re] 'day'	: ['r.re] 'good day (greeting)'

Figure 2.12 – Two examples of minimal pairs showing phonemic stress.

These examples do need some qualification. In the first pair of words [fi.'na] 'like' is only one of several alternative realizations of concatenated /fi/-/na/, which can also surface as [fna ~ fena ~ vena ~ vna], depending on the speaker and context. The form [fi.'na] occurs in careful speech. In casual speech, the form *fna* 'like' [f^əna], with vowel reduction, is far more common. [f] and [v] are free variants of /f/, and both words, 'like' and 'fire', have been observed with [f] and [v]. A further consideration is the morphological status of /fi/-/na/ 'like/as', in which /fi/ is an active verb marker or nominalizer, and /na/, which, in this instance, can be analyzed as a habitual/infinite marker. It has characteristics of both word and affix. It behaves like a word, because it can be inflected by other prefixes and stand alone (*n-f-na*), but it can also be used in the derivation of other words (/f-na-//*nen*o/ [fnane'no] 'how').

In the second pair of words in **FIGURE 2.11**, it is rare to observe the stress pattern of [r^ə.re] 'good day' because it is one of the only examples that has been noted for stress on a minor syllable, and the contrast it provides is not a productive semantic alternation. This characterization also clashes with the fundamental assumption that stress is a property of major syllables rather than minor syllables. In essence, to claim a stress pattern like [r^ə.re] is to claim that stress could fall on [r] or even non-phonemic material [ə]. In actuality, stress falls on the initial consonant [r] despite the fact that there is a considerably better, i.e., more sonorous, candidate to bear stress, i.e., the [+syll] /e/. This issue will be left for future work on prosody because it is peripheral to claiming that Irarutu has phonemic stress. Although the stress system appears complicated, stress itself is a good diagnostic for investigating allomorphy such as compounding, reduplication, and cliticization, discussed below (§2.6).

On a related point, Matsumura (1991) noticed that a syllable preceding a stressed syllable tends to have extra phonetic length. This is observed in the Fruata dialect words such as /wagari/ 'frog', which surfaces as [wa.^ŋga'.ri]. However, Matsumura's account was closely tied to his analysis of schwa, which the present analysis does not support. The importance of Matsumura's observation is that stress in Irarutu is not specifically correlated with increased duration. This agrees with the characterization above, in which stress is expressed through relatively greater quantities of intensity, volume, and pitch.

2.4.2.3 Resyllabification

Where syllabification is an operation that assigns segments to syllables at the word level, resyllabification is a phrase-level operation that reassigns segments, in particular, those that violate the sonority sequencing principle, across word boundaries to adjacent syllables in connected speech, often-times avoiding an SSP violation. Resyllabification reflects the preference in syllabification to distribute segments as evenly as possible between onsets and codas by improving the match between segments and syllables following a preference for simple codas (VC) over complex onsets (CCV) when words are strung together in a phrase. **FIGURE 2.13** illustrates how V.CCV is resyllabified as VC.CV.

a)	b)
<u>Syllabification:</u>	<u>Resyllabification:</u>
[ja. ^u ga. fa. m ^u gyr. wer]	[ja. ^u ga.fam. ^u gyr.wer]
σ σ σ σ σ σ	σ σ σ σ σ
/ / / / \ / \	/ / / \ / \ / \
CV CV CV CCVC CVC	CVCVCVCCVCCVC
/ja ga fa mgyr wer/	/ja ga fa mgyr wer/
1S DES go buy water	

Figure 2.13 – Example of resyllabification of the phrase /ja ga fa mgyr wer/ ‘I’m going to buy water’.

An additional property of resyllabification, indicated in the syllable tier of the example above, is that it reduces – but does not always eliminate – the presence of seemingly extra-syllabic material, which is relegated to minor syllables in citation form, from connected speech. An SSP violation motivates resyllabification because it presents a sequence with zero change. However, evidence from ‘type-III’ initial consonant clusters and three-member medial consonant clusters (see §2.4.1) indicates that some minor syllables cannot be redistributed into major syllables, even after resyllabification.

2.4.1.5 Derived diphthongs

Resyllabification is also responsible for producing derived diphthongs from sequences of post-morphologically adjacent vowels, which constitute separate syllable peaks /u, i, y/ under syllabification.

Out of twenty-four possible combinations of adjacent vowels illustrated in **TABLE 2.9**, there are eleven attested vowel sequences that surface as either ‘rising’ (V̆V) or ‘falling’ (VV̆) diphthongs. The phonemes /i, u/ frequently desyllabify when one is unstressed adjacent to a stressed vowel /nia/ [nja] ‘what’, /ad-e-i/ [aⁿdej] ‘the (lit. the-distal-known)’. No known examples show that /i/ desyllabifies word initially. In examples where [ɪ] is observed as the non-syllabic element in a diphthong, it is interpreted as a realization of /i/, e.g., /fier/ [fier ~ fjer] ‘speak’. There are also two examples that suggest derived triphthongs are possible: [wajn] ‘itch’ and [njawnt] ‘k.o. ant’.

	i	ɪ	y	u	e	o	a
i	—	—	—	[ⁿ s ^ŋ gwifr] ‘move’	[ⁿ d ^m bej] ‘he hit it’	—	[faj] ‘circle’
ɪ	—	—	—	—	—	—	—
y	—	—	—	—	—	—	—
u	[nju.'en] ‘he weaves’	—	—	—	—	—	[^m gaw ⁿ g.r] ‘lively’
e	[mjef] ‘even’	—	—	[^m bwek] ‘betel nut’	—	—	—
o	[sjo.ru] ‘k.o. knife’	—	—	[wo] ‘line’	—	—	—
a	[pjar] ‘nourishment’	—	—	[pwar] ‘slap’	—	—	—

TABLE 2.9 – Examples of derived diphthongs.

Not all instances of adjacent vowels in Irarutu become derived diphthongs. This is especially the case in words with penultimate stress. In the short phrase /buek pua/ [^mbuek 'pu.a] ‘a lot of betel nut (lit. betel-nut many)’, the vowel-vowel sequence in the quantifier /pua/ ['pu.a] ‘much/many’, behaves differently than the vowel-vowel sequence in the noun /buek/. /pua/, which has penultimate stress, is always disyllabic, regardless of speech tempo. To reduce the number of syllables by altering the location of stress [pu.'a ~ pua] is considered ill-formed. On the other hand, the noun /buek/ ‘betel nut’, which has final stress, usually surfaces as a monosyllabic form with a derived diphthong [^mbuek], even though when several different Fruata dialect speakers were asked to say the word slowly, it was broken-down into four parts [^mb], [u], [e], [k]. Another example that illustrates ‘derived diphthongs’ is the word /nia/ ‘what’. In casual speech, the stress falls on the final syllable of this word [ni.'a], allowing the preceding vowel /i/ to desyllabify [i] and subsequently coalesce with the stressed syllable nucleus, resulting in the form [n̩a]. In terms of syllabification, the two vowels are assigned to a single syllable.

A generalization based on the preceding examples is: words that have two adjacent vowels at the lexical level, and final accent, tend to be realized as derived, rising diphthongs in casual speech, whereas sequences of adjacent vowels at the lexical level that have penultimate stress tend to occupy two syllables, rather than derived falling diphthongs, although a few rare examples of derived falling diphthongs are attested. The low vowel [a] tends to retain its syllabicity, even when it is unstressed. This is true in words of more than two syllables, for example [so.a.'so.a] ‘lizard’, *[sq̩a.'sq̩a]. However, the generalization appears to have some exceptions. The word for ‘say/speak’, /fier/ [fi.'er] (careful tempo) varies between [fjer ~ fir] (casual tempo).

At the phrase level, resyllabication also interacts with vowel reduction, as noted in the discussion of vowel reduction, above, and together they are responsible for deleting

vowels in certain contexts, such as an unstressed vowel adjacent to a word that is constituted by only a stressed vowel.

/te/ /o/ /m-tit/ + Vowel Reduction → [tə.o.m.tit] + Resyllabication → [tom.tit]
 INCID 2S 2S-fall
 ‘You will fall’

2.4.2.4 Nasalization

The feature [+nasal] plays an important role in vowel allophony but is not confined to the domain of individual words. Vowels become nasalized preceding a nasal consonant, including the nasals /m/ and /n/, but also /b/ [ᵐb], /d/ [ᵐd], /g/ [ᵑg], and, for speakers who prenasalize the voiced palatal stop, /j/ [ᵑj]. From an articulatory perspective, anticipatory nasalization in vowels is motivated by a gesture that lowers the velum to prepare for [+nasal] consonants. This accounts for the fact that [-nasal] consonants block nasalization from spreading towards the beginning of a word, except for the laryngeal continuant /h/, which does not involve a gesture that conflicts with a lowered velum, e.g., /loruhama/ [ˈlo.ru.hã.ma ~ ˈlo.rũ.hã.ma] ‘woman’s name’. In Irarutu, vowel nasality is predictable, not contrastive.

V → Ṽ / __ N
 /sum/ → [sũm] ‘fish’
 /san/ → [sãn] ‘house’
 V → Ṽ / __ C_[-continuant, +voice, +nasal]
 /seba/ → [sẽᵐba] ‘month/moon’
 /kadfa/ → [ˈkãᵐd.fa] ‘pants’
 /aje/ → [ãᵑje] ‘my father’
 /muᵑg/ → [mũᵑg] ‘much/big’

In contrast, vowels following /m/ or /n/ are imperceptibly nasalized, if at all, at their beginning, and only in cases where the timing of velic closure is not coordinated with release of the oral constriction. Vowels after a voiced stop are not nasalized.

/nof/ → [nof] ‘wind’
 /a-/ma/ → [ã.ˈma] ‘I come’
 /dap/ → [ᵑdap] ‘can/able’

The examples of vowel nasalization in the words for ‘month/moon’, ‘my father’, and ‘I come’, above, show that nasalization can cross a syllable boundary. However, nasalization can cross a word boundary after application of resyllabication. For example, /te o mtit/ [te.o.mtit] (careful speech) becomes [to.ˈmtit ~ tõm.ˈtit] (casual speech) ‘you will fall’.

2.5 Loanword phonology

A few comments on loanword phonology are in order on account of the fact that Irarutu is an endangered language and one sign of a language’s vitality is its ability to nativize loanwords. The description below focuses on how phonemes from other languages, such

as [p, t, c, ʃ, h, l, ŋ], are nativized, and how the phonological structures of loanwords are adapted, such as disyllables to monosyllables.

Under the supposition that Irarutu is an Austronesian language (see Chapter 4), all historic instances of PMP *h > IRH Ø and PCEMP *p > IRH f, meaning the presence of these phonemes in modern-day Irarutu must have been reintroduced from an external source. As mentioned previously, nearly all Irarutu speakers today also have good command of Indonesian, and Irarutu speakers who live in and near to Kaimana also possess some knowledge of Arabic. Therefore, multilingualism in Indonesian and Arabic provides likely sources for contemporary Irarutu *p* and *h* (examples are given in the descriptions of /p/ and /h/, above). There is also considerable likelihood that one or more of Irarutu's neighboring languages also contributed to reintroducing /p/, as in /pua/ [ˈpu.a] 'much', which probably paved the way for words borrowed from Indonesian. Based on what is known about the languages that surround Irarutu, this hypothesis is tentative, meaning that the source of such words is presently unknown, but IRH /baba/ [ˈmba^mba] appears to come from Kamrau [baba] 'elder sibling'.

Indonesian/Papuan Malay is the source for several other loanwords, such as [tbur] 'grave' (IND kubur) and *ʃut* 'form' (IND bentuk), in which there appears to be a correlation between Indonesian k and Irarutu *t*. The interpretation of PCEMP *k > IRH /t/ is paralleled by sporadic reversal of coronals and velars in several widely separated Austronesian languages (Mayer 2001, Blust 2005, Donohue 2006, and see Jackson 2004 for a phonetic motivation). Presumably, Indonesian and Irarutu use different phonetic cues to distinguish [t] and [k]. The identification of such cues is an integral part of the phonetics-phonology interface in Irarutu that deserves further research.

The Indonesian palatal stops [c] and [ɟ] generally merge as Irarutu [c], but IND [c] can also become [s]. In loanwords from Indonesian, IND 'j' [ɟ] is interpreted in Irarutu as either [ɟ], or devoiced [c]. For example, IND mengajar [mə.ˈɲa.ɟar] 'to teach' is borrowed as IRH [fi.'car] 'to teach'. By contrast, the voiceless palatal stop [c] is sometimes reanalyzed as a voiceless continuant, which can be plain [s] or palatalized [s^j ~ ʃ^j]: IND coba ['co.ba] 'try/sample' borrowed as IRH [fi.cō^mb] ~ [fi.ʃō^mb], or IND cukur ['cu.kur] 'shave' borrowed as IRH ['su.kur].

Irarutu has no velar nasal phoneme. As a consequence, Indonesian ng [ŋ] is interpreted in Irarutu as [n], IND sembahyang is realized as IRH ['s^mba.jan]. The Fruata dialect also lacks /l/. In nativized loanwords, [l] from another language is reinterpreted as [r]. For example, IND bola 'ball' becomes IRH *bora*.

The naturalization of loanwords pertains, not only to the segmentation of Irarutu phonemes, but also to the phonotactic and word shape constraints in Irarutu phonology. The data that is available in this regard mostly originates from Indonesian. In general, the process of nativization often involves reducing the number of syllables in a loanword. One example, IND pintar 'smart' to IRH [pintr], which is the only instance of a three-member final CC, shows that Irarutu phonotactics is flexible enough to accommodate an otherwise unattested sequence, but that the SSP violation between [t] and [r] can be

accounted for under the general pattern of allowing word-final minor syllables. In that example, and others, the reduction of number of syllables takes place through vowel deletion processes such as syncope, and deletion of consonants in various locations. The result is more often than not a monosyllabic form. For these types of words there are many examples: IND /meN-//**paku**/ me'maku 'to nail' to IRH [n-fi-**pa**], IND me-**'nampar** 'to knock' to IRH [fi.**pwar**], IND /meN-//**pelihara**/ memeli'hara 'to nourish' to IRH [fi **pyar**], IND men-**dapat** 'able/can' to IRH [fi **dap**], IND **'bisa** 'can/able' to IRH [**bis**], and IND **bi'asa** 'usual' to IRH [**bjas**], and so forth. However, there are also examples where two or more syllable words in Indonesian attain an equal number of syllables in Irarutu. An example is Indonesian **kapal** 'a ship' to Irarutu [**kap.ri**], where IND **l** was borrowed as IRH **r**, **a** was deleted in the final syllable, but the addition of [-i] is unexplained. An example of a loanword that is ultimately from a European source, [tə.'ba.ko] 'tobacco', but which was probably filtered through Indonesian, is IRH [**ke^m.ba.ku** ~ '**k^mba.ku**] 'tobacco'. As with the generalizations concerning speed of speech, the number of syllables in this item varies between two and three.

Unfortunately, nativization of loanwords appears to be on the decline, whereas the role of code-mixing (which is indicated, for example, by the segment [l], and is suggested by [j] in a syllable onset) is increasing, especially in domains where Irarutu does not have naturally occurring terminology, such as [al.'lah] 'Allah (God)' and ['s^mba.jan] 'prayer/church service'. In addition, access to the internet via cell phones and personal computers has introduced terminology from global languages, such as English ['i.meɪl] 'email'. The increasing use of code-mixing, even in conversations between native speakers of Irarutu, is also observed in the use of phonologically nativized loan words such as [suk] 'like' (cf. IND suka), despite the presence of an equally adequate Irarutu term ['tu.fi] 'like'.

2.6 Morphophonemics

Four types of allomorphy are discussed in this section because their effects are seen frequently in the following two chapters: verbal inflection, reduplication, compounding, and cliticization,

2.6.1 Allomorphy in verbs

Allomorphy in verbs is observed in the set of subject markers and in the stative marker. Both types of markers are prefixes that are conditioned by the segment they precede. One generalization is that identical segments across an inflectional morpheme boundary tend to reduce to a single segment in casual speech.

2.6.1.1 Subject markers

Two disparate substitutions are conditioned by the agent marking morphemes, nasal place assimilation involving [-labial] consonants, and voiceless continuant fortition. The subject markers are /a-/ 'first person' (which is considerably more frequent in relative clauses than main clauses), /m-/ 'second person', and /n-/ 'third person': /a-//fror/ [a'fror] 'I run', /m-//fror/ [^mbror] 'you run', /n-//fror/ [nfror] 'she runs'.

2.6.1.1.1 Nasal place assimilation

In broad terms, sequences of a homorganic nasal-voiced obstruent as well as sequences of homorganic nasal-nasal segments that result from morphological processes merge. For example, /m-/ ‘second person subject’ plus /mi/ ‘live’ becomes [mi] ‘you live’, and /m-/ + /brif/ becomes [mbrif-] ‘you laugh’. The sequences [mm-] and [m^mb-] only occur in hyperarticulated speech. As mentioned in the phonetic description of /n/ above, the [-labial] [+nasal] phoneme /n/ has an optional allomorph [ŋ] in casual speech where it occurs before a dorsal segment, but does not do so in careful speech, for example /n-/ ‘third person subject’ plus /kka/ ‘swim’ [n.k.'ka ~ ŋ.k.'ka] ‘he swims’. Furthermore, assimilation of /n-/ to [+dorsal] before g-initial roots as [ŋg] is more probable in casual speech, as in /n-//gigr/ [n^ŋgi^ŋgr ~ ŋgi^ŋgr] ‘ask’. This is not observed for [+labial] consonants, allophonically or allomorphically. That is, the [+labial] /m/ does not assimilate to a following [+dorsal] consonant, either at the lexical level /mgor/ ‘to bite’ [m^ŋgor] *[ŋgor], or at the inflectional level /m-/ ‘second person subject’ /kka/ [mkka] *[ŋkka] ‘you swim’, /m-/ /gigr/ ‘ask’ [m^ŋgi^ŋgr] *[ŋgi^ŋgr]. This type of morphophonemic anticipatory place assimilation can be summarized as follows.

$$C_{[+coronall][+nasal]} \rightarrow [+dorsal] / __ C_{[+dorsal]}$$

Matsumura (1991) comments that the morpheme /n-/ ‘third person subject’, can assimilate to a following [+labial], summarized as /n-/ → [m] / __ C_[+labial]. This does not appear to be true in the Fruata dialect. For instance, /n-//pa/ [npa] ‘he hammers/nails’, whereas /n-//gor/ becomes [n^ŋgor] ‘he bites’. In addition, Matsumura (1991) and Voorhoeve (1989) independently propose that /n-/ merges as [s] before /s/ initial roots in the dialects they describe. This is true in the Fruata dialect for the third person subject marker, depending on speech tempo. Citation forms such as [ŋsi] ‘he sees’ were recorded in the research for this description.

/n-/ + /si/ → [ŋsi] (careful speech) ~ [si] (casual speech) ‘see’

/n-/ + /su/ → [su] ‘return’ (Voorhoeve 1989)

2.6.1.1.2 Voiceless obstruent fortition

When a verb that begins with the voiceless labial continuant /f-/ is inflected with the second person subject /m-/ , or when the voiceless anterior stop /t/ is inflected with the third person subject marker /n-/ , the sequences surface as prenasalized voiced stops [m^b] and [n^d] respectively. In both cases, the voicing of the nasal influences the voicing value of the following segment but the nasal vanishes as a separate segment. In the case of /f/, it is also changed into a stop. The allomorphy of /m-f/ → [b] at the beginning of verbs was noticed by both Voorhoeve (1989) and Matsumura (1991). Voorhoeve (1989) also noticed that /n-t/ → [n^d].

/m-fa/ [m^bba] ‘you go’
 /m-fun/ [m^bbun] ‘you make’
 /m-fi/ [m^bbi] ‘you (active verb...)’
 /m-fnogr/ [m^bbno^ŋgr] ‘you hear’
 /m-fier/ [m^bbjer] ‘you say’

/n-tfe/ [ˠdfe] ‘he spits’
 /n-tfur/ [ˠdfur] ‘he sinks’

It seems relevant that [f] and [v] are in free variation because the allomorphy of /m-f/ to [ᵐb] entails three processes: voicing assimilation ([m-f] ~ [m-v]), manner change ([mv] to [mb]), and finally, reduction of the nasal phoneme [mb] to prenasalization in the following segment [ᵐb]. In the alternation of /n-t/ to [ˠd], the manner change is not observed, but the process nevertheless involves assimilation to [+voiced]. Notably, fusion of /-mf-/ → [ᵐb] occurs for second person referents, but not in all cases of a first person plural exclusive subject [amfa] ‘we (exclusive) walk’, see **TABLE 2.10**. The irregularity of the alternation for first person exclusive subjects suggests that this is a morphophonological substitution.

		1n	[tfa] ‘we (inclusive) walk’
1	[a fa] ‘I walk’	1x	[amfa] ‘we (exclusive) walk’
2	[o ᵐba] ‘you walk’	2	[ᵐba] ‘you all walk’
3	[i fa] ‘he walks’	3	[nfa] ‘they walk’

TABLE 2.10 – Verbal paradigm for ‘walk’ showing allomorphy in subject marking pronominal prefixes. The 1x pronoun /am-/ sporadically causes *f*-initial verb stems to become *b*-initial.

Importantly, the /m-f/ → [ᵐb] morphophonemic alternation does not occur in cases that do not involve subject marking, such as stative marking, which is indicated by /m-/, e.g., *ja m-fun* (I STAT-do) ‘I am doing’ ***ja bun*, and see §2.6.1.2.

2.6.1.2 Stative marker

Following the generalization given above that morphological sequences of homorganic [+nasal] segments merge, the stative verb marker (denoting a state rather than a dynamic event) /m-/ merges with verbs that start with /m/ as a single [m], or with /b-/ as [mb-].

/m-/brfun/ → [mᵐbrfun] → [ᵐbɾ.fun] ‘to hide/be hiding’

Because the third person subject marker /n-/ can be added in front of the stative marker, [n-m-tür] ‘he sits’, and not vice-versa, it can be said that the stative marker plus verb gives rise to a stative verb stem, to which the subject marker is added. Merger of homorganic nasal sequences, as in /m-/ second person subject marker plus /m-/ stative marker, plus a verb such as /brfun/ prevents forms such as *[ᵐᵐᵐbrfun], which surfaces as [ᵐbɾ.fun] ‘you are hiding’. Additional examples of stative verbs are given in **Figure 3.5** in §3.1.2.2.2.4.

2.6.2 Reduplication

There are three broad patterns observed for the phonological treatment of reduplicated roots in Irarutu: full reduplication and two types of partial reduplication: prefixed partial and suffixed partial reduplication. Both forms of partial reduplication entail vowel deletion in the reduplicant.

2.5.2.1 Full reduplication

Full reduplication can apply to monosyllabic as well as disyllabic roots. This form of reduplication does not undergo any segmental changes in the reduplicant. It can be found in several word classes and often involves a morphosyntactic change. Based on the location of stress in the final syllable in a reduplicated CVC monosyllable (CVC.'CVC), or the penultimate syllable in a reduplicated CVCV disyllable (CV.CV.'CV.CV), full reduplication appears to prefix the reduplicant to the base, as shown in the following examples. [rã^m.bi.'rã^m.bi] is the low prestige register equivalent of high register /kobis/ [kõ^mbis]. Stress appears to be retained by the underlying root.

/met/ 'half' + Full Reduplication → [met.'met] 'little bit'

/tım/ 'to close' + Full Reduplication → [tım.'tım] 'close (agentless verb)'

/rabi/ (no independent meaning) + Full Reduplication → [rã^m.bi.'rã^m.bi] 'k.o. knife'

2.6.2.2 Prefixed partial reduplication

There are two types of prefixed partial reduplication. One reduplicates the initial consonant only. This form of reduplication adds one consonant to the surface representation of monosyllabic CV lexical items, for example /re/ → [r^ə.re] 'day' (cf. /re+fefa/ [re.fe.fa] 'afternoon', /re+dir/ 'good' [rẽ.^mdir] 'good day'). Other items that appear to be products of this type of prefixed partial reduplication are /ke/ [k^əke] 'jaw', /kor/ [k^əkor] 'chicken', /ka/ [k^əka] 'swim', /fu/ → [f^əfu] 'top', and possibly also in the disyllabic CVCV form /maje/ → [m^ə.'ma.je] 'tongue', but evidence from compounding that confirms the underlying form for many of these items is not available. The semantic contribution of this type of reduplication is not certain, but it correlates most closely with citation form.

The other type of prefixed partial reduplication (abbreviated PPR) targets CVC(C) roots. The consonants that are adjacent to the vowel are copied, merged into a consonant cluster, and prefixed to the root, resulting in a form such as CC-CVC(C). This process produces casual speech monosyllables and is correlated with a range of added semantic components, such as nominalization, plurality, and distributivity.

/gir/ 'story-tell' + PPR → [g^əgr.^mgir] 'story'

/bar/ 'lung' + PPR → [b^mbr.^mbar] 'lungs'

/met/ 'half, part' + PPR → [m^ət.^mmet] 'little by little'

Some words that appear to be products of this type of reduplication suggests stems that are no longer independently meaningful, such as /kur/, /fadr/, and /tagr/. Although the form [kur] has no independent meaning, [kr.^mkur] 'young/green', with hypothesized CC-reduplication, contrasts with [k^əkur] 'shave', with hypothesized C- reduplication. The forms [m.kr.^mkur] 'green/unshelled/unprocessed' and [m,t^əg.ta^əgr] 'k.o. onomatopoeia' (below) suggest that reduplication precedes verbal inflection, such as addition of the stative marker /m-/. On a final note, CC- PPR does not include any C that is not adjacent to the syllable nucleus: *[fⁿdrfaⁿdr] 'floor on stilts'.

/kur/ ‘(no independent meaning)’ + PPR → [kr.'kur] + /m-/ ‘stative’
 → [m.kr.'kur] ‘young/green’
 /fadr/ ‘(no independent meaning)’ + PPR → [fⁿd^o.faⁿdr] ‘floor on stilts’
 /tagr/ ‘(no independent meaning)’ + PPR → [tⁿg.taⁿgr] + /m-/ ‘stative’
 → [m,tⁿg.taⁿgr] ‘k.o. onomatopoeia’

2.6.2.3 Suffixed partial reduplication

There are putatively three varieties of suffixed partial reduplication. One variety targets roots that end in open syllable shapes, copies the C_{adjacent} and the V and suffixes it to the root, resulting in a form such as CCVCV, such as /tbe/ ‘hit’ to [tⁿd^mbe.^mbe] ‘he hit repeatedly’. This form of suffixed partial reduplication correlates with iterativity. The second form of suffixed partial reduplication targets stems that end in closed syllables, such as /esu/ ‘one’ /-em/ ‘just’ becoming [e'su^mem] ‘just one’, copies the VC rime and produces a form that ends as VCVC: [e'su^mem-em] ‘various’. This type of suffixed partial reduplication adds plurality, or variegation, to the meaning of the stem.

A putative third variety of suffixed reduplication appears to target a monosyllabic root with a CC onset and no coda, i.e., CCV. This form of reduplication copies the CC- and suffixes it to the root, as in /fra/ ‘hand’ to [fra-**fr**] ‘wing’. But an example such as /fri/ ‘cost/value’ compared with [frifr] ‘wide/ocean’ is coincidental and cannot be considered an example of this type of suffixed reduplication. A notable problem in asserting this type of reduplication is the lack of associated semantic correlation between the proposed root and the derived stem. It could add an element of abstraction or generalization (‘hand’ to ‘wing’), but this correlation needs additional support that cannot be cited at the present time, due to lack of relevant data.

2.6.3 Compounding

Compounding interacts with two phonological phenomena, stress assignment and resyllabication. Stems in compounds have one stress peak, which usually occurs on its second member (/sirnie/ [sir.'nje] ‘potato’ + /fe-/ ‘thing’ /tambr/ ‘plant’ → [sir.,nje.fe.'tã^mbr] + penultimate vowel reduction → [sir.,nje^f.tã^mbr] ‘sweet potato’), as opposed to stress falling on each constituent, which is the case for a phrase composed of two independent words, [sũn.sũn 'mmat] ‘green shirt’. Only the vowel in the word /fe/ deletes in the word for ‘sweet potato’. Stress is retained on the final syllable of the word for ‘potato’ in the compound, but it is reduced to secondary stress.

In inalienable possession, the first member of a compound noun is circumfixed (i.e., prefixed and suffixed) with possessive morphemes that agree with the person and number of the possessor. In the example below, the consonant cluster /mt/, which belongs to the root /mta/ ‘eye’, is resyllabified as the coda of the preceding possessive morpheme. The word-final location of stress on the second syllable of the compound is retained in the trisyllabic inflected form.

/mta/-/rum/ ‘face’ + /a-...-ⁿg/ → /a.mtaⁿg.rum/ + Resyll → [ãm.tãⁿg.'rũm] ‘my face’
 eye house 1SPOSS.INAL

On a point related to inalienable possession, suffixes hosted by stems with final vowels are typically a single consonant, /-g/ ‘first person possessive’ or /-m/ ‘second person possessive’, but suffixes hosted by stems with final consonants appear as [-ug] and [-um], with a vowel to separate the stem final consonant and the suffix. Voorhoeve (1989) proposed the vowel could be either [u] or [o]. Only forms with [u] were collected for this research.

2.6.4 Cliticization

There are three clitics in Irarutu, /=ti/ ‘negative marker’, /=ro/ ‘topic marker, and /=e/~/=o/ ‘politeness marker’. These elements are important for the description of phonology in the sense that they are phonologically dependent syntactic elements (Zwicky 1985, but see also Pike 1947), therefore they can be tacked onto a word with no effect on stress placement. For example, the verb /tufi/ [‘tu.fi] ‘like’ plus /=ti/ ‘negative’ becomes [‘tu.fi.ti] ‘dislike’ and has stress on the same syllable [tu] as the plain verb, although this means that stress is on the antepenult, which is fairly rare in Irarutu, rather than on the penultimate syllable, which is the more common location.

As a general rule, none of the three clitics bears stress /gin/ ‘drink’ plus /=ti/ *[gĩn.‘ti] ‘I/you do not drink’; however, in an instance where a speaker wants to clarify negation, /=ti/ *can* stand on its own, in which case, it must bear stress [gĩn_‘ti] ‘I do NOT drink’.

Stress placement also helps to distinguish between the copula /rau/, which is realized as [ro], and the topic marker /=ro/ [ro]. This is useful because either /rau/ or /ro/ can surface in otherwise seemingly identical positions in a clause (see Ch. 3). The copulas can bear (secondary phrasal) stress, whereas the topic marker cannot.

2.7 Summary

This chapter described several topics in the synchronic phonology of Irarutu, starting with the phoneme inventory, based on minimal pairs. Several types of allophonic variation were described. The description then turned to examining phonotactic constraints using attested consonant clusters and the process of syllabication. These two forms of evidence motivated identifying two types of syllables for descriptive purposes and formulating a generalized template for Irarutu syllables. Above the level of the syllable, the canonical shape of Irarutu words was summarized, followed by a description of suprasegmental phenomena, including stress and accent, resyllabication, and nasalization. A brief discussion of how segments and word shape are nativized in loanwords was also presented, as were several types of allomorphy that are crucial in the description of the morphosyntax of the language, which is the topic of the next chapter.

Chapter 3. Morphology and syntax

Irarutu is a rigid word-order, or strongly ‘configurational’, SVO language with no morphological case marking. Noun phrases in naturalistic language data tend to be fairly compact relative to the verb phrase, which tends to be relatively elaborate. A notable feature for Irarutu as a VO language is that adjectives follow the nouns they modify, but, against certain typological generalizations, the genitive precedes a possessed noun. This chapter describes the basic syntactic properties of Irarutu following common templates for descriptive syntax (Comrie and Smith 1977, Shopen 1985, Payne 1997, Dixon 2011–12, O’Grady 2013). The overall description is organized into four main parts. Syntactic categories are identified and discussed in §3.1, which is followed in §3.2 by descriptions of four interrelated phenomena: grammatical relations, word order, transitivity, and agreement. Findings about these phenomena are then used to describe complex clauses, with particular attention to relative clauses, in §3.3. In the final section, §3.4, speech act distinctions and their correlations with syntax are presented. A summary, §3.5, concludes the chapter.

3.1 Syntactic categories

The description of syntactic categories, or parts-of-speech (also called lexical categories), can be divided into two major classes: ‘open’ and ‘closed’ (Baker 2003). ‘Open’ categories have an unlimited number of items, whereas the members of ‘closed’ categories can easily be delimited. Both classes are further subdivided according to several criteria, including distribution, syntactic function, and morphological or syntactic category specification (Schachter 1985). One test proposed by Croft (1991), to determine the category of a word is by the application of bound morphology. The next few sections describe Irarutu word-formation and overviews its basic morphological properties. The actual names for the various parts of speech are relatively traditional, being drawn conveniently from traditional semantic or notional correlation. A list of Irarutu parts-of-speech is presented in **TABLE 3.1**. In this chapter, ‘y’ represents [j], ‘j’ represents [ɟ], and ‘ü’ represents [y], which is similar to first-language-speaker orthography of the language, except for the added vowel distinction.

Open classes

Nouns (common: *san* ‘house’, *figrgir* ‘(a) story’; proper: *jek* ‘Jack’)

Verbs (active *fun* ‘do’, *figrgir* ‘to (talk) story’; stative: *mtür* ‘sit’; copular *rau* ‘to be/is’
REPEAT)

Adjectives (descriptive: *mo kokon* ‘small child (baby)’; predicative *ja tni trro* ‘I am tall’)

Adverbs (sentential *najiro* ‘if its like that’; directional *mri* ‘come up’; degree *fade* ‘very/big’; manner *fissiet* ‘fast’; time *rre* ‘today’)

Closed classes

Proforms:

pronouns (personal: *ja* ‘I’; reflexive; reciprocal; demonstrative; indefinite; relative; interrogative)

pro-adverbs (*neno* ‘where’; temporal noun + *nya* ‘when’)

pro-verbs (*dir* ‘it is good...’)

Noun adjuncts:

role markers

discourse markers (contrast markers); limiting adjectives

preposition (*nene* ‘to’)

quantifiers (quantities: *timbe* ‘all’; numerals: *tor* ‘three’)

classifiers (*fu* ‘fruit’)

articles/determiners (*adei* ‘the’)

definite/indefinite

demonstrative (limiting)

Verb adjuncts:

auxiliaries (TAM *du* ‘already’, *ga* ‘I want’, polarity = *ti* ‘no/not’)

Conjunctions:

coordinating (*-ir* ‘and’, *te* ‘or’, *temo* ‘but’)

subordinating (*nirge* ‘because’, *we* ‘although’, *ge* ‘that’, *temo* ‘whether’)

complementizers (*ge* ‘that’)

relativizers (*uf* ‘that’, *fî* ‘which’)

adverbializers (after...)

Other:

clitics (enclitic mood markers, negator and imperative, and politeness markers)

copulas/predicators (*rau* ‘copula’)

interjections (*woyi* ‘oh!’)

TABLE 3.1 – Syntactic categories in Irarutu with examples, based on Schachter 1985.

3.1.1 Word formation processes

As a result of the cluster of features approach (Schachter 1985) adopted here, instances of complete overlap between categories such as noun (e.g., *figrgir* ‘(a) story’) and verb (e.g., *figrgir* ‘to (talk) story’) are seen as evidence that these may actually be subcategories of a single larger category of morpheme that are sometimes deployed as nouns and sometimes as verbs. Several morphological processes: reduplication, compounding, borrowing, and relexification, are discussed in greater detail in §3.1.1.1-3. Various types of Irarutu bound morphology are listed in TABLE 3.2.

Bound morphology

Prefix (alienable possessive *a-* ‘my’; verbalizer *f-* ‘Active’; agreement marker *n-* ‘he’)

Circumfix (inalienable possession *a-* *-g* ‘my (inalienable)’)

Suffix (demonstratives *-mai* ‘Distal’)

Reduplication (full reduplication *rabi-rabi* ‘k.o. self-protection knife’; long prefixed partial reduplication *kr~kor* ‘wind (rope)’, short prefixed partial reduplication *r~re* ‘day’, suffixed partial reduplication *dbe~be* ‘hit repeatedly’)

Precategorical root (*-gir* ‘story’)

Frozen morphology (*-rie* in *fta-rie* ‘butt (stomach-posterior?)’)

Clitic (negative marker *=ti*; topic marker *=ro*, politeness marker *=e/o*)

TABLE 3.2 – Varieties of Iraputu bound morphology. Equals sign ‘=’ indicates a clitic rather than an affix, which is indicated with a hyphen ‘-’.

Roots can occur independently (preposition *nene* ‘to, at, in, on’) or can be modified by bound morphology (noun *a-mtie* ‘my eye’). Iraputu has both prefixes and suffixes. At least one construction, i.e., inalienable possession, appears to be a circumfix. Of the three types of affixes, prefixation is predominant. It has the most complexity, and occurs most frequently. Not all bound morphemes are affixes. There are some precategorical roots (Broschart 2000), and frozen morphology in compounds, which never occur alone. Furthermore, numerous nouns, especially body part terms are always obligatorily possessed, and many verbs are obligatorily inflected for their subject. There are two important special clitics, i.e., grammatical elements that adjoin to units larger than words. The enclitic *=ro* occurs with NP topics, whereas the negative enclitic *=ti* follows the last element in a clause. Word boundaries are especially blurry within the verb complex. For example, native speaker intuitions in representing word boundaries vary widely. It is difficult to determine where these elements, especially TMA markers, are located along the cline from free morpheme to affix. Two additional word-formation processes are relexification, also called semantic shift (e.g., PCEMP **qinəp* ‘lie down’ > IRH *m* ‘sleep’), and borrowing (e.g., Arabic *hukum* > Indonesian *hukum* > IRH *hukum* ‘justice/judgment’).

3.1.1.1 Reduplication

Reduplication is a fairly productive word formation process and is observed with a wide range of syntactic categories, including nouns, verbs, adjectives, adverbs, and prepositions. Reduplication does have a range of semantic values for each of the categories, but it does not specify any particular part-of-speech. Three types of reduplication were identified according to morphophonological grounds in Chapter 2: full reduplication, two types of prefixed partial reduplication, and four types of suffixed partial reduplication. One type of prefixed partial reduplication was argued to be motivated solely by phonological conditions on word shape. Examples of the other three types of reduplication are provided in 1)–3).

Full reduplication

- 1a) *tim~tim*
PASS~close
'closed'
- 1b) *met~met*
INCREMENTAL~little
'little-by-little'
- 1c) *rabi~rabi*
NMLZ~unknown root
'k.o. knife'

Prefixed partial reduplication

- 2a) *br~bar*
PL~lung
'lungs'
- 2b) *fr~fier*
TRANSITIVIZER~say
'answer'

Suffixed partial reduplication

- 3a) *esuem~em*
one~VARIEGATION
'various (ones)'
- 3b) *dbe-be*
hit~REPETITION
'hit repeatedly'

The semantic contribution of reduplication in the examples above includes: a (passive-like) agentless construction (1a), incremental (1b), nominalization (1c), plurality (2a), transitivity (2b), variegation (3a), and repetition (3b).

Prefixed partial reduplication is particularly productive in verbs. About one in fifty lexical items (2% of the vocabulary) from a 1000+ item word-list demonstrates this type of reduplication. Suffixed partial reduplication occurs in only about one in a hundred and fifty items (<1%). On account of the fact that Irarutu tends to have monosyllabic roots, the existence of disyllabic roots that look like reduplicated CV monosyllables, but that do not have a meaning of their own, suggests that a small set of lexical items were formed by a type of prelexical reduplication, such as *gege* 'ladder' and *toto* 'purity'. This type of reduplication occurs once in about eighty words (1.5%). Reduplication applies before other bound morphology: *n-fr-fier* 'he speaks', **nf-nfier*.

3.1.1.2 Compounding

Compounding is not limited to nouns, but it does play a role in nominalization. Compounding is merger of two roots through adjunction and it results in a form that is distinct from its parts in phonology and semantics; however, due to the location and lexical distribution of possessive morphology, compounding also appears to involve

stems. The most frequent parts-of-speech observed in compounds are noun plus noun (N+N), although noun plus adjective (N+Adj), or some other element (see discussion of numerals in §3.1.3.2), can also be compounded. Words can also be formed through crystallization. Irarutu has ‘frozen morphology’ (FM), that is, morphemes that do not (any longer) occur on their own and do not have clear semantics.

- 4) *san* + 'rimta
house + mouth
‘door’
- 5) *tgra* + 'garn
ear + inside
‘earhole’
- 6) *mu* + 'mtie
FM+ eye
‘morning’

Compounding is a productive process that is frequently used to improvise a word when a speaker does not know or remember a more appropriate term for a concept. This is especially common among younger speakers, but it is also observed in older speakers. The compound *san+rimta* ‘door (lit. house+mouth)’ can be substituted for *frrufn* ‘door’. Examples of compounds are provided in **TABLE 3.3**.

Noun + Noun

<i>kad+fa</i>	‘pants’ (lit. covering?+leg)’
<i>man+fru</i>	‘feathers (bird+hair/feather)’
<i>rguin/-tbon+fru</i>	‘headhair (head+hair/feather)’
<i>fra+fu</i>	‘shoulder (arm+boney protrusion)’
<i>-wi(-)+fu(-)</i>	‘knee (leg(frozen morpheme?)+boney protrusion)’
<i>mo/matu+kapri</i>	‘ship hand (child/person+ship)’
<i>rre+mce</i>	‘clock (day+eye)’
<i>wer+mta</i>	‘spring of water (water+eye)’
<i>-mta-+re</i>	‘cheek (eye+frozen morpheme(?))’
<i>-fta-+rie</i>	‘buttocks (stomach+end portion(?))’
<i>-rfu-+riri</i>	‘tooth (tooth+seed)’ ~ dialectal <i>-rfu-+ru</i>
<i>e+rit</i>	‘cloth/bark’ (tree+skin)’
<i>-rimta-+rit</i>	‘lip (mouth+skin)’
<i>-mte-+rue</i>	‘tears (eye+liquid)’
<i>-mta-+rum(-)</i>	‘face (eye+house)’ NB: <i>rum</i> ‘living place, i.e., house’
<i>tni+rum</i>	‘feature’ (body+house)
<i>e+ro</i>	‘leaf’ (tree+leaf)
<i>-fra+si(-)</i>	‘elbow (arm+crook/bend(frozen morpheme?))’
<i>fi+tbi</i>	‘drinking glass (thing+drinking vessel)’
<i>mce+wer</i>	‘goggle-eyed (eye+water)’

TABLE 3.3 – Examples of compound nouns. A hyphen indicates the site where possessive morphology occurs.

The relationship between elements in a compound can be based on possession (*mta+wer* ‘tear (water of eye, lit. eye’s water)’, attribution (*mo+sot* ‘daughter (lit. child female)’), or composition through metaphorical extension of one element (*wer+mta* ‘spring (of water)’ i.e., ‘eye of water’, lit. ‘water eye’).

There are instances of multi-root compounds. Some examples are given in 7)–9). The internal structure of these forms shows asymmetry through binary branching (Kayne 1994), where the first two elements are merged before the subsequent element, as indicated by square brackets.

- 7) [*ira+rü*]+*tü*
 3P+voice+true
 ‘their true language’
- 8) [*fra+ntu-g*]+*rir*
 [arm+offspring-1SPOSS]+seed
 ‘fingernail’
- 9) [*fta+rie*]+*gur*
 [stomach+end.portion]+orifice
 ‘anus’

3.1.1.3 Borrowing and semantic shift

Two additional dimensions that shape syntactic categories involve borrowing and semantic shift. The technical term ‘borrowing’ is somewhat misleading (cf. Sebba 1997), but it is essentially used to mean that a word from one language is used in another to refer to a new or prestigious concept. The term ‘transfer’ is sometimes used by some linguists for ‘borrowing’ of grammatical resources. This phenomenon occurs for a variety of reasons, most of which are attributable to the nature and duration of contact of speakers of different languages in contact with one another. Semantic shift, which is here used as a cover term including relexification and back formation, is the process of changing the meaning or use of a word and can involve change of syntactic category. Neither borrowing nor semantic shift is exclusive to a particular syntactic category, but together they show how categories interact on an organizational level, how novel items are interpreted, and how some items can float between categories. There are numerous instances of recent loans from Indonesian. The overall quantity of ‘borrowed’ words in Irarutu is rather large.

3.1.2 Open classes

Several properties distinguish the open class parts-of-speech from one another. The open class syntactic categories are: nouns, verbs, adjectives, and adverbs.

3.1.2.1 Nouns

According to the traditional notional definition, a noun is ‘a name for a person, place, or thing’. Based on grammatical properties, nouns can be distinguished by possessive morphology, may co-occur with a determiner and other adjuncts, and function syntactically as arguments (subject or object) or the head of an argument (Schachter 1985). Frozen morphology is observed in Irarutu inalienable possessive constructions

involving lexicalized compound noun phrases (see also van den Berg and Matsumura 2008). A noun may also function predicatively in an equational construction 10).

- 10) *matu (rau) guru*
 person (COP) teacher
 ‘that person is a teacher’

Typical categories that nouns are specified for, either morphologically or having an adjoined word, are number (single: plural), class (classifiers), and definiteness (*fne* ‘pig’ vs. *fne ad* ‘the pig’). There is no case marking in Irarutu. Subclasses are common versus proper. Proper nouns for specific human referents can occur with the vocative morpheme *-o* (or variant *-a*).

3.1.2.1.1 Function of nouns

Andrews (1985) identifies three major functions of noun phrases. These are: semantic function, pragmatic function, and grammatical function. Grammatical function, i.e., the semiotic (semantic plus pragmatic) function of a noun as it relates to a predicate, is distinguished for the purpose of description, from grammatical relation, which is involved in determining the syntactic form of any given construction. Three broad grammatical functions, as defined above, are: core, oblique, and external. Grammatical relations, which are discussed further in §3.2.1, are only concerned with the core functions S (subject of intransitive), A (agent of transitive), and O (object of transitive).

Grammatical relations and semantic roles, also called thematic roles (θ -roles), have been shown cross-linguistically to be closely related (Bickel 2011). Thematic roles concern what arguments – and non-arguments – are selected by a particular predicate.

Thematic roles that take part in the description below are: agent, patient/theme, dative-benefactive (selected by ‘give, bring’ etc.), associative/reciprocal, instrumental (selected by ‘cut (with), hit (with)’, etc.), associative/reciprocal (selected by ‘kiss, meet’, etc.), locative, time, duration, repetition, frequency, cause, purpose/intent, source, and goal. Nouns with agent and patient/theme semantic functions are the most prominent and frequently correspond with A, S, and O. Givón (1990) also points out that dative-experiencers (selected by ‘listen (to)’, ‘look (at)’, ‘think (about)’, etc.) are also prominent cross-linguistically. This is true in Irarutu. Examples of theme, agent, patient, dative-experiencer, and dative-benefactive are given below.

- 11) [*o-mo adi=ro*]_{theme} *dridn*
 2S-child DET=TOP sick
 ‘your child is sick’
- 12) [*matu*]_{agent} *n-fruet* [*sum*]_{patient}
 person 3-search fish
 ‘the person is fishing’
- 13) [*mo adi*]_{DAT-EXP} *n-fi-drbun* [*finnar adej*]_{theme} *rau dir*
 child DET 3-ACT-think joke DET COP good
 ‘the child thinks the joke is funny’

- 14) [matu]_{agent} *n-fi-grgir* [snan igris]_{patient} *na-ne* [i-mo]_{DAT-BEN}
 person 3-ACT-give words English 3-give.to 3S-child
 ‘the person teaches English to her child’

Oblique and external thematic roles, including location, time, source, and goal, can be introduced by a role marker, such as a preposition.

- 15) [nyafi=ro]_{theme} [nene san rimta]_{location}
 gift=TOP PREP house door
 ‘the gift is in front of the house’
- 16) [rre admi]_{time} [nof nabad]_{cause} *n-tim* [frrufn adi]_{patient}
 day DET wind big 3-close door DET
 ‘today a big gust of wind closed the door’
- 17) [ja]_{agent} *ma* [nene refefa]_{time}
 1S come PREP yesterday
 ‘I arrived yesterday’
- 18) [ja]_{agent} *bu mi* [pasar]_{source} *ma*
 1S TAM stay market come
 ‘I came back from the market’

Instrumentals can be adjoined with a preposition such as *nir* ‘with’, as in ‘I split wood with an axe’, but occur more frequently in naturally occurring language data in a core role created by switch reference, where the Object of first verb is the Subject of the second verb, 20–21).

- 19) *ja tar ema* [nir suri]
 1S split wood with axe
 ‘I split wood with an axe’
- 20) [i]_A *n-ar* [waranda]_{INS} *na-n-tut* [sum]_P
 3S 3-take knife INF-3-cut fish
 ‘he uses a knife to cut fish’
- 21) [ja]_A Ø [suri]_{INS} *ra-tar* [e+ma]_P
 1S take axe 1-split dry+wood
 ‘I axe-split the firewood’

3.1.2.1.2 Noun Morphology

Nouns and only nouns can host possessive morphology. A subset of nouns occurs with the nominalizing prefix *wen-*, and proper nouns can take a vocative suffix.

3.1.2.1.2.1 Possessive morphology

Nouns are the only syntactic category that accept possessive morphology. There are two basic types of morphological possession: alienable and inalienable. Alienable possessive morphology is less complex than inalienable possessive morphology. The following three examples show alienable genitive morphemes prefixed to a possessed noun.

- 21) *a- jie*
1S.POSS-father
'my father'
- 22) *o- deno*
2S.POSS-mother
'your mother'
- 23) *i- san*
3S.POSS-house
'her house'

The use of the alienable possessive prefixes is extensive; but, there is also inalienable possession, which is expressed through a circumfix, i.e., co-occurring prefix and suffix, for speech act participants, first and second persons. Examples are provided in 24)–26). There is no third person possessive suffix, meaning the alienable-inalienable relationship is not marked for third person possessive. Therefore the type of possessive relationship must be read off of the behavior of the first or second person possessive marking of the same noun, some of which show exceptions such as where the first person possessive is marked as inalienable but the second person possessive is alienable.

- 24) *a- rui -g*
1S.POSS-elder.sibling-1.POSS
'my elder sibling'
- 25) *o- rui -m*
2S.POSS-elder.sibling-2.POSS
'your elder sibling'
- 26) *i- rui*
3S.POSS-elder.sibling
'her elder sibling'

On account of the fact that inalienable nouns in Irarutu are limited to body parts (such as *-mtie* 'eye'), some kin terms (*-rui-* 'elder opposite gender sibling'), and the word *nü* 'name', it appears that alienable possession is the default case, a subset of which needs to be obligatorily possessed, such as 'eye' and 'elder sibling', but are not inalienable. Categorization of possessive relationships, especially within the elaborate kinship system, as alienable or inalienable, is idiosyncratic in Irarutu. Compare 21–22) with 25) or 26). Furthermore, the circumfix and close kin relationship in 24) and 25) suggest the same relationship may be expressed by a parallel construction (a circumfix) in 26), but that example follows instead the prefix only pattern, from 21)–23), but has a semantically closer kin relationship (parent vs. elder sibling).

3.1.2.1.2.2 Nominalizer *we(n)-*

The morpheme *we(n)-* is infrequent but it does occur in a handful of forms: *wen-ga* 'food', *we-dud* 'work', *we-num* 'land', *wentrur* 'loss (decrease)', *ma wen* 'welcome', *nanrun wen* 'know (familiar)', and possibly *wen-i* 'plural determiner'. As seen with *wen-ga* 'food', the nominalizer *we(n)-* converts a verb root *ga* 'eat' into a noun.

3.1.2.1.2.3 Vocative -o/a

Another type of morphology that occurs with nouns is the vocative suffix *-o/-a*. This morpheme only occurs with a proper noun that uniquely identifies a human referent and functions as an alternative to using a second person pronoun when addressing someone. It is also used in contexts where the speaker addresses a specific person, possibly from a set of referents of known size, which causes that particular person to be drawn into a speech act event as a participant. In this regard, the vocative is an alternative to using a second personal pronoun,

- 27) *Arnol-o / Arnol-a*
 Arnol-VOC
 ‘Arnold (I’m talking to you)’, ‘Hey, Arnold’

3.1.2.1.3 Distribution of Nouns

Nouns function as arguments and non-arguments, or the head of arguments and non-arguments. The clause-level distribution of nouns as arguments and non-arguments of verbs, as well as the phrase-level distribution of nouns in relation to the elements that modify head nouns, are described below.

3.1.2.1.3.1 Clause-level distribution of nouns

At the clause level, nouns occur with two basic types of verbs. Monadic verbs select one argument. Dyadic verbs select two arguments. In constructions with monadic verbs, the single argument precedes the verb. This is SV order.

- 28) *Jek i(=ro) n-in*
 Jack 3(=TOP) 3-sleep
 ‘Jack is sleeping’
29) *Jek n-tür*
 Jack 3-sit
 ‘Jack is sitting’
30) **n-tür Jek*
 3-sit Jack
 ‘Jack is sitting’

However, there are two types of monadic verbs: unergatives and unaccusatives, discussed further in §3.1.2.2.1. By and large, these two types pattern the same: SV. A limited number of unaccusatives, particularly those that select an argument that undergoes some type of change, action, or state, i.e., theme or patient, allow that argument to be post-verbal, VS, reflecting a difference in structure from unergatives, which typically have volitional subjects.

- 31) *ja m-tit*
 1S STAT-fall
 ‘I fell’

- 32) *m-tit ja*
STAT-fall 1S
'I fell'
- 33) *gerfar it*
return 1Pn
'we have returned'

In constructions with two arguments, the subject (S), which is usually an agentive argument, precedes the verb, whereas the object (O), which is usually a patient/theme follows the verb: SVO (Givón 1990).

- 34) *matu (i) (n-)fruet sum*
person 3S 3-search fish
'the person catches fish'
- 35) *Jek n-the fun*
Jack 3-hit dog
'Jack hit the dog'

3.1.2.1.3.2 Phrase-level distribution of nouns

Nouns usually precede their dependents, including deictic determiners and demonstratives, which indicate the discourse or spatial status of a referent. Importantly, there is one construction – the possessive – where the head noun follows its dependent. Evidence shows that the Irarutu noun phrase has the structure shown in **FIGURE 3.1**, below.

(POSS) **Noun** (S_{REL}/ADJ (DEG)) (CLF) (NUM) (QUANT) (DET)

FIGURE 3.1 – Template of the Irarutu noun phrase.

3.1.2.1.3.2.1 N precedes adjunct

Nouns can be specified in several ways, such as by adjoining a determiner, classifier, numeral, quantifier, adjective, relativized clause, conjunction and another noun, or various combinations of these elements. The noun precedes those adjuncts:

- 36) *termus [ad]*
thermos DET
'the thermos'
- 37) *matu [tni]*
person CLF
'people'
- 38) *san [tuen]*
house long.time
'old camp'

- 39) *mo* [ru]
child two
'two children'
- 40) *matu* [nabad]
person old
'big person'
- 41) *mtarum* [dir sfar]
face good very
'beauty'
- 42) *a-matu* [uf n-tim frufn adi]
my-friend REL 3-close door DET
'my friend who closed the door'
- 43) *fne* [nir fun]
pig CONJ dog
'pigs and dogs'
- 44) *umagi fu tur adini*
coconut CLF NUM DET
'these three coconut [fruits]'

3.1.2.1.3.2.2 N follows adjunct

In contrast to the general pattern where the noun precedes the adjunct, possessive constructions place the noun after its adjunct. The syntactic possessive occurs frequently, and is paralleled in compounds (*rguin fru* 'head's hair'), and morphological possession (*a-nefut* 'my-little brother').

- 45) [mo] rü
child voice
'child's voice'
- 46) [matu] (i-) san
person 3S.POSS house
'person's house'
- 47) [a-jie] (i-) san
1S.POSS-father 3S house
'my father's house'

3.1.2.2 Verbs

Following Schachter (1985), the notional definition of verbs is to 'express actions, processes, states, and the like'. Verbs characteristically function as predicates. Non-verbal predicates include adjectives, locatives, commentative nouns (nominal predicates/nominal predicate complements), and so forth. In Irarutu, there is an optional copula, *rau*, that occurs with predicate nominals or adjectives, but not with verbs. There is no distinct class of verbal nouns (e.g., '[his eating] in the middle of the night is unusual'), but there are serial verb constructions (for more on the topic, see Bradshaw 1993; and Crowley 1987, 2002) and incorporation. Categories that can be indicated for verbs are tense (time relative to time of utterance), aspect (complete or incomplete action), mood (indicative vs. subjunctive; declarative vs. interrogative), and polarity

(affirmative vs. negative). The Irarutu verb is also typically marked to indicate the person of its subject. The two major subclasses of verbs are transitives, a major property of which is that they occur with an object, and intransitives, which have no object. Intransitive verbs are further divided into unergatives and unaccusatives. Constructions that are treated as ditransitives in other languages, where a transitive verb selects two objects, are accomplished by a serial verb construction in Irarutu.

3.1.2.2.1 Function of verbs

Verbs function as predicates and include intransitive verbs and transitive verbs. Transitivity is discussed in §3.2.3. According to Perlmutter and Postal (1983), intransitive predicates fall into two classes, unergative or unaccusative. Unergatives denote willed or volitional acts ('run, talk, retire'), manner of speaking verbs, and involuntary bodily processes. In general, unergatives have a subject that is semantically perceived as actively initiating or being responsible for the action expressed by the verb. Directional verbs such as *mri* 'come in', *bri* 'come up', *iet* 'go up' etc. are a noteworthy subclass of unergatives (see also Matsumura and Matsumura 1991 for a description of directionals in E. Arguni dialect). In contrast, unaccusatives tend to be: adjectival predicates, predicates with an undergoer subject, inchoatives, existential/occurrence predicates, verbs that refer to phenomena that impinge on the senses, aspectual predicates, and duratives. In general these are constructions where the single syntactic argument is not a semantic agent for the action of the verb, expressed for example, by treating it the same as the accusative argument of a transitive verb.

Prototypical transitive verbs have objects which are created (*nyuen dub* 'weave a basket'), destroyed objects (demolish a house, smash a glass), physical change in the object (*tot ema* 'chop the firewood'), change in object's location (*mud mama re* (you-hold you-take-you-come leaf) 'pass the scratchy leaf'), change by means of an implied instrument (*tar ema* 'split the wood, i.e., cut with axe'), surface change (*sukur kke* 'shave the jaw'), internal change (*fun werwin* 'make hot water'), and change with implied manner (*dru* 'shoot with arrow').

3.1.2.2.2 Verb Morphology

Several affixes occur with verbs, including subject agreement prefixes, *ta-* 'Incidental stative', *m-* 'Stative', *fi-* 'Active/Causative marker', and *na-* 'Habitual/Inchoative', as well as two types of suffixes, *-fe* 'Detransitivizer' and a set of pronominal object markers.

The relative order of these elements is presented as a template in **FIGURE 3.2**, below. Combinations of basic verbalizing morphology fill a range of grammatical functions. Some sequences of prefixes have idiosyncratic meanings. Three examples are *na-n-* 'Gerundive/suppositional infinitive', *f-na-* 'regarding/like', and *n-fi-* 'Causative/Inchoative'.

HAB-AGR-ACT(=)AGR-INCI-Verb-DTR

FIGURE 3.2 – Template for Irarutu verbs. Reduplication occurs adjacent to the verb and is complementary to AGR

A notable characteristic of the verb is the occasional repetition of subject markers, represented as ‘AGR’ for agreement. Some auxiliaries also host the subject markers.

3.1.2.2.2.1 Subject marking on the verb

Within the verb complex and across clauses, the most frequent inflectional morpheme is the subject marker, illustrated in **FIGURE 3.3**. It agrees in person with the subject of the clause and is complementary with the stative marker *ta-* as well as full reduplication.

1 st person	<i>a-</i>
2 nd person	<i>m-</i>
3 rd person	<i>n-</i>

FIGURE 3.3 – Irarutu subject markers. The first person subject marker is seldom used in main clauses.

Each of the three grammatical person features behaves somewhat differently. The first person is perhaps the hardest to analyze. In several cases, first person agent referents are represented as *a-* without an independent first person pronoun such as *ja* ‘1S’. It is due to this fact that although it is listed in **FIGURE 3.3** as *a-*, in main clauses it tends to surface as \emptyset , allowing the independent pronoun form *ja* to contract to the verb (/ja a-frro/ ‘I I-run’ becomes *ja frro* and finally *ja=frro*). The distribution of these elements, *ja* and *a-*, appears to be determined by discourse constraints. However, some lexical items are more likely to occur with a zero marker than *a-*, depending, for example, on the initial segment (*a-a...* is dispreferred, see Chapter 2). Furthermore, the independent pronoun *ja* and an inflectional *ja-* never co-occur. The combination of *ja a-* in main clauses is considered awkward, even ill-formed; but, *ja=ro a-* is acceptable (though rare), depending on the context. The use of *ja* ‘1S’ in coordinate and subordinate clauses is considered redundant, however, subordinate clauses often but not always contain *a-*. To present a condensed view of relevant subject marking phenomena in four different settings, syntax data is organized according to the pronoun paradigm.

The second person prefix *m-* is the most regularly distributed. Only in some imperatives and with a few exceptional roots is it omitted. The second person prefix co-occurs with the independent pronouns *o* ‘2S’ and *e* ‘2P’, which may attach to the first element of the verb complex or stand alone. The topic marker frequently occurs with second person independent pronouns, but never with inflectional prefixes. The presence of second person inflection can even be determined phonologically in some cases where it seems to be absent, such as before *f-* initial roots, where /m-/ + /f/-initial stem becomes a *b*-initial stem [ᵐb] (see **TABLE 3.4** and Chapter 2). The agreement prefix *m-* also occurs regularly in coordinate and subordinate clauses.

The third person agent marker, *n-*, occurs with most verb roots in main clauses; however, there are several examples that do not have this subject marker. Some of these exceptions can be explained as allomorphy, but others cannot. For roots that host the third person prefix, it can be used in subordinate clauses for clarity. Not only does the

independent pronominal subject co-occur with the inflectional prefix, but it is also free to host the topic marker. If it is not hosting the topic clitic =*ro*, an independent pronoun may attach to the verb. Examples are given in TABLE 3.4, above.

No inflection for adjectival predicates or copula

<i>ja bitr</i> ‘I am hungry’	<i>it bitr</i> ‘we (including addressee) are hungry’
<i>o bitr</i> ‘you are hungry’	<i>am bitr</i> ‘we (excluding addressee) are hungry’
<i>i bitr</i> ‘he is hungry’	<i>e bitr</i> ‘you (plural) are hungry’
	<i>ir bitr</i> ‘they are hungry’

Subject Inflection with monovalent verb

<i>ja Ø-ud</i> ‘I hold’	<i>it Ø-ud</i> ‘we (inclusive) hold’
<i>(o) m-ud</i> ‘you hold’	<i>am Ø-ud</i> ‘we (exclusive) hold’
<i>i n-ud</i> ‘he holds’	<i>e m-ud</i> ‘you (plural) hold’
	<i>ir n-ud</i> ‘they hold’

Subject Inflection with detransitivized verb

<i>ja Ø-tbe-fe</i> ‘I hit something’	<i>it=ro Ø-tbe-fe</i> ‘we (inclusive) hit something’
<i>o=ro m-tbe-fe</i> ‘you hit something’	<i>am Ø-tbe-fe</i> ‘we (exclusive) hit something’
<i>i=ro dbe-fe</i> ‘he hit something’	<i>e=ro m-tbe-fe</i> ‘you (plural) hit something’
	<i>ir=ro dbe-fe</i> ‘they hit something’

Subject Inflection with transitive verb

<i>ja Ø-fun imel</i> ‘I am doing email’	<i>it=ro Ø-fun imel</i> ‘we (inclusive) are doing email’
<i>o=ro bun imel</i> ‘you are doing email’	<i>am=ro Ø-fun imel</i> ‘we (exclusive) are doing email’
<i>i n-fun imel</i> ‘he is doing email’	<i>e=ro bun imel</i> ‘you (plural) are doing email’
	<i>ir=ro n-fun imel</i> ‘they are doing email’

TABLE 3.4 – Inflectional paradigms for several example verbs. When the third person subject marker /n-/ is added to a verb that starts with /t/, the result is a verb that starts with *d* [ʰd], whereas when the second person subject marker /m-/ is added to a verb that starts with /f/, the result is a verb that starts with *b* [ʰb] (see Chapter 2). The morpheme =*ro* is a topic marker.

Van den Berg and Matsumura (2008) comment that agreement prefixes and free pronouns could also be analyzed as proclitics. The evidence cited above supports the analysis that the pronominal agreement markers are bound and selectional, whereas the independent pronouns can contract to the verb under certain circumstances.

A handful of verbs, such as *-in* ‘drink’, *-in* ‘sleep’, *-a* ‘eat’, constitute a subclass of subject marking. These very frequent predicates mark speech-act-participant subjects with *g-* but non-speech-act-participant subjects with *n-*, as shown in FIGURE 3.4.

<i>g-in</i>	‘I drink/you drink’
<i>n-in</i>	‘(s)he drinks’
<i>g-m</i>	‘I sleep/you sleep’
<i>n-m</i>	‘(s)he sleeps’
<i>ga</i>	‘I eat/you eat’
<i>na</i>	‘(s)he eats’

FIGURE 3.4 – Irregular verb subject marking differentiates speech-act-participants from non-speech-act-participants.

The subject marker *n-*, can function as a generalized agreement marker (AGR) to indicate a generic subject, which often results in an interpretation along the lines of ‘one does...’ or ‘you...’ (lit. ‘he/she/it does...’). This property is particularly common when *n-* precedes the active marker/relativizer *f(i)-*, examples of which are given in 49)–50).

- 48) *n-m-tür* *ffu*
 AGR-STAT-sit top
 ‘win (lit. he is in the state of sitting on top)’
- 49) *na-f-tür*
 HAB-REL-sit
 ‘indication (lit. the thing that regularly sits)’
- 50) *na-n-f-tür*
 HAB-AGR-ACT-sit
 ‘to indicate (lit. it is actively sitting)’

The pair of examples 49)–50) differs in possessing or lacking the subject agreement morpheme *n-* and the function of *f-* as a relativizer or an active verb marker. These differences are reflected in the interpretation of 49) *na-f-tür* as ‘indication’ and 50) *na-n-f-tür* as ‘to indicate (something)’.

3.1.2.2.2 Incidental stative *t(a)-*

The morpheme *ta-* (contractable to *t-*) is an incidental stative marker that reduces transitivity by deflecting attention away from the agent and connoting a sense of ‘later’ by indicating the action is not fully realized or complete. Its semantic specification and usage are somewhat similar to ‘stative’ reduplication, focusing on a state rather than the subject. Due to its role as a demotion device, *t(a)-* is translated effectively in some constructions as a passive. The morpheme *ta* can contrast with desiderative *g-*, as shown in 52)–53).

- 51) *o=ro* *t-bsufn* *t-fyer* *fi*
 2S=TOP INCI-study INCI-say things
 ‘You study how to speak (Irarutu)’

- 52) *ja ta a-gin=ti*
 1S INCI 1-SAP.drink=NEG
 ‘I don’t want to drink yet (i.e., I don’t happen to be thirsty yet)’
- 53) *ja ga a-gin=ti*
 1S DES-1 1-SAP.drink=NEG
 ‘I do not want to drink’

3.1.2.2.2.3 Active *fi-*

Verbalizer *fi-* and relativizer *fi-* are primarily distinguished by context and often attain a simultaneous sense of relativizer (nominalizer) and active marker (verbalizer). As a verbalizer, *fi(-)* is an active marker that contributes both causative and adjectival senses. Indeed, *fi* can mean ‘make/cause’, but this is only part of its range of meaning. A survey of a broader range of contexts, e.g., 54)–60), reinforces the interpretation of *fi* as an active marker. In 54) and 55), the conjugated form of *fi* for a second person subject is *bi* (/m-/ ‘second person subject’ + /fi/ ‘active marker’ → [ʰbi] ‘second person subject with an active verb, see Chapter 2).

- 54) *ba-bi-dir*
 2.go-2.ACT-good
 ‘goodbye (lit. go well)’
- 55) *mi-bi-dir*
 2.stay-2.ACT-good
 ‘goodbye (lit. stay well)’
- 56) *ir nasan n-fi sgffo ir*
 3P RECP 3-ACT kiss 3P
 ‘They are kissing (lit. they kiss each other)’
- 57) *ir nasan n-fe kkuk ir*
 3P RECP 3-ACT box 3P
 ‘They are boxing (lit. they box each other)’
- 58) *ja fi-ar wen-ga ra-ge o-matu uf m-tim frrufn*
 1S ACT-give food 1S-to 2-person REL STAT-close door
 ‘I give food to your friend who closed the door’
- 59) *n-fi gr-gir-fe*
 3-ACT REDUP~tell-DTR
 ‘tell a story’
- 60) *matu adeni i-den=ro buna n-fi-tifn-i fade*
 person DET 3SPOSS-mother=TOP SUPR 3-ACT-love-3S just
 ‘that person’s mother loves her very much’

The roots that appear in these constructions tend to be precatatorial, i.e., not inherently a noun or verb. Consider 61), where it can have either reading, verbal ‘I am same-with her’ or relative ‘I who same-with her’.

- 61) *ja fi sama-nir i*
 1S ACT same-CONJ 3S
 ‘She and I’

3.1.2.2.2.4 Stative *m-*

The stative verb marker is *m-*. It occurs immediately prefixed to the verb, without any intervening element. In examples 62–64), the verbs are overtly marked with *m-*. Example 64) shows a stative verb overtly marked with *n-* ‘third person subject’ agreeing with its third person subject, ‘the child’. The second person subject marker *m-* becomes Ø before *m-* (see Chapter 2), whereas the first person subject marker *a-* is generally omitted in main clauses. Example 65) shows that the stative marker *m-* can optionally occur with the verb *fun* ‘to do’ to add nuance ‘what I am (in the state of) doing (right now)’.

- 62) *ja Ø-m-tür*
1S 1S-STAT-sit
‘I am sitting’
- 63) *te o m-tit*
later 2S 2;STAT-fall
‘You will fall/you will be falling’
- 64) *mo kosi n-m-tit*
child small 3-STAT-fall
‘the child fell/is falling’
- 65) *ja (m-)fun-nya nyati*
1S (STAT-)do-what nothing
‘I am (in the state of) doing nothing.’

Although *m-* is a productive morpheme that, based on context, appears either obligatorily (for stative meaning) or optionally (to add a stative sense), there are several non-stative verbs, whose shape suggests initial *m-* in some verbs is a frozen morpheme. It no longer conveys a stative meaning, see FIGURE 3.5 (and further discussion in Chapter 4). In the cases where *m-* is not segmentable, or no longer segmentable, it is considered lexicalized, i.e., part of the root, as in the words for ‘buy’ and ‘salty’, not a separate morpheme.

obligatory stative

<i>m-tür</i>	‘sit’
<i>m-rur</i>	‘sit cross legged’
<i>m-tugr</i>	‘sit with one leg out’
<i>m-sutr</i>	‘sit holding knees up’
<i>m-tim</i>	‘close v.’
<i>m-rir</i>	‘stand’

optional stative

<i>(m-)fun</i>	‘do’
----------------	------

non-stative

<i>mgür</i>	‘buy’
<i>mgenir</i>	‘salt(y)’

FIGURE 3.5 – Examples of *m*-initial verbs.

3.1.2.2.2.5 Infinitival/Habitual *na-*

This morpheme has the least grammatical salience for native speakers and the least detectable impact on translation. In these regards, *na-* is difficult to analyze. The meaning as well as the grammatical function of *na* is further obscured by its distribution in a wide range of contexts, such as purposive, infinitival, habitual, suppositional, inchoative, conditional, as well as being attached to several adverbs and clause linkers. Based on some tendencies in glossing, one analysis is to assert that *na-* is a causativizer. However, *na-* can combine with the subject marker *n-* to form *nan-* ‘habitual/inchoative third person subject’, and it also combines with *fi-* to form *fna* ‘like’, without a causative meaning. In a few contexts, *na-* appears to function as an inchoative marker ‘begin/start’. The widest function appears to be habitual/infinitival ‘to be’. The fourth meaning is conditional/suppositional ‘if’.

- 66) *matu na-ge-rofar o*
person HAB-DES.3-mislead 2S
‘the person tricked you’
- 67) *na-mse*
HAB-until
‘until’
- 68) *o ba san na-se nate o byer rad*
2S 2.go village INF-if new 2S 2.say smooth
‘(if) you go to a village then surely you will speak fluently’
- 69) *i=ro n-ar korrar fi na-si i=n-tati?*
3S=TOP 3-use what ACT INF-see 3S=3-REFL
‘he uses a mirror to see himself’
- 70) *ja fier ge m-tbe bnisr n-mse na-mtat*
1S say COMP 2-hit mosquito HAB-until HAB-dead
‘I said to hit the mosquitos until (they are all) dead’
- 71) *na-jiro*
INF-EXPL
‘if it’s like that (lit. to be that way)’

Na- only occurs with the non-speech-act-participant subject marker *n-*, which results in the complex prefix *nan-* ‘habitual/inchoative third person subject’. One verb in particular, *nanrun* ‘think’ nearly always occurs with initial *na-* and conveys a sense of incomplete or durative. There seems to be some significance to the fact that a cognitive verb such as ‘think’ is regularly conjugated with an infinitive/habitual prefix.

- 72) *i=n-ar waranda na-n-tut sum*
3S=3-take knife HAB-3-cut fish
‘he uses a knife to cut fish’
- 73) *warada fi i=nar=ad na-n-tut sum adi=ro bu nsir*
knife REL 3S-3.take=DET HAB-3-cut fish DET-TOP already dirty
‘the knife which she used (it) to cut the fish is dirty’

- 74) *Mikel n-tiin uce na-n-fun tafad*
Michael 3-cook papeda HAB-3-make elder.brother
'Michael cooked papeda for his big brother'
- 75) *Mikel=ro n-tiit wagt mug adi namse na-n-bssi*
Michael=TOP 3S-pound taro QUANT DET until INF-3-smooth
'Michael must pound taro many times so that it becomes smooth'
- 76) *matu n-fi-jual adi=ro na-n-run=ge nyaro andeji=ro du-metyet*
person 3-ACT-sell DET=TO HAB-3-know=that vegetables DET=TOP TAM-fresh
'The vendor knows if the vegetables are fresh'

3.1.2.2.2.6 Verbal suffixes *-fe*, *-i*, and *-nya*

In addition to verbalizing prefixes, there is one suffix that occurs with verbs, detransitivizing *-fe*. It occurs with lexically transitive verbs and decreases the valence of verbs that without it would otherwise require two arguments. The presence of *-fe* prevents the occurrence of an overt object. In free translations of *-fe*, it can be glossed as non-specific definite 'things', roughly correlated with a definite category, or type, but a general, i.e., non-specific, number of entities within that type parallel to English phrases such as 'the *thing* for me is cars'/'cars are my *thing*'. The *-fe* in *ga-fe* 'eat-things' refers to 'things that are eaten together as a meal', which usually includes some type of starch (rice, taro, cassava, papeda), several types of vegetables or fruit (papaya leaves, water spinach, jackfruit, fried bananas), a spicy sauce, and most times some type of protein (fish, tempeh, chicken, pork, or goat). A few examples of verbs that can occur with *-fe* are listed in FIGURE 3.6.

<i>ga</i>	'eat'	<i>ga-fe</i>	'eat-things'
<i>nut</i>	'stab/pierce'	<i>nut-fe</i>	'sew-things'
<i>ntiin</i>	'cook'	<i>ntiin-fe</i>	'cook-things'

FIGURE 3.6 - Examples of detransitivizing suffix *-fe*.

The syntactic function of *-fe* typologically resembles an antipassive, a device that decreases valency and indicates a non-specified object. Transitive verb roots that take *-fe* require only one argument, the subject. Example 77) shows a transitive construction, whereas example 78) shows that the detransitivizing suffix lowers the transitivity of a clause that is built with a semantically transitive verb root 'eat' by deleting the overt object in favor of a morpheme on the verb that indicates a non-specific 'eat-things'. Furthermore, it is ill-formed for an overt object to co-occur with the marker *-fe* (79).

- 77) *it-ga uce*
1Pn-eat papeda
'we eat papeda'
- 78) *it-ga-fe*
1Pn-eat-DTR
'we eat (food)'

- 79) **it-ga-fe uce*
 1Pn-eat-DTR papeda
 ‘we eat papeda’

Similar to TAM markers, the status of *fi*, whether suffix or separate word, is not absolute in the minds of native speakers. For example, *gafi* ‘eat(detransitived)’ was described also as *ga fe* ‘eat things’. The morpheme *-fe* does not occur with a [+human] object, even though this is a possible category that could be specified by various verbs. The verb *ntün-fe* ‘shoot-things’ does not include people, but *ntün-i* ‘shoot him/them’ does. At least one base, *fier* can alternately take the suffix *-fe* (*n-fier-fe* ‘speak (lit. he says things)’ or a third person object marker *-i* (*n-fier-i* ‘lecture (lit. he says it)’). In this context, *-i* represents a definite, third person pronominal object, and can have a human referent. Two more verbs are also observed with the third person object marker, *dbe* ‘hit n./v.’ *dbe-i* ‘hit v.’, *sagwe* ‘breath n.’, *sagwe-i* ‘breathe v.’. However, these verbs correlate with *-fe*, but with a human, i.e., actor, referent. Although both the object marker and the overt argument that it refers to can co-occur, it usually only does so to increase clarity in a complex construction. If the overt argument is non-specific, such as *nyati* ‘nothing’, the verb is marked with the object marker *-nya* ‘(indefinite) something’. Object agreement has not been verified for definite objects, so it is mentioned only in passing.

- 80) *ja Ø-fun-nya nyati*
 1S [AGR_S]-do-AGR_O nothing
 ‘I’m not doing anything.’

3.1.2.2.3 Distribution of Verbs

Verbs function as the predicate of clauses and head the verb complex. Because of these functions, the distribution of verbs at the clause-level and the phrase-level is important in defining verbs.

3.1.2.2.3.1 Clause-level distribution

The clause-level distribution of verbs was foreshadowed in the discussion of nouns in §3.1.2.1.3, above. To recapitulate, monadic verbs, and non-verbal predicates, select a single argument and are distributed primarily in relation to that argument, the subject, most frequently following it (SV), although a minor pattern (VS) is observed for unaccusatives. Monadic verbs and non-verbal predicates are similar with respect to clause-level distribution. Dyadic verbs select two arguments and add the object relation in the order VO.

- 81) *i dbe-i fun*
 3S AGR_S.hit-AGR_O dog
 ‘he hit the dog’

A second VO pattern is also found in double object constructions, which are expressed by a serial verb construction (SVC). A variety of constructions involve two, or more, serialized verbs. In some contexts, serial verbs constitute a single complex, but discontinuous, predicate. These types of elements are considered to be part of the verb

complex, described in the next section. In contrast to conjoined structures, which cross-linguistically tend to occur without an overt conjunction, serial verbs constitute a single predication. The following example shows that verbs tend to be serialized rather than conjoined by an overt coordinator, which – based on the structure of conjoined elements (described in §3.1.3.4) – should occur between *fa pasar* ‘go market’ and *fa mgür* ‘go buy’, but that cannot in this context.

- 82) *Mikel n-ir ja tese (am-)fa pasar fa mgür rimun*
 M. 3-with 1S must (1Pn-)go market [1Pn-]go buy oranges
 ‘Michael and I will go to the market to buy oranges.’

Another construction, the double object construction (expressed as a ditransitive construction in other languages) always marks dative/indirect objects with a special set of inflected serial verbs. At least one, *ge*, is suppletive: *ra-ge* ‘I (give s.t.) to s.o.’, *ma-ge* ‘you (give s.t.) to s.o.’, and *na-ne* ‘She (gives s.t.) to s.o.’. Other verbs in this context are *fun* ‘do’, and *run* ‘know’.

- 83) *[o]_A bi m-ar [wen-ga]_{DO} ma-ge [o-matu uf m-tim frrufn]_{IO}*
 2S ACT 2-give food 2.give-to 2-person REL STAT-close door
 ‘You gave food to your friend who closed the door’

An alternate strategy to express an object with the dative-benefactive role in a double object construction is to relativize the direct object. The relative clause functions to express the semiotic relationships between the arguments, non-arguments, and verbs. In the example below, the relative clause contains a directional verb *fa* ‘go’ and is required in order for the location, here encoded as an oblique, i.e., non-argument, not an IO, to be expressed in the clause:

- 84) *ja tu [surti fi=n-fa a is]*
 1S send letter REL=3-go U.S.
 ‘I sent a letter to the U.S. (lit. I sent a letter which went to the U.S.)’

In a serialized verb construction, the subject of the second verb is usually not expressed overtly if it is identical with the subject of the first verb or auxiliary. If the subject of the second verb is discontinuous with the subject of the first verb, such as being the object of the first verb, that argument can be overt, and the main/second verb agrees with it.

- 85) *Willy=ro n-fir=ge Moses tese n-ir-i mse na-n-fa Fakfak*
 Willy=TOP 3-say=COMP Moses must 3-with-3 until HAB-3-go Fakfak.
 ‘Willy wants Moses to go (with him) to Kaimana’

Matsumura and Matsumura (1991) provide a handful of examples of serial verb constructions that have switch reference, in which the object of the first verb is the subject of the second verb, such as *bi-n-iet* (2.CAUS-3-go.up) ‘you put it up (lit. you do it goes up)’. In this context, *fi-* functions as a causativizer. This phenomena can also be seen in 86), *mbi nfa* ‘you cause it go’.

- 86) *bi wanggt and-i bi n-fa kurni*
 2;CAUSE taro DEF-DET 2;ACT 3-go pot
 ‘put the taro in a pot’

Serial verb constructions often have a causative sense, which may have contributed to the Matsumura’s analysis of *fi*. Example 87), literally ‘come lemon you come’ means ‘cause a lemon to come to me’ or simply ‘bring me a lemon’. The main verb, intransitive *ma* ‘come’ is inflected according to the main pattern of agreement /m-ma/ → *ma* ‘you come’, whereas the second *ma* verb in the serial verb construction takes the SVC minor pattern of agreement marker *ma-*, i.e., *ma-ma* ‘you come’, which supplements the directional information and sublimates the meaning of the predicate. The noun *rimun* ‘lemon’ is treated as an oblique.

- 87) *ma rimun ma-ma*
 come lemon 2-come
 ‘come bring a lemon (you come causing lemon to come here)’

In both double object and directional serial verb constructions, the second verb cannot occur immediately after the first verb. It must occur after the first object/oblique, if any is present.

- 88) *ja bu-mi [pasar] ma*
 1S TAM-stay market come
 ‘I returned (came up) from market’

3.1.2.2.3.2 Phrase-level distribution of verbs

At the phrase-level, verbs are distributed in relation to adverbs and verb adjuncts: auxiliaries and verbal particles. There are a number of auxiliaries that are integrated into the Irarutu verb complex, covering such categories as Tense, Aspect, and Mood, as well as ‘helping’ verbs, some of which are familiar from their role as main verbs, e.g., *fa* ‘go’ and *fun* ‘do’. The verb complex can optionally begin with an adverb, followed by an auxiliary, and the verb plus inflectional morphology occurs last.

Epistemic adverbs, such as *tese* ‘must’ and *bis* ‘can’, generally occur between the subject and the verb (89b). In the following example, the main verb (*mamrun* ‘know/remember’) of the complement clause in 89b) has been deleted under conjunction reduction with 89a).

- 89a) *adi=ro catatan met fnote, sefu o mam-run*
 DET=TOP note little just, so.that 2S 2-know
 ‘This is just a little instruction so that you remember’
 89b) *mo ja ra-run ge tese o bis*
 CONJ 1S 1S-know COMP ADV 2S 2-can
 ‘But, I know that you are surely able.’

An example of a ‘helping verb’ auxiliary, *suk* ‘like’, is shown in 90). Reversing the order of head and auxiliary is ill-formed: **ga suk*.

- 90) *ja suk ga fas taf*
 1S like SAP.eat rice yellow
 ‘I like to eat yellow rice’
- 91) *ja tufiti ga fas taf*
 1S like.NEG SAP.eat rice yellow
 ‘I don’t like to eat yellow rice’

3.1.2.3 Adjectives

The third open class in Irarutu is labeled adjective. Defined according to notion, an adjective identifies words that denote qualities or attributes. Functionally they can either modify nouns, or can be adjective predicate complements in a clause. Irarutu has a subclass of descriptive adjectives for degree (positive, comparative, superlative). The copula in Irarutu optionally occurs with adjectival predicate complements. At the clause-level, predicative adjectives (in contrast with descriptive adjectives) function as unaccusative verbs and follow the subject. At the phrase-level, adjectives can be modified by a following degree word, to specify extent, or they can combine with other adjectives according to various ordering constraints. Some adjectives accept verbal morphology such as the subject marker, but this is rare. Most adjectives do not occur with additional bound morphology. Several categories of adjectives are presented in **FIGURE 3.7**.

An example of an adjective phrase with a degree word, which modifies a noun, is given in 94).

- 92) *bin srwer*
 ground wet
 ‘the ground is wet/wet earth’
- 93) *mo mran*
 child male
 ‘the child is male/son’
- 94) *sansun [nsir nabad]*
 shirt dirty very
 ‘the shirt is very dirty/a very dirty shirt’

<u>Gender:</u>		
<i>sot</i>	‘female’	<i>matu sot</i> ‘woman (lit. person female)’
<i>mran</i>	‘male’	<i>matu mran</i> ‘(gentle)man (lit. person male)’
<u>Appearance (Color terms):</u>		
<i>wams</i>	‘red’	
<i>bfut</i>	‘white’	
<i>grmutn</i>	‘black’	
<i>taf</i>	‘yellow’	
<u>Size:</u>		
<i>kokon/kosi</i>	‘small’	<i>san kokon</i> ‘small house/(the) house is small’
<i>nabad</i>	‘big’	<i>matu nabad</i> ‘judge (lit. person big)’
<i>trro</i>	‘tall’	
<i>frran</i>	‘short’	
<u>Distance:</u>		
<i>nero</i>	‘far’ (<i>ne-iro</i>)?	
<i>nbabr</i>	‘near (close)’	
<i>fumta</i>	‘next’	
<u>Subjective quality:</u>		
<i>dir</i>	‘good/well/nice’	<i>matu dir</i> ‘good person (lit. person good)’
<i>fit</i>	‘strong/brave/evil/bad’	<i>matu fit</i> ‘crook (lit. person bad)’
<i>kabia</i>	‘bad/broken’	<i>radni kabia</i> ‘un-traversable road’
<i>bias</i>	‘usual/ordinary’	<i>matu bias</i> ‘lay person (lit. person usual)’
<i>bunat</i>	‘new/young/fresh’	<i>matu bunat</i> ‘novice (lit. person new)’

FIGURE 3.7 – Some adjective categories in Irarutu, with examples.

As modifiers of a noun, no element from a dissimilar grammatical category can intervene between the head noun and an adjective. Examples are given in 95)–97).

- 95) *wer mud nsir*
water murky dirty
‘dirty murky water’
- 96) *sansun mmat nabad*
shirt green big
‘big (and) green shirt’
- 97) *sansun mmat nabad nsir (fade)*
shirt green big dirty only
‘The big green shirt is dirty’

Example 97) shows that heavy noun phrases, such as a noun plus three adjectives, are possible – albeit rare in naturally occurring data due to a tendency to have compact NPs (reflecting the situation commented on by Foley 2003) – contrary to a claim in the

previous literature (van den Berg and Matsumura 2008). There do appear to be ordering tendencies, such as visual appearance (*mmat* ‘green’) before relative size (*nabad* ‘big’), and relative size before physical condition (*nsir* ‘dirty’). Degree words occur at the end.

- 98) *sansun mmat buna nsir nabad*
 shirt green most dirty big
 ‘a green shirt that is very dirty (IND baju hijau sama sekali kotor sekali)’
- 99) *sansun mmat buna nsir fade*
 shirt green most dirty just
 ‘The green shirt that is very dirty (IND baju hijau sama sekali kotor sekali)’
- 100) [*sansun mmat adi*] *buna nsir fade*
 shirt green DET most dirty just
 ‘The green shirt is really dirty (Indo baju hijau itu sama sekali kotor sekali)’
- 101) **sansun nabad nsir mmat*
 shirt big dirty green
- 102) **sansun mmat [nsir nabad]*
 shirt green [dirty big]

Some of the adjectives listed in **FIGURE 3.7**, can function as degree words if they occur after an adjective. For example, *nabad* can mean ‘big’, or, used as a degree word, it can mean ‘very’.

- 103) *nabad rarur*
 big very
 ‘very big’
- 104) *nsir nabad*
 dirty big/very
 ‘very dirty (a big mess)’

Comparison between adjectives is indicated by the comparative *sfar*, superlative *fi buna* ‘which most’ (*buna* ‘a lot/very much’), or adverbs *sufusr* ‘as much’, *nbergin* ‘enough’.

- 105) *matu adini=ro trro sfar matu adeji*
 person DET=COP tall more person DET
 ‘this person is taller than that person’
- 106) *motor adini=ro dir mo motor refiden mai fi buna dir sfar*
 motorcycle DET=TOP good DISJ motorcycle side DIST ACT most good more
 ‘This motorcycle is better than other motorcycles’
- 107) *i fi buna fit*
 3S REL most strong
 ‘He is the strongest’
- 108) *ad=ro matu tninya mo i fi buna tni trro*
 DET=COP person QUANT DISJ 3S REL most body tall
 ‘there are several people but he is the tallest’

One indication that *f* serves multiple functions is that it precedes habitual/inchoative *na-*, and *na-* precedes the subject marker *n-*, but subject markers precede verb-adjacent *fi-*

- 109) *na-f-na-fe-j*
 HAB-similar-thing-DEM
 ‘about/concerning’

3.1.2.4 Adverbs

Traditionally, adverbs ‘modify verbs, adjectives, or other adverbs by specifying intensity, manner, temporal information, and so forth’. Adverbs tend to occur without affixes in this language. If an adverb does occur with any bound morpheme, it is likely to be the third person subject marker. Functionally, adverbs modify constituents other than nouns, i.e., sentences (to express the speaker’s attitude), verbs (to express time, place, direction, manner, etc.), adjectives, other adverbs, and VPs (to express temporal information and so forth). Five subclasses of adverbs are described below: sentential, directional/location, degree, manner, and temporal. Some adverbs, such as *fade* ‘only/just’, have functions that range over several categories.

		<u>before VP</u>	<u>after VP</u>
<i>bunat</i>	‘newly (recently)’	Yes	Yes
<i>fade</i>	‘just/only’	Yes	Yes
<i>is</i>	‘beforehand’	Yes	Yes
<i>besu</i>	‘only’	Yes	No
<i>namse</i>	‘until/later on’	Yes	No
<i>nase</i>	‘maybe’	Yes	No
<i>nate</i>	‘new’	Yes	No
<i>te</i>	‘later/consequently’	Yes	No
<i>tese</i>	‘must/surely’	Yes	No
<i>fene</i>	‘only’	No	Yes
<i>mse</i>	‘soon’	No	Yes
<i>nbergin</i>	‘enough’	No	Yes

TABLE 3.5 – Examples of Adverbs comments about tendencies in their distribution in relation to VP.

There are no categorizations common to the entire class of adverb. Often no category is specified at all (possible exception: manner adverbs can be specified for degree); however, cross-linguistically, some adverbs can be derived from adjectives. In some languages, adverbs are phonologically distinct from other parts-of-speech, but this is not true in Irarutu. A list of some frequently used adverbs is given in **TABLE 3.5**. It is not clear if *na-* is segmentable or not.

In languages without an open class of adverbs, a phrase consisting of a preposition and a noun or NP can be used to express an adverbial meaning (time, place, direction, manner). This strategy is used in Irarutu to express a variety of adverbial meanings. Furthermore,

expressions involving adjectives can paraphrase certain adverbs. Another strategy that is used cross-linguistically, and is found in Irarutu, is to employ a verb to express a meaning that could be captured by an adverb. This is especially relevant for comparative and superlative constructions. For example, the verb *sfar* means ‘surpass’, and is found in comparatives because Irarutu does not use special affixes on adjectives to express the comparative or the superlative. A final strategy that is relevant to describing how Irarutu handles adverbial information is the serial verb construction, which involves auxiliaries.

3.1.2.4.1 Sentential adverbs

Some adverbs, such as *awesi* ‘suddenly’, *dir* ‘fortunately’, and *fit* ‘unfortunately’ and epistemics meaning ‘probably, maybe, necessarily, most likely’, occur at the edges of a clause in Irarutu. The words *moe* ‘sorry’ and *fade* ‘only’ are examples.

- 110) *ja besu m-tiir fade*
 1S just STAT-sit only
 ‘I’m just sitting’

Clausal adverbs are distinguished by the fact that they take scope over entire clauses. They occur before the subject.

- 111) *te o m-tit*
 later 2S STAT-fall
 ‘later you’ll fall’

Te ‘later/consequently’ is one of a few clause-level adverbs in Irarutu. Another example is *dugag* ‘already’, illustrated in 112).

- 112) *dugag ja=du gin=ti*
 already 1S=already SAP.sleep=NEG
 ‘I have not sleep yet’

The adverb *fade* ‘only/just’ occurs at the very end of a sentence, after the verb, object, and other adverbs.

- 113) *ja m-tiir fade*
 1S STAT-sit just
 ‘I’m just sitting’
 114) *ja ga fa fi tatatu fade*
 1S DES-1 go REL/ACT around just
 ‘I’m just strolling around’

3.1.2.4.2 Directional/Locational adverbs

Some of the work performed by adverbs in other languages is accomplished by directional verbs such as *niet* ‘go up’, *briet* ‘come up’, and *msru* ‘go down’. Prepositions are also differentiated from locative objects *atifro* ‘on top’. Irarutu has a well-developed system for specifying spatial information. Matsumura and Matsumura 1991 make a

similar claim, but base their assertion only upon evidence from directional verbs. Four locational adverbs are: *gan* ‘inside’, *ffu* ‘top’, *tfu* ‘under’, and *tuifa* ‘underneath/beneath’.

- 115) *ge fafu*
PREP top
‘to the top’
- 116) *nene ffu*
PREP top
‘above/at’
- 117) *(nene) gan*
PREP inside
‘inside’
- 118) *wer gweg*
water enveloped
‘In the water’
- 119) *nene tug gan adi*
PREP kitchen inside DET
‘inside the kitchen’

3.1.2.4.3 Degree adverbs

Adjectives that express extent, such as ‘extremely, overly, very, big, more, too, little, much, less, little bit, ease, and quite’ are not distinguished by morphological patterning, but usually occur after the verb. A list of some degree words is given in **FIGURE 3.8**.

<i>dir</i>	‘very (much)’	<i>rarur</i>	‘much’
<i>nabad</i>	‘big’	<i>fusr</i>	‘much’
<i>bidi</i>	‘big’	<i>friemn</i>	‘much (repeatedly)’
<i>s(o)far</i>	‘more (very)’	<i>sufusr ti</i>	‘less (not as much)’
<i>na sfar</i>	‘too’	<i>moru</i>	‘less’
<i>kokon</i>	‘little’	<i>na moru dir</i>	‘less’
<i>kosi</i>	‘little’	<i>n-genaf</i>	‘less’
<i>ntu</i>	‘offspring’	<i>met</i>	‘little bit, half, shortened’
<i>mug</i>	‘many’	<i>bi met</i>	‘ease’
<i>pua</i>	‘many’	<i>fitu</i>	‘quite’

FIGURE 3.8 – Examples of degree words

- 120) *ir fade n-fun kder*
3P just 3-make chair
‘they just made a chair’

The adverb *nbergin* ‘enough’ occurs after the verb. This adverb is notable because it hosts the third person subject marker, and derives as ‘it (is) enough’.

- 121) *ja bu gan n-bergin*
 1S TAM eat 3-enough
 ‘I already ate enough’

Three adverbs are semantically related: *fade* ‘only/just’, *fene* ‘just’, and *besu* ‘only’. *Fade* and *fene* occur after a clause. In addition, *fade* and *besu* can occur between S and the verb.

- 122) *ja ga g-in te (fade)*
 1S DES.1 SAP-drink tea (only)
 ‘I want to drink tea only’
- 123) *ir fade n-fun kder*
 3P only 3-make chair
 ‘they just made a chair’
- 124) *ja besu m-tür fene*
 1S only STAT-sit just
 ‘I’m just sitting’
- 125) *ja besu m-tür fene (ja) fun-nya nyati*
 1S just STAT-sit only (1S) do-what nothing
 ‘I’m just sitting, not doing anything.’

The degree adverb *namse* ‘until/later on’ is illustrated in 126):

- 126) *n-tiit wagt mug adi namse na-n-bssi*
 3S-pound taro QUANT DET until HAB-3-smooth
 ‘he pounds taro until it becomes smooth’

3.1.2.4.4 Manner adverbs

These adverbs indicate the manner, e.g., ‘hard’, or ‘fast’, in which an action or state is achieved.

- 127) *tbe friemn*
 hit repeatedly
 ‘hit repeatedly’

Some adverbial meanings for manner are expressed through periphrastic constructions, such as a preposition and an adjective, ‘come quick’.

- 128) *ma nir ssyet*
 come PREP fast
 ‘come quick (lit. come with speed)’

3.1.2.4.5 Temporal adverbs

Temporal adverbs corresponding with e.g., ‘yesterday’, or ‘tomorrow’, occur in a variety of locations. Example 129) shows that the temporal adverb *rror* ‘tomorrow’ occurs at the

beginning of the sentence and takes scope over the entire clause, both clauses being set in the time-frame of ‘tomorrow’.

- 129) *rror si taje kabia mo Jek=ro tese o=bu-ba*
 tomorrow even.if weather bad DISJ Jack=TOP must 2S=TAM-2.go
 ‘tomorrow seems like it will have bad weather, but, Jack, you must depart’

The word *bunat* ‘new’ functions as an adjective (shown in 130) or an adverb meaning ‘newly/just’ (131). In 131), *bunat* occurs outside the scope of the determiner *adi*, but adjacent to the verb *nbriet* ‘it-rises’.

- 130) *taun porna fi bunat weni (je)-feni n-mi dapur*
 sago bread REL new DET:PL (contents)-REL.DET 3-live kitchen
 ‘the bread which is freshly made is in the kitchen’
 131) *Belandina=ro (bu-)n-fan mo ne(ne) rre adi bunat n-briet*
 Belandina=TOP (TAM)-3-go CONJ PREP sun DET newly 3-rise
 ‘Belandina began walking at sunrise’

The adverb *nate* ‘new’ also contains the morpheme *na-*. It occurs at the beginning of the verb complex, as illustrated in 132).

- 132) *o ba san nase nate o byer rad*
 2S 2.come village RES new 2S 2.speak smooth
 ‘(if) you go to (an Irarutu) village then you will newly become a fluent speaker’

The adverb *is* ‘ahead of time/beforehand’ is a temporal adverb. It does not precede the subject. This word is considered to make an utterance more formal, in the sense of grammatically correct. If the optional pronoun does not occur, *is* ‘prior’ contracts to the verb, as in 133).

- 133) *a-sagwe ja is*
 1-breathe 1S prior
 ‘I’m going to rest first’
 134) *a-sagw=is*
 1-breathe=prior
 ‘I’m going to rest first’
 135) *(ja) is=ga-bu*
 1S prior=DES.1-not.yet
 ‘I do not want to (right now)’

3.1.3 Closed syntactic categories

The number of closed parts-of-speech is determined roughly by the degree of synthesis in a language. Irarutu leans towards the isolating end of the spectrum. Therefore, several closed classes are expected. There are five in Irarutu: pro-forms, noun adjuncts, verb adjuncts, conjunctions, and a miscellaneous category.

3.1.3.1 Pro-forms

Proform is a cover-term for several closed classes of words which, depending on circumstance, can be used as a substitute for words belonging to one of the four open classes, or a larger constituent. Irarutu has pronouns, pro-adverbs, and pro-verbs.

3.1.3.1.1 Pronouns

Cross-linguistically several subtypes of pronouns have been identified: personal, reflexive, reciprocal, demonstrative, indefinite, and relative. A key cross-linguistic property of pronominal prefixes is that they usually occur with independent pronouns of matching person and number. This is the case for Irarutu, but only grammatical person is marked. The independent pronoun, like other nouns, can host the topical clitic =*ro*. They can occur in main clauses, but independent pronouns are less frequent in subordinate clauses. For example, the agreement marker *a-*, which is seldom seen in main clauses, is preferred over the independent pronoun *ja* ‘I’ in subordinated clauses.

Pronouns are, for a large part, discursive devices. They can be used in place of a variety of nouns in appropriate contexts. There are several types of pronouns in Irarutu, personal (I, you, he), possessive (my, your, his), reflexive (himself), reciprocal (each other), demonstrative (this/that), indefinite (one, someone, anyone), and relative (who, which). The first person independent subject pronoun is usually identifiable from discourse context. It is often deleted. Subjects are also recoverable to a certain extent based on the agreement marker on the verb.

3.1.3.1.1.1 Personal Pronouns

Irarutu has seven independent personal pronouns marked for person (first, second, third) and number (singular versus plural). The first person plural is further divided in two (inclusive versus exclusive). Free pronouns are illustrated in TABLE 3.6. In some Austronesian languages, the inclusive-exclusive distinction is said to be addressee centered because it takes into account the relationship between the speaker and the addressee (Wilson 2012). This distinction also highlights a contrast between speech act participants’ and non-speech act participants’ degree of involvement in a given proposition. However, Laycock (1977) presents an alternative perspective of inclusive and exclusive first person plural pronouns. He proposed aligning first person pronouns with high rank in the topicalization hierarchy, followed by second and third person pronouns. To contrast with topicalization, he proposed a referentiality hierarchy that gives third person pronouns the highest rank, followed by second then first persons. Based on these two dimensions, Laycock asserted that the inclusive first person plural pronoun (i.e., the first person plus the second person) is the unmarked category, whereas the exclusive first person plural pronoun (i.e., the first person plus the third person) is marked because it combines high ranking topicalization and referentiality. There is no grammatical gender in Irarutu. The two-way number distinction singular : plural can be refined further by suffixing a specific number, *e-ru* ‘you (PL) two’. Case marking is not distinguished in the pronouns. Highly grammaticalized verbs only make a two-way distinction between first and second person (Speech Act Participant, or SAP) versus third person (Non Speech Act Participant, or NSAP) subjects. Subject pronouns are the same form as object pronouns.

	Singular	Plural
1n		<i>it</i>
1x	<i>ja/a</i>	<i>am</i>
2	<i>o</i>	<i>e</i>
3	<i>i</i>	<i>ir</i>

TABLE 3.6 – Irarutu free pronouns.

- 136) *ja g-m*
1S SAP-sleep
'I am sleeping'
- 137) *o m-tbe fun adei*
2S 2-hit dog DET
'You hit the dog'
- 138) *nya n-gor ja*
something 3-bite 1S
'I'm getting bitten/something bit me.'

3.1.3.1.1.2 Possessive pronouns

Two possessive relationships are distinguished in Irarutu: alienable and inalienable. Possession is one grammatical phenomenon where Irarutu treats speech act participants, i.e., first and second persons (Bickel 2011), differently from third person non-speech act participants.

3.1.3.1.1.2.1 Alienable possessive pronouns

The alienable possessive relationship is the major, default pattern. It is achieved by a pronominal possessor prefix on the possessed noun. The pronominal prefixes are shown in TABLE 3.7.

	Singular	Plural
1n		<i>it-</i>
1x	<i>a-</i>	<i>am-</i>
2	<i>o-</i>	<i>e-</i>
3	<i>i-</i>	<i>ir-</i>

TABLE 3.7 – Alienable possessive morphemes.

The possessive pronouns optionally occur in coreference with an independent noun when possession is simultaneously expressed syntactically, as in 139), where 'house' is morphologically and syntactically possessed by 'father', and 'father' is morphologically possessed by the first person singular referent 'me'.

Singular	Plural
1n	<i>it-den</i> ‘our mother’
1x <i>a-den</i> ‘my mother’	<i>am-den</i> ‘our, but not your, mother’
2 <i>o-den</i> ‘your mother’	<i>e-den</i> ‘your (plural) mother’
3 <i>i-den</i> ‘his/her mother’	<i>ir-den</i> ‘their mother’
1n	<i>it skripsi</i> ‘our theses’
1x <i>(j)a skripsi</i> ‘my thesis’	<i>am skripsi</i> ‘our, but not your, theses’
2 <i>o skripsi</i> ‘your thesis’	<i>e skripsi</i> ‘your (plural) theses’
3 <i>i skripsi</i> ‘his thesis’	<i>ir skripsi</i> ‘their theses’

TABLE 3.8 – Paradigms for two alienable possessives, -den ‘mother’ and skripsi ‘thesis’.

- 139) *a-jie* *i-san*
1S.POSS-father 3S.POSS-house
‘my father’s house’

Paradigmatic examples for alienable possession of *-den* ‘mother’ and *skripsi* ‘thesis’ are given in **TABLE 3.8**. In the class represented by ‘mother’, possession is expressed morphologically by a prefix, however in the class represented by ‘thesis’, possession is expressed syntactically by an initial independent personal pronoun.

Even in the emphatic construction *a-den ja-ri* ‘mother of mine (lit. my mother I have)’, the relationship of ‘mother’ is obligatorily marked by a possessive prefix, e.g., *a-* ‘my’.

Irarutu	Gloss	Free translation
<i>a-den ja-ri</i>	1SPoss-mother 1S-have	‘mother of mine’
<i>o-den o-ri</i>	2SPoss-mother 2S-have	‘mother of yours’
<i>i-den i-ri</i>	3SPoss-mother 3S-have	‘mother of hers’
<i>a-den it-ri</i>	1Poss-mother 1Pn-have	‘mother of ours’
<i>a-den am-ri</i>	1Poss-mother 1Px-have	‘mother of ours, not your’
<i>e-den e-ri</i>	2PPoss-mother 2P-have	‘mother of you all’s’
<i>ir-den ir-ri</i>	3PPoss-mother 3P-have	‘mother of theirs’
<i>skripsi ja-ri</i>	thesis 1S-have	‘thesis of mine’
<i>skripsi o-ri</i>	thesis 2S-have	‘thesis of yours’
<i>skripsi i-ri</i>	thesis 3S-have	‘thesis of his’
<i>skripsi it-ri</i>	thesis 1Pn-have	‘thesis of ours’
<i>skripsi am-ri</i>	thesis 1Px-have	‘thesis of ours, not yours’
<i>skripsi e-ri</i>	thesis 1Pn-have	‘thesis of you all’s’
<i>skripsi ir-ri</i>	thesis 1P-have	‘thesis of theirs’

TABLE 3.9 – Two emphatic possessive constructions, using ‘mother’ and ‘thesis’.

The personal pronouns in the relationship to ‘thesis’ are optional in the emphatic possessive construction *(ja) skripsi ja-ri* ‘thesis of mine (lit. (my) thesis I have)’, and

often dropped. The difference between the prefix and the free pronoun is subtle, but it can be seen in the form of the first person singular, *ja/a-*. *A-* is seldom used as a subject marker with main verbs. On the other hand, the first person pronoun *ja* often contracts to the verb. The use of *ja* with *-den*, **ja-den* for ‘my mother’ is not native-like.

The elements *ja-ri* ‘mine’, *o-ri* ‘yours’, etc. are demonstrative possessive pronouns and can stand independently, for example as the answer to a question such as ‘who’s is this?’ An optional relativizer can occur after the head noun, suggesting that the emphatic possessive constructions illustrated in TABLE 3.9 are fundamentally relativized structures with the possessed noun originating as the object of *ri* ‘possess’, i.e., ‘my mother that I have’.

The following pair of examples shows that a syntactic possessive construction such as *ja-ri* in 141) occurs in the usual position for a relative clause in relation to a head noun.

- 140) *ja tese a-ra-fa-n-fa-g* [nirge motor adi kabia]
 1S must 1S.POSS-1-leg-3-go-STEM because motorcycle DET broken
 ‘I must walk because the motorcycle has a problem (is broken)’
- 141) *ja tese a-ra-fa-n-fa-g* [nirge motor ja=ri adi kabia]
 1S must 1S.POSS-1-leg-3-go-STEM because motorcycle 1S=have DET broken
 ‘I must walk because my motorcycle has a problem (is broken)’

3.1.3.1.1.2.2 Inalienable possessive pronouns

Although the inalienable possessive construction is observed in a set of relationships that is smaller than that of alienable relationships, it is a structure that occurs frequently. Inalienable possession partially resembles alienable possession morphology. The pronominal possessive prefixes are identical in both types of possessive relationships, but inalienable relationships also require a possessive pronominal suffix (see TABLES 3.10–11), which makes inalienable possessive morphology appear on the surface to be a circumfix, especially considering that the possessive suffixes does not occur elsewhere.

	Singular	Plural
1n		<i>it- __-(o)g</i>
1x	<i>a- __-(o)g</i>	<i>am- __-(o)g</i>
2	<i>o- __-(o)m</i>	<i>e- __-(o)m</i>
3	<i>i- __-Ø</i>	<i>ir- __-Ø</i>

TABLE 3.10 - Inalienable possessive circumfixes.

- 142) *a-mtie-g*
 1S.POSS-eye-1.POSS
 ‘my (inalienable) eye’

The prefix portion of the inalienable possessive morphology agrees in person, number, and inclusivity for all pronominal referents. The inalienable possessive suffixes only agree in person with speech act participants, not number (142). Inalienable possessive

relationships with third person possessors are not marked by an overt morpheme. There is a zero allomorph in the third person possessive suffix. Therefore, the inalienable relationship must be determined by the behavior of the first and second person forms. A sample paradigm for the inalienable relationship ‘neck’ is given in TABLE 3.11.

	Singular	Plural
1n	—	<i>it-drbon-og</i> ‘our throats’
1x	<i>a-drbon-og</i> ‘my throat’	<i>am-drbon-og</i> ‘our, but not your, throats’
2	<i>o-drbon-om</i> ‘your throats’	<i>e-drbon-om</i> ‘your (plural) throats’
3	<i>i-drbun</i> ‘his throats’	<i>ir-drbun</i> ‘their throats’

TABLE 3.11 – Example inalienable possession paradigm using the relationship of ‘throat’.

Root	Gloss	1SPOSS	2SPOSS	3SPOSS
<i>-din-</i>	‘plate’ ((?))	<i>adinug</i>	<i>odinum</i>	<i>idin</i>
<i>-doyo-</i>	‘forehead’	<i>adojog</i>	<i>odojom</i>	<i>idoje</i>
<i>-drbon-</i>	‘throat’	<i>adrbonog</i>	<i>odrbonom</i>	<i>indrbum</i>
<i>-fa-</i>	‘leg’	<i>afag</i>	<i>ofam</i>	<i>ifa</i>
<i>-fgie-</i>	‘back (anat.)’	<i>afgieg</i>	<i>ofgiem</i>	<i>ifgie</i>
<i>-fi-kasar</i>	‘vagina’	<i>afig</i>	<i>ofim</i>	<i>ifi</i>
<i>-fir-</i>	‘bile’	<i>afirug</i>	<i>ofirum</i>	<i>ifir</i>
<i>-fra-</i>	‘hand’	<i>afrag</i>	<i>ofram</i>	<i>ifra</i>
<i>-fta-</i>	‘belly’	<i>aftag</i>	<i>oftam</i>	<i>ifta</i>
<i>-garn-</i>	‘mouth (dialectal)’	<i>aganug</i>	<i>oganum</i>	<i>ingan</i>
<i>-gigr-</i>	‘armpit’	<i>agigrug</i>	<i>ogigrum</i>	<i>igigr</i>
<i>-jo-</i>	‘bodily feeling’	<i>ajog</i>	<i>ojom</i>	<i>ije</i>
<i>-rguin-</i>	‘head’	<i>arguinug</i>	<i>orgunum</i>	<i>irguin</i>
<i>-mtie-</i>	‘eye’	<i>amceg</i>	<i>omcem</i>	<i>imce</i>
<i>-nmu-</i>	‘thigh’	<i>anmu(nu)g</i>	<i>onmu(nu)m</i>	<i>inmu(num)</i>
<i>-nu-</i>	‘name’	<i>anug</i>	<i>onum</i>	<i>inu</i>
<i>-srfir-</i>	‘back of neck’	<i>asrfirug</i>	<i>osrfirum</i>	<i>isrfir</i>
<i>-ri-</i>	‘wrist’	<i>arirug</i>	<i>orirum</i>	<i>irir</i>
<i>-rimta-</i>	‘mouth’	<i>arimtag</i>	<i>orimtam</i>	<i>irimta</i>
<i>-rit(ru)-</i>	‘skin’	<i>aritug</i>	<i>oritum</i>	<i>irit</i>
<i>-rü-</i>	‘voice’	<i>arug</i>	<i>orum</i>	<i>iru</i>
<i>-rui-</i>	‘sibling’	<i>aruig</i>	<i>oruim</i>	<i>irui</i>
<i>-tgra-</i>	‘ear’	<i>atgrag</i>	<i>otgram</i>	<i>itgra</i>
<i>-ti-kasar</i>	‘penis’	<i>atig</i>	<i>otim</i>	<i>iti</i>
<i>-wa-</i>	‘widow’	<i>(j)awag</i>	<i>owam</i>	<i>iwaman</i>

TABLE 3.12 – Twenty-five inalienable-relationship nouns. Words with subscript *kasar* are low prestige forms. Van den Berg and Matsumura (2008) also cite *a-kki-ug* ‘my chin’ and *a-ti-ug* ‘my liver’. One other form attested in the literature is *tñi-ug* ‘my body’.

TABLE 3.12 shows the inalienable possessive relationship is expressed with at least twenty-nine word roots. Twenty-five of these are observed in inalienably possessed monomorphemic stems.

The specific relationships that are encoded as inalienable are idiosyncratic. Twenty-one are body part relationships. Two are abstract nouns ('name' and 'bodily feeling'). Only one term is a kin term ('widow'). Of these terms, ten of them occur as the first element in possessed compounds. The best published data on this phenomenon is Smits and Voorhoeve (1992), which cites approximately thirty inalienable stems. Twenty-five of which are body parts. There is one abstract noun, 'name' and four kin terms 'younger sister, older sister, aunt, wife'. That data was collected in the 1950s and is valuable because it allows a shallow diachronic perspective. The set of elements which occurs with obligatory inalienable morphology appears to be diminishing. This is most evident in inalienable relationships with compound nouns, discussed below.

Some of the relationships are signified by inalienable morphology and derive from the high prestige variety of the language, while alternate terms for the same relationship from the prestige variety are marked as alienable. For example, *wa* 'widow' is inalienable whereas *wa+man* 'husband', *wa+bfin* 'wife', and *waniku* 'widower (man who has passed away)', all appear to contain the root *wa* but parallel the terms *sabatdu* 'female widow (husband passed away)' and *masyan* 'male widow (wife passed away)' as being coded as alienable relationships. Similar paradoxes are found in terms for siblings and genitalia. Alienable include the basic register term *baba* 'elder sibling' (*ababa* 'my elder sibling', *obaba* 'your elder sibling', and *ibaba* 'her elder sibling'), and the high register word *nifut* 'younger sibling' (*anifut*, *onifut*, *inifut*). This contrasts with the root *nfut* 'younger sibling' (*agfut*, *onfut*, *infut*), which has an irregular paradigm (the first person contains the inalienable morphology whereas the second person is marked as alienable), and any terms with *rui* '(elder) sibling' or high register *ta+fad* 'sibling'.

	Singular	Plural
1		<i>it-fra-g+ntu i-den</i> 'our (incl.) thumbs'
1	<i>a-fra-g+ntu i-den</i> 'my thumb'	<i>am-fra-g+ntu i-den</i> 'our (excl.) thumbs'
2	<i>o-fra-m+ntu i-dena</i> 'your thumb'	<i>e-fra-m+ntu i-den</i> 'your (PL) thumbs'
3	<i>i-fra-Ø+ntu i-dena</i> 'his thumb'	<i>ir-fra+ntu i-den</i> 'their thumbs'

TABLE 3.13 – Paradigm for inalienable possession involving a compound noun.

A further complication arises because some body part terms are compounds. Alienability seems to be determined by the first member of compound body part terms. Inalienable possessive morphology is observed on the first element in several compounds. Some forms such as *-fa* 'leg' and *-rit* 'skin' occur without the inalienable suffix when referred to on their own, but in a compound like *fa+ntu* 'toe (lit. foot offspring)', the first element in the compound is marked by inalienable morphology: *a-fa-g+ntu* 'my toe'. Conversely, when a noun that is generally coded as inalienable is used as the second element in a compound, such as *sus+mce* 'nipple', the compound is coded according to alienability of the first member of the compound (*asus* 'my breast, *osus* 'your breast',

isus ‘her breast’). A paradigm for the relationship to the complex noun *fra+ntu i-den* (hand+offspring 3SPoss-mother) ‘thumb’ is provided in TABLE 3.13.

-Ø-+ <i>nfut</i>	‘younger sibling’	<i>agfut</i>	<i>onbut</i>	<i>infut</i>
- <i>fa</i> -+ <i>drbun</i>	‘heel’	<i>afagdrbun</i>	<i>ofamdrbun</i>	<i>ifdrbun</i>
- <i>fa</i> -+ <i>jarf</i>	‘sole of foot’	<i>afagjarf</i>	<i>ofamjarf</i>	<i>ifajarf</i>
- <i>fa</i> -+ <i>saffu</i>	‘top of foot’	<i>afagsaffu</i>	<i>ofamsaffu</i>	<i>ifasaffu</i>
- <i>fa</i> -+ <i>war</i>	‘crotch’	<i>afagwar</i>	<i>ofamwar</i>	<i>ifawar</i>
- <i>fa</i> + <i>sme</i>	‘ankle’ ((check))	<i>afagsme</i>	<i>ofagsme</i>	<i>ifasme</i>
- <i>fgie</i> -+ <i>rüri</i>	‘spine’	<i>afgiegrur</i>	<i>ofgiemrur</i>	<i>ifgierur</i>
- <i>fi</i> -+ <i>fon</i> _{halus}	‘vaginal canal’	<i>afigfon</i>	<i>ofimfon</i>	<i>ififon</i>
- <i>fi</i> -+ <i>gur</i>	‘vaginal canal’	<i>afiggur</i>	<i>ofimgur</i>	<i>ifigur</i>
- <i>fra</i> -+ <i>fü</i>	‘wrist’	<i>afragfü</i>	<i>oframfü</i>	<i>ifrafü</i>
- <i>fra</i> -+ <i>fu</i>	‘shoulder’	<i>afragfu</i>	<i>ofrabu</i>	<i>ifrafu</i>
- <i>fra</i> -+ <i>jarf</i>	‘palm’	<i>afragjarf</i>	<i>oframjarf</i>	<i>ifrajarf</i>
- <i>fra</i> -+ <i>natu</i>	‘finger’	<i>afragntu</i>	<i>oframntu</i>	<i>ifrantu</i>
- <i>fra</i> -+ <i>si</i>	‘elbow’	<i>afragsi</i>	<i>oframsi</i>	<i>ifrasi</i>
- <i>fra</i> -+ <i>tgrir</i>	‘fingernail’	<i>afragtgrir</i>	<i>oframtgrir</i>	<i>ifratgrir</i>
- <i>fta</i> -+ <i>ntu</i>	‘intestines’	<i>aftagntu</i>	<i>oftamntu</i>	<i>iftantu</i>
- <i>fta</i> -+ <i>rie</i>	‘bottocks’	<i>aftagrie</i>	<i>oftamrie</i>	<i>iftarie</i>
- <i>fta</i> -+ <i>rie</i> + <i>ntu</i>	‘butt cheeks’	<i>aftagrientu</i>	<i>oftamrientu</i>	<i>iftarientu</i>
- <i>je</i> -+ <i>fut</i>	‘calf (anat.)’	<i>ajegfut</i>	<i>ojefut</i>	<i>ijefut</i>
- <i>mta</i> -+ <i>re</i>	‘cheek’	<i>amtagre</i>	<i>omtamre</i>	<i>imtare</i>
- <i>mta</i> -+ <i>ru(m)</i>	‘face’	<i>amtagrug</i>	<i>omtamrum</i>	<i>imtarum</i>
- <i>mte</i> -+ <i>rue</i>	‘tears’	<i>amcegrue</i>	<i>omcemrue</i>	<i>imcerue</i>
- <i>rfu</i> -+ <i>rir</i>	‘tooth’	<i>arfugrir</i>	<i>orfumrir</i>	<i>irfurir</i>
- <i>rimta</i> -+ <i>rit</i>	‘lip’	<i>arimtagrit</i>	<i>orimtamrit</i>	<i>irimtarit</i>
- <i>ro</i> -+ <i>fut</i> _{halus}	‘penis’	<i>arogfut</i>	<i>orobut</i>	<i>irofut</i>
- <i>rwi</i> -+ <i>bfin</i>	‘sister’	<i>arwig</i>	<i>orwim</i>	<i>irim</i>
- <i>rwi</i> -+ <i>mran</i>	‘brother’	<i>arwigmran</i>	<i>orwim</i>	<i>irwimran</i>
- <i>ta</i> -+ <i>fad sot</i>	‘older sister’	<i>atagfadsot</i>	<i>otabadsot</i>	<i>itafadsot</i>
- <i>ta</i> -+ <i>fad</i> _{halus}	‘elder sibling’	<i>atagfad</i>	<i>otabad</i>	<i>itafad</i>
- <i>tbon</i> -+ <i>fru</i>	‘head hair’	<i>argunugfru</i>	<i>orgunumfru</i>	<i>irgunfru</i>
- <i>tgra</i> -+ <i>gur</i>	‘ear hole’	<i>atgraggur</i>	<i>otgramgur</i>	<i>itgrarur</i>
- <i>tgra</i> -+ <i>ro</i>	‘external ear’	<i>atgragro</i>	<i>otgramro</i>	<i>itgraro</i>
- <i>ti</i> -+ <i>rir</i>	‘testicles’	<i>atigrir</i>	<i>otimrir</i>	<i>itirir</i>
- <i>tñi</i> -+ <i>tu</i>	‘body’	<i>atnigtu</i>	<i>otnimtu</i>	<i>itnitu</i>
- <i>we</i> -+ <i>guru</i> -	‘nose’	<i>aweggugur</i>	<i>owegmgur</i>	<i>iweguru</i>
- <i>wi</i> -+ <i>fu</i>	‘knee’	<i>awigfu</i>	<i>owibu</i>	<i>iwifu</i>
- <i>wi</i> -+ <i>rür</i>	‘shin bone’	<i>awigrür</i>	<i>owimrür</i>	<i>iwirür</i>

TABLE 3.14 – Thirty-seven inalienable compound relationships based on nineteen inalienable roots. Some items are also found with inalienable roots. *halus* are high prestige forms.

Some inalienable possessive constructions cited in early research (Anceaux 1961, Grace 1957, and Smits and Voorhoeve 1992) show a compound lexical item, such as *wi+fu*

‘knee’ (*wi* is a frozen morpheme, *fu* means ‘leg’), with an inalienable suffix on each element in the compound, *a-wi-g+fu-g* ‘my knee’. By contrast, attested contemporary forms include *wifu*, *wigfu*, and *wifug*, but not *wigfug* ‘my knee’ and *amtagrug* ‘my eye’. The second occurrences of the suffix are optional.

- 143) *a-mta-g+ru(-g)*
 1S.POSS-eye-1POSS+house(-1POSS)
 ‘my face’

Van den Berg and Matsumura 2008 cite three relevant forms: *-fra-+fu-* ‘shoulder’ with a recurring suffix, *-Ø-+tat* ‘grandparent/ grandchild’ with initial empty slot allowing *a-* and *-g* to occur next to one another, and *-fta-+rie gur* ‘anus’ with independent, final *gur*. Anceaux (1961) cites *-fra-ta-+rire* ‘fingernail’ instead of *-fra-+tgrir*, as well as *fta irie* ‘buttocks’ with initial *i* in *irie*. Van den Berg and Matsumura (2008) report *rie* in phrases such as *rie su* ‘tail end (of bird)’, and *rie ftu* ‘be in heat’. These facts suggest that the productivity of inalienable possession is receding while simplified lexical structure is increasing. Nine elements introduced in TABLE 3.12 also occur in the compound forms listed in TABLE 3.14. In addition, four of the first elements in the compounds from TABLE 3.14 are found in inalienable compound nouns, but not in uncompounded inalienable stems listed in TABLE 3.12.

Most of the second members in compounds are found elsewhere, such as *mta* ‘eye’, *rit* ‘skin’, *fu* ‘bone protrusion (where the bone is anatomically close to the skin)’, but five of the second members in these compounds are either frozen morphemes or roots that do not occur alone, such as *-si*, *-re*, *-(i)rie*, *-wi*, *-tgrir*. On a final point, some inalienable compounds are marginally acceptable with the suffix portion of the inalienable morphology at the end of the whole compound: **awifug* ‘my knee’.

Inalienable possession has been discussed in previous literature in greater detail than other areas of Irarutu grammar. The most recent discussion is van den Berg and Matsumura 2008, in which six possessive patterns in the East Arguni dialect are presented. The first two patterns divide roots according to the presence or absence of a vowel in the inalienable suffix, which is largely determined by the shape of the final syllable of the host. They identify three patterns for compounds, which are described as having an infixable inalienable morpheme after the first element in the compound, or having both an infix and a suffix, or a null first element causing the prefix and suffix to be adjacent. The sixth pattern is for irregular forms.

No data that they cite, or that has been identified independently, shows that inalienable morphology is ever inserted within a root, which would be the case for true infixation. From a semantic point of view, van den Berg and Matsumura rightly point out that inalienable relationships generally pertain to internal organs (they cite: ‘lung, brain, heart, vein, breast, Adam’s apple, blood’) as well as eight kinship terms (‘opposite-gender sibling, spouse, younger sibling, child-in-law, older sibling, parent-in-law, friend/brother, and grandchild’). Consultants for the present description do not code the word ‘breast’ as inalienable. Van den Berg and Matsumura also point out that the full set of morphology

(prefix and suffix) is obligatory in the emphatic inalienable possessive construction, *afrag jari* ‘hands of mine’.

3.1.3.1.1.3 Reflexive and reciprocal pronouns

Both reflexive and reciprocal pronouns are coreferential with an antecedent in the same construction, either the subject or another noun. Examples of each are given below.

Reflexive and emphatic are sometimes formally similar. Emphatics are discourse markers. There is a person distinction in reflexives (SAP vs. N-SAP) as well as use of an invariable form *ntati*. Reflexives involve a head nominal modified by a pronominal possessive that agrees with the subject, i.e., reflexively interpreted noun or NP.

Reciprocals are used to express mutual actions, conditions, etc. It is not uncommon, cross-linguistically, for ambiguity to arise between reflexive and reciprocal interpretations because the same form/construction is used for both.

There are three constructions that achieve a reflexive interpretation (‘do ... to self’). Each involves some kind of anaphoric pronoun following the verb. Reciprocals (‘do ... to each other’), which are described after reflexives, parallel one of these patterns.

The first strategy is to use the reflexive (adverb) *ntati* ‘self’ to modify the object, shown in 144). *Ntati* can also be translated as ‘original/inherent’.

- 144) *i=ro i si i n-tati [fi nene kirrar adie]*
3S=TOP 3S see 3S 3-self/REFL REL PREP mirror DET
‘he sees himself in the mirror (lit. he sees just himself who is in the mirror)’
- 145) *matu wenum (n-tati)*
person land 3S-REFL
‘native (lit. person land original)’

The second reflexive strategy is to mark the object as emphatically possessed using the ‘of mine’, i.e., ‘X (which) personal pronoun has’, relative clause construction.

- 146) *ja si a-mtagru-g ja=ri*
1S see 1S.POSS-face-1POSS 1S=have
‘I see myself (lit. I see my face)’
- 147) *ja si a-gmon-g ja-ri*
1S see 1S.POSS-1.POSS 1S-have
‘I see myself (lit. I see my reflection)’
- 148) *i si gabar i-ri gue*
3S see picture 3S-have alone
‘he sees himself (lit. he sees his picture only)’

The third strategy that can achieve a reflexive meaning is to simply repeat the subject pronoun after the verb in the object slot. In an example like 150), where the pronouns are non-speech act participants, arguments can be coreferential. The verb potentially implies

reciprocal action due to ambiguity between a standard transitive ‘she hit him’, a reflexive ‘she hit herself’, and a reciprocal ‘they hit each other (he (Jack) hits him (Michael) / he (Michael) hits him (Jack))’.

- 149) *ja gor ja*
 1S hit 1S
 ‘I hit myself’
 150) *i dbe i*
 3S hit 3S
 ‘she hit him / she hit herself / they hit each other’
 151) *ir fi-smsamt ir (fiesuem)*
 3P ACT-embrace 3P as.one
 ‘they hug each other (lit. they hug them)’

The word *fiesuem* ‘as one’ can be added at the end of 151) to strengthen the reciprocal sense of the predicate *fismsamt* ‘hug/embrace’.

The reciprocal marker is *nasan*. It occurs immediately after the subject. The object position in a construction with *nasan* is coreferential with the subject.

- 152) *ir nasan tbe friemn*
 3P RECP hit repeatedly
 ‘they hit each other (lit. they themselves hit repeatedly)’
 153) *ir nasan n-fi kkuk ir*
 3P RECP 3-ACT box 3P
 ‘they box each other’
 154) *ir (nasan) n-fi sgffo ir*
 3P RECP 3-ACT kiss 3P
 ‘They kiss (each other)’

3.1.3.1.1.4 Demonstrative pronouns

This category is concerns words like ‘this, that, these, those’ in constructions such as ‘this is the same as that’ and ‘do you like these or those?’ Sometimes languages do not distinguish non-SAP personal pronouns and demonstrative pronouns, as in morphologically related demonstrative articles.

Demonstrative pronouns ‘this’ and ‘that’ are identical in form with demonstrative determiners, but the distributional traits of these elements are slightly different. Demonstrative determiners are obligatory, whereas demonstrative pronouns are frequently omitted, as in the question-answer pair in 155)–156).

- 155) *gata i=ri Ø?*
 who 3S=have DEM.PRO
 ‘who’s is this?’

- 156) *a-matu adi=ro i=ri Ø*
 1S.POSS-friend DET=TOP 3S=have DEM.PRO
 ‘It’s my friend’s (lit. my friend owns the (thing))’

3.1.3.1.1.5 Indefinite pronouns

Indefinite pronouns are translatable as ‘someone, something, anyone, anything’, and cross-linguistically are composed of an indefiniteness marker and ‘person/thing’. In Irarutu, there is an indefinite/interrogative pronoun ‘someone/anyone’. As an alternative, a generic noun can be used, such as *matu* ‘person/someone’, example 157). This evidence suggests that nouns are, by default, both non-specific and indefinite. One alternative in Irarutu, which is common cross-linguistically, is to use a *wh*- word, *nia* ‘what’, to express an indefinite referent.

- 157) *matu fide*
 person house
 ‘someone’s house’

3.1.3.1.1.6 Relative pronouns

Relative clauses are discussed in §3.3.2. There are two relativizers in Irarutu, *uf* and *fi*, but these do not encode a difference between human (or animate) and non-human (or inanimate) referents. The two morphemes are in complementary distribution. Neither is considered a relative pronoun, but pronouns are not necessary to form a relative clause in Irarutu. Coincidentally, the word *gata* ‘who’ is not used for this purpose. As it was described by Irarutu speakers, the fundamental difference between *uf* and *fi* is whether the information is assumed—or presumed—from prior context, in which case *uf* occurs, or if it is being asserted, in which case *fi* occurs. The following example contains both relativizers.

- 158) *wenum ad-uf ro dir nu adi-fi Fruat*
 land DET=REL COP good name DET=REL Fruata
 ‘there is a place that is very beautiful, which is called Fruata’

Another common cross-linguistic tendency in forming relative constructions that is witnessed, to an extent, in Irarutu is to delete N_{REL} from S_{REL} .

3.1.3.1.2 Other pro-forms: question words

The types of interrogative pro-forms vary cross-linguistically. They cross-cut other syntactic categories: interrogative pronouns (who, what), interrogative adverbs (where, when), and interrogative articles (‘which’ in ‘which book’ represent discourse articles). Furthermore, interrogative pro-forms occur in interrogative sentence types (see §3.4). In Irarutu, there are five pro-form question words: *gata* ‘who’, *nia* ‘what’, *neno* ‘where’, *nano* ‘why’, and *fnaneno* ‘how’. Temporal (‘when’) questions are constructed with a temporal noun, e.g., *rre* ‘day’ or *seba* ‘month’, followed by the interrogative pronoun *nia* ‘what’. A grammatical, albeit constructed, example sentence that contains several question words is 159).

- 159) *nano gata n-fun nya nene gata?*
 why who 3-do what PREP who
 ‘why did who do what to who?/why did someone do something to someone.’

The indefinite/interrogative pronoun *gata* ‘who’ is used in questions about human referents. To inquire about possession, the morpheme *-ri* ‘have’ is suffixed to the interrogative pronoun, producing *gata-ri* ‘whose’.

- 160) *gata i=ri?*
 who 3S have
 ‘who’s is this’
- 161) *gata m-tim frrufn?*
 who STAT-close door
 ‘who closed the door?’

The most frequent, and functionally diverse, interrogative pro-form is the pronoun *nia* ‘what’. It is used to request information about non-human referents, or when combined with a time word, it requests information about temporal state. *Nya* also functions as an indefinite pronoun in declarative clauses, see 165–6).

- 162) *o g-o ba mgür nya?*
 2S DES-2 2.go 2.buy what
 ‘what are you going to buy?’
- 163) *o=ro bun nya?*
 2S=TOP 2.do what
 ‘What are you doing?’
- 164) *ja fun(-nya) nyati*
 1S do(-AGR_o) nothing
 ‘I am not doing anything’
- 165) *i, nya gor ja*
 EXCL something [3-]bite 1S
 ‘I am bitten (lit. something bit me)’
- 166) *matu n-taf nya fi i si*
 person 3-hunt something REL 3S see
 ‘the hunter shoots what he can see’

The interrogative pro-adverb *neno* ‘where’ requests information about absolute spatial reference as well as relative terms. One of the most common components of greetings in Irarutu, after exchanging salutations such as *mumcie/rre/rrefefa/gisie dir* ‘good morning/day/afternoon/night’, is to ask where someone is going, 167). A common type of answer is 168).

- 167) *o=ro ba [neno]?*
 2S=TOP 2.go where
 ‘where are you going?’

- 168) *ja fa [ne pasar]*
 1S [1-]go PREP market
 ‘I am going to the market’

The interrogative pro-forms *nano* ‘why’ (alternately *nfnanuf*) and *fnano* ‘how’ are closely related forms. They are also both pro-clauses. *Nano* is used to inquire about the cause of some particular action or state, whereas *fnano* is used to request information about the manner or means that result in an action or state. The question word *nano* often occurs at the beginning of a clause. *Fnano* occurs in commonplace phrases such as 169).

- 169) *o=ro fnano?*
 2S=TOP how
 ‘how are you?’

Temporal information must be requested using a periphrastic construction, such as 170)–171).

- 170) *erre nya?*
 day what
 ‘when (lit. what day)?’
 171) *ja mi mse rre nya?*
 1S stay until day what
 ‘how much longer may I stay?’

Matsumura and Matsumura (1991) cite a relevant example, *oro gin gun nia se oro bubae?* ‘When are you leaving?’ but with dialectal *gun* instead of *rre* for ‘day’. One of my consultants corrected this example by removing the topic marker from the dependent clause.

- 172) *o=ro g-in gun nia se o bu-ba-e?*
 2S=TOP SAP-sleep day what then 2S TAM-2.go-Q
 ‘When (after you sleep how many days then) are you leaving?’

3.1.3.1.3 Other pro-forms: pro-sentences, pro-clauses, and pro-verbs

Pro-sentences *e* ‘yes’ and *ee* ‘no’ are understood as equivalents for corresponding affirmative and negative sentences in response to certain kinds of questions. *I ma?* ‘is he coming?’ versus *i ma=ti?* ‘isn’t he coming?’ can show different equivalents for ‘yes’. Existential questions such as *werwin nene termus* ‘is there hot water in the thermos?’ can be answered by a negative assertion *werwin=ti=o* ‘There is no hot water’. Pro-sentences can be used as pro-clauses. Pro-clauses, for example, tag question: *i ma, (fitu)ti* ‘he’s coming, not true?’. This is an exceptional use of *ti* as a superficially independent element. An example of a pro-verb is *dir* ‘(it is) good’, shown in 173).

- 173) *dir noti fi=ge it=ro fi-ttut*
 good EXPL REL=COMP 1Pn=TOP ACT-meet
 ‘it is good that we met’

The aspectual morpheme *du* ‘already’ in 175) is a possible answer to the question in 174) and therefore functions with pro-sentence *e* ‘yes’. (Note: bathing during the middle of the day is a cultural trait of Indonesia; asking someone if they bathed is similar to asking someone if they ate lunch in English.)

- 174) *o=ro du kka?*
 2S=TOP already bathe
 ‘have you bathed?’
- 175) *e, du*
 yes already
 ‘yes, (I) already (bathed)’

3.1.3.2 Noun Adjuncts

Noun adjuncts, in the sense of Schachter 1985, typically form phrasal constituents with nouns and have clear semantic import not contained in the noun itself, such as the role of the noun in relation to the action expressed by the verb, or single or plural number, etc. In some cases, a noun adjunct, such as a classifier, is semantically empty and solely required by the syntax of the language. From a typological perspective, four noun adjuncts can be distinguished: role markers (which cross-linguistically can include case markers, discourse markers, and other adpositions), quantifiers, classifiers, and articles.

3.1.3.2.1 Role markers

Role markers include case markers (which indicate syntactic ‘subject’ or semantic role ‘agent’), discourse markers (discourse role ‘topic’ in the sense of Chafe 1976), and other adpositions. Irarutu does not have case markers, but it does have several prepositions that function as discourse markers. These are involved in various locative and temporal relationships. The prepositional location of the case markers (occurring before their heads) correlates with VO word order. Some adpositional relationships are not overtly marked with a separate word, in which case the information or function they provide and their position allows them to still be interpreted accurately. The role of one NP, the subject, nearly always appears on the verb. Roles can also be indicated by special syntactic construction (clefting) or intonation.

Irarutu has a relatively small number of prepositions, see **FIGURE 3.9**. Some also function in other categories, such as *nir* ‘with’, which can be a conjunction or preposition. One fairly unusual trait is that several of the prepositions are inflected for material outside the prepositional phrase, such as a preceding second person subject (‘you’). The most frequent prepositions are *nene* ‘to, at, on’, inflected *-ir* ‘with’ (*mir* ‘addressee with’ and *nir* ‘non-addressee with’), and inflected *-e* ‘to’ (*ge* ‘SAP to’ and *ne* ‘non-SAP to’). In some contexts *nir* ‘with/because’ can also function the same as the clause linker *nirge* ‘because’.

Irarutu	Gloss	Inflected forms
-ir	‘and, with, about, because’	<i>nir</i> ‘I CONJ’, <i>mir</i> ‘you CONJ’, <i>nir</i> ‘he CONJ’ <i>nir firati</i> ‘without’
-e	‘to’	(ra-)ge ‘I to’, (ma-)ge ‘you to’, (na-)ne ‘she to’
<i>mine no</i>	‘from’	
<i>ne</i>	‘for, in, at’	
<i>nene</i>	‘for, in, at’	

FIGURE 3.9 – Irarutu prepositions

Examples of the prepositions *nir* ‘concomitive, instrumental’ and *ne/nene* ‘locational, temporal’ are illustrated below, (176)–(182). Prepositional phrases can stand on their own as answers to *wh*- content questions such as ‘where is X?’

- 176) *ja fa Fruat [nir Mikel]*
1S go Fruata PREP Michael
‘I went to Fruata with Mikel’
- 177) *ja tar ema [nir suri]*
1S split firewood PREP axe
‘I cut firewood with an axe’
- 178) *it fa fun mir [ne dedan]*
1Pn go do tend PREP garden
‘We are going to do gardening (lit. we go do work in garden)’
- 179) *ja m-tür [nene ben]*
1S STAT-sit PREP floor
‘I am sitting on the floor’
- 180) *ja g-a fa [(nene) pasar]*
1S DES.1 go (PREP) market
‘I am going to the market’
- 181) *ja ma [nene refefa (uf fifi)]*
1S come PREP yesterday (REL previous)]
‘I arrived [yesterday] (lit. I came to evening which previous)’
- 182) *o=ro m-si gata fi [nene kirrar]*
2S=TOP 2-see who REL PREP mirror
‘who do you see in the mirror?’

Some common prepositional phrases in Irarutu, listed in **FIGURE 3.10**, below, show how they function to provide adverbial information, such as location.

<i>ne rfun</i>	‘between (lit. to middle)’
<i>ne tfu dir</i>	‘within (lit. to within good)’
<i>nene trro</i>	‘vertical (lit. to height)’
<i>nene figie</i>	‘behind (lit. to back)’
<i>nene refid</i>	‘(be)side (lit. to side)’
<i>nene tfu</i>	‘under (lit. to below)’
<i>nene adini</i>	‘those (lit. to the here/this)’
<i>nene amai</i>	‘there (lit. to there/those)’
<i>nene neno</i>	‘everywhere (lit. to from)’
<i>nene bunat adini</i>	‘latest/now (lit. to new the.this)’
<i>nene wenum uf</i>	‘elsewhere (lit. to land which)’
<i>nene oro nfanno</i>	‘according to (lit. to you do polite)’

FIGURE 3.10 – Examples of prepositional phrases.

Prepositional phrases, which are always obliques, can occur at the beginning of a sentence when they are topicalized:

- 183) [*nene kirror*]=*ro i=ro si gata?*
 PREP mirror=TOP 3S=TOP see who
 ‘In the mirror, who does he see?’

In ditransitives, a serial verb construction that includes a second, inflected, prepositional verb *ge* ‘to’ to introduce the second (i.e., indirect) object, as in 184).

- 184) *ja fi-ar wenga [ra-ge o-matu uf m-tim frufn]*_{VP2}
 1S ACT-give food 1-to 2S.POSS-friend REL [3-]STAT-close door
 ‘I gave food to your friend who closed the door’

The second ‘verb’ in these types of serial verb constructions follows a minor pattern of subject agreement. The inflectional elements *ra-* ‘first person subject’ and *ma-* ‘second person subject’ plus the preposition *ge* ‘speech-act-participant-subject to’ and *na-* ‘third person subject’ plus *ne* ‘non-speech-act-participant-subject to’ do not occur with main verbs. Another construction where the same minor pattern of inflection is also found is a type of instrumental causative construction, where the agentive subject uses an instrument to cause it to do something to an object, e.g., ‘I (cause/use the instrument) axe to cut wood’. 185) illustrates a canonical SVO ‘I cut wood’ plus oblique ‘with an axe’ construction. By contrast, 186) illustrates the instrumental causative construction. My primary language consultant considered the SVO plus oblique example sentence to be a translation of an Indonesian structure, whereas the instrumental causative sentence is more naturalistic.

- 185) *ja tar ema [nir suri]_{INS}*
 1S split firewood PREP axe
 ‘I cut firewood with an axe’
- 186) *ja Ø suri ra-tar ema*
 1S (take) axe 1-split firewood
 ‘I axe-split the firewood’/‘my axe splits the firewood’

Because prepositions function to introduce non-arguments into clauses, prepositions can be used to introduce a second argument into a construction with a monadic verb such as *tufi* ‘like’ or *tufi=ti* ‘dislike’.

- 187) *ja tufiti [(nene) wenum fi a-fa mi]_{OBL}*
 1S dislike PREP land REL1-go [1-]stay
 ‘I don’t like where I live’

A few verbs in Irarutu that are monadic in other languages, such as *mi* ‘live/stay’ are able to be dyadic in Irarutu. A preposition cannot be added without compromising grammaticality in 188a), shown by the ungrammatical 188b).

- 188a) *ja (fa) mi Bintuni*
 1S go stay Bintuni
 ‘I live in Bintuni’
- 188b) **ja fa mi [nene Bintuni]*
 1S go stay PREP Bintuni
 ‘I live in Bintuni’

3.1.3.2.2 Determiners

Based on distributional grounds, definite ‘the’ and indefinite ‘a’ articles and demonstrative adjectives/modifiers (this man, that woman) group together. These elements occur in the same position in relation to the noun and other elements of the NP, but do not co-occur in a single NP; however, there may be exceptions for suffixed articles. Demonstrative modifiers are like definite articles because they indicate reference (‘this’ means close at hand, ‘the’ means already established in the discourse). ‘The’ and ‘that’ can be the same word. Demonstrative modifiers are also semantically and morphologically closely related to demonstrative pronouns. Articles may show agreement with the nouns they modify. Indefiniteness can be expressed by the lack of a demonstrative suffix. This follows a cross-linguistic trend where definite is correlated with accusative, and indefinite with nominative.

The presence and composition of a determiner can indicate number (single vs. plural), definiteness (definite vs. indefinite), specificity (specific vs. non-specific), and referentiality (referential vs. nonreferential) in a noun phrase, see **FIGURE 3.11**.

Relative	Determiner	Specificity	Demonstrative	Gloss
	<i>ad</i>			‘the’
	<i>ad</i>	<i>-i</i>		‘the (specific)’
	<i>ad</i>	<i>-i</i>	<i>-ni</i>	‘this/here/proximal’
	<i>ad</i>	<i>-i</i>	<i>-mai</i>	‘over there/distal’
	<i>ad</i>	<i>-i</i>	<i>-fi(fi)</i>	‘the which (previous)’
	<i>ad</i>	<i>-i</i>	<i>-ro</i>	‘that’
	<i>ad</i>	<i>-e</i>	<i>-i</i>	‘there’
	<i>ad</i>	<i>-e</i>	<i>-j</i>	‘aforementioned (specific)’
	<i>ad</i>	<i>-e</i>	<i>-ji</i>	‘the aforementioned’
	<i>ad</i>	<i>-uf</i>		‘the which’
<i>f-</i>	<i>ad</i>	<i>-i</i>		‘the (resumptive SG)’
	<i>wen</i>	<i>-i</i>		‘the (PL)’
	<i>wen</i>	<i>-i</i>	<i>-ni</i>	‘these’
<i>f-</i>	<i>(w)en</i>	<i>-i</i>		‘the (resumptive PL)’

FIGURE 3.11 – Template for determiners.

3.1.3.2.2.1 Definiteness and specificity

There is no indefinite determiner in Irarutu. Unmodified nouns have indefinite and non-specific interpretation. An example is given in 189).

- 189) *matu fide*
 person house
 ‘someone’s house’

Example 190) provides another instance of an indefinite. Similar to a temporal question, however, the indefinite pronoun *nia* ‘what’ is adjoined to a quantifiable noun *fri* ‘price’. As a question word it presupposes specificity and requests information about definiteness.

- 190) *fri nia*
 price what
 ‘how much (*lit. what price*)?’

There are two definite determiner roots, singular *ad* and plural *weni*. When *ad* occurs without any bound morphology, it indicates definiteness (the referent is identifiable based on previous discourse or context) but not specificity (a particular instance of the referent from the pool of possible candidates is not asserted). When *ad* is modified by either of two suffixes, *-i* or *-e*, specificity is made concrete. The suffix *-i* marks specific proximal, while *-e* marks specific distal. Determiners can, and frequently do, occur with a demonstrative suffix. Distributionally, a definite determiner is the last modifier of a head noun if modifiers are present, 191)–195).

- 191) *termus ad nene dapur tena werwin n-mi?*
 thermos DEF PREP kitchen CONJ hot.water 3-stay
 ‘Does the thermos in the kitchen have any hot water?’
- 192) *mo adi*
 child DET
 ‘the child’
- 193) *sansun mmat adi*
 shirt green DET
 ‘the green shirt’
- 194) *man fi i si adi*
 bird REL 3S [3-]see DET
 ‘a bird which he saw’
- 195) *uce fi a-tün adi ne refefa (je-)fadi n-mi san gan*
 papeda REL 1-cook DET PREP yest. (contents-)REL.DET 3-live house inside
 ‘the papeda which I made yesterday is in the room’

In contrast to the singular determiner, examples 196)–197) show the plural determiner *weni*, which can be glossed as ‘some’.

- 196) *matu weni*
 person DET
 ‘the people (lit. several/some people)’
- 197) *rre weni*
 day DET
 ‘daily (lit. several days)’

A definite determiner can function something like a pronoun in contexts where the head noun is deleted, shown in 198–199).

- 198a) *Mikel=ro na sum*
 mikel=TOP eat fish
 ‘Michael ate fish’
- 198b) *Mikel=ro na adeni*
 mikel=TOP eat DET
 ‘Michael ate it’
- 199a) *ja fe-tün fras*
 1S ACT-cook rice
 ‘I cooked rice’
- 199b) *ja fe-tün adeji*
 1S ACT-cook DET
 ‘I cooked it’

In complex clauses, especially those with apparently redundant determiners, this property appears to be employed as a resumptive strategy to indicate the location that corresponds to the relativized element in a relative clause.

- 200) [matu ade_o-mai ja si ade_o]=ro rau gun bar
 person DET_o-DIST 1S see DET_o=TOP COP head bald
 ‘the man that I saw is bald (lit. the man over there I saw him is bald)’

There are contexts where a determiner and preceding relativizer phonologically contract, producing forms such as *feni* ‘the (which are plural)’ and *fadi* ‘the (which is singular)’, see examples 201)–203).

- 201) *i=ro n-fi frmi weni n-fi sog mese na-fe rrer*
 3=TOP 3-ACT clothes DET 3-ACT hang after NSAP.eat-DTR noon
 ‘he hung all the clothes after lunch’
- 202) *taun porna je-feni n-mi dapur*
 sago bread contents-REL.DET 3-stay kitchen
 ‘baked sago bread is in the kitchen’
- 203) *taun porna fi bunat weni (je-)feni n-mi dapur*
 sago bread REL new DET (contents)- REL.DET 3-stay kitchen
 ‘the bread which is freshly made is in the kitchen’

3.1.3.2.2 Demonstratives

Deictic demonstrative morphemes (*-ini*, *-ei*, and *-imai*) and discourse demonstrative morphemes (*-ej*, *-eji*) are both found in Irarutu. Both types of demonstrative suffix to a determiner (*ad* ‘singular’ or *we-* ‘plural’) in Irarutu, see **FIGURE 3.11**, following certain constraints.

There are three deictic demonstrative morphemes: *-ini* ‘here (by speaker)’, *-ei* ‘there (by hearer)’, *-mai* ‘over there (far from speech act participants)’. There is also a discourse demonstrative *-j* ‘aforementioned (presumed from discourse)’. A template for demonstrative determiners is provided in **FIGURE 3.11**. Examples cited in Matsumura and Matsumura (1991) also include *adiro* ‘that (lower)’, *adiri* ‘that (across the path)’, *adimari* ‘that (across the water)’, *wenimari* ‘those (across the water)’.

Matsumura and Matsumura (1991) lists two temporal demonstratives, but these are compositional, consisting of a determiner plus a relativizer *adi+fi* ‘previous/last’ and a determiner plus an adverb *adi+fifi* ‘two previous’. The adverb *fifi* can be seen to occur independently in 204).

- 204) *ja ma nene refefa (uf fifi)*
 1S come PREP yesterday (REL previous)
 ‘I arrived the other day’

The data shows that Irarutu determiners are generally modified for specificity and definiteness by either of two specificity suffixes, *-i* or *-e* and three deictic demonstrative suffixes, including *-i* ‘here’, *-ini* ‘there’, and *-mai* ‘way over there’. Three discursive demonstratives, *-uf*, *-fi*, and *-j(e)* can occur as the last suffix on a determiner. A few examples show that the determiner can be prefixed by *r-* or *f-*. The role of these elements needs further research, but *r-* occurs with a first person subject, similar to a minor pattern

of agreement, whereas *f-* occurs on determiners at the end of relativized clauses, in particular those introduced with *fi*.

- 205) *a-matu adi=ro i=ri.*
 1S.POSS-person DET=TOP 3S=have
 ‘It’s my friend’s (lit. my friend owns the (thing))’

The preverbal location of the determiner and topic marker in 205) deserves some clarification. The clause is syntactically a possessive. The possessed pronoun is relativized by the emphatic possessive construction.

3.1.3.2.2.3 Deictic demonstratives

Three spatial deictic demonstratives, in contrast to discourse demonstratives (following Schachter 1985), occur frequently. Deictic demonstratives are in complementary distribution with discursive demonstratives. All demonstratives are bound morphemes, so their distribution is determined by their determiner host.

- 206) *bariku adi-ni*
 lighter DET-PROX
 ‘this lighter’
 207) *fun bidi ru adeni*
 dog big two DET-PROX
 ‘these two big dogs’
 208) *fide adi-mai*
 house DET-DIST
 ‘that house (over there)’

3.1.3.2.2.4 Discourse demonstratives

There is only one dedicated discourse demonstrative morpheme, *-j* ‘aforementioned’. It attaches to the stem *ade-*; however, the resulting word can also host another suffix, *-i*.

- 209) *finnar adej rau dir*
 joke DET COP good
 ‘the joke is funny’
 210) *wen-ga-fe adeji*
 NMLZ-SAP-eat-DTR DET
 ‘the food (that was just mentioned)’
 211) *Jek-a o(=ro) m-tür adeji rau nano?*
 Jack-VOC 2S(=TOP) STAT-sit DET COP how
 ‘Jack, how are you doing, sitting there?’

3.1.3.2.3 Quantifiers and numerals

Quantifiers indicate quantity and are involved in scope: numerals, and words meaning ‘many’, ‘much’, ‘few’, ‘all’, ‘some’, ‘each’ etc. Quantifiers can be required if plurality is explicitly indicated, without which, nouns can default to single or plural number

depending on how they are conceptualized. Irarutu can encode quantifier information in nouns (e.g., people with abundance) or verbs (e.g., be enough).

The Irarutu number system is compositional, based upon a limited set of elements. Terms for ‘one’ to ‘five’ are non-compositional. There are also non-compositional words for ten, twenty, and one hundred. Mass nouns, such as *enyefu* ‘sand’, behave different from count nouns, such as *omage* ‘coconut’. Count nouns occur with numerals. Mass nouns are plural by default and are not directly modified by a numeral. For example, *enyefu* ‘sand’ is inherently plural. A single granule of sand needs to be specified as such: *enyefu rrir esuem* ‘one grain of sand’. On the other hand, count nouns, such as *nyaunt* ‘k.o. ant’, must be interpreted based upon context, see 213). *Pua* ‘many’ and *mug* ‘many’ can be used to specify a large but non-specific quantity.

212) *nyaunt*
ant
‘(there are) ants!’

213) *nyaunt pua*
ant QUANT
‘many ants’

By contrast, proper nouns are singular since they have individual referents.

214) *Mikel=ro n-tiit wagt mug adi_i namse nanbssi*
Michael=TOP 3-pound taro much DET until smooth
‘Michael must pound taro many times so that it becomes smooth’

In general, nouns default to a plural interpretation. There is a clear sense that most referent sets contain more than one member. A more concrete interpretation can be indicated by means of a numeral, determiner, or demonstrative. Variegation, i.e., many different types of a general thing, can be specified by reduplication.

3.1.3.2.3.1 Quantifiers, mass nouns, and plurality

Quantifiers are a closed, but rather loosely associated set of items. Quantifiers are able to occur with mass nouns as well as count nouns, e.g., *kkor su tur timebe* ‘all three tail of chicken’. Examples follow, and a list of quantifiers that occur frequently is given in **FIGURE 3.12**.

215) *matu fusr*
person QUANT
‘cluster of people’

216) *bwek fu pua*
betel.nut fruit QUANT
‘a lot of betel nut’

217) *matu nyati*
person QUANT
‘no one’

- 218) *Mikel=ro n-tiit wagt mug adi namse na-n-bssi*
Michael=TOP 3S-pound taro much DET until HAB-3-smooth
‘Michael must pound taro many times so that it becomes smooth’
- 219) *er du bitr e-ga timadbe*
if TAM hungry 2P-SAP.eat QUANT
‘If you are still hungry, eat everything’
- 220) *ja=mgür sirin funya fi a-si si-dir*
1S=buy papaya QUANT REL 1S-see PL-good
‘I bought some papaya which looked good.’

<i>nyati</i>	‘none/empty’
<i>moru/met</i>	‘little/few/minus’
<i>fu-nya</i>	‘some (for fruit; lit. CLF-what)’
<i>n-genaf</i>	‘less’
<i>na morudir</i>	‘less’
<i>pua</i>	‘many’
<i>mug</i>	‘many’
<i>fusr</i>	‘many’
<i>sfar</i>	‘more (than)’
<i>timebe</i>	‘all/every’
<i>timadbe</i>	‘everything’
<i>rarur</i>	‘much’
<i>si</i>	‘plural marker’

FIGURE 3.12 – Irarutu quantifiers.

Another strategy to indicate plurality is through the use of a conjunction, such as 221)

- 221) *mgir nir mgir*
mountain CONJ mountain
‘mountains (lit. mountains and mountains)’

3.1.3.2.3.2 Numerals and count nouns

Countable nouns can be modified by a numeral. The numeral follows the noun it modifies, as shown in 222–223). There is a clear pattern for forming numbers up to twenty using the basic numbers one through five. There appears to be a morphological ligature *-d-* that occurs in two important numbers, *fradfid* ‘five’ and *fradru* ‘ten’.

- 222) *mo ru*
child two
‘two children’
- 223) *seba tur*
month three
‘quarter (of a year)’

A more comprehensive set of numerals is provided in TABLE 3.15, below. Cardinal numbers ‘first’, ‘second’, and ‘third’ are the same as plain numbers *esuem*, *ru*, and *tor*. The Irarutu counting system above ‘twenty’ has been forgotten by most speakers, who generally use Indonesian for higher counting and mathematics. During elicitation, novel forms such as ‘136’ received no consensus among the consultants. The Kuri counting system is more intact among younger speakers, and appears to be nearly identical to Irarutu. Therefore, it is cited along side the Fruata dialect data.

Below twenty, Irarutu is arguably a base-five language. The form *esu* ‘one’ was noted, but the more general term, i.e., the one used for counting, is *esuem* ‘just one’. The numerals *esu* ‘one’, *ru* ‘two’, *tor* ‘three’, and *gigti* ‘four’ are not derived from smaller parts. *Fradfid* ‘five’ is composed of the elements *fra* ‘arm’, a ligature *-d-*, and *(re)fid* ‘side/five’, which literally means ‘hand-side’ or ‘hand-five’, i.e., ‘five’.

The numerals six through ten (*treso* ‘six’, *treru* ‘seven’, *trator* ‘eight’, *tregigti* ‘nine’) are optionally introduced by *fradfid* ‘five’. These numerals also share the element *tre* which is translated as ‘more than five’. The numeral ‘ten’ is composed of *fra* ‘arm’, the ligature *-d-*, and *ru* ‘two’, which means ‘two hands’ or ‘ten (fingers)’.

The pattern from eleven to twenty is similar to that established for six through ten. A quinary stem (5X2) is followed by a word, *risi*, that means ‘plus’ and a numeral from one through ten. The word *risi* ‘plus’ itself can be analyzed as *ri* ‘have/possess’ plus *si* ‘Plural marker’.

The third quinary stem is *matutni* ‘twenty’ (5X4). This word is compound of *matu* ‘person’ and *tni* ‘body’, literally ‘person’s body’, and refers to the fact that – generally speaking – a person’s body has twenty countable extremities: ten fingers and ten toes. It has been pointed out that using a word for ‘person’ to mean ‘twenty’ is a common areal trait in Papua and some parts of Southeast Asia. From twenty to one hundred, Irarutu uses base-20; however, the counting system from ‘twenty-one’ to ‘thirty’ and beyond needs further research because there are numerous inconsistencies. One example is *matutni esu* means ‘twenty-one’ (20+1) but *matutni ru* was given for ‘forty’ (20X2), not ‘twenty-two’ (20+2), and *ratu esu* ‘one hundred’ but *ratu risi esu* ‘101’. An additional characteristic that seems to be retained in Kuri and also plays a role in Irarutu, is that of referring to half a person by means other than *fradru* ‘ten’. Kuri has the word *nemet* ‘half’, as in *tmatutni ru nemet* ‘fifty’ and *tmatutni tor nemet* ‘seventy’. The structure of these numerals is 20(*n*)+10. For ‘fifty’, the phrase is equivalent to ‘two and a half people’: 20(2)+10=50. The structure of the term *tmatutni ru nemet* ‘thirty’ is unclear. There are no explicit terms for multiplication or division, but in TABLE 3.15, there are several instances of implicit multiplication, e.g., 40 = ‘2 x 20 (two people)’.

Gloss	Irarutu	Kuri
1	<i>esu (esuemem/esunat '1 satu)</i>	eso
2	<i>ru/rifu</i>	ru
3	<i>tur</i>	tor
4	<i>gigiti</i>	gegte
5	<i>fradfid (fid '5')</i>	fradfid
6	<i>tresu</i>	teso
7	<i>treru</i>	treru
8	<i>tretur</i>	tretor
9	<i>tregigiti</i>	tregegte
10	<i>fradru</i>	fradru
11	<i>fradru risi esu</i>	fradru rir eso
12	<i>fradru risi ru</i>	fradru rir ru
13	<i>fradru risi tur</i>	fradru rir tor
14	<i>fradru risi gigiti</i>	fradru rir gegte
15	<i>fradru risi fradfid</i>	fradru rir fradfid
20	<i>matutni</i>	tmatutni
21	<i>matutni esu</i>	tmatutni eso
26	<i>matu tni fradfid tresu?</i>	
30	<i>matutni risi esu</i>	tmatutni ru nemet (?) (**tmatutni eso nemet)
31	<i>matutni risi esu risi esu?</i>	
40	<i>matut tni ru</i>	tmatutni ru
50	<i>matut tni ru risi esuem</i>	tmatutni ru nemet
60	<i>matut tni ru risi ru</i>	tmatutni tor
70	<i>matut tni tor</i>	tmatutni tor nemet
80	<i>matut tni fradfid</i>	tmatutni gegte
90	<i>matut tni fradfid ri si esuem</i>	tmatutni gegte nemet
100	<i>ratu esu</i>	ratu eso
101	<i>ratu risi esu</i>	
110	<i>ratu risi fradru</i>	
200	<i>ratu rifu</i>	
1000	<i>rifu esuem (←rifunú)(?)</i>	

TABLE 3.15 – Irarutu numerals. Plain font items were collected in 2010, items in *italics* were confirmed in 2013

3.1.3.2.4 Classifiers

Classifiers are required when a noun is modified by a numeral, and occur in some instances when a generic noun occurs without a numeral. The classifier that occurs with a given noun is selected by that noun. Classifiers may be semantically arbitrary, especially when more than one classifier may occur with a given noun.

CLF	GLOSS	Example	Gloss
<i>brimt</i>	‘load (of firewood)’	<i>ema brimt rifu</i>	‘two loads of firewood’
<i>fin</i>	‘school (of fish)’	<i>sum fin</i>	‘school of fish’
<i>jagffu</i>	‘bunch’	<i>fud jagffu</i>	‘hand of bananas’
<i>je</i>	‘meat (lit. contents)’	<i>je mug</i>	‘a lot of meat’
<i>fu</i>	‘fruit’	<i>umagi fu</i>	‘coconut fruit’
<i>fu</i>	‘piece’	<i>efute</i>	‘a piece of wood’
<i>tni</i>	‘body’	<i>fun tni tur</i>	‘three dogs’
<i>fru</i>	‘tree’		
<i>rue</i>	‘liquid’		
<i>fatn</i>	‘stick’		
<i>tbban</i>	‘packet’		
<i>sü</i>	‘tail’		
<i>bge</i>	‘pair’		
<i>rre</i>	‘date’	<i>rre tur fimta</i>	‘three days hence’
<i>gun</i>	‘day’		

TABLE 3.16 – Fifteen Irarutu classifiers.

Classifiers occur after a countable noun when one is quantified by a specific number or determiner, but can also occur without a quantifier or determiner; however, not all count nouns need to occur with a classifier, e.g. *mo* ‘child’ (see example 224) and *seba* ‘month’. To give a sense of the range of semantic noun classes that occur with classifiers, TABLE 3.16 lists some Irarutu classifiers and example phrases.

- 224) *mo* Ø *ru*
 child [CLF] two
 ‘two children’

There are fifteen classes of countable nouns that select classifiers, for example, *fu* ‘fruit-like’, *sü* ‘tail (as in three tail of cattle)’, and *tni* ‘body’, examples are shown in 225)–229).

- 225) *kkor sü tur adini*
 chicken tail three DET
 ‘these three [tail of] chicken’
- 226) *matu tni ru*
 people CLF two
 ‘two people’
- 227) *umagi fu esuem*
 coconut CLF one
 ‘one coconut fruit’
- 228) *umagi fru adi*
 coconut CLF DET
 ‘the coconut tree’

- 229) *san nabad sguifr timebe adei*
 house ADJ DEG QUANT DET
 ‘all [of] the really big houses’

3.1.3.3 Verb adjuncts

In Irarutu, two verb adjuncts form phrasal constituents with verbs, auxiliaries (some researchers prefer ‘auxiliary verbs’) and verbal particles. Cross-linguistically, verbal particles are selected to co-occur with certain verbs and contribute to the meaning (‘wake-up’, ‘turn-off’); but the particles are not always required to be adjacent to the verb (‘switch [the light] on’). They have distinguishable locative or directional semantics, but can also be idiomatic (‘hurry up’). In different languages verbal particles can have lots of different semantics, e.g., position, evidentiality, and so forth. They are frequently derived from historical adpositions, but in Irarutu, they appear closely related to predicates in serial verb constructions.

3.1.3.3.1 Tense, aspect, and modality auxiliaries

Auxiliary verbs are traditionally seen as helping verbs such as ‘be, do, have, can, could, must’ in English, that fulfill grammatical functions, such as expressing contrast in tense (past, present, future), aspect (completed, ongoing), and modality (possibility, probability, obligation). Auxiliaries can also express the polarity (positive or negative) of the verb they are associated with. Voice is not distinguished in Irarutu. Sequences of two or more auxiliaries are allowed in prescribed order. Both inflected and uninflected auxiliaries occur before the verb reflecting the order of verb before object. Position in relation to the verb can be different in main versus subordinate clauses. Negation belongs to a distinct category of negator.

The Fruata dialect does not have a distinct tense marking system but relies instead on combinations of aspect and modality. This can be construed to mean that verbs are inherently unspecified for tense. For example, the term *n-in* can mean ‘he drank’, ‘he drinks’, or ‘he is drinking’. Tense, aspect, modality (TAM) markers indicate properties of the temporal-conceptual structure of verbs and some non-verbal predicates.

Occurrence accompanied by TAM markers is a sure sign that a novel word is a verb. Due to the elusive nature of the TAM system, some typological resources are outlined below. The reader is reminded that the description is focused on the behavior of the tense, aspect, and modality morphemes with the goal of formulating means to evaluate verbs as a syntactic category.

The analysis by Matsumura and Matsumura 1991 identifies five TAM markers *e* ‘Past’, *bu* ‘already (Progressive)’, *du* ‘not yet ()’, *su* ‘then (Completive)’, and *g* ‘want (Volitional)’. The status of *e* ‘Past’ is questionable. It is not included in TABLE 3.17. The primary consultant for this research was not familiar with this element and never used it on his own. It seems likely that it is a dialectal feature of the East Arguni variety researched by the Matsumuras. Data suggests that *mo* ‘non-completive/irrealis’ and *te* ‘possibility(?)’ are also TAM markers. The progressive *bu* and the *du* are in complementary distribution. These two never co-occur. The remaining TAM markers are combined in a variety of ways to specify temporal-conceptual information.

Aspect		Modality
<i>du-</i>	Non-progressive 'as yet/not yet'	<i>g-</i> Desiderative
<i>bu-</i>	Progressive 'already'	<i>mo-</i> Irrealis
<i>su-</i>	Completive 'then'	

TABLE 3.17 – Aspect and Modality markers. Matsumura (1991) calls the desiderative 'volitional'. Voorhoeve (1995) calls it 'intentional'.

There is disagreement in the literature regarding the status of TAM markers. They have been analyzed as both prefixes (Matsumura and Matsumura 1991) and particles (Voorhoeve 1989). There is sufficient data to propose TAM markers are independent words, in particular, auxiliaries. It is clear that tense-mood-aspect markers occur before the verb.

- 230) *g-a=gin kopi fade*
DES-1=SAP.drink coffee only
'I just want to drink coffee'

In 229), the desiderative morpheme, the subject marker, and the verb are spoken and orthographically represented as a single word by the primary native speaker consultant. By contrast, the general trend is to write TAM markers as separate words. Only the desiderative morpheme *g-* is inflected for subject by suffixes, *ga* 'I want/will', *go* 'you want/will', and *ge* 'he wants/will'. The verb is also conjugated to agree with the subject by prefixes, *a-* '1S', *o-* '2S', and *i-* '3S' when it occurs with the desiderative marker. In several contexts, the prefixes are considered optional. The full structure of the verb complex in 229) is therefore *g-a a-gin*, which, if it were only one word, *g-a-a-gin* is unexpected because there are two adjacent elements that mark the same argument. The third person plural form *ire genin kopi* 'they want to drink coffee' also shows an agent marker *-e* after the desiderative *g-*, and an agent marking prefix *n-* hosted by the verb *in* 'drink'.

The TAM markers should be considered independent elements based on phonological behavior and two grammatical properties: inflection and distribution. At least one TAM marker, *g-* 'Desiderative', is inflected – although in an exceptional way – to match the person feature of the subject. Due to the location of the inflectional element as a suffix, for example first person *-a*, which can precede a verb inflected for a first person subject by *a-*, there is a pause between the two (adjacent) inflectional elements. On the other hand, there is some evidence that TAM markers could be considered prefixes. Phonologically, TAM markers tend to be parsed together with the verb, but this appears to be due to contraction. Insofar as distribution, TAM markers have specific ordering requirements in relation to one another, and they can only be separated from the verb by auxiliaries. In addition, Irarutu speakers variably write TAM markers as part of the verb or as a separate word containing all the TAM markers that are present in the clause, but

representing each TAM marker independently is rare. In this description, TAM markers are analyzed as words, not affixes.

3.1.3.3.1.1 Tense

Verbs, by default, are not marked for past or present tense. Matsumura and Matsumura (1991) and Voorhoeve (1989) cite a past tense marker *e*, which precedes the verb and any other TAM markers, as in 231), or *fa* ‘I go’ and *e fa* ‘I went’.

- 231) *i e-samt [sum weni]*
 3S PST-catch fish DET:PL
 ‘He caught some fish.’ (Matsumura & Matsumura 1991:100)

Irarutu consultants for the present research did not consistently recognize this element in discussing language data, nor could they use it in spontaneous speech, when asked about it. It is sufficient to mention that a past tense marker has been described in the literature, but the data collected in the field does not lend itself to a closer description. Future tense is sometimes indicated with the irrealis marker *mo-*. Habitual tense ‘always’ is achieved through verb reduplication or the verbalizing morpheme *na-*.

3.1.3.3.1.2 Aspect

There are three aspect markers, *du-* ‘already am (progressive)’, *bu-* ‘not yet ()’, and *su-* ‘done (completive)’. *Du* and *bu* never co-occur, but both can co-occur with, and precede, *su*. Aspect markers occur before the verb, and tend to precede the modality marker, if one is present. Although interpreting *du-* ‘Non-progressive’ and *bu-* ‘Progressive’ is sometimes problematic, aspect markers occur frequently, and so they are a fairly reliable diagnostic for identifying a verb. Exceptions can be explained as examples of dropping in conjoined clauses, i.e., conjunction reduction.

3.1.3.3.1.2.1 Non-Progressive *du*

The first aspect marker to be described is the non-progressive, *du-* ‘already’, which specifies that an act or state has occurred prior to the time frame of the discourse. Compare *rga* ‘life’ and *du-rga* ‘living (lit. already alive)’. As mentioned previously, *du* does not co-occur with *bu-*, the progressive, but on the other hand, aspect markers for progressivity can occur with aspect markers such as completive aspect *su*.

- 232) *ja du suft-fe*
 1S still smoke-DTR
 ‘I am still smoking’
 233) *ja du tim frufn*
 1S already close door
 ‘I closed the door already (the door is closed)’

When *du-* appears in the same phrase with *=ti* ‘Negative’, as in *du-ntün=ti* ‘uncooked’, the resultant meaning is ‘not yet’. *Du ntün=ti* literally means ‘already cooked not’. Additional examples are given in 234–235), below. In the present context, it is worth noting that Irarutu uses a VP to express concepts that could be seen as adjectives. This

shows overlap between verbal and adjectival predicates, but distinctness from modifying adjectives.

- 234) *du-si=ti*
 already-see=NEG
 ‘unseen (lit. not yet seen)’
- 235) *ja du-kka=ti*
 1S already-showered=NEG
 ‘I haven’t yet showered’

TAM appears to occur with less prototypical predicates, such as the preposition *nene* ‘to’, or the conjunction *nir* ‘with/and’. In example 236), below, the TAM marker occurs between the agent marker and the preposition as a single word. This representation aligns with an analysis where TAM are bound to morphemes. But, based on this characterization, TAM markers appear to be clitics rather than affixes, which would be considerably more selectional, in which case they could not attach to non-verbs.

- 236) *ja-du=nene matu wenum*
 1S-TAM=PREP person land
 ‘I’ve already gone to other territories/lands’

A final property that must be addressed is that the progressive aspect marker reveals one area of overlap between verbs and adverbs. *Du* can combine with *-gag* to form *dugag* ‘still/not yet’, an alternate of *mse* ‘still/not yet’.

3.1.3.3.1.2.2 Progressive (realis) *bu*

In contrast with non-progressive *du*, the progressive *bu* ‘still/not yet’ indicates that a state or event is occurring during the time frame of the speech event, although it may have begun before that time frame. The progressive and the non-progressive occupy the same pre-verbal slot in the verb complex and are in complementary distribution. Another constraint on *bu* is that it cannot co-occur with a reduplicated root such as **butımtım* ‘not yet closed’, suggesting that some uses of reduplication in verbs contain aspectual information.

There is an option for combining the progressive and the completive. This combination of aspect markers is discussed below. Several examples containing the progressive aspect marker *bu* are given below.

- 237) *a bu tım frufn*
 1S still close door
 ‘I am still closing the door’
- 238) *o=ro bu kka?*
 2S=TOP not.yet shower
 ‘Haven’t you showered yet?’

- 239) *sbajan bu rfan*
 prayer PROG begin
 ‘Prayer service has already begun (and is still going-on)’
- 240) [*frrufn ad-ini*]=*ro bu m-tim ad*
 door DET-DEM=TOP PROG STAT-close DET
 ‘The door is still staying closed’

One use of the progressive, the prohibitive, stands out because it violates the general distributional trend of aspect markers that occur before a mood marker. The progressive idiosyncratically follows the inflected desiderative mood marker to form negative desire: *gabu* ‘I don’t want to...’, *gobu* ‘you don’t want to...’, and *gebu* ‘s/he doesn’t want to...’.

3.1.3.3.1.2.3 Completive *su-*

The ‘Completive’ (punctual) marker *su-* follows any other aspect morpheme, if one is present. Whereas *du-* indicates an action is on-going, and *bu-* indicates that something is not happening presently, *su-* indicates that an activity, state, or event has come to completion/is punctual.

- 241) *ja-su tim frrufn*
 1S-CPL close door
 ‘I already closed the door’
- 242) *bis ta- su rifr?*
 may INCI-CPL request
 ‘may I ask?’
- 243) *rre fi e=bi-janji nene refefa adi=ro bu-su nenot*
 time REL 2S=2.ACT-promise PREP yest. DET=TOP PROG-CPL change
 ‘the time which you promised yesterday has changed’

Multiple aspect markers can co-occur in a single construction. One constraint is that the non-progressive marker *du* and the progressive marker *bu* are in complementary distribution. The completive marker always occurs after one of the progressivity markers. The observed combinations of aspect markers are non-progressive plus completive, *dusu*, and progressive plus completive, *busu*. The progressive completive is roughly glossed ‘later’, and is illustrated in 244) below.

- 244) *it ga-fe is, si it bu-su fier-fe*
 1Pn eat-DTR prior, then 1Pn PROG-CPL speak-DTR
 ‘Let’s eat first, then later we can speak’

The combination of *bu* ‘progressive’ with *su* ‘completive’ requires a particular type of context. The presence of the linker *si* ‘if’ indicates that the TAM in the linked clause stands in relation to the temporal structure of the main verb *gafe*, and attains a completive semantic, something like ‘(the process is) over and done’.

3.1.3.3.1.3 Modality

Irarutu has two modality markers: ‘desiderative’ *g-*, which follows a minor pattern of subject agreement that is unlike any other inflectional paradigm, and ‘irrealis’ *mo*.

3.1.3.3.1.3.1 Desiderative *g-*

To indicate desiderative mood, the element *g-* prefixes to the verb, or suffixes to a preceding tense or aspect marker, if one is present. Examples 245–6) represent the primary pattern, but structures such as 247), where the agreement marker *a-* ‘first person subject’ prefixes rather than suffixes to the desiderative TAM auxiliary, also exist, but are rare.

- 245) *ja g-a ra-run fun nino fi bu=n-tim frrufr adi?*
1S DES-1 1-know dog where REL PROG=3-close door DET
‘I want to know which dog caused the door to close.’
- 246) *Jek g-o [fa] m-ir ir?*
Jack DES-2 [go] 2-with 3P
‘Jack, do you want to [go] with (i.e., accompany) them?’
- 247) *a=ge tim frrufn*
1S-DES close door
‘I want to close the door’

3.1.3.3.1.3.2 Irrealis *mo*

In addition to the progressive/non-progressive, desiderative, and completive, there appears to be an irrealis marker, *mo*, that is used similarly to an auxiliary ‘want’, but this morpheme is challenging to analyze. Its use is illustrated in 248)–249).

- 248) *ja mo bitr=ti*
1S IRR hungry=NEG
‘I’m not yet hungry’
- 249) *ja mo-rarun ge radni adeji rote kabia*
1S IRR-know COMP road DET COP ruined
‘I do not (yet) know why the road is closed’

An example cited from Matsumura and Matsumura 1991 shows four TAM markers (past, progressive, completive, and desiderative), 250). The sequence *e-bu-su-g-n-ma* is translated ‘he had already wanted to come’, but the Fruata dialect does not have *e*.

- 250) *i-ro e-bu-su-g-n-ma kampun*
3S-FOC PST-PROG-COMP-DES-3S-arrive village
‘He has already then desired he (i.e., decided to) come to the village.’
(He has not arrived in the village. His decision was in the past. Maybe he has already left for the village.) (Matsumura and Matsumura 1988:87)

3.1.3.3.2 Other auxiliaries and verbal particles

In addition to TAM auxiliaries, Irarutu also has other auxiliaries, such as ‘helping-verbs’ (FIGURE 3.13). The four most frequent helping verb auxiliaries are *fa* ‘go’, *fun* ‘cause’, -

ar ‘give’, and *-ud* ‘take’. All four of these elements host agreement morphology and are therefore seen as fully-inflecting auxiliaries. Other auxiliaries, many of which, *bis* ‘can’ or *suk* ‘like’, do not host agreement morphology.

<u>‘Helping-verb’ auxiliaries</u>			
<i>fa</i>	‘go’	<i>tram</i>	‘can’
<i>fun</i>	‘cause/make/do.to’	<i>mtat</i>	‘can’
<i>-ar</i>	‘take’	<i>suk</i>	‘like/want’
<i>-ud</i>	‘give’	<i>suk ti</i>	‘dislike’
<i>-mi</i>	‘be/stay’	<i>fisuk</i>	‘want’
<i>fidap</i>	‘can’	<i>suk dir</i>	‘eager’
<i>nanfidap</i>	‘get’	<i>tufi</i>	‘like’
<i>bis</i>	‘can’	<i>tufiti</i>	‘dislike’
		<i>-ge</i>	‘give.to’

Idiomatic Auxiliary Phrases

fta mtut (nir) ‘want’ (lit. ‘stomach wants (with)’)
nfa nfi du ‘he wants’ (*n-fa n-fi du* ‘he-goes he-which already’)

FIGURE 3.13 – Irarutu auxiliaries illustrating ‘helping-verb’ auxiliaries and idiomatic auxiliary phrases.

A notable feature of some verbs that can also function as an auxiliary, such as *ma* ‘come’, is their frequent occurrence in serial verb constructions. Because they are part of a discontinuous predicate, there is an ordering constraint that locates them after objects (251a). Reversing the order, as in 251b) is ill-formed.

- 251a) *Arnol ba gud rimun ma-ma*
 A. 2.go SAP.take lemon 2-come
 ‘Arnol, go bring me a lemon’
- 251b) **Arnol ba gud ma-ma rimun*
 A. 2.go SAP.take 2-come lemon
 ‘Arnol, go bring me a lemon’
- 252) *m-ar buku ma-mri mtu*
 2-bring book 2-COME 2.save
 ‘Go put your book away (in the house)’
- 253) *g-ud tab ma-ma ja isi*
 SAP-take add 2-come 1S see
 ‘bring it here for me to see’

3.1.3.3.2.1 Other auxiliaries

Fully inflecting auxiliaries can accept agent marking prefixes as well as the verbalizer *fi*. These auxiliaries can be divided into two classes, one that hosts the irregular agreement markers, and one that accepts the standard agreement markers, see **FIGURE 3.14**.

auxiliary uses of *nar* ‘to give’

<i>nar fisigr</i>	‘edge (give edge)’
<i>nar inantut</i>	‘lift/elevate (give push)’
<i>nar finanei</i>	‘aid (give help)’
<i>nar nan fa</i>	‘lead (give go)’
<i>nar rrir sigt</i>	‘lift/boost (give energy)’
<i>nar nane</i>	‘giver (give give.to)’
<i>nar nanei</i>	‘exchange/submit/leave (give give.to.it)’

verbal uses of *nar* ‘to give’

<i>nar fdfadr</i>	‘litter (give rubbish <u>alas</u>)’
<i>nar mufni</i>	‘nurse/medicate (give medication)’ <i>ja mar mufni</i> ‘I nurse’
<i>nar snan frfier</i>	‘describe/view’ (give words speak)
<i>nar sus</i>	‘nurse/suckle (give breast)’
<i>nar wenga</i>	‘nourish (give food)’

FIGURE 3.14 – Examples of *nar* ‘to give/bring/take’ as an auxiliary and as a verb.

The root *-ar* means ‘use/bring.to’ and can be observed as both an auxiliary as well as a main verb. As an auxiliary (254–258), it behaves similarly to a serialized verb by occurring as the first element in the verb complex; but, as a main verb, it can stand independently and accept *fi-* (258).

- 254) *ruguru mumri n-ar=ja na-n-fa n-fi (nene) matu wenum*
k.o.wave season 3-bring=1S HAB-3-go 3-ACT (to) person land
‘the season caused me to go to their land’
- 255) *j=ar kirrar ra-si ja*
1S=use mirror 1-see 1S
‘I use a mirror to see myself’
- 256) *o=ro m-ar nya fi o ma-msi o adej?*
2S=TOP 2-use what REL 2S 2.come-2-see 2S DET
‘what do you use to see yourself?’
- 257) *i n-ar warada na-n-tut sum*
3S 3-use knife HAB-3-cut fish
‘he uses a knife to cut fish’
- 258) *ja fi-ar wenga ra-ge o-matu uf m-tim frrufn*
1S ACT-give food 1S-to 2-person REL STAT-close door
‘I give food to your friend who closed the door’

Another auxiliary that is inflected using an irregular pattern for agreement is *-ud* ‘take/give’. Example 259) shows that the subject is marked on the auxiliary following the irregular, speech act participant distinction pattern (cf. ‘eat’, ‘drink’, ‘sleep’) when it occurs after ‘go’; compare *-ud* conjugated for ‘you’ *g-* versus ‘he’ *n-*. However, there are contexts such as 260–61), where *-ud* follows the main pattern for inflection when it is the first verb.

- 259) *Arnol ba=gud rimun ma-ma*
 A. 2.go=SAP.take lemon 2-come
 ‘Arnol, go bring a lemon here’
- 260) *n-ud-n-fign*
 3-take-3-plant
 ‘He is planting (lit. ‘he-give-he-plant)’
- 261) *m-ud=bign*
 2-take=2.plant
 ‘You are planting (lit. ‘you-give-you-plant)’

Examples of ‘helping-verbs’ that are not inflected include *tufi* ‘like’, *tufiti* ‘dislike’, and *bis* ‘can’.

- 262) *ja tufi g-a mi Bintuni*
 1S like DES-1 stay Bintuni
 ‘I like to live in Bintuni’
- 263) *i=ro bis dru man*
 3S=TOP can 3.shoot bird
 ‘he can shoot birds’

3.1.3.3.2 Verbal particles

In addition to auxiliaries that can also function as main verbs, there are verbal particles that relay a strong, albeit idiomatic, directional component, such as *fa* ‘go’, *iet* ‘go up’, *ri* ‘go across’. Irarutu has an elaborate directional/location/lative system (i.e., directional predicates/adverbials that indicate the origin and direction of certain activities) that uses elements listed in **FIGURE 3.15**. These elements can be analyzed as auxiliaries in some contexts, but must be seen as verbal particles in others. For example, unless there are extraordinary circumstances, all verbs of movement from location A to location B are marked with directional information.

<i>fa</i>	‘go’
<i>iet</i>	‘up’
<i>ri</i>	‘go across’
<i>ro</i>	‘go down’
<i>mri</i>	‘go in (inessive)’
<i>ma</i>	‘come’
<i>briet</i>	‘come up’
<i>bri</i>	‘come across (come in)’
<i>bro</i>	‘come down’

FIGURE 3.15 – Some Irarutu directional verbs.

There are inflected auxiliaries that accept the standardized agreement markers, in particular, *fa* ‘go’, *fun* ‘do/cause’, and several epistemic auxiliaries, such as *bis* ‘can’.

The verb *fa* ‘go’ is a main verb that frequently also functions as an auxiliary and can occur as a verbal particle. As an auxiliary, it expresses progressive aspect/immediate future tense. In examples such as 264), *fa* is optional, but its use is considered more grammatically complete.

- 264) *it (fa) dritn dedan*
 1Pn (go) clean.up garden
 ‘We are going to clean up the garden’
- 265) *g-a-su ma fa-ra-fi adeji*
 DES-1-COMPL come go-1-ACT DET
 ‘please permit me to pass’
- 266) *ja tufiti nene fi (j)a fa mi adi*
 1S dislike PREP REL 1S go stay DET
 ‘I dislike where I live’
- 267) *ja du nene matu wenum*
 1S NPROG PREP person land
 ‘I already [went] to their land’

Another main verb that can also function as an auxiliary is *fun* ‘do/make/cause’. It accepts agreement markers used for the main pattern and can combine with auxiliary *fa* ‘go’.

- 268) *i=ro n-fun kapar nene wenum fade*
 3S=TOP 3-do chief PREP land just
 ‘he is a village leader’
- 269) *it fa=fun mir ne dedan*
 1Pn go=do clear PREP garden
 ‘We are going to clear forest to make a garden’
- 270) *ja fun-(nya) nyati*
 1S do-AGR nothing
 ‘I’m not doing anything’

3.1.3.4 Conjunctions

Conjunctions join words, phrases, or clauses. Iraputu has coordinating and subordinating conjunctions. Coordinating conjunctions indicate equal rank between constituents, e.g., *nir* ‘and’, *te* ‘or’, *mo* ‘but’ (different from *te* ‘possibility’ and *mo* ‘irrealis’).

Subordinating conjunctions indicate unequal rank and include words such as *ere* ‘whether/if’, *ge* ‘that’, and *esi* ‘conditional’.

3.1.3.4.1 Coordinators

Coordinating conjunctions occur between conjoined elements. The conjunction *nir* ‘and/with/accompany’ can be used as a coordinating conjunction or a preposition ‘with’ to introduce an oblique argument. *Nir* is more closely associated with the constituent that follows it, shown by being separated from the preceding element by slight pause. This location is correlated with SVO word order, and is expected for all but verb final languages. Correlative (paired: both–and, *esi* ‘either’– *mo* ‘but/or’) elements each

precede the associated conjunct. Correlatives typically repeat same conjunction (*esi* ‘either’ – *esi* ‘either’). Nouns and noun phrases can be conjoined, as well as verbs and VPs, adjectives, and adverbs. The more common variant *n-ir* ‘I accompany/he accompanies’ is conjugated for a second person agent as *m-ir* ‘you accompany’. An alternate strategy is concatenation, which is especially common with verbs and VPs. Coordination by verb ‘be with’ involves serial verbs.

Three coordinating conjunctions are important, *nir* ‘and’, *mo* ‘but’, and *te* ‘or’. Several examples suggest that conjunctions can be morphologically complex. For example, the default conjunction is *nir* ‘and/with/by’; however, when the initial conjunct is a second person referent, the conjunction is conjugated *m-ir*.

- 271) *fa nir fra*
feet CONJ hand
‘feet and hands (extremities)’
- 272) *met (nir) met*
little (CONJ) little
‘bit-by-bit (lit. little and little)’
- 273) *Mikel nir ja*
Michael CONJ 1S
‘Michael and I’
- 274) *Jek, o te nir kopi nir wer*
Jek, EXST tea CONJ coffee CONJ water
‘Jack, there is tea, and coffee, and water’

In the following example, *-ir* functions as a predicate: it is preceded by the modal adverb *g-* and is inflected for a second person agent *m-*.

- 275) *g-o m-ir ir?*
DES-2 2-accompany 3P
‘do you want to go with them?’

It is possible to use the verb *fi sama* ‘be with/along with/same with’ to form a construction that is equivalent to conjunction.

- 276) *ja fi sama nir i*
1S ACT be.with CONJ 3S
‘he and I (lit. I along with him)’

In questions, there is a difference between ‘inclusive-or’, that is, cases where each of the conjoined elements as well as combinations of them are conceptually allowed answers, and ‘exclusive-or’, where only one of the conjoined elements – to the exclusion of the others – is allowed (McCawley 1981). Example 277) shows ‘exclusive-or’ used in a list. The disjunctive, *te* is phonologically packaged with the preceding noun. Because *te* occurs after each element in a string of conjoined nouns, including the final noun, ‘exclusive-or’ occurs in a postpositional location.

- 277) *te te, kopi te, werwin te*
 tea DISJ coffee DISJ hot.water DISJ
 ‘tea, coffee, or hot water’

Inclusive-or uses *te* between the first two elements, but coordination uses *nir* before the final conjunct: *Damianus, Willy, nir Moses* ‘Damien, Willy, and Moses’.

The disjunct is *mo* ‘but’. It tends to link clauses and indicates that some aspect of the dependent clause is contrary to the main clause.

- 278) *motor adini=ro dir mo motor refiden mai fi buna dir sfar*
 moped DET=TOP good CONJ moped side DIST REL most good more
 ‘this motorcycle is better than other motorcycles.’

3.1.3.4.2 Subordinators

Subordinating conjunctions integrate a subordinate clause into a larger construction.

They are prepositional in Irarutu. There are three classes of subordinators:

complementizers (mark clause as complement of a verb, often means ‘say’, not restricted to indirect quotation), relativizers (mark relative clauses, abbreviated RCs), and adverbializers (mark that the clause as having an adverbial function: time, purpose, result, etc.). Cross-linguistic alternatives to complementation are to not mark the complement clause, or to nominalize it (nominalize verb and mark subject as possessive).

Relativizers, which according to Keenan (1985) just indicate that there is a RC, are not the same as relative pronouns, which function as a nominal in a RC. Alternative strategies are to use special relative verb forms or just not mark the RC. An adverbializer in a subordinate clause can be paired with a conjunction in the main clause. Examples of apparent juxtaposition to indicate subordination can be attributed to deletion of the overt subordinator.

There are three important subordinators, the complementizer *ge* ‘that’, and the relativizers *uf* ‘that/who’ and *fi* ‘which’ (which never have a nominal function within a RC). The complementizer *ge* is used to introduce most complement clauses, including indirect speech, i.e., quoting another’s words. The use of *ge* for indirect speech is shown in the first instance of the word in 279).

- 279) *Matu polisi n-fier ge [ja mo-ra-run [ge radni adeji rote kabia]].*
 person police 3-say COMP 1S IRR-1-know COMP road DET COP broken
 ‘The policeman said “I do not know why the road is closed”’

In addition to a complementizer and a relativizer, Irarutu has several resources to link subordinate clauses to main clauses. A detailed discussion of these elements and the nuances in their use remains for future research. Single word clause subordinators include *nirge* ‘because’, *sefu* ‘so that (because)’, *na* ‘resultative’, *nate* ‘just/newly’, *si* ‘then after’, and *esi* ‘even if’.

- 280) *ja tese a-ra-fa-n-fa-g [nirge motor adi kabia]*
 1S must 1S.POSS-1-go-3-leg-1.POSS because motorcycle DET broken
 ‘I must walk because the motorcycle has a problem (is broken)’
- 281) *ja g-a kka [sefu a-gin]*
 1S DES-1 bathe so.that 1-SAP.sleep
 ‘I want to bathe so that I can go to sleep’
- 282) *ja g-a-bu [na nyaunt gor ja uf]*
 1S DES-1-PROG if ants bite 1S PROH
 ‘I don’t like to be bitten by ants’
- 283) *ja fa trfat fud nene pasar [nate ja mgür funya fi a-si si-dir]*
 1S go find banana PREP market CONJ 1S buy some REL 1-see PL-good
 ‘I looked for bananas at the market and I bought some which looked good’
- 284) *Jek=ro rror ro-ge o=ro bu=ba [esi taje kabia]*
 Jek=TOP tomorrow COP-COMP 2S=TOP PROG=2.go even.if weather bad
 ‘Jack, tomorrow you must be going even if the weather is bad’
- 285) *it ga-fe is, [si it bu-su fier-fe]*
 1Pn SAP.eat-DTR prior, COND 1Pn PRO-CPL talk-DTR
 ‘let’s we eat first, then we talk (some more)’

The subordinators *na* ‘resultative’ and *te* ‘later’ do not appear to be related. However, there is a subset of clause linkers that appear to be composed from a few basic roots, such as *se* ‘temporal linker’ and *mo* ‘disjunction’, the meaning of which is integral to the meaning of the resulting word, for example, *mse* ‘until’, *tese* ‘must/surely’, *temo* ‘consequently but’, and *femo* ‘unexpectedly but’.

- 286) *Willy=ro n-fir=ge Moses tese nir-i [mse na-n-fa Fakfak]*
 Willy=TOP 3-say=COMP Moses must with-3S later HAB-3-go Fakfak.
 ‘Willy wants Moses to go (with him) to Fakfak’
- 287) *[mo adi]=ro n-gigr ge i-den [tese n-fi-tab sen na-n-fun i]*
 child DET=TOP 3-ask COMP 3S.POSS-mother must 3-ACT-add money HAB-3-make 3S
 ‘the child_i asks his mother to give him_i some money’
- 288) *Wer-kuri=ro i=ro safatero t-rir, [temo tot awesi n-gerfa]*
 water-Kuri=TOP 3S=TOP well.up INCI-wrap CONJ moment suddenly 3-return
 ‘The flowing Kuri river meanders away but at a later moment suddenly returns.’
- 289) *ja=mgür sirin funya fi a-si si-dir,*
 1S=buy papaya QUANT REL 1S-see PL-good
[femo a-ra-ma tut a-si nyaunt rarur (nene) gan]
 DISJ 1S-1-come cut 1S-see ant many (PREP) inside
 ‘I bought some papaya which looked good, but when I cut them open there were a lot of ants inside them.’

Another strategy to link clauses is to use a discontinuous marker. Several of the elements that can be used individually to conjoin a dependent clause can co-occur in a complex clause to introduce the main clause, with or without another conjunction to introduce the dependent clause. For example, the disjunctive coordinator *mo* can introduce a dependent clause that is introduced by *we* or *si*, to form a conditional ‘if...then’

construction. The word *ere* ‘whether/if’ can be used in conjunction with the adverb *msse* ‘then/until’ as a discontinuous marker *ere...msse* ‘if...then’ to function as a temporal conditional.

- 290) *we taje kabia [mo Jek rror ro-ge o bu-ba]*
 although weather bad DISJ Jack tomorrow COP-COMP 2S PROG-2.go
 ‘although the weather is bad, but, Jack/, tomorrow you must depart’
- 291) *rror si taje kabia [mo Jek=ro tese o=bu-ba]*
 tomorrow COND weather bad DISJ Jack=TOP must 2S=TAM-2.go
 ‘tomorrow, even if the weather is bad, but Jack you must depart’
- 292) *mo e=re bitr=ti, [ro e-ga-of]*
 COND 2P=TOP hungry=NEG, then 2P-SAP.eat-PROH
 ‘If you all are not hungry, then don’t eat’
- 293) *ere du-bitr=ti, [msse e=bu-ga-fe]*
 if NPROG-hungry=NEG later 2P=PROG-SAP.eat-DTR
 ‘If you all are not hungry yet, then eat later’
- 294) *ge e=re bitr, [ro-ge e=ga wenga fi a-tiin adi]*
 COMP 2P=TOP hungry COP-COMP 2P=SAP.eat food REL 1-cook DET
 ‘if (it’s the case that) you are hungry, then eat the food I cooked’

In some cases, the dependent clause marker can be deleted.

- 295) *er du bitr [e-ga timadbe]*
 if NPROG hungry 2P-SAP.eat QUANT
 ‘If you are still hungry, eat everything’

3.1.3.5 Miscellaneous parts-of-speech

According to Schachter (1985), there are three miscellaneous parts-of-speech that can be identified in Irarutu: 1) clitics (=ro, =ti, and =e/o), 2) the copula *rau* and existential marker *o*, and 3) interjections such as *i* ‘ouch’.

3.1.3.5.1 Clitics

The current understanding of clitics is that there are two major types: simple clitics and special clitics. Clitics tend to be phonologically light, i.e., short and unstressed (see Chapter 2). A simple clitic is a frequently contracted, or phonologically dependent, version of a full syntactic element. Cross-linguistically, it is common for personal pronouns to cliticize more than other types of elements. This is true for the Irarutu subject pronoun *ja*, because it is paradigmatically an independent pronoun, corresponding to *o* ‘you’ and *i* ‘(s)he’, but tends to attach phonologically to the following verb complex. Importantly, this type of contraction is semi-independent of speed-of-speech (casual versus careful). Van den Berg and Matsumura (2008) even analyzed the agreement marker on the verb as a clitic rather than a prefix.

Special clitics, on the other hand, occur in a fixed position in relation to some other sentence element, often a phrasal constituent, but are non-selectional, that is, they can attach to many different types of hosts, at the word level (Zwicky 1985). The clitics (=ro

‘Topic Marker’, =*ti* ‘Negative marker’, and =*e/o* ‘Politeness marker’) in Irarutu are enclitic, meaning they follow their host. Special clitics have fixed position in relation to one another. For example, the politeness clitic always occurs after a negative clitic. Clitics are known to cross-cut other syntactic categories that are defined largely by function because of the unique positional constraints that define clitics. In other words, clitics are challenging to define on functional grounds. There is a distinct negative imperative -*of*, but this appears to suffix to verbs. Politeness markers indicate a deferential stance towards the addressee and support the cross-linguistic tendency to have a polite form for ‘you’.

3.1.3.5.1.1 Topic marker =*ro*

The morpheme =*ro* is a discourse marker that indicates ‘topic’. It seldom occurs with first person pronouns, but nearly always occurs with second person pronouns. The variant ‘=re’ is found with second person plural subjects. It is enclitic to the last element in a noun or noun phrase that bears the ‘topic’ role and can be doubled, i.e., occur more than once in a clause. Topics are usually fronted. The examples below show that =*ro* can attach to a range of elements, including nouns, determiners, and verbs.

- 296) [*mikel*]=*ro na adeni*
 mikel=TOP N-SAP.eat DET
 ‘Michael ate it’
- 297) [*rre adi*]=*ro bu didirie*
 sun DET=TOP PROG lean
 ‘the sun is setting’
- 298) [*matu n-fi-jual adi*]=*ro na-n-run=ge nyaro adeji rau du-metyet*
 person 3-ACT-sell DET=TOP HAB-3-know vegetables DET COP NPROG-fresh
 ‘The vendor knows that the vegetables are fresh’
- 299) [*mumce fi i n-yuen dub adi*]=*ro buna win*
 morning REL 3S 3-weave noken DET=TOP most hot
 ‘the morning when he made a noken (k.o. hand woven shoulder bag) was hot’
- 300) [*wega uf a-den n-tiin*]_o=*ro ar ra-ge a-je*
 food REL 1S.POSS-mother 3-cook= TOP give 1-to 1S.POSS-father
 ‘The food which my mother cooked I gave to my father’

The topic marker can optionally attach to the head noun of a topicalized NP, as illustrated in the following pair of sentences.

- 301) [*nyafi ar ra-ge it-den adi*]_s=*ro nene san+rimta*
 gift [1S-]give 1-to 1Pn.POSS-mother DET=TOP PREP house+mouth
 ‘the gift which I gave to our mom is in front of the house’
- 302) [[*nyafi*]=*ro ar ra-ge it-den adi*]_s *nene san+rimta*
 gift=TOP [1S-]give 1S-to 1Pn.POSS-mother DET PREP house+mouth
 ‘the gift which I gave to our mom is in front of the house’

Examples of topicalized obliques show they can occur at the beginning or the end of a clause, consistent with the behavior of obliques in general.

- 303) *[i]_A=ro n-ar kirror fi na-si i=n-tati*
 3S=TOP 3-give mirror ACT HAB-see 3S=3-REFL
 ‘He uses a mirror to see himself’
- 304) *nene [kirror]_{OBL}=ro i=ro si gata?*
 PREP mirror=TOP 3S=TOP see who
 ‘In the mirror, who does he see?’
- 305) *m-ro bguen wer+win nene [termus adi]_{OBL}=ro*
 2-run 2.look hot+water PREP thermos DET=TOP
 ‘see if there is some hot water in **the thermos**.’

In constructions with repeated topics, such as 306–7), a topic NP is introduced with =ro and followed by a coreferential pronoun also marked with =ro in the same clause. This supports Rizzi (1997), who noted that there can be multiple topics marked in a clause. By contrast, focus is thought to be unique. Only one element can be focused (Rooth 1992).

- 306) *[Wer+kuri]_S=ro [i]_S=ro safatero t-rir*
 water+Kuri=TOP 3S=TOP well.up INCI-wrap/wind
 ‘the flowing Kuri river meanders away’
- 307) *[Jek]_i=ro rror ro-ge [o]_i=ro bu-ba esi taje kabia*
 Jek=TOP tomorrow COP-COMP 2S=TOP PROG-2.go even.if weather bad
 ‘Jack, tomorrow you leave even if the weather is bad.’

The preceding example uses a proper pronoun to refer to a speech act participant. This strategy, combined with the frequent occurrence of the topic marker with second person pronouns reflects the politeness system in Irarutu in which there is a dispreference for bare second person pronouns. This is reflected in the contrastive meaning of the pair of examples below.

- 308) *o se*
 2S how
 ‘(but) how are you?’
- 309) *o=ro se*
 2S=TOP how
 ‘you may’

In contrast with second and third person pronouns, the first person singular pronoun *ja* virtually never occurs with the topic marker. My consultants explained that it is considered redundant to add =ro to *ja*. Topicalized first person plural pronouns, such as *am* ‘we (exclusive)’ are more likely to occur with the topic marker.

- 310) *[am]=ro dud-fe du=ra-si matu weni n-radwen bor*
 1Pn=TOP work-DTR NPROG=1-see person DET:PL 3-watch ball
 ‘We work while they watch soccer’

Subjects (*'The Kuri river meanders away'*), objects (*'The food that mom cooked, I gave to dad'*), indirect objects, and obliques (*'In the mirror, he sees someone'*, *'The morning he made a noken was hot'*) can all be topicalized.

- 311) *matu adimai i-den=ro buna n-f-tifn-i fade*
 person DET 3S.POSS-mother=TOP most 3-ACT-love-TR just
'that person's mother loves him very much'

3.1.3.5.1.2 Negative marker =*ti*

Negators are words that correspond to 'not'. In Iraputu, these are the morphemes =*ti* and =*of*. Negative elements negate a sentence, clause, verb, or other constituent. Cross-linguistically there are three strategies to form a negative: use of a negative word (which may be analyzed as an adverb because it often combines with the verb), use of a bound morpheme that attaches to the verb, or finally, use of a special negative verb or auxiliary that can combine with TAM markers (Givón 1990). Iraputu uses a combination of the first and second strategies, that is, a bound morpheme, =*ti*, at the end of a clause (see also Matsumura and Matsumura 1991 and Voorhoeve 1989 for felicitous claims). In intransitive clauses, the negative marker frequently combines with the verb, but in transitive clauses – due to its final position – it occurs after the object. The following examples show that =*ti* can cliticize to a range of elements.

- 312) *nof ssis (nya)=ti*
 wind light (what)=NEG
'windless'
- 313) *ja g-a a-gin=ti*
 1S DES-1 1-SAP.drink=NEG
'I do not want to drink'
- 314) *matu drbudi*
 person think.NEG
'idiot'
- 315) *ja du kka=ti*
 1S NPROG bathe=NEG
'I haven't bathed yet'
- 316) *ja du tim frrufn [matu n-mi san=ti]*
 1S NPROG close door person 3-stay house=NEG
'I closed the door because nobody is home'

Another example illustrates that when the negative marker occurs at the end of a main clause, it takes narrow scope, affecting only the content of the main clause, not the dependent clause.

- 317) [*ere du-bitr=ti*], *mesese e=bu-ga-fe*
 if NPROG-hungry=NEG, later-RED 2P=PROG-eat-DTR
'If you are not hungry yet, then eat later'

However, the negative element ‘=ti’ also occurs in negative polarity words such as *nyati* ‘nothing’, from the indefinite pronoun *nya* ‘something’ and =ti ‘negative’, and auxiliary *tufi=ti* ‘dislike’. (Matsumura and Matsumura 1991 also cite *nfnate* ‘don’t’, which appears to be composed of *nfnate*- ‘[do] like’ and *te* ‘negative’). Words such as *sagwe=ti* ‘breathless’, and *sufusr=ti* ‘less’ follow the main pattern, where =ti occurs last in the clause: *ja sagwe nene bunat adini=ti* ‘Right now I’m breathless’ vs. *ja sagwe=ti nene bunat adini* ‘(ibid)’.

- 318) *ere ja fi ra-run=ti*
 FACTUAL 1S ACT 1-know=NEG
 ‘I don’t know (what bit me)’

Negative imperatives, also called ‘prohibitives’, are not formed with the negative =ti, but rather with a special ‘prohibitive’ marker, -of.

- 319) *e ga=of*
 2P SAP.eat-PROH
 ‘(You all) don’t eat’
 320) *bi-mi-nir bier-fe(=of)*
 2.ACT-stay-with 2.speak-DTR(-PROH)
 ‘silence (enough talking)!’

A special case exists for negative desire, which is usually expressed by the inflected desiderative *g-* affixed with progressive *bu*, forming a word that means ‘not want to’ (see §3.1.3.3.1, above).

3.1.3.5.1.3 Politeness marker =e/o

The politeness clitic =e/o attaches at the end of sentences. The politeness markers are always the final element and can attach to a wide variety of hosts. In the examples below, the politeness marker can be seen attached to a verb, a pro-form, a noun, the end of a question (321), an answer (322), or a salutation (323). In 323), the politeness marker indicates the difference between a simple declarative *gisie* ‘(it is) evening’ and a greeting *gisie-o* ‘(good) evening’.

- 321) *o=ro bun nya fi mo jie=e?*
 2S=TOP 2.do what REL IRR complete=POL
 ‘(you there,) what are you doing?’
 322) *ja fun nyati=o*
 1S do nothing=POL
 ‘I am not doing anything’
 323) *gisie-o*
 evening-POL
 ‘[good] evening’
 324) *termus adi-mo wer+win=ti=o*
 thermos DET-CONJ water+hot=NEG=POL
 ‘The thermos has no hot water.’

The preceding example illustrates an ordering requirement that places the politeness clitic after the negative clitic.

Matsumura and Matsumura 1991 identify a similar looking device, =*e*, in the East Arguni dialect that they say signals an interrogative utterance. This element is located at the end of a sentence, and is supported by examples such as 325), cited below.

- 325) [*Rut nia*] *se* [*i*] *bu-na n-ma=e?*
 time what then 3S PROG-HAB 3-come=Q
 ‘What time is she going to arrive?’ (Matsumura & Matsumura 1991)

This is apparently a dialectal feature. My Fruata consultants do not use a clausal clitic =*e* as a question marker. When asked about it, the speakers could not use this element spontaneously. However, there are several instances, as well as metalinguistic data, that supports the analysis of sentence final =*e* (alternatively =*o*) as a politeness marker.

3.1.3.5.2 Copula *rau* and existential *o*

Copulas indicate the relation between a subject and predicate nominal or a subject and an adjective. In Irarutu, neither the copula nor the existential marker are treated the same as verbs. For example, neither is inflected for subject agreement. If there is no overt subject, the predicate nominal occurs by itself and the subject is implied by the context, as in 326). A commonplace alternative is to juxtapose the elements, as in 327).

- 326) (*rre*) *win*
 today hot
 ‘(Today is) hot.’
 327) *taje kabia*
 weather bad
 ‘The weather is bad.’

The copula, *rau*, which optionally occurs in equational constructions (‘NP is NP’) or predicate adjective constructions (‘NP is ADJ’), does not usually occur with verbal predicates or oblique locationals. It also tends to occur more frequently in elaborate structures, such as a complex clause (328), rather than a simple structure, such as 327), above. Furthermore, its location between the subject NP and the predicate superficially overlaps with the location of the topic marker, =*ro*, on topicalized subjects, compare (328) and (329). Both the topic clitic =*ro* and the copula *rau* can co-occur; however, when one of these elements is missing, it is usually the copula (329), but there are also constraints on the topic marker that can prevent it from occurring (see §3.1.3.5.1.2).

- 328) [*matu adimai ja si ade*]=*ro rau* *gun bar*
 person DET:DIST 1S see DET=TOP COP head bald
 ‘the man that I saw is bald’
 329) [*nyafi ar ra-ge it-den adi*]=*ro* nene san rimta*
 gift [1S-]give 1-to 1Pn.POSS-mother DET=TOP PREP house door
 ‘the gift which I gave to our mom is in front of the house’

Equational constructions are divided into topic-comment or focus-presupposition. Examples of these types of copula-free constructions are shown in 330–3).

- 330) *finnar adej=ro dir*
joke DET=TOP good
‘the joke is funny’
- 331) *nyaro adeji=ro du-metyet*
vegetables DET=TOP NPROG-fresh
‘The vegetables are fresh’
- 332) *mumce fi i=nyuen dub adi=ro buna win*
morning REL 3S=3.weave satchel DET=TOP most hot
‘the morning [when he made a satchel] was hot’
- 333) *warada [fi i=n-ar=ad na-n-tut sum] adi=ro bu nsir*
knife REL 3S=3-take=DET HAB-3-cut fish DET=TOP PROG 3.dirty
‘the knife which she used to cut the fish with is dirty’

A morphologically complex copula does occur in sentences such as 334), where it presumably combines with the element *-te* ‘incidental stative’.

- 334) *ja mo-ra-run ge radni adeji rote kabia*
1S IRR-1-know COMP road DET COP broken
‘I do not know why the road is closed’

In addition to the copula, Irarutu also has an existential marker *o* ‘there is’, shown below. The existential marker is used to express an expletive subject.

- 335) *Jek, o te nir kopi nir wer*
Jek, EXST tea CONJ coffee CONJ water
‘Jack, there is tea, and coffee, and water’

3.1.3.5.3 Interjections

Interjections are universal. They have an exclamatory character, which can constitute an utterance in itself, similar to a pro-sentence. Cross-linguistically, interjections can be phonologically distinctive (such as lack a nucleus). In Irarutu, the exclamation for ‘ouch!’ is phonetically long [i:], despite the fact that length is not phonemic (compare this generalization with Pike 1947). As a class, interjections have an informal, colloquial character. Several interjections are presented in the examples below

- 336) *afa! 'nia fe adige?*
EXCL what REL happen
‘I’m surprised! What happened?’
- 337) *i, nya gor ja*
ouch, what bite 1S
‘Ouch, I am getting bitten (what bit me)’
- 338a) *nia fna deye*

- what like reason
 ‘how (what’s the cause)’
 338b) *fna bu fna deye*
 like PROG like reason
 ‘how (what was the reason)’

3.2 Grammatical system

The syntactic categories identified in §3.1, above, are fundamental to describing the grammatical system of four interrelated phenomena in Irarutu: grammatical relations, word order, transitivity, and agreement, which are presented, as much as possible, as separate phenomena, but their interrelation is fundamental. The grammatical system serves as a set of tools for understanding other syntactic phenomena, such as relative clauses and the form of interrogative sentences.

3.2.1 Grammatical relations

The traditional view of grammatical relations is that they relate arguments to a clause. However, cross-linguistic data on grammatical relations suggest that relevant phenomena are not so simple. Following Bickel (2011), a grammatical relation is viewed as the set of arguments that is selected by a construction for a particular syntactic purpose, e.g., Subject or Object. Furthermore, grammatical relations are different from thematic roles, which relate arguments to their predicate (§3.1.2.1). The view that grammatical relations are structure- (and function-)specific is supported by Irarutu data on the word order and agreement properties of S (the sole argument of a monadic verb), A (the subject of a dyadic verb), and O (the object of a dyadic verb), in main and relative clauses.

3.2.1.1 Grammatical relations in main clauses

There are no case markers in Irarutu. Grammatical relations are established by word order and agreement. In a clause with a single argument, S usually precedes the predicate, SV (339), however some unaccusatives allow VS order (340).

- 339) [ja]_S firro
 1S run
 ‘I run/am running/ran.’
 340) *tim~tim* [firruⁿ ad]_S
 PASS~close door DET
 ‘The door is (already) closed’

In constructions with second and third person referents, the verb is usually inflected to agree with the person (first, second, or third) of A. Agreement in the first person in main clauses does occur, but it is rare, and considered by native speakers to sound redundant. In cases where SV *ir n-mtit* and VS *n-mtit ir* ‘they fell’ are both possible, the verb is inflected to match S. Because VS is the exception rather than the norm with monadic verbs, SV appears to be the canonical word order.

In a canonical two-argument clause, A precedes the predicate, whereas O follows it. When agreement morphology occurs, that morpheme matches the person feature of A.

- 341) [ja]_A tbe [o]_O
 1S hit 2S
 'I hit you.'
- 342) [o]_A m-tbe [ja]_O
 2S 2-hit 1S
 'You hit me.'

When the behavior of constructions that have a single argument is compared with the behavior of constructions that have two arguments, word order and agreement generally align with the nominative-accusative pattern.

Transitivity also plays a role in the distribution of grammatical relations. For example, obliques, which are introduced by an (optional) preposition and are therefore not A, S, or O, occur in the periphery, but their occurrence can satisfy the patient role required by a dyadic verb such as *rit* 'clear/tend' expressed in a construction that only accepts preverbal S. In other words, constructions such as 343) have a high degree of semantic xtransitivity but are treated as intransitives according to grammatical relations.

- 343) [it]_S fa rit [(ne) dedan]_{Obl}
 1Pn go tend PREP garden
 'Let's go tend (to) the garden'

Of course, obliques can occur with verbs that only select one argument, such as *mi* 'live/stay', and *ma* 'come',

- 344) ja nenebe nmi [ne Kiman]
 1S before live PREP Kaimana
 'I lived in Kaimana'
- 345) ja te-na-su nmi [ne Kiman] [fi ebe]
 1S INCI-HAB-CPL live PREP Kaimana REL again
 'I will live in Kaimana again'
- 346) ja ma [nene refefa uf fifi]_{Obl}
 1S come PREP yest. REL previous
 'I arrived the other day'

or, they can occur with to a two-place, AVO, construction such as 347).

- 347) [nene kirrar=ro]_{Obl} i=ro (n-)si gata?
 PREP mirror=TOP 3S=TOP (3-)see who
 'In the mirror who does he see?'

Double object constructions in Irarutu do not code the second object as an oblique, and it cannot simply be adjoined to an AVO construction. Instead, the second object is integrated into the clause through a serial verb construction which employs a special

‘directional-give’ verb plus indirect object (IO) adjoined to a basic two-place construction represented as: A V O [V IO]. Agreement shows there is only one A for the whole clause, such as first person Ø- with *fi-ar* ‘give’, but first person minor pattern *ra-* with the secondary verb *ge* ‘(give) to’ in 348).

- 348) [ja]_A *fi-ar* [wen-ga]_{DO} *ra-ge* [o-matu uf m-tim frufn]_{IO}
 1S ACT-give NMLZ-food 1-to 2-person REL STAT-close door
 ‘I gave food to your friend who closed the door’

3.2.1.2 Grammatical relations in relative clauses

Concerning main clause primary verbs, overt agreement morphology for first person subjects is rare (349). In serial verbs and subordinate clauses, this agreement morphology is more frequent.

- 349) *ja (a-)ar kirrar a-si ja*
 1S (1-)give mirror 1-see 1S
 ‘I use a mirror to see me’
 350a) *ja bu kka sefu a-gm*
 1S PROG bathe so.that 1S-SAP.sleep
 ‘I did not bathe so that I can sleep’
 350b) **ja bu kka sefu ja=gim*
 ‘I did not bathe so that I can sleep’
 351) *wen-ga [fi a-tiin] adi*
 NMLZ-eat REL 1S-cook DET
 ‘the food which I cooked’

However, there seems to be increased variability in the third person, which is almost always marked in main clauses, but less frequently in a dependent clause.

- 352) [ja]_A *fi-ar* [wen-ga]_{DO} *ra-ge* [o-matu uf m-tim frufn]_{IO}
 1S ACT-give NMLZ-EAT 1-to 2S.POSS-person REL STAT-close door
 ‘I gave food to your friend who closed the door’

3.2.2 Word order

Word order in Irarutu is rigid Subject-Verb-Object. Not only is word order one of two primary means of signaling grammatical relations, from a typological perspective, Irarutu word order patterns also illustrate a violation to Hawkins’ (1983) word-order universals.

A common assumption in contemporary linguistic theory in the ‘configurational’ (Andrews 1985) vein of research is that clauses are headed by an (abstract) inflectional category that may surface as a tense marker (Chomsky 1995, Givón 1990). This appears to be true in Irarutu, although the inflectional category is represented by a TAM marker.

In canonical structures, the preverbal slot is for subjects, that is, S in monadic constructions and A in dyadic constructions, whereas the postverbal slot is for O. S also occurs before the predicate in equational constructions and with adjectival predicates.

- 353) [taun porna je-feni]_s [n-mi]_v [dapur]_{OBL}
 sago bread contents-REL.DET 3-stay kitchen
 ‘sago bread is in the kitchen’
- 354) [a-den]_s [ggun]_v [ja]_o [nene kiman]_{OBL}
 1S.POSS-mother birthed 1S PREP Kaimana
 ‘I was born in Kaimana (lit. my mother birthed me in Kaimana)’
- 355) [ja]_s [tiin]_v [nyaro]_o
 1S cook vegetables
 ‘I cook vegetables’

In several examples, a clause-initial full-NP subject is accompanied by a matching pronoun in the slot immediately before the verb complex. In the examples below, the subject NP is repeated as a pronoun *ir* ‘they’ in the preverbal slot.

- 356) [Damianus, Willy, nir Moses]_s [ir]_s n-fier-fe nfntagre
 D. W. CONJ M. 3P 3-say-DTR about
 na-n-fun matu weni Fruat nene rror adini
 HAB-3-do people DET Fruat PREP yesterday DET:DEM
 ‘Damianus, Willy, and Moses spoke about life in Fruata yesterday’
- 357) [wer+kuri]_{s=ro} [i]_{s=ro} safatero t-rir
 water+Kuri=TOP 3S=TOP well.up INCI-wrap/wind
 ‘the flowing Kuri river meanders away.’

The word order patterns in Irarutu support the majority of the generalizations proposed in Hawkins’ (1983) ‘Prepositional noun modifier hierarchy’ (PrNMH), see **FIGURE 3.16**, which is an elaborate unilateral implicational universal that can be interpreted to mean that if a language has prepositions, then if it has Noun-Demonstrative order or Noun-Numeral order then it will have Noun-Adjective order, and if it has Noun-Adjective order then it will have Noun-Genitive order, and if it has Noun-Genitive order then it will have Noun-Relative order.

$$\supset \text{Prep } ((\supset (\text{NDem} \vee \text{NNum})(\text{NA})) \wedge (\supset \text{NA NG}) \wedge (\supset \text{NG NRel}))$$

FIGURE 3.16 – Prepositional Noun Modifier Hierarchy (adapted from Hawkins 1983).

One Irarutu word order strategy consistently violates the PrNMH. Irarutu genitives precede their head N, which conflicts with the second conjunct in the resultant clause in **FIGURE 3.16** ($\supset \text{NA NG}$). Examples of Irarutu data are given in **FIGURE 3.17**.

basic WO: SV	<i>ja trro</i> ‘I am tall’
basic WO: SVO	<i>sirion ntut sum</i> ‘the village leader cut the fish’
N+ADJ: N ADJ	<i>wer win</i> ‘hot water (the water is hot)’
N+GEN: GEN N	<i>matu san</i> ‘someone’s house’
N+ADPOS: PREP N	<i>nene pasar</i> ‘to market’
N+Dem: N DEM	<i>termus adi</i> ‘the thermos’
N+NUM: N NUM	<i>mo ru</i> ‘two children’
N+REL: N REL	<i>refefa uf fifi</i> ‘day which previous’

FIGURE 3.17 – Examples of various word order patterns.

3.2.3 Transitivity

Transitivity is a global property of clauses. Since Hopper and Thompson 1980, it has been viewed as a scalar phenomena based on component parts, including number of participants, potency of A, affectedness and individuation of O, and six properties that in Irarutu belong to the verb complex: degree of kinesis, perfective aspect, punctuality, volitionality, realis mode, and affirmative or negative. Although the other components can override it, the component ‘number of participants’ can be used broadly to determine if a clause is transitive or intransitive. At one end of the spectrum, intransitive clauses with monovalent, monadic verbs, do not carry an action over from one participant to another. Notably, intransitive verbs in Irarutu show a ‘fluid split-S system’ (Dixon 1994). At the other end of the scale, transitive clauses with bivalent, dyadic verbs, do transfer action between participants. There is also a double object construction, which is expressed by a type of serial verb construction.

3.2.3.1 Intransitives

For the most part, unergatives and unaccusatives are morphologically very similar. The underlying arguments in both of these types of intransitive, i.e., monovalent, constructions surface in the canonical subject slot and occur with an inflected verb. This is shown in the partial paradigms presented in **TABLE 3.18**.

<u>Unergatives</u>	<u>Unaccusatives</u>
<i>ja=sur</i> ‘I jump’	<i>ja=^mbrif</i> ‘I laugh’
<i>msur</i> ‘you jump’	<i>o=^mbrif</i> ‘you laugh’
<i>i=nsur</i> ‘he jumps’	<i>i=ⁿbrif</i> ‘he laughs’
<i>ja=mrir</i> ‘I stand’	<i>ja=mtit</i> ‘I fell’
<i>o=mrir</i> ‘you stand’	<i>o=mtit</i> ‘you fell’
<i>i=nmrir</i> ‘she stands’	<i>i=nmtit</i> ‘she fell’

TABLE 3.18 – Verb conjugation for unergatives and unaccusatives with singular subjects.

A fairly large subset of intransitive predicates are marked with stative *m-*. With or without the stative marker, predicates are generally inflected for the subject. They can also occur with the active morpheme *fi-*.

- 358) [ja]_{sp} *f-rran*
 1S ACT-short
 'I am short'
- 359) *ere* [ja]_{sp} *fi ra-run=ti*
 unseen 1S ACT 1-know=NEG
 'I don't know (what bit me)'
- 360) [ja]_{sa} *ma fi jemat*
 1S come ACT late
 'I arrived late'

However, a significant number of unaccusatives, such as the verb *mtit* 'fall', allow the sole argument to appear after the verb, providing counter-examples to the generalization that S occurs before V. Since Dixon (1994), it has been conventional to label the sole argument in a single-argument construction as 'S', regardless of whether it is thematically an agent or a patient. However, in Irarutu, S occurs in either the pre-verbal or the post-verbal slot. The variation can be accounted for as a difference between unergative and unaccusative constructions. Therefore, because these two types of S behave differently, they represent an important grammatical distinction in alignment. Native speakers consider some single-argument constructions to be ill-formed if the argument occurs in preverbal position. Other constructions allow stylistic alternation between pre-verbal and post-verbal position.

- 361) *a-sagwe* [ja]_{sp} *is*
 1-breathe 1S prior
 'I want to take a break (lit. I'm going to breathe for a while)'
- 362) *m-tit* [ja]_{sp}
 STAT-fall 1S
 'I fell (lit. fell I)'
- 363) [ja]_{sp} *m-tit*
 1S STAT-fall
 'I fell'
- 364) [ja]_{sa} *t-rir*
 1S INCI-stand
 'I stand'

The difference between VS constructions in 361–2) and examples of the more common SV, pre-verbal subject only in 363–4), is that VS constructions have a theme S (_{sp}), whereas examples of SV have an agentive subject (_{sa}). This difference supports the distinction between un-ergatives and un-accusatives.

3.2.3.2 Transitives

Two of the ten components identified by Hopper and Thompson (1980), degree of affectedness of O, and individuation of Os are particularly important for clauses with two participants. Prototypical individuated objects are proper, human-animate, concrete, singular, count, or referential-definite, whereas prototypical non-individuated objects are

common, inanimate, abstract, plural, mass, and non-referential; but not all nouns fall neatly into one of these categories. Combinations of these features occur, meaning that some objects share one or more properties with the individuated category and other properties with non-individuated objects.

Some additional factors that are known cross-linguistically to affect the transitivity of a clause are: having a locative direct object (enter the house, swim the channel), having a recipient/locative DO and implied patient (feed, paint, dust), movement of a part of the referent of the subject (hit, kick, slap, punch), dative-experiencer subject (cognition, sensation, volition), reciprocal/associative object, stative ‘have’ (get, take, obtain), cognate objects (sing a song, dance a dance; “bleached” verbs: give a talk, take a breath, take a turn), non-human/inanimate agents (liquor killed him), and object suppression (anti-passive or object incorporation).

- 365) *ja ga-bu [na [nyaunt]_A gor [ja] uf]*
 1S DES-1-PROG if ants bite me PROH
 ‘I don’t like to be bitten by ants’
- 366) *Mikel=ro n-tiit wagt mug adi namse nanbssi*
 Michael=TOP 3-pound taro much DET until HAB-3-smooth
 ‘Michael must pound taro many times so that it becomes smooth’

Following the proposition that transitivity is a scalar property, the occurrence of certain elements in Irarutu can decrease the transitivity of an otherwise transitive predicate. For example, the detransitivizer *-fe* denotes that the patient is non-specific, indefinite, and not expressed overtly in the clause. Examples are *ga-fe* ‘eat (something)’, *nut-fe* ‘sew (lit. pierce something)’, *ntünfe* ‘cook things’.

- 367) *[it]_{Sa} ga-fe is*
 1Pn SAP.eat-DTR prior
 ‘Let’s eat first’

3.2.3.2.1 Nominative-Accusative alignment

When canonical intransitive constructions are compared with canonical transitive constructions, there is overwhelming evidence that Irarutu follows a nominative-accusative alignment pattern, shown in **FIGURE 3.18**, below. In particular, the majority of Ss occur preverbally; and, all As occur preverbally. Agreement morphology always matches the person feature of S, regardless of the location of S in intransitives, and it matches the person feature of A in transitives.

Alignment Pattern					
Intransitive					
S _{a/p}		V	S	V	
<i>aje</i>	(i)	<i>n-tag</i>			
1POSS-father (3S)		3-cry			
'My father cries'					
Transitive					
S	V	O	S	V	O
<i>a-den</i>	<i>n-tiin</i>	<i>nyaro</i>			
1POSS-mom	cook	vegetables			
'My mother cooks vegetables'					

FIGURE 3.18 – The nominative-accusative pattern, with example sentences and abstract representation of the alignment pattern.

3.2.3.2.2 Split-S alignment

Contrary to the characterization of Irarutu as a pure nominative-accusative language, unaccusative Ss behave differently from unergative Ss. This is illustrated in **FIGURE 3.19**. Split S poses a complication for the analysis of Irarutu case, because there is an ergative-absolutive sub-pattern, where S_p aligns with O (thematic patient) by following the verb.

Alignment Pattern					
Intransitive					
	V	S _p		V	S _p
	<i>m-tit</i>	<i>mo</i>			
	STAT-fall	child			
'The child fell'					
Transitive					
S	V	O	S	V	O
<i>matu</i>	<i>m-tbe</i>	<i>fun</i>			
2S	2-hit	1S			
'The person hit the dog'					

FIGURE 3.19 – The unaccusative S pattern, with example sentences and abstraction. S_p represents an unaccusative S (commonly a patient).

The contrast between **FIGURES 3.18** and **3.19**, indicates that Irarutu uses a 'fluid split-S' system (Dixon 1994). The dominant system is nominative-accusative, where S and A pattern the same in word order and agreement. There is also a subset of examples that follow an ergative-absolutive pattern, where unaccusative S patterns the same as O insofar as word order (although it patterns with A in terms of agreement). Therefore, despite the existence of a 'fluid split-S' pattern in the otherwise nominative-accusative language, agreement in the subsystem supports the category of Subject.

3.2.3.3 Double object pattern

There are three patterns for constructions with three ‘core’ participants, *ar* ‘give’ type, null main verb type, and a relativized O type. These participants are referred to here as A O₁ (or DO) and O₂ (or IO). In non-relativized structures, the second object, O₂, is introduced by a serialized, often directional, secondary verb V₂. An abstraction is: A V₁ O₁ V₂ O₂. Some directional-predicate constructions, e.g., source-goal constructions, parallel the structure of the double object construction, that is, use serial verbs. According to Givón (1990), double object constructions occur cross-linguistically with locatives (‘put...on’, ‘take...off’, ‘remove...from’, ‘move...to’, ‘send...to’, ‘carry...to’, ‘bring...to’), datives/benefactives (‘give, bring, receive, tell, show, ask’), and instrumentals (‘fill Y with X’, ‘empty Y of X’). See TABLE 3.19.

Indirect Object Construction

	<i>itar wenga rage amatu</i> ‘we gave food to my friend’
<i>jar wenga rage amatu</i> ‘I gave food to my friend’	<i>omar wenga mage amatu</i> ‘we, not you, gave food to my friend’
<i>mar wenga mage amatu</i> ‘you gave food to my friend’	<i>emar wenga mage amatu</i> ‘you all gave food to my friend’
<i>nar wenga nane amatu</i> ‘she gave food to my friend’	<i>irnar wenga nane amatu</i> ‘they gave food to my friend’

Source-Goal Construction

	<i>itro mi warnet briet san</i> ‘we came up from the internet café’
<i>ja mi warnet briet san</i> ‘I came up from the internet café’	<i>amro mi warnet briet san</i> ‘we came up from the internet café’
<i>oro mi warnet (o) briet san</i> ‘You came up from the internet café’	<i>ero mi warnet briet san</i> ‘you all came up from the internet café’
<i>i nmi warnet nbriet san</i> ‘he came up from the internet café’	<i>irro nmi warnet nbriet san</i> ‘they came up from the internet café’

TABLE 3.19 – Distribution of pronouns and inflectional prefixes in SVCs.

In Irarutu ‘give’-type dative-benefactive constructions, O₂ follows V₂, which is inflected by a special subset of agreement markers for A: *ra-* ‘1’, *ma-* ‘2’, and *na-* ‘3’. The third person also coincides with a suppletive verb form, such as *ne* ‘to’ whereas the basic form is *ge*. The set of subject markers with V₁ *ar* ‘give/take’ is different than the set of subject markers with V₂ *-ge* ‘to’, but both agree with A. Coincidentally, second objects (IO)

cannot precede primary objects (DO). The cross-linguistically common alternant, A V IO DO, e.g., **ja ar [o]_{IO} [pipisi]_{DO}* ‘I gave you money’ is ill-formed.

- 368) [*ja*]_A *ar* [*pipisi*]_{O1} *ra-ge* [*o*]_{O2}
 1S give money 1-to 2S
 ‘I give money to you.’
- 369) [*ja*]_A *ar* [*wenga*]_{O1} *ra-ge* [*a-mo*]_{O2}
 1S give food 1-to 1S.POSS-child
 ‘I give food to my child’
- 370) [*ja*]_A *fi-ar* [*wenga*]_{O1} *ra-ge* [*o-matu* *uf m-tim* *frrufn*]_{O2}
 1S ACT-give food 1-to 2S.POSS-person REL STAT-close door
 ‘I gave food to your friend who closed the door’
- 371) [*matu adei*]_A *n-ar* [*wenga*]_{O1} *na-ne* [*i-dena*]_{O2}
 person DET 3-give food 3-to 3S.POSS-mother
 ‘That person gave food to his mother’

Several other constructions are similar to the ‘give’-type constructions, with a range of possible secondary verbs, such as *fun* ‘do’ to derive the meaning ‘teach’, and *tut* ‘cut’.

- 372) [*ja*]_A *fi ar* [*snan ffrfier bahasa igris*]_{O1} *ra-fun* [*a-mo*]_{O2}
 1S ACT give word speak language English 1-do 1S.POSS-child
 ‘I teach my child English’
- 373) *i=n-ar* [*warada*]_{O1} *na-n-tut* [*sum*]_{O2}
 3S=3-take knife INF-3-cut fish
 ‘he uses a knife to cut fish’

Instrumental O₁ can also occur without a first verb. Without the context of the double object construction, examples such as 374) would appear to be double nominative constructions because two participants occur preverbially, the agent and the instrument.

- 374) [*ja*]_A [*suri*]_{O1} *ra-tar* [*ema*]_{O2}
 1S axe 1-split firewood
 ‘I split firewood’

The third option to express three participants in a clause is to use a verb that implies an object and relativize that object. The third example below shows that *surti* ‘(to write a) letter (to)’ functions as a predicate with an implied object *surti* ‘a letter’, because it occurs with the non-progressive *du* ‘already’.

- 375) *ja fi surti* [*ra-ge a-den* *ja=ri*]
 1S ACT letter 1-to 1S.POSS-mother 1S=have
 ‘I am sending a letter to my mother (for my mother to have)’
- 376) [*ja*] *fi surti fi n-fa a-den*
 1S ACT letter REL 3-go 1S.POSS-mother
 ‘I am sending a letter to my mother’

- 377) *a-den du surti fi n-ma ja*
 1S.POSS-mother NPROG letter REL 3-come 1S
 ‘my mother sent me a letter’

3.2.4 Agreement

Agreement (Baker 2008) is an important dimension of Irarutu grammar. In addition to rigid word order, agreement is one of the primary means to indicate grammatical relations in Irarutu. Three crucial properties of agreement are: i) monadic verbs and other single-place predicates often have agreement morphology that matches features of S (378), even if word order is not SV; ii) dyadic (non-serialized) verbs nearly always have agreement morphology that matches A (379); and iii), dyadic verbs are not typically marked to agree with O, but there are some contexts where object agreement does occur.

- 378) *mo+ntu n-in*
 child+new SAP.AGR_S-sleep
 ‘the baby sleeps (the baby is sleeping).’
 379) *o=ro m-si gata?*
 2S=TOP 2.AGR_S-see who
 ‘who do you see?’

Agreement morphology matches either the participant status (378) or the grammatical person (379) of Subjects. Number is not relevant to agreement in Irarutu. There are four patterns for agreement: agreement in regular verbs and auxiliaries, agreement in irregular verbs, agreement in certain serial verb constructions, and agreement with the modality marker *g-*. Factors that play a role in the form and location of agreement morphology are grammatical relations (S and A but not O), discourse context (three-way person marking vs. binary speech act participant marking), serialization (main vs. serial verb), and relativization (overtly marking first person subjects).

3.2.4.1 Agreement in regular verbs

The main pattern for agreement is the use of a marker on the verb that matches the person of the Subject. There are three relevant morphemes, *a-* ‘first person subject’ (nearly always dropped in msain clauses), *m-* ‘second person subject’, and *n-* ‘third person subject’. The first person agreement marker, *a-*, only occurs in cases where the Subject is not sufficiently clear from context and needs to be identified specifically, as found in certain intransitive constructions and relative clauses.

- 380) *ja fa=o*
 1S [1-]go=POL
 ‘I’m going (moving on)’
 381) *(o) m-guir i*
 2S 2-call 3S
 ‘(you) call him’
 382) *(o) g-o ba neno?*
 2S DES-2S 2.go where
 ‘Where are you going?’

- 383) *matu n-taf nya fi i=(n-)si*
 person 3-hunt what REL 3S=(3-)see
 ‘the hunter shoots things that he can see’
- 384) *m-ar buku ma-mri m-tu*
 2-take book 2-enter 2-send
 ‘Go put your book away’
- 385) *Jek-a o ma bɪ ji-uf tɪfn f-adeji*
 Jack-VOC 2S 2.come 2.ACT EXPL-PROH bee REL-DET
 ‘Jack! Come, don’t be in there, there’s a bee’

3.2.4.2 Agreement in irregular verbs

A minor pattern of agreement, found with a small set of frequently-used verbs such as *-in* ‘sleep’, *-in* ‘drink’, *ga* ‘i/you eat’ (*nan* ‘he eats’), only distinguishes the participant status of the subject and can be characterized as a two-way distinction grouping first and second person ‘speech act participant’ subjects (*g-*) separately from third person ‘non-speech act participant’ subjects (*n-*). The person category of the subject can be specified by an independent pronominal subject.

- 386) *gin*
 SAP;sleep
 ‘I/you sleep’
- 387) *ja ga uce*
 1S SAP;eat papeda
 ‘I eat papeda’
- 388) *o=ro ga uce*
 2S=TOP SAP;eat papeda
 ‘You eat papeda’
- 389) *nin*
 NSAP.sleeps
 ‘she sleeps’
- 390) *i nan uce*
 3S NSAP;eat papeda
 ‘He eats papeda’

3.2.4.3 Agreement in serial verb constructions

Another minor pattern of agreement (*ra-* ‘1’, *ma-* ‘2’, *na-* ‘3’), is also found with serial verbs as well as (‘axe-split’ type) incorporation verbs.

- 391) *ja ar wega ra-ge o-matu*
 1S [1-]give food 1-to 2S.POSS-person
 ‘I gave food to your friend’
- 392) *ja suri ra-tar ema*
 1S axe 1-split firewood
 ‘I axe-split firewood (I cut firewood)’

3.2.4.4 Agreement in auxiliaries and other categories

The third, and most limited, minor pattern of agreement (-*a* ‘1’, -*o* ‘2’, -*i* ‘3’) is found with the desiderative modality marker *g-*. The pattern appears to be the only element to take an inflectional suffix, shown in 393).

- 393) *ja g-a (a-)mgür wer*
1S DES-1 (1-)buy water
‘I’m going to buy water (said while walking)’
- 394) *i=ro g-e n-mgür wer*
3S=TOP DES-3 3-buy water
‘He is going to buy water (said while referent is walking)’

3.3 Complex clauses

Two grammatical phenomena that have been avoided as much as possible in the previous sections, but which are pertinent to the description of Irarutu syntax are complementation and relativization, both types of subordination.

3.3.1 Complement clauses

In Irarutu, a subset of verbs can take complement clauses as their object instead of an NP. This set includes, for example, ‘say, think, see, believe, try, hear, feel, know, want, call, and ask’ (Noonan 1985). Complement clauses are sentences introduced by a complementizer, Irarutu *ge* ‘that/if’. The complementizer is not used in all instances of constructions with complement clauses, such as when it is the complement of *drbun* ‘think’. The word *fier* ‘say’, which is used cross-linguistically as a complementizer, is only used for indirect and direct, i.e., quoted, speech in Irarutu. Examples of complements taking the predicates ‘see, think, say, and know’ briefly introduce the construction, below. ‘Hope’- and ‘intend’-type complement constructions are expressed by auxiliaries such as *g-* ‘desiderative (wish)’, and *mo* ‘irrealis (want)’, or serial verb constructions such as *fa* X ‘go (intentional) X’ and *fun* X ‘do (causative) X’. A more detailed analysis of Irarutu complement clauses is set aside for future research.

- 395) *matu n-ttan fi-si ge [man adi n-m-tür erum]*
person 3-hunt ACT-see COMP bird DET 3-STAT-sit tree.top
‘the hunter saw the bird land in the tree’
- 396) *mo adi n-fi-drbun [finnar adej ro dir]*
child DET 3-ACT-think joke DET COP good
‘the child thinks the joke is funny’
- 397) *con n-fir ge [rror ro-ge it t-m-tür]*
pastor 3-say COMP tomorrow COP-COMP 1Pn INCI-STAT-sit
‘The pastor said that there will be service tomorrow’
- 398) *matu polisi n-fier ge [ja mo-ra-run ge radni adeji rote kabia]*
person police 3-say COMP 1S IRR-1-know COMP road DET COP broken]
‘the policeman said “I do not (/want to?) know why the road is closed”’

3.3.2 Relative clauses

Following Keenan (1985), a relative clause (RC) is viewed as a type of sentence (S_{rel}) that identifies a particular member NP_{rel} from a set of possible referents. Keenan also proposes NP_{rel} is fully specified and definite. Cross-linguistically, relative clauses can be internally headed, i.e., NP_{rel} is *in situ* in S_{rel} , or externally headed, i.e., where NP_{rel} is outside S_{rel} . Irarutu relative clauses are externally headed. The S_{rel} can occur before an external head or after it. Irarutu has postnominal relative clauses. These two characteristics are illustrated in 399), below. In the example, the subject [*amatu uf ntim frrufn adi*] is a topicalized NP that contains a relative clause. The subject of the relative sentence is a gap that correlates with the subject of the matrix clause, i.e., *amatu*. The gap in 399) is indicated by an underscore.

- 399) [[*a-matu*] $_{NP_{rel}}$ *uf* [____ *n-tim frrufn adi*] $_{S_{rel}}$]=*ro tni trro*
 my-friend REL 3-close door DET=TOP body tall
 ‘My friend who closed the door is tall’

There are two elements in Irarutu that can introduce S_{rel} , that is, the relativizers, *uf* and *fi*. These elements are in complementary distribution and both occur before S_{rel} . No distinction is made between human and non-human heads of relative clauses in the language. (The interrogative pronoun *gata* ‘who’ cannot be used to introduce a relative clause.) A cross-linguistically common alternative strategy to express information that could go in a relative clause is the correlative construction, not found in Irarutu.

3.3.2.1 Functions of RCs

Irarutu relies on relativization to realize a number of functions. Therefore, it is essential in describing RCs in Irarutu to point out some of these functions. Although nouns can be modified by adjectives, one widespread use of relativization is to express a range of (adjectival) modification that is dispreferred to occur as simple adjunction. For example, *uf nar rga* ‘life-giving’ is a headless relative clause ‘which gives life’. The novel phrase *war uf nar rga* ‘life-giving water’ is composed of the NP_{rel} [*war*] ‘water’, a relativizer [*uf*], and S_{rel} [____ *n-ar rga*]. Additional examples are listed in **FIGURE 3.20**.

<i>uf bis</i>	‘bravery/excess’	<i>uf (bu)nmur (tu)</i>	‘final/last’
<i>uf buna barie</i>	‘one who cleans’	<i>uf nmi</i>	‘lag’
<i>uf buna dir</i>	‘better’	<i>uf nanmi</i>	‘lasting’
<i>uf bunat</i>	‘new’	<i>uf ntagre bero</i>	‘next’
<i>uf dir</i>	‘best’	<i>uf riwe</i>	‘liquid’
<i>uf nar rga</i>	‘life-giving’	<i>uf rot</i>	‘key’
<i>uf nmur</i>	‘finish/last’	<i>uf ru</i>	‘second’
<i>uf bunmur</i>	‘end (of story)’	<i>uf wartim</i>	‘previous’

FIGURE 3.20 – Examples of relativized modifiers.

A related use of relativization in Irarutu is to form NPs that correspond to morphologically derived nouns in other languages. In some instances, such as when NP_{rel} corresponds with S or A in S_{rel} , and the optional relativizer is absent, full sentences

and NPs that contain a RC are superficially identical. Only the broader context determines how these constituents are interpreted. Of course, if the relativizer is present, the ambiguity disappears.

- 400) *matu (uf) sis ron*
 person (REL) sing song
 ‘singer/a person sings songs’
- 401) *matu (uf) fruet sum*
 person (REL) seek fish
 ‘fisherman/A person searches for fish (to catch)’
- 402) *matu (uf) m-tür hakim*
 person (REL) STAT-sit justice
 ‘derelict/a person who disrespects the law’

Another frequent use of relativization is to express possession. This is based on the predicate *ri* ‘have/possess’ and the relativized NP originates as O. In this context, the relativizer *fi* is optional.

- 403) *motor (fi) [ja=ri ____]*
 motorbike (REL) 1S=have
 ‘my motorcycle (lit. the motorcycle which I own)’
- 404) *motor (fi) [matu i=ri ____]*
 motor-bike (REL) person 3S=have
 ‘that person’s motorcycle (lit. the motorcycle which some person owns)’

3.3.2.2 The relativization hierarchy

NP_{rel} corresponds to a specific grammatical function within S_{rel} such as S, A, O, or Oblique (Keenan 1985). It has been shown cross-linguistically that there is an implicational hierarchy involved in the ability for the various elements to be relativized, as diagrammed in **FIGURE 3.21**.

Keenan and Comrie’s (1977) ‘Accessibility Hierarchy’ S>DO>IO>Obj of adpos>Possessor states that ‘the lower an NP_{rel} is on the Hierarchy, the more common it is to find it expressed by a personal pronoun’ i.e., it is uncommon for a relativized S to be expressed by a personal pronoun.

S < A < O < OBL < POSS < COMPARATIVE

FIGURE 3.21 – Implicational hierarchy for raising from relative clauses. The symbol ‘<’ can be read as ‘implies member to the left’ (Keenan 1985).

An element to the right of ‘<’ in **FIGURE 3.21** implies that the preceding element, that is, the one to the left, can also be relativized. In other words, if O can be relativized, so can A, and so can S, but it does not imply that obliques, possessives, or comparatives can be relativized. In Irarutu, S, A, O, and Obl can be relativized. 405) shows that the relativized NP can originate in a clause as S.

- 405) *matu uf bu-nanun serne=ti*
 person REL PROG-3-know God=NEG
 ‘heathen (lit. person who knows not God)’

An A can also be relativized. The next set of examples shows an adjectival predicate intransitive construction 406), a transitive construction 407), and a complex clause 408), in which the structure of the main clause corresponds to 406) and its NP_{rel} correlates with A in an S_{rel} that corresponds to 407).

- 406) *a-matu=ro tni trro*
 1S.POSS-friend=TOP body tall
 ‘My friend is tall.’
- 407) *a-matu n-tim frrufn adi*
 1S.POSS-friend 3-close door DET
 ‘My friend closed the door.’
- 408) *[[a-matu]_{NPrel} uf [____ n-tim frrufn adi]_{Srel}]=ro tni trro*
 my-friend REL 3-close door DET=TOP body tall
 ‘My friend who closed the door is tall.’

An O can also be relativized. The first example below illustrates another adjectival predicate intransitive construction. The next example is a transitive construction in which the person features of the subject and object are different (first versus third person). The third example shows that S of the main clause is an NP_{rel} that corresponds with the grammatical relation of O in S_{rel}.

- 409) *matu ad=ro rgun bar*
 person DET=TOP head bald
 ‘The person is bald.’
- 410) *ja si matu*
 1S see person
 ‘I saw a person’
- 411) *matu ad-uf [ja si ____] (ad=ro) rogun bar*
 person DET-REL 1S see DET=TOP head bald
 ‘the person that I saw is bald’

The complex clause that shows relativized O is unusual because of the occurrence of an extra determiner between the head noun and the relativizer. (The occurrence of a determiner plus topic clitic after S_{rel} is expected, see 412) below.) Other examples also attest that this element is optional, an observation that is compatible with the assumption that NP_{rel} is fully specified and definite. A determiner that modifies an NP_{rel} redundantly marks specificity and definiteness.

- 412) *man fi [i=si ____] adi*
 bird REL 3S=see DET
 ‘birds that he sees’

In some contexts, the presence of a determiner would be inappropriate, as in 413), which shows that it is possible for an NP_{Rel} such as *nya* ‘something’ to be non-referential but specific and definite (the person shoots identifiable entities, i.e., things he saw, but the set of possible entities is rather large, including, for example, various types of birds, deer, and so forth).

- 413) *matu n-taf nya fi [i=si ____]*
 person 3-hunt what REL 3S=see
 ‘the hunter shoots things that he can see’

In constructions with three arguments, the first object (O₁), as well as the second object (O₂) can be relativized. Notably, a determiner is used as a resumptive strategy and occurs in the slot where the first object originated.

- 414) *warada fi [i n-ar ad na-n-tut sum] adi=ro bu nsir*
 knife REL 3S 3-take DET INF-3-cut fish DET=TOP PROG 3.dirty
 ‘the knife which she used to cut the fish with is dirty’

The next three examples show how S of an intransitive can be an NP_{Rel} that corresponds with the second O, the position of which is indicated by PRO, of a clause with two objects.

- 415) *matu na-fe*
 person 3.eat-DTR
 ‘the person eats’
- 416) *ja ar uce ra-ge o-matu*
 1S give papeda 1-to 2S.POSS-friend
 ‘I gave papeda to your friend.’
- 417) *matu fi [ja ar uce adi ra-ge ____] g-e-su na-fe n-fi- tab*
 person REL 1S give papeda DET 1-to PRO DES-3-CPL 3.eat-DTR 3-ACT-add
 ‘the person who I gave papeda wants to eat more’

The final elements that can be reasonably shown to undergo relativization in Irarutu are obliques. The following three examples show how the simple structures correlate with the complex clause.

- 418) *rre win*
 day hot
 ‘today is hot/a hot day’
- 419) *i n-yuen dub adi nene mumce*
 3S 3-weave noken DET PREP morning
 ‘He made the noken (kind of hand woven shoulder bag) in the morning’
- 420) *mumce fi [i n-yuen dub ____] adi=ro buna win*
 morning REL 3S 3-weave noken DET=TOP most hot
 ‘the morning when he made a noken was hot’

Possession expressed through syntactic means alone (*ja skripsi* ‘my thesis’), in contrast with morphological possession (*a-fra-g* ‘my-hand-my’), is not subject to relativization due to the role of relativization to express emphatic possession, e.g., *a-den (fi) ja-ri* ‘my-mother (which) I-have’. In effect, the participants in a possessive relation are only relativized in the same way as A or O in a standard relativization construction. Relativization of comparatives is also not possible, but this claim could be overturned by future research.

Overall, the data supports the relativization hierarchy presented in **FIGURE 3.21**. Obliques and all lower-ranking (to the left of ‘<’) grammatical functions can be relativized in Irarutu. In addition to originating in a variety of grammatical relations in S_{rel} , NP_{rel} can occur in a variety of grammatical relations in the main clause, including S (408, above), A, O, as well as an Oblique. An example is provided below.

- 421) *bariku adi fi=ni*
 lighter DET REL-DEM
 ‘that lighter (lighter the which here)’

3.3.2.3 Differences between the relativizers *uf* and *fi*

There are distributional-functional constraints on the occurrence of the relativizers *uf* and *fi*. In some regards, *uf* appears to correlate more closely with a referential, restrictive relativizer than *fi*, which correlates more closely with an assertive, non-restrictive relativizer. However, according to use, *uf* is used in contexts where the referent that is relativized is assumed to be known to the speaker or the hearer from discourse context, whereas *fi* is used in contexts where the relativized referent is being asserted, or introduced, into the discourse context. The following two examples use non-restrictive relative clauses.

- 422a) *a-matu uf n-tim frrufn adi=ro tni trro*
 my-friend REL 3-close door DET=TOP body tall
 ‘my friend who closed the door is tall.’
 422b) **amatu fi n-tim frrufn adi=ro tni trro*
 423a) *matu fi n-tim frrufn adi=ro tni trro*
 person REL 3-close door DET=TOP body tall
 ‘the person who closed the door is tall.’
 423b) **matu uf ntim frrufn adiro tni trro*

In the first pair of examples above, *amatu* ‘my friend’ is a known referent to the speaker (because the referent is already known to me and is my friend) and occurs with *uf*. It is not well-formed to use *fi* in this context. In the second pair of examples, *matu* ‘person’ is a referent that is not presumed to be identifiable based on previous discourse and occurs with *fi*. It is not well-formed to use *uf* in this context.

- 424a) *matu uf m-tim frrufn-ni=ro a-matu ja=ri*
 person REL STAT-close door-PROX=TOP 1.POSS-friend 1S=have
 ‘the person who closed the door is my friend’
- 424b) **matu fi m-tim frrufn-ni=ro a-matu ja-ri*

Based on examples 422–3), the equational sentence in 424) should use *fi* to introduce the RC [*mtim frrufnni*] because *matu* ‘person’ is not identifiable from previous discourse, but the predicate *amatu* ‘my friend’ is construed to mean that *matu* ‘person’ is, in fact, identifiable (‘the person is my friend’). Therefore, positive identification of a referent as belonging to the discourse context can be achieved by NP_{rel} or elsewhere in the construction. For this reason, using *fi* to relativize *matu* in 424b) is ill-formed. An example where *uf* is selected by the context of the relative clause, i.e., seeing a person presupposes that person’s existence, is given below.

- 425) *matu ad-uf [ja si ____] ad=ro rgun bar*
 person DET-REL 1S see DET=TOP head bald
 ‘the person that I saw is bald’

There are some contexts where either *uf* or *fi* can be used, such as citing qualities that are usually expressed by a relative clause. This again appears to be due to differences in how a speaker views a particular linguistic context at a particular time.

3.3.2.3 On relativizers, determiners, and resumption

Several of the examples above, repeated below for simplicity of reference, show that relativizers and determiners (as well as the topic marker) interact.

- 426) *matu ad-uf ja si (ad=ro) rau gun bar*
 person_i DET-REL 1S see DET=TOP COP head bald
 ‘the person that I saw is bald’
- 427) *bariku adi fi=ni*
 lighter DET REL=DEM
 ‘that lighter (lighter the which here)’

The determiner associated with a relativized noun usually occurs at the end of NP, after S_{Rel}. However, a determiner (*ad*) can also occur between the head noun and a relativizer (*uf*). In this position, the determiner is clearly adjoined to the head noun, but it contracts to the relativizer, e.g., *ad=uf*, *adi=fi*. The role of determiners in other locations is less clear-cut.

- 428) *wenum ad=uf ro dir nü adi=fi Fruat*
 land DET=REL COP good name DET=REL Fruata
 ‘there is a place that is very beautiful called Fruata
 (lit. the land which is good has the name which is Fruata)’

One related phenomenon that is particularly pertinent for the description of relativization is the use of determiners in resumption. Resumption has two typologically common

solutions. One is to leave a gap in the structure where the raised element originated. The other is to insert a resumptive element where the gap occurs. In contexts where O is relativized, the determiner *ad* ‘the’ appears to function as a resumptive pronoun. Example 429) shows a mono-clausal construction to compare with 430), which has a topicalized N_{Rel} that originated as O₁ of S_{Rel}. This use of the determiner in relativization is only attested for relativized (primary) objects.

- 429) *i=n-ar warada na-n-tut sum*
 3S=3-take knife HAB-3-cut fish
 ‘he uses a knife to cut fish’
- 430) *warada [fi i=nar=ad na-n-tut sum adi]=ro bu nsir*
 knife REL 3S=take=DET HAB-3-cut fish DET=TOP PROG 3.dirty
 ‘the knife which she used to cut the fish with is dirty’

3.4 Speech act distinctions and syntax

The grammatical system described above also interfaces with various types of speech acts. Sadock and Zwicky (1985) proposed three types of sentences are used to express universal functions: declaratives, imperatives, and interrogatives. Two dimensions involved in the system of sentence types in Irarutu concern specificity of functions and higher order affinities. Declaratives serve a wide range of functions. Negatives are formed by adding a negative clitic at the end of a clause. In Irarutu, the three sentence types are strikingly similar, which supports a view that there is higher order affinity among them. Declaratives and imperatives are similar in having S before V, but imperatives tend to lack an overt S. Declaratives and interrogatives have strikingly similar word order, except that content interrogatives have an interrogative pro-form. Minor sentence types (such as exclamations, imprecatives, and optatives) will not be explored here.

The following, multiple paired-part example is from a typical husband-wife greeting and illustrates each of the three types of sentences. An analysis of Irarutu discourse structure is deferred to future research, but Appendix 2, which presents two short examples of discourse, will be of particular value to readers who are interested in exploring relevant data. Two paired-part examples are cited below.

- 431a) *refefa dir, o=ro bun nya fi mo jie-e?*
 evening good 2S=TOP 2.do what REL IRR happen-POL
 ‘Good evening, what are you doing there?’
- 431b) *ja besu m-tür fene (ja) fun-nya nyati*
 1S just STAT-sit only (1S) do-AGRO nothing
 ‘I’m just sitting, not doing anything.’
- 432a) *najiro m-ro beguen wer+win nene termus adi=ro*
 if.so 2-run 2;see hot+water PREP thermos DET=TOP
 ‘if that’s so, go see if there is some hot water in the thermos.’
- 432b) *termus adi=mo wer+win=ti=o*
 thermos DET=DISJ hot+water=NEG=POL
 ‘the thermos doesn’t have hot water.’

3.4.1 Declaratives

Irarutu has unmarked declaratives. These are the basic sentence type. Declaratives convey assertions, make announcements, express beliefs or doubt, report, conclude, narrate, assess likelihood, and so forth. This type of sentence is subject to judgments of truth and falsehood. Cross-linguistically, subtypes include: indefinite, inferential, dubitative, and emphatic. Formally, declarative sentences (433) have the same word order as subordinate clauses such as relative clauses (434), complement clauses (435), and conditional clauses (436).

- 433) *ja besu m-tür fene (ja) fun-nya nyati*
 1S just STAT-sit only (1S) do-AGRO nothing
 ‘I’m just sitting, not doing anything.’
- 434) *[[a-matu]_{NPre} uf [____ n-tim frrufn adi]_{Srel}]=ro tni trro*
 1S.POSS-friend REL 3-close door DET=TOP body tall
 ‘My friend who closed the door is tall.’
- 435) *mo adi n-fi-drbun [finnar adej ro dir]*
 child DET 3-ACT-think joke DET COP good
 ‘the child thinks the joke is funny’
- 436) *ja fier ge m-tbe bnisr nmse na-mtat*
 1S say COMP 2-hit mosquito until HAB-dead
 ‘I said to hit the mosquitos until (they are all) dead’

Negatives are generally formed by adding the negative clitic =*ti* at the end of a corresponding declarative, without altering word-order. The negative marker can interact with TAM markers, but, for example, a verb complex with *du* ‘progressive’ and the negative marker (negative progressive) is not equivalent to *bu* ‘non-progressive’.

- 437) *ja du kka=ti*
 1S NPROG bathe=NEG
 ‘I haven’t bathed yet’

3.4.2 Imperatives and Vetatives

The second sentence type that reflects a specific pragmatic function is the imperative. In relation to declarative sentence types, there is a distinct syntactic form to indicate the speaker’s desire to influence future events: make requests, suggestions, and entreaties, as well as give commands, orders, instructions, and so forth. To indicate the speaker’s wish to effect the addressee’s action, the imperative is restricted to second person subjects. Imperatives are notionally future. Sadock and Zwicky (1985) point out that agreement is redundant in imperatives because the subject naturally refers to the addressee; and furthermore, aspect is redundant because the prototypical request is to do some task to completion. In Irarutu, imperatives frequently lack an overt subject, but agreement morphology is always present although the verb in an imperative does tend to have fewer than the normal number of affixes.

Also reflecting the use of this sentence type to effect the actions of an addressee, many Irarutu imperatives are based on serialized-verb, directional constructions with very little additional morphology.

- 438) *gud tab ma-ma ja isi*
 SAP.take add 2-come 1S see
 ‘bring it here for me to see’
- 439) *ba m-ud suri ma-ma adini*
 2.go 2-get axe 2-come DET
 ‘go and get the ax and bring it over here!’ (Matsumura & Matsumura 1991)
- 440) *m-ar buku ma-mri m-tu*
 2.take book 2.take-enter 2-save
 ‘Go put your book away (in the house)’

The two preceding examples illustrate that objects in imperative sentences are treated in the same way they are treated in declaratives. Imperatives can include relative clauses; but the ability for dependent clauses to function as imperatives is not explored here.

- 441) *bier fi bergin*
 2.speak REL enough
 ‘quiet down (enough talking)!’

One subtype of the imperative is the hortative, ‘let’s...’. In Irarutu, this type of function is easily achieved by the use of the first person inclusive pronoun subject *it* ‘we’.

- 442) *it fa rit dedan*
 1P go tend garden
 ‘Let’s go tend the garden (lit. we go tend garden)’

Another important subtype of the imperative is the vetative/prohibitive, a negative imperative that is different from the negative declarative, which uses the clitic *=ti*. In the prohibitive, the marker *=of* ‘do not’ adds the negative semantic. The next example contains both *=ti*, and *=of*, which allows a good picture of their use and distribution.

- 443) *mo e=re bitr=ti, [ro e=ga=of]*
 COND 2P=TOP hungry=NEG, then 2P=SAP.eat=PROH
 ‘If you are not hungry, then don’t eat’
- 444) *bi-minir bier-fe(=of)*
 2.ACT-stop 2.speak-DTR(=PROH)
 ‘silence (enough talking)!’

A cross-linguistically common alternate strategy to alter the addressee’s action is to use a declarative sentence beginning with *te o...* ‘you will...’, which serves as an assertion of a possible consequence such as ‘don’t do that (or...)’.

- 445) *te o m-tit*
 later 2S STAT-fall
 ‘you will fall’

3.4.3 Interrogatives

The third universal sentence type is the interrogative, which is used to elicit a verbal response from the addressee as a way to gain information. Typologically, two basic question types are universally recognized: yes/no (also called ‘nexus’) questions, and *wh*- (also known as ‘content’) questions. These two question types do not form a unified class. This is due to the fact that the two basic types have distinct syntactic and phonological properties. Nexus questions are the most universal. Coincidentally, none of the available data suggest that Irarutu uses dependent interrogative clauses, which would appear as the complement of ‘ask, say, know’ etc. Biased questions and rhetorical questions are not explored here.

3.4.3.1 Yes/No Questions

This subtype of interrogative sentence has a characteristic rising-final intonation pattern (cross-linguistically similar to the intonation of a non-final disjunct), which marks a clause’s interrogative force. These types of sentences are generally used to determine if an addressee can confirm or negate the content of a proposition. Formal indications of a nexus question resemble the antecedent in conditional sentences ‘if...’. They have the same word-order as declaratives, and intonation is often the only feature that distinguishes a construction as a nexus question.

- 446) *o=ro du kka?*
 2S=TOP NPROG bathe
 ‘haven’t you bathed? (are you going to bathe?)’
- 447) *ir g-e nin kopi*
 3P DES-3 NSAP;drink coffee
 ‘They want to drink coffee/Do they want to drink coffee?’

Occasionally a clause final element *fe* (possibly the same as Matsumura and Matsumura’s *e*) signals a construction is a nexus interrogative. Apart from intonation, which is effectively identical between the pair of examples below, the morpheme *fe* can be used to signal a question.

- 448) *matu adesai=ro kapar wenum*
 person DET=TOP chief land
 ‘That man is a village chief’
- 449) *matu adesai ro kapar wenum fe?*
 person DET TOP chief land Q
 ‘Is that man a village chief?’

Alternative questions, in which the speaker provides a list for which the right, i.e., mutually exclusive, answer might be selected – minimally a proposition and its negation,

or manner other than polarity – can be grouped formally with nexus questions. For example, rising intonation occurs on each alternative except the last.

- 450) *Jek, g-o gin te te, kopi te, wer te*
 Jek, DES-2 SAP.drink tea DISJ, coffee DISJ, water DISJ
 ‘Jack, you will drink tea, coffee, or water/
 Jack, do you want to drink tea, coffee, or water?’

Alternative questions are distinct from leading questions, i.e., questions where the answer is only ‘no’ if none of the clauses are true, but ‘yes’ if one or more is true, with rising intonation on every alternative, including the last.

When the topic clitic occurs in a nexus interrogative, the resulting structure can be ‘focused’ on the phrase that hosts the topic marker. This type of sentence is similar to a biased question. Both display a belief that some proposition is true. A negative adverb added to an interrogative sentence forms a positively biased nexus question.

- 451) *o=ro du n-tim frufn=ti?*
 2S=TOP NPROG 3-close door=NEG
 ‘why did you not close the door already/why isn’t the door closed?’

The final type of nexus-group interrogative sentence is the confirmative, which demands that the addressee express agreement or disagreement with a proposition or its intended function. This type of sentence lacks question intonation, but often ends with a tag.

3.4.3.2 Information Questions

Speakers use information/question-word questions to present alternatives in an open-ended way. These constructions can solicit a response that involves quantification. Semantically, information questions are similar to alternative questions because they specify a range from which to select an answer. Cross-linguistically, this type of interrogative uses interrogative proforms as well as inversion and special morphology, but these structures sometimes also use the indicators of nexus questions, such as particles and intonation. In Iraputu, declaratives and interrogatives tend to be structurally identical.

- 452) *motor matu i=ri ____*
 motorbike person 3S=have
 ‘that person has a motorcycle’
 453) *gata i ri ____ ?*
 who 3S have
 ‘who’s is this?’

Functionally, information questions interrogate part of a proposition and imply that the rest of the proposition is old/presupposed information. New information is requested for the identity of the interrogated part of the sentence, its ‘focus’ or ‘topic’ (in the sense of being what the sentence is about). Cross-linguistically, interrogative proforms are

commonly found in focus/topic position. In Irarutu, an interrogative form can occur in the same syntactic positions as a non-interrogative form, including a structural topic.

- 454) *nya fi n-gor eje?*
 what REL 3-bite type
 ‘what bit you?’
- 455) *Jek, g-o gin nya?*
 Jek, DES-2 SAP.drink what
 ‘Jack, what do you want to drink?’
- 456) (o) go ba=nenó?
 (2S) DES-2 2.go=where
 ‘where are you going?’
- 457) *o=ro m-ar nya fi o ma-msi o adej?*
 2S=TOP 2-give what REL 2S 2-2.see 2S DET
 ‘what do you use to see yourself?’

Cleft constructions, such as for topicalization, can be used to stress the new information/old information dichotomy by focusing on an interrogated constituent. These constructions are the same for declarative sentences. According to Rizzi (1997), in the topic-comment structure, the topic is given and followed by new information in the comment. In Irarutu, these often have a resumptive element, especially for topicalized direct objects. Multiple elements can be topicalized, and they can precede *wh*- words. Focus-presuppositions on the other hand present new information followed by given information. These constructions do not have a resumptive element. Quantificational elements are allowed to be focused. In other words, focus must be unique because it provides a unique piece of information.

An interrogative word, such as an interrogative pronoun in Irarutu (possibly combined with another strategy), can indicate what part of a proposition the addressee’s attention is being guided towards. From a typological perspective, interrogative words also limit the field the unknown is a part of and have the same function as ‘whoever, whenever, wherever’, etc., but indicate interrogative force. The minimum number of attested question words in a language is three: ‘who’, ‘who/what’, and ‘where’, whereas some languages have a dozen or more interrogative forms. Irarutu has five question words, listed in **FIGURE 3.22**, below. Speakers distinguish between personal *gata* ‘who’ and impersonal *nia* ‘what’ interrogative pronouns, and use interrogative proadverbials for ‘where’, and ‘how’.

<i>gata</i>	‘who’
<i>nia</i>	‘what’
<i>nenó</i>	‘where’
<i>fna(ne)no</i>	‘how’
<i>nfnanuf</i>	‘why’

FIGURE 3.22 – Irarutu question words.

Although it does not appear to have a direct bearing on their usage, the words that correlate to ‘how’ and ‘why’ appear to be morphologically complex. On best approximation, *fnaneno* appears to be composed *fna* ‘like’ or ‘causative marker’ and the pro-adverbial *nen* ‘where’. The word *nfnanuf* appear to contain a third person subject marker, *n-*, followed by the short form of the word for ‘how’, i.e., *fna*, plus a grammaticalized relativizer *uf*.

The word '*nia* (~*nya*) is a general interrogative proform. In conjunction with an auxiliary, it can be used as a pro-verb (‘to do what?’).

- 458) *O=ro bun nya?*
 2S=TOP 2.do what
 ‘what are you doing?’

A periphrastic construction based on a temporal word (such as *rre* ‘day’) followed by *nya* ‘what’, is used for temporal adverb, i.e., ‘when’, questions.

- 459) *ja mi mse rre nya?*
 1S stay until day what
 ‘how much longer may I stay?’

Two more alternatives that cross-linguistically achieve the same result as an information question, and are found in Irarutu, are indefinite statements and indefinite nexus questions. Example 460), ‘why did who do what to who’ can also be interpreted as a multiple wh- question. The dual function of certain words as indefinite pronouns and interrogative pronouns is known to be cross-linguistically very common.

- 460) *nano gata n-fun nya nene gata?*
 how who 3-make what PREP who
 ‘why did who do what to who?/why did someone do something to someone?’
- 461) *te i=ro bis dru?*
 later 3S=TOP can 3.shoot
 ‘what can he shoot?’

3.5 Summary

This chapter presented four aspects of Irarutu syntax and morphosyntax: the parts-of-speech found in Irarutu, a description of how four semi-independent components constitute the grammatical system in Irarutu, an introduction to complex clauses with particular attention to relative clauses, and a brief survey of three sentence types. Some noteworthy points include: the language uses highly ‘configurational’ SVO word order, there are no case markers, a cross-linguistically exceptional structure is given to the genitive in relation to the head noun provided ordering constraints on other phrasal elements, the language uses a ‘fluid split-S’ alignment system, and a serial verb construction is used to introduce indirect objects. In addition, several grammatical phenomena indicate that the distinction between speech act participants and non-speech act participants is important in the grammar. These include having an inalienable

possessive suffix, inflection in irregular verbs, and suppletion in second verbs in serial verb constructions. Areas identified for future research include topicalization, complementation, and serial verb constructions.

Chapter 4. Historical phonology and genetic affiliation

A large obstacle for describing the historical phonology of Irarutu, and making progress in assessing the genetic affiliation of the language, and, more generally, the internal and external interrelationships of languages in the South Halmahera-West New Guinea subgroup of the Austronesian (An) language family, has been the low demonstrable retention rate of Austronesian etyma, which has been – and still is – due primarily to lack of accurate information. This has caused uncertainty in pin-pointing reliable sound correspondences. An overview of the situation is captured in the 200-word list presented in Appendix 3a. In Irarutu, the retention rate is approximately thirty-five percent (up to 67 retentions/190 lexical items), see **TABLE 4.1**. Proto-Central Eastern Malayo-Polynesian (PCEMP) is the closest node that dominates Irarutu, and which is reconstructed sufficiently for comparative purposes (Blust 1993). This chapter examines, in detail, sound correspondences in eighty-six lexical items that are contained in the 200-word list and that appear to correspond with PCEMP. For the purpose of identifying recurrent correspondences, sixty-four lexical items are grouped into seven general categories in the table below. There are also nineteen lexical items that have a high enough degree of similarity with a reconstructed PCEMP form to be evaluated for status as cognate, but show ‘sporadic’, i.e., only one token in the data set, and aberrant, i.e., conflicts with the ‘recurrent’, or more than two tokens in the data set, reflexes of proto-segments. ‘Semi-recurrent’ correspondences, or those that are instantiated by only two tokens, are also taken into account. By all indications, as the sample set is increased (1500 lexical items have been collected for the language), the percent of demonstrable cognates decreases significantly.

No.	PMP	PCEMP	IRH	Gloss
I. Linear deletion 1: Final syllable rime deletion (32 examples, 2 are exceptional)				
1.	susu	susu	sus	‘breast’
2.	taŋis	taŋis	tag	‘cry’
3.	kaen	kan	ga	‘eat’
4.	inum	inum	in	‘drink’
5.	qinep	qenəp/qinəp	m	‘lie down’ (IRH <i>m</i> > ‘sleep’)
6.	ma-tudan	tudan/todan	tür	‘sit’
7.	ma-diRi	diRi	m-rir	‘stand’ (PMP *ma-; IRH <i>m</i> - ‘stative’)
8.	ma-Ruqanay	ma-Ruqanay	mran	‘male/man’
9.	bahi/ b-in-ahi/	bai/b-in-ay/		
	ba-b-in-ahi	ba-b-in-ay	bfin	‘female/woman’ (IRH <i>sot</i> ; bfin > ‘wife’)
10.	Rumaq	Rumaq	rum	‘house’ (IRH <i>san</i> ; <i>rum</i> > ‘village’)
11.	bunuq	bunuq	n-fun	‘kill’ (IRH <i>n</i> - is morphological)
12.	matay	matay	n-mat	‘die (dead)’ (IRH <i>n</i> - is morphological)
13.	kahiw	kayu	e	‘wood’ (IRH ‘tree’; <i>e-ntü</i> ‘wood’)
14.	tanem	tanəm	n-tan	‘to plant’ (IRH <i>n</i> - is morphological)
15.	baReq	baRəq	br-bar	‘swell’
16.	kali	kali/keli	gar	‘dig’
17.	tuktuk	tutuk	m-tut	‘to pound’ (IRH <i>m</i> - is morphological)
18.	manuk	manuk	man	‘bird’
19.	kutu	kutu	ut	‘head louse’
20.	dahun	daun	ro	‘leaf’ (IRH <i>e-ro</i> ‘leaf’)

21.	buah	buah	fu	‘fruit’
22.	hiup	upi	uf	‘to blow’
23.	tunu	tunu	tün	‘burn’ (IRH sfrer; <i>tün</i> > ‘cook’)
24.	ma-putiq	ma-putiq	bfut	‘white’
25.	mataq	mataq	m-mat	‘green (unripe)’ (IRH <i>m</i> - ‘stative’)
26.	i-kahu	i-kau	o	‘you’
27.	si-ia	s-ia	i	‘he/she’
28a.	i-k-ita	k-ita	it	‘we (inclusive)’
28b.	i-k-ami	k-ami	am	‘we (exclusive)’
29.	si ida	sida	ir	‘they’
30.	duha	dua	ru	‘two’
31.	tuqela(n,ŋ)	zuRi	<u>ru</u>r	‘bone’
32.	telu	təlu	<u>to</u>r	‘three’

II. Linear deletion 2: Final syllable deletion (5 ex.s, 1 is exceptional)

33.	panaw	panaw	fa(n)	‘walk, go’
34.	um-aRi	mai/maRi	ma	‘come’
35.	buja	buja	fü	‘flower’
36.	i-aku	i-aku	ja	‘I’
37.	qalejaw	qaləjaw	re	‘day’ (IRH <i>rre</i> ‘day’)

III. Linear deletion 3: Stress shift and deletion of the penult (3 ex.s)

38.	kulit	kulit	rit	‘skin’
39.	qulu	qulu	rü	‘head’ (IRH: ‘source’)
40.	ma-bener	bənər/tuqu	tü	‘correct, true’

IV. Targeted deletion 1: Final syllable onset deletion and vowel coalescence (4 ex.s, 1 excep)

41.	daqan	daqan	ran	‘branch’
42.	uRat/wakaR	wakaR	war	‘root’
43.	wahiR	waiR	w <u>er</u>	‘water’
47.	ma-takut	ma-takut	mtat	‘afraid’

V. Targeted deletion 2: Final syllable nucleus deletion (4 ex.s, 1 is exceptional)

44.	zalan	zalan	rarn	‘road’
45.	laŋit	laŋit	ragt	‘sky’
46.	ma-diŋdiŋ	ma-dindiŋ	dridn	‘cold’
48.	deŋeR	deŋəR	f-nogr	‘hear’ (f- is morphological)

VI. Targeted deletion 3: Stress shift and deletion of the penult nucleus (10 ex.s, 2 excep.s)

49.	qabaRa	qabaRa	fra-fu	‘shoulder’ (IRH <i>fra</i> > ‘arm’; <i>fu</i> ‘bony’)
50.	lipen/nipen	lipən	rfo	‘tooth’
51.	tawa	tawa/malip	brif	‘laugh’
52.	taliŋa	taliŋa	tgra	‘ear’
53a.	anak	anak	—	‘child’ (IRH <i>mo</i>)
53b.		natu	ntü	‘young’
54.	qateluR	qatəlur	tru	‘egg’
55.	bulu	bulu	fru	‘feather’
56.	labaw	kanzupay	sfe	‘rat’ (IRH: ‘mouse’)
57.	mata	mata	mtie	‘eye’
58.	a taqas/ i babaw	atas/i babaw	ffu	‘above’

VII. No deletion in penult or final syllable (6 ex.s, 3 exceptional)

59.	ini/ani	ini	-ini	‘this’
60.	buni	buni	<u>br</u> -funi	‘hide’ (IRH <i>br-</i> is morphological)
61.	dilaq	maya	m-maje	‘tongue’
62.	tau/taumataq	tau/taumataq	matü	‘person’
63.	qatəp	qatəp	atīf	‘thatch’ (IRH <i>atīf-ro</i> ‘thatch’)
64.	esa/isa/tasa	əsa/isa/tasa	esu	‘one’

Table 4.1 – Sixty-four proposed cognates in Irarutu, cited from Appendix 3a and following the order there, but renumbered for clarity. Items are grouped according to seven changes (I.–VII.) in word shape in the last two syllables. ‘Linear deletion 1’ may have first involved final consonant deletion followed by apcope. ‘Linear deletion 2’ appears to have involved a second stage of final consonant deletion. ‘Targeted deletions’ 1 and 2 involve deletion of vowels in historically final syllables. Fifteen exceptional forms, identified by **boldface** contain a sporadic sound correspondence (indicated by underlining).

Twelve Irarutu items that are cited in **TABLE 4.1** are noted as ‘exceptional’ (the boldface items), but within the margin for error. These forms contain sound correspondences that are attested by sub-recurrent, i.e., fewer than three instantiations in the data set, but have probabilistic support, i.e., the proportion of exceptional changes versus the number of segments that show recurrent sound correspondences suggests cognacy, see **FIGURE 4.1**.

All segments

- 48 sound correspondence sets reflect 22 PCEMP protosegments
- 15 correspondence sets only attested by one token
- 12 by two tokens
- 22 recurrent

Consonants

- 30 consonant correspondence sets reflect 16 PCEMP phoneme categories
- 8 by one token: *d/n_>d, *ŋ>n, *z>r, *z>s, *t>s, *r>Ø, *j>Ø, *w>f
- 9 by two tokens: *s>s, *w>w, *b>b, *k>g, *m>b, *R>Ø, *ŋ>Ø, *m>Ø, *t>Ø
- 13 recurrent (3 are retentions attested by more than 2 tokens)

Vowels

- 19 vowel correspondence sets reflect 6 PCEMP phoneme categories
- 6 by one token: *i/_CV_[-tense]>I, *u>ü, *ə>o, *ə>i, *a>ie, *i>j
- 4 by two tokens: *-aw>Ø, *a.u>o, *-a>u, PMP*ə>e
- 9 recurrent (3 PCEMP retentions)

FIGURE 4.1 – Summary of sound correspondences in proposed cognates, illustrating sub-recurrent correspondences, that is, those that occur in more than 1% of the data set. There are sixteen innovations (13–3=10 and 9–3=6, 10+6=16). See Appendix 3b for details.

The twenty-two recurrent sound correspondences referred to in **FIGURE 4.1** are described in **TABLE 4.2** on account of the fact that several Irarutu segments appear to correspond to multiple reconstructed PCEMP proto-segments (e.g., IRH *b* < PCEMP *m or *b), and many proto-segments have multiple reflexes in Irarutu (*b > IRH *f*, *b*, or Ø), several types

of conditioning environments need to be identified to help explain the variability (but see Blust 2005 and Boretsky 1984 for alternative explanations of historical change that do not rely on conditioning environment). The most important factor appears to be distance from the end of a word, followed by location of stress, and two patterns of deletion ('linear' and 'targeted'). Both types of deletion can occur in a single word at opposite ends. Of the twenty-two recurrent sound correspondences, six are retentions and sixteen are innovations. In conjunction with proportion of recurrent sound correspondences identified in each word, these twenty-two correspondences function as primary criteria for evaluating if a lexical item reflects an Austronesian proto-form.

Retentions (6)

	PREPENULT		PENULT		FINAL			
	*__'σσ ₁	*__'σσ ₂	*'__σ ₁	*'__σ ₂	*__'σ	*'σ__ ₁	*'σ__ ₂	*σ'__
*t>t (22)		1	11		1	4	1	4
*m>m (12)	1	4	4		1	2		
*n>n (10)					1	7	1	
*a>a (18)	—	—	13	3		—	—	2
*u>u (17)	—	—	13			—	—	5
*i>i (9)	—	—	7	2		—	—	2

Innovations (16)

	*__'σσ ₁	*__'σσ ₂	*'__σ ₁	*'__σ ₂	*__'σ	*'σ__ ₁	*'σ__ ₂	*σ'__
*p>f (5)			1			1		3
*b>f (7)			5		2			
*ŋ->g (3)					1	1	1	
*q>Ø (14)	3			1	2	7	1	
*k>Ø (11)			6	1	1	2	1	
*l>r (11)			1	1	1	2		6
*R>r (9)		1	3			2	2	1
*d>r (7)			3	2		2		
*n>Ø (6)			1			3	1	1
*s>Ø (3)			2			1		
*ay>'e (3)	—	—	1	1				1
*-ay>Ø (3)						3		
*a>Ø (21)	1	6			4	8	2	
*u>Ø (17)	1	1			5	10		
*i>Ø (13)		1			2	8	2	
*ə>Ø (5)					1	4		

TABLE 4.2 – Recurrent reflexes of specific PCEMP segments. Each change occurs three or more times in its row, totals in parentheses. Bold vertical lines separate prepenultimate, penultimate, and final syllable positions. Subscript ₁ indicates a class of words that underwent 'linear' erosion of segments from the edge of a word, whereas subscript ₂ indicates deletions that target a specific position.

In addition to the individual innovations cited in **TABLE 4.2**, there are two classes of superficially overlapping multisegmental changes, which are represented as abstractions in **FIGURE 4.2**. This is an important point because it indicates that there are two types of ‘linear’ deletions. Furthermore, conflict arises because the Irarutu reflex of a proposed etymon cannot be reliably predicted based on reconstructed word shape.

Final syllable rime deletion:	$*CV_1.CV_2(C) > *CV_1.CV_2 > *CV_1C$
Final syllable deletion:	$*CV_1.CV_2 > *CV_1C > *CV_1$

FIGURE 4.2 – Summary of multisegmental changes: final syllable rime deletion and final syllable deletion.

The sources of the secondary Austronesian forms, borrowed non-Austronesian forms, or possibly invented forms, are not very clear; but, there is evidence that surrounding languages have contributed to the Irarutu vocabulary (Anceaux 1958, 1961; Barr and Barr 1978b; Flaming 1983b; L. Peckham 1978, 1982a, 1982b, 1991; N. Peckham 1981a, 1981b, 1983; Peckham and Peckham n.d.; Smits and Voorhoeve 1994; Visser 1989).

The goal of this chapter is to evaluate and refine the current understanding of Irarutu’s genetic affiliation based on comparison with two proto-languages, PMP and PCEMP, as well as the preceding grammatical description (Chapters 2–3). Despite the concrete resources, classifying Irarutu is not a simple matter. The first section (§4.1) describes in as much detail as possible the historical phonology of Irarutu based on sporadic, semi-recurrent, and recurrent correlations (e.g., retentions and innovations) between Irarutu and PCEMP forms. It also presents a brief discussion of evidence that suggests there are lexical strata in Irarutu. The following section (§4.2) reviews literature that is pertinent to determining the genetic affiliation of Irarutu, starting with the oldest branch of the Austronesian family tree (Proto-Austronesian: PAn) and working down through the language family (Proto-Malayo-Polynesian: PMP; Proto-Central Eastern Malayo-Polynesian: PCEMP; and Proto-Eastern Malayo-Polynesian PEMP) to the youngest relevant major subgroup (Proto-South Halmahera West New Guinea: SHWNG) to include as many attested historical correlations as possible in the discussion. The third section (§4.3), following Malkiel 1968, synthesizes the linguistic data with metalinguistic and non-linguistic data (e.g., Anthony 1990), in an effort to formulate a hypothesis about the migration history of the Irarutu people. A conclusion (§4.4) ends the chapter.

4.1 Historical Phonology

In addition to the specific monosegmental and multisegmental changes identified in **TABLE 4.2**, several processes in the historical phonology need to be described to provide a sufficiently adequate picture of relevant factors in Irarutu, such as ‘stress shift’, ‘breaking’, and ‘unpacking’. In relation to PCEMP, deletion is obviously the most frequent innovation. There are two types of deletions, labeled ‘linear’ and ‘targeted’. The majority of deletions can be viewed as proceeding in a linear order from the periphery of a word eliminating multiple adjacent segments, e.g., $*C_1V_1.C_2V_2C_3 \rightarrow C_1V_1.C_2V_2 \rightarrow C_1V_1C_2$. A smaller set of deletions targets specific segments, e.g., $*q > \emptyset$, or specific positions, such as syncope, which targets the penultimate vowel: $C_1V_1.C_2V_2C_3 \rightarrow$

C₁C₂V₂C₃. Stressed penultimate syllables are thought to have been the default pattern in the history of the Austronesian language family (cf. Zorc 1983). A substantial number of Irarutu lexical items do reflect penultimate stress in inherited An words; however, both linear and targeted deletions result in monosyllabic forms, some retaining the historically stressed syllable, some involving stress-shift to reconstructed final syllables. Retentions are therefore as important as innovations in examining Irarutu historical phonology.

4.1.1 Retentions

The first set of sound correspondences that allows cognates to be identified are retentions. Three consonant retentions that occur in more than two instances between PCEMP and Irarutu are free of positional constraints *t>t, *m>m, and *n>n. The other three retentions that occur repeatedly are the cardinal vowels in stressed syllables *a>a, *u>u, and *i>i. There are no examples which show that any of these vowels, when they are stressed, have been innovated. For ease of reference, portions of TABLE 4.2 are repeated in their own tables.

	* <u> </u> 'σσ ₁	* <u> </u> 'σσ ₂	*' <u> </u> σ ₁	*' <u> </u> σ ₂	* <u> </u> 'σ	*' <u>σ</u> <u> </u> ₁	*' <u>σ</u> <u> </u> ₂	*σ' <u> </u>
*t>t (22)		1	11		1	4	1	4
*m>m (12)	1	4	4		1	2		
*n>n (10)					1	7	1	
*a>a (18)	—	—	13	3	—	—	—	2
*u>u (17)	—	—	13		—	—	—	5
*i>i (9)	—	—	7	2	—	—	—	2

TABLE 4.3 – Six PCEMP segment retentions in Irarutu, three are consonants and three are vowels.

4.1.1.1 Consonant retentions *t, *m, *n

It is worth mentioning that of the three consonant retentions that occur repeatedly, two alveolars, *t and *n, and two nasals *n and *m were retained, especially in stressed syllables, because it suggests relative perceptive-productive stability across time for at least one point of articulation, and one manner of articulation. A hyphen in an Irarutu form indicates the position of obligatory synchronic inflectional morphology, such as subject agreement or possession. Additional comments are given in parentheses.

PCEMP *t > IRH t

*taŋis > -tag ‘to cry’

*tunu ‘to burn’ > m-tiin ‘to cook’ (m- is lexicalized; semantic shift)

*matay > -mat ‘die’

PCEMP *m > IRH m

*ma-Ruqanay > mran ‘man/male’

*matay > n-mat ‘to die’

*mataq > m-mat ‘green/unripe’

but, *m > IRH b (in forms with targeted deletion)

*malif > brif ‘laugh’

*ma-putiq > bfut ‘white’

PCEMP *n > IRH *n*

***ini** > *-ini* ‘this/here’

***bunuq** > *-fun* ‘to kill’

***tanəm** > *-tan* ‘to plant’

4.1.1.2 Other consonant retentions *s, *w, *b, *d

Four more retentions are referred to in **FIGURE 4.1**, but with too few tokens to be called recurrent. These are *s > *s*, *w > *w*, *b > *b*, and *d > *d* (when it occurs after *n).

Although there are three citable examples of *b > *b*, this change is considered semi-recurrent because it is shadowed by the more well-attested shift *b > *f*. Under the assumption that only word-final glides can be reconstructed for PCEMP and historically prior proto-languages, initial *w-* is treated as an approximant (glide) consonant, but in coda position, *-w* [ɰ] is treated as the non-syllabic off-glide in a diphthong. (This positional difference also applies to *y [i], although it has undergone a separate type of innovation.)

Two of the three instances where *s > *s* are found in a single word, which could easily have been borrowed from a neighboring SHWNG language. The other instance is found in a three segment word where two of the segments reflect sub-recurrent innovations. Nevertheless, it is assumed the recurrent reflex of *s is a retention, *s*.

PCEMP *s > IRH *s*

***susu** > *sus* ‘breast’ (initial and medial positions)

***əsa** > *esu* ‘one’

but, *s > IRH Ø

***sida** > *ir* ‘3P (they)’

***sia** > *i* ‘3S (he/she/it)’

***taŋis** > *n-tag* ‘he cries’

It is possible that the third person pronouns were derived from proto-forms that lacked initial *s-, that is, PEMP *ida ‘3P’ and *ia ‘3S’, but, in relation to the reconstructed forms listed in Appendix 3a, it must be pointed out that *s- > Ø in the personal pronoun series did occur, even if it happened before the formation of Irarutu.

PCEMP *w > IRH *w*

***wakar** > *war* ‘root’

***waiR** > *wer* ‘water’

but, PCEMP *w > IRH *f*

***waqay** > *fa* ‘leg/foot’

PCEMP *b > IRH *b* (in initial position, and after a homorganic nasal)

***baRəq** > *-bar* ‘swell’

***ba-b-in-ay** > *bfin* ‘wife’

but, PCEMP *b > IRH *f*

***ba-b-in-ay** > *bfin* ‘wife’

PCEMP *d > IRH *d* (after a homorganic nasal)

*ma-dindin > *dridn* ‘cold’

but, PCEMP *d > IRH *dr* ‘rhoticization’

*ma-dindin > *dridn* ‘cold’

4.1.1.3 Vowel retentions *a, *u, *i

Segments in stressed syllables, and syllables that became stressed due to syncope, are considerably more likely to be retained than in unstressed, or de-stressed, syllables. This is especially relevant for vowels. Zorc (1983) reconstructs a handful of PAn forms with final stress: PAn *Ca'zem ‘sharp’, PAn *ti'aN ‘belly’, PAn *la'ɲuy/*na'ɲuy ‘to swim’, and PAn *ba'tu ‘stone’. Unfortunately, based on what is presently known about the Irarutu lexicon, none of these forms are relevant.

PCEMP *a > IRH 'a

*panaw > *fa* ‘go’

*manuk > *man* ‘bird’

PCEMP *u > IRH 'u

*susu > *sus* ‘breast’

but, PCEMP *u > IRH 'ü

*tunu ‘burn’ > *-tün* ‘cook’ (relexification)

PCEMP *i > IRH 'i

*inum > *-in* ‘drink’

*diRi > *rir* ‘stand’

4.1.1.4 Vowel retention *e

In addition to the retentions of *a, *u, and *i, it is worth mentioning that Irarutu *e* can generally be traced to PMP *e [ə] through PCEMP *ə, but there are two notable exceptions in the data set, PMP *telu > PCEMP təlu > *tor* ‘three’ and PMP *deɲer > PCEMP *deɲər > *nogr* ‘hear’ which both show correspondence of PMP *e [ə] > PCEMP *ə/e > IRH *o* (see also §4.2.3, below). Compare PMP *qateluR > PCEMP *qatəlur > IRH *tru* ‘egg’, which shows stress shift and syncope.

PMP *e > PCEMP *ə > IRH *e*

*qalejaw > qaləjaw > *re* ‘day’

*esa > əsa > *esu* ‘one’

but, PMP *e > PCEMP *ə > IRH *o*

*telu > təlu > *tor* ‘three’

*lipen > lipən > *rfo* ‘tooth’

also problematic, PMP *e > PCEMP *e > IRH *o*

*deɲer > *nogr* ‘hear’

The best hypothesis, albeit based on a semi-recurrent sound correspondence that has notable exceptions (in particular palatalization), for PCEMP stressed schwa is that it became labial before a (non-liquid; see, for example, possible labialization of *a/___l in §4.1.3.1) coronal consonant, PCEMP *ə > IRH 'e, as shown in the words for ‘day’ and ‘one’.

4.1.2 Innovations

Innovations in the historical phonology of Irarutu, of which sixteen are recurrent (see **FIGURE 4.1**), greatly outnumber its retentions, of which there are only six. The sixteen individual innovations altered eleven PCEMP protosegments: six consonants and five vowels (four monophthongs and one diphthong). To preview §4.2, four generalized innovations that are found in Irarutu have played a role in the literature for classifying a language as a member of the SHWNG group: 1) a consonant shift where PCEMP *p>f, which is hereafter called ‘fricativization’; 2) a vowel shift, which is hereafter called ‘labialized mid vowels’, i.e., (PMP) *e > o in the penult; 3) a consonant merger that will be called ‘dorsal stop deletion’, i.e., the stops *q and *k > Ø; and 4) a second class of consonant merger that will hereafter be called ‘rhoticization’, i.e., four segments *d, *z, *l, *R > r (Blust 1978, Jackson 2008). These changes are attested in Irarutu, but there are some exceptions in ‘labialized mid vowels’ (mentioned above) and ‘rhoticization’.

4.1.2.1 Recurrent consonant innovations

The ten innovations in the consonant inventory that occur in more than one percent of the data in the sample set include shifts and mergers falling into seven categories of context-free innovations: 1) fricativization, 2) ‘unpacking’, 3) dorsal stop deletion, 4) rhoticization, 5) ‘anticipatory voicing assimilation’, 6) four types of nasal-related phenomena, and 7) ‘glide fortition’, as well as one context-dependent innovation, ‘final consonant deletion’.

The differential treatment of *w and *y in the discussion of Irarutu historical phonology interacts with the recurrent process of final consonant deletion and the semi-recurrent process of glide fortition.

	*__'σσ ₁	*__'σσ ₂	*'__σ ₁	*'__σ ₂	*__'σ	*'σ__ ₁	*'σ__ ₂	*'σ'__
Fricativization								
*p>f ⁽⁴⁾			1			1		2
*b>f ⁽⁷⁾			5		2			
Unpacking								
*ŋ->g ⁽³⁾					1	1	1	
Dorsal Stop Deletion								
*q>Ø ⁽¹⁴⁾	3			1	2	7	1	
*k>Ø ⁽¹¹⁾			6	1	1	2	1	
Rhoticization								
*l>r ⁽¹⁰⁾			1	1		2		6
*R>r ⁽⁹⁾		1	3			2	2	1
*d>r ⁽⁷⁾			3	2		2		
Other Deletions								
*n>Ø ⁽⁶⁾		1	1			3	1	
*s>Ø ⁽³⁾			2			1		

TABLE 4.4 – Ten recurrent consonant innovations. The first two categories are shifts, the remaining three categories are mergers.

4.1.2.1.1 Fricativization of labial stops *p, *b > f

The fricativizing consonant shift of *p > f is paralleled by *b > v, but was putatively followed by a loss of phonemic voice distinction in the fricatives, v > f. For the voiceless bilabial, ‘fricativization’ was comprehensive, save for final stops, which deleted under final consonant deletion in some words; but for the voiced stop, it was not. A significant number of cognates show the retention *b > b, but with phonetic prenasalization.

PCEMP *p > IRH f

*upi > uf ‘blow’

*malip > brif ‘laugh’

*kanzupay > sfe ‘rat’

but, PCEMP *-p > IRH Ø

*qinəp > -m ‘sleep’

PCEMP *b > IRH f

*bunuq > fun ‘kill’

*buaq > fu ‘fruit’

*bulu > fru ‘feather’

*qabaRa ‘shoulder’ > fra ‘hand’

*ba-b-in-ay > bfin ‘female’ (rare; but cf. wa-bfin ‘wife’)

but, PCEMP *b > IRH b [ᵐb]

*baRəq > br-bar ‘swell’ (br- is reduplication)

*ba-b-in-ay > bfin ‘female’

Regarding the reflexes of *b and *p, it is important to point out that synchronically b is allomorphically conditioned by the subject agreement marker m- ‘second person subject’ on roots that begin with f, periodically where the first person pronoun am cliticizes to f-initial verb roots, and in some possessive forms, because, for example, ba ‘you go/walk’ has been erroneously cited as evidence that *b > b, but the root in the paradigm is fa ‘go’ (ja=fa ‘I walk’, ba ‘you walk’, i nfa ‘he walks’), which occasionally surfaces as fan in more formal, or grammatically complete, contexts. In other words, PCEMP *p > IRH f, not b, therefore PCEMP *panaw > fa(n), not ba(n) ‘you walk’; furthermore, final -n, even though it is optional, shows fa(n) originated from *panaw, not *ba.

4.1.2.1.2 ‘Unpacking’

An adequate description of ‘unpacking’ involves two subheadings, anticipatory voicing assimilation, and prenasalization of voiced stops.

4.1.2.1.2.1 Anticipatory voicing assimilation

A precursor for ‘unpacking’ is ‘anticipatory voicing assimilation’ (AVA), whereby voiceless velar stops became voiced word-initially preceding a vowel, and sporadically between vowels. This change is recurrent and supported, synchronically, by the relative scarcity of singleton k in monosyllables (cf. kkor ‘chicken’, kke ‘jaw’, and kka ‘bathe/swim’). AVA also interacted with nasal-related phenomena discussed in the next section.

PCEMP *k > IRH *g* [ʔg]

***kan** > *ga* [ʔga] ‘eat’

***kali** > *gar* [ʔgar] ‘dig’

***i-aku** > *-g* [ʔg] ‘first person inalienable possessive’

The only conditioned reflex of *t that has been proposed to date is *s*, but this putatively occurred before *i* in PSHWNG (Blust 1974). However, this correlation is not observed in any known Irarutu lexical items (but see §4.1.3 for a possible example of *t > *s*).

On the other hand, AVA is a general process that matches the voice value of a consonant to the following segment. Under the premise that syncope occurred historically prior to AVA (see §4.1.2.2.5.3), this process explains why *z lost voicing in the word for ‘rat’.

PCEMP *z > IRH *s* / __ C [-voice]

***kanzupay** (>...> *zfe*) > *sfe* ‘rat’

4.1.2.1.2.2 Prenasalization of voiced stops

Four interrelated phenomena, which will be hereafter called ‘unpacking’, ‘fusion’, nasal assimilation, and nasal substitution, are tied to the feature nasal. ‘Unpacking’ is used in this discussion to describe cases where word-initial nasals became phonetically prenasalized voiced oral stops (*N- > ^NC[+voiced][-continuant]), and resembles mistiming (e.g. insertion of [b] as a transition between [m] and [r]) in several regards. This use of ‘unpacking’ is somewhat incongruent with the normal use of the term, which is generally used to describe cases where nasalized vowels became vowel–nasal sequences, but the term has been appropriated here due to its association with nasality and the ability for nasal segments to develop into sequentially distinct articulatory gestures.

PCEMP *m > IRH *b* [ᵐb]

***malif** > *brif* ‘laugh’

but, PCEMP *m > IRH *m*

***ma-Ruqanay** > *mrان* ‘man/male’ (occasionally cited as *bran*)

PCEMP *ŋ > IRH *g* [ʔg]

***taliŋa** > *tgra* ‘ear’

***taŋis** > *-tag* ‘cry’

but, PCEMP *-ŋ > IRH *n*

***ma-dindiŋ** > *dridn* ‘cold’

Metathesis of the velar nasal and the liquid in the form *taliŋa ‘ear’ is widespread. It is observed in languages as distant as the Formosan branches of An and elsewhere (Blust 2009b). This sporadic change can therefore be attributed to drift or doubleting. The reader is reminded that synchronically there is no velar nasal phoneme ŋ in Irarutu. All reconstructed instances of *ŋ have become either *g*, *n*, or Ø.

No clear-cut examples in the 200-word list (or in additional data) of *n unpacking are attested. On the contrary, there is one example of a verb that reflects *d as *n*, rather than

expected *r*, and it is strongly supported by the proportion of recurrent sound correspondences in relation to number of segments.

PCEMP *deŋəR > IRH *f-nogr* ‘to hear’ (*f-* is inflectional)

This unusual reflex suggests that ‘nasal substitution’ took place in the history of Iraputu. Austronesian languages are thought to have utilized a verb-inflecting ‘active voice’ prefix PMP *meN- (N represents a ‘placeless’ nasal that receives its place specification from the following segment, usually a consonant) that was shortened to PCEMP *N- (Blust 2004, Ross 2002). Under nasal substitution, the stem-initial consonant essentially provides the place of articulation for the ‘active voice’ prefix, but is itself subsequently deleted. This can be illustrated as:

PCEMP *N-dəŋəR (> *n-dogr*) > IRH *nogr*.

In medial position, there is a pattern whereby homorganic nasal-voiced oral stop sequences fuse as (phonetically) prenasalized voiced stops. This accounts for the sporadic retention of *d > *d*. Furthermore, this sort of ‘fusion’ provides a phonological context in which to expect prenasalization, and reinforces its phonetic motivation. In contrast to ‘unpacking’, which lacks an alveolar correlate, there are no clear-cut examples for homorganic clusters of nasal-velar segments.

PCEMP *nd > IRH *d* [ⁿd]

*ma-dindɪŋ > *dridn* ‘cold’

and possibly, PCEMP *b/___ > IRH *b* [^mb]

*tumbuq > *sub* ‘grow’

The historic ‘active voice’ marker is connected to another phenomenon called ‘nasal assimilation’. Voicing of words with initial *k has already been attributed to AVA, but interaction of *N with velar stop-initial verbs is worth exploring. Furthermore, a form such as *N-kan ‘eat (active)’, or *N-gan if preceded by AVA, would have provided a context where *N- assimilated to the place of the velar, producing a hypothetical form such as *ŋ-gan, which substantiates the missing homorganic nasal-stop environment outlined for ‘fusion’.

PCEMP *N-kan ‘eat (active)’ > *ga* [^ŋga] ‘eat’

The historically context-dependent processes of ‘unpacking’ and ‘fusion’ account for the synchronic prenasalization of Iraputu voiced stops. The historical remnants of nasal assimilation and nasal substitution account for the exceptions in the data.

4.1.2.1.3 Dorsal stop deletion *q, *k > Ø

Two changes constitute ‘dorsal stop deletion’, *q > Ø and *k > Ø. The innovation *q > Ø is absolutely regular in inherited An forms. It occurred in all words in all positions where *q is reconstructed. Evidence from a range of Austronesian languages suggests that loss of *q began with erosion in the antepenult before the emergence of PEMP (Blust 2009a), followed by loss of medial and final instances of the segment in Iraputu.

PCEMP *q- > IRH Ø in the antepenult

*qabaRa ‘shoulder’ > *fra* ‘hand’

*qatəluR > *tru* ‘egg’

*qaləjaw > *re* ‘day, today’

PCEMP *q- > IRH Ø in the penult

*qinəp/qenəp > *-in* ‘sleep’

Possibly also

*qənay > *enye-fu* ‘sand’

PCEMP *-q- > IRH Ø

*waqay > *fa* ‘foot/leg’

*ma-Ruqanay > *mran* ‘male/man’

PCEMP *-q > IRH Ø

*bunuq > *-fun* ‘shoot (kill)’

*buaq > *fu* ‘fruit’

*tau/taumataq > *matü* ‘person’

*baRəq > *br-bar* ‘swell’

Loss of *k is the other part of ‘dorsal stop deletion’. The most common reflex of PCEMP *k is Irarutu Ø. This reflex is frequently observed in initial position and in all of the free pronouns; however, there are some contexts where *k is reflected as something other than Ø, such as g (see discussion of AVA). On a side note, PCEMP *-k was preserved in some dialects, such as Kuri *manik* ‘bird’. Anceaux (1961) also cites *manik*.

PCEMP *k- > IRH Ø

*kutu > *ut* ‘louse’

*kulit > *rit* ‘skin’

*kayu > *e* ‘wood’

*k-ita > *it* ‘1Pn’

*k-ami > *am* ‘1Px’

PCEMP *-k- > IRH Ø

*i-kau > *o* ‘2S’

*wakaR ‘root’ > *war* ‘root’

PCEMP *-k > IRH Ø

*manuk > *man* ‘bird’

but, PCEMP *k > IRH g

*kan > *ga* ‘eat (SAP subject)’

*kali > *gar* ‘dig’

*i-aku > *-g* ‘first person inalienable possessive’

Possibly also

*kaRat > *m-gor* ‘bite’

Blust (p.c. 2014) has suggested that the first person inalienable possessive suffix could have been conditioned by a preceding nasal, e.g. *-ni-ku > nku > ŋku > ŋgu, which is consistent with the treatment of nasal-stop sequences becoming Irarutu prenasalized voiced stops (see §4.1.2.1.2).

4.1.2.1.4 Rhoticization *l, *z, *d, *R > r

The liquids *l and *R, as well as *z and *d, seem have generally merged as Irarutu *r*. On the other hand, there is one instance of *z > s, *d following *n was retained as *d*, and *R was deleted in one lexical exception. There are no examples of PCEMP *r > *r* in the sample data set.

PCEMP *l > IRH *r*

*kulit > *rit* ‘skin’

*bulu > *fru* ‘feather’

*taliŋa (> *tanila*) > *tgra* ‘ear’

*zalan > *rarn* ‘road’ (kaimana dialect; alternates: *radni* and *rarum*)

*qaləjaw > *re* ‘day’

PCEMP *z > IRH *r*

*zuRi > *rur* ‘bone’

but, PCEMP *z > IRH *s*

*kanzupay > *sfe* ‘rat/mouse’

PCEMP *d > IRH *r*

*diRi > *mrir* ‘stand’

*dua > *ru* ‘two’

*todan > *m-tür* ‘sit’

*si-ida > *ir* ‘3P (they)’

*ma-dindiŋ > *dridn* ‘cold’

but, PCEMP *d > IRH *d* (context dependent)

*ma-dindiŋ > *dridn* ‘cold’

PCEMP *R > IRH *r*

*diRi > *m-rir* ‘stand’

*baRəq > *br-bar* ‘swell’ (prefixed reduplication)

*waiR > *wer* ‘water’ (guna assimilation)

*zuRi > *rur* ‘bone’ (apocope)

*qabaRa ‘shoulder’ > IRH *fra* ‘hand’

but, PCEMP *R > IRH Ø (exceptional item)

*maRi > *ma* ‘come’

4.1.2.1.5 Other deletions

Two other recurrent deletions are identified in TABLE 4.4, *n > Ø and *s > Ø. Like dorsal stop deletion, some instances of these changes are included under final consonant deletion. The remaining examples of *n > Ø and *s > Ø will be, hereafter, called collectively alveolar continuant deletion.

4.1.2.1.5.1 Final consonant deletion

The closest thing to a regular, not just recurrent, innovation in Irarutu is word-final consonant deletion (FCD). Incidentally, FCD overlaps with the context-free innovation ‘dorsal stop deletion’. FCD was a context-dependent change that is observed in the majority of demonstrable cognates. It is also part of both types of multisegmental linear deletions, final syllable rime deletion and final syllable deletion. Final consonants in the coda position were deleted when they occurred after vowels (final syllable rime deletion),

and final consonants in the onset position were also deleted in some instances provided the reconstructed final syllable nucleus was deleted (final syllable deletion), giving the impression that final consonant deletion was applied repeatedly. Consonants that followed a neutral vowel, i.e., a vowel not specified for palatal or labial color: *ə or *a, appear to have been particularly susceptible to this process.

PCEMP *-n > IRH Ø

*tudan/todan > -tür ‘sit’

*daun > ro ‘leaf’

*kan > -ga ‘eat’

but also, PCEMP *n > IRH Ø (final syllable deletion)

*panaw > fa ‘go, leave, walk away’

(compare PCEMP *ba > fa ‘go’)

PCEMP *-s > IRH Ø

*tañis > -tag ‘cry’

As a recurrent process, FCD also caused *-p, *-m, and *-R to merge with Ø.

PCEMP *-p > IRH Ø

*qinəp ‘lie down’ > -m ‘sleep/lie down’ (semantic shift)

but, PCEMP *-p > IRH f

*malip > mbrif ‘laugh’

PCEMP *-m > IRH Ø

*tanəmə > n-tan ‘plant’

*inum > -in ‘drink’

PCEMP *-R > IRH Ø

*qatəluR > tru ‘egg’

but, PCEMP *-R > IRH r ‘rhoticization’

*waiR > wer ‘water’

*dəŋəR > f-nogr ‘hear’

Notably, the voiced alveolar stop *t appears to be unaffected by FCD.

PCEMP P(CE)MP *-t > IRH t

*lañit > ragt ‘sky’

but, *-t > IRH Ø (but *a > o in the following item is also problematic)

*kaRat > m-gor ‘to bite’ (m- is lexicalized)

There is a trend for final consonants to be retained in words that show additional deletions which often, but not always, resulted in consonant clusters (wer ‘water’ has simple CVC shape).

4.1.2.1.5.2 Alveolar continuant deletion

There are two additional environments for *n > Ø and *s > Ø. One instance of context-dependent *-n- > Ø was discussed for ‘fusion’ (*ma-dindñ > dridn ‘cold’), but due to lack of interaction with ‘fusion’, this process is not applicable for

PCEMP *-n- > IRH Ø

*kanzupay > *sfe*,

which may have deleted due to its location in a prepenultimate syllable, following the deletion of the dorsal stop and the prepenultimate vowel (discussed below). This deletion is the only instance of consonant cluster reduction that does not involve a stop as the second consonant in the consonant cluster.

The two remaining instances of alveolar continuant deletion, i.e., *s > Ø, are both found in pronouns. But, as pointed out in §4.1.1.2, it is important to note that this segment may have been deleted prior to the formation of Irarutu, which is supported by the fact that many languages reflect the third person pronouns without the person marking morpheme that contains *s, but this deletion cannot be assumed outright. In fact, Blust 1995 proposes *si_ida for PMP, WMP, CMP, and SHWNG and *si_ia for PAN as well as WMP and CMP. It is therefore safest to take a neutral stance and explore both options.

It has already been shown that Irarutu is a syncretic language, having reduced the use of prior more morphologically complex resources, which would be further attested by

PMP *s (> PCEMP *s/Ø) > IRH Ø

PMP *si-ia > PCEMP *s-ia/ia > *i* ‘3S’

PMP *si ida > PCEMP *sida/ida > *ir* ‘3P’.

The change in the plural pronoun could be considered an instance where an Irarutu form shows a more feasible connection with PMP than PCEMP (see discussion of PMP *e retention, in §4.1.1.4, above). In both PMP pronouns, *si, which is reconstructed as a bound morpheme for the third person singular but a free – although adjacent – morpheme for the plural form, appears to have been dropped in the historical development of Irarutu. By contrast, relying on just the PCEMP forms would entail both a morphological reduction (the morpheme *s- > Ø in the singular), and a phonological deletion (initial *s > Ø in the plural). Alternatively, assuming *s > Ø in these two pronouns prior to the formation of Irarutu, the only consequence is losing these instances of *s > Ø as a recurrent correspondence, is a simplification, because a separate category of deletion, i.e., ‘alveolar continuant deletion’ need not be proposed.

4.1.2.1.6 Semi-recurrent innovation: Glide fortition *y > j [j]

For historical purposes, fortition is assumed to be a process of constricting the phonation type of a particular segment. Blust (2008b) identifies two types of glide fortition, light and heavy, in conjunction with high vowel centralization (HVC) as a type of complex sound change. For instance, light glide fortition (LGF), in combination with HVC reflects penultimate *iw and *uw as *əb*, and *iy and *uy as *əz*. The products of heavy glide fortition (HGF) for the same series are commonly *ək^w/əg^w* and *əj*, where *j* identifies a voiced alveolar affricate. A caveat is stress shift from the penultimate syllable to the ultima.

One piece of sporadic evidence suggests that Irarutu has a variety of LGF for the labial approximant, $*w > f$ ($*waqay > fa$ ‘leg’), the result of which overlaps with ‘fricativization’, a type of lenitive change, e.g., $*b > f$. On the other hand, there is recurrent evidence that HGF was triggered in Irarutu by the palatal glide $*y-$, which, due to the analysis of this segment in initial position as a consonant, is reflected as j . This is a fairly distinctive change, but one that resembles fortition of glides in the Austronesian languages discussed under complex sound change in Blust 2008b.

From a phonetic standpoint, the change of $*y > j$ can be seen as constricting the initial articulation by altering the onset to a voiced plosive, such as $^d y$ and thereby causing the segment to become complex. Transition from the plosive beginning to the approximant target surfaces phonetically as affrication [$^d y$]. The glide y may then either syllabify, resulting in $^d i$ ($a-^d ie$ [$a^d ie$] ‘my father’), or cause the plosive to become an affricate and then itself merge with zero, resulting in j ($a-je$ [$a-ɟe$] ‘my father’). To conform to phonological requirements, the affricate j was first phonetically, then phonemically, reinterpreted as a palatal stop [ɟ]. In a handful of lexical items in the Fruata dialect, and all words in some dialects (and for some speakers), y [j] and j [ɟ] alternate freely. The more frequent of the two realizations in the Fruata dialect is j ($ja \sim ya$ ‘1S’, and Indonesian *sembayan* > IRH *sbajan* ~ *sbayan* ‘prayer’). However, a further complication, discussed below, is y – which is also a synchronic allophone of i – never becomes j (*ssiet* ~ *ssyet*, $**ss^d iet/ssjet$). In other words, in present day Irarutu, the synchronic semi-vowel y [i] does not alternate with j [ɟ].

In relation to published data from the past century, glide fortition appears to be a change in progress. The attested patterns are: 1) the consonant y surfaces as [ɟ] with no further alternation, 2) y persists in flux with [ɟ] and [$^d i$], or 3) y disappeared from the set of possible surface forms and [$^d i$] alternates with [ɟ]. Furthermore, j can be prenasalized [$^n j$], an essential property of voiced stops. In the first person pronoun $*i-aku$, the $*i-a$ sequence is assumed to have become $*ya$, paralleling vowel coalescence in PMP $*si-ia > PCEMP *s-ia$, and PMP $*si ida > PCEMP *sida$, rather than $*i- > \emptyset$, which is a viable, but presently unprovable alternative.

PCEMP $*y > IRH j$

$*maya > mmaje$ ‘tongue’

$*i-aku > ja$ ‘I’ (ya rare)

and, in a borrowed form (possibly from PEMP or PSHWNG)

Wandamen $a-yai$ ‘my father’ > IRH $a-die \sim a-je \sim a-^n je$ ‘my father’

In contrast to glide fortition as part of complex sound change, which Blust (2008b) associated exclusively to conditioning with high vowel centralization in the languages of Sarawak and Tunjung (SE Borneo), the Irarutu forms do not interact with HVC; but, glide fortition is implicated in the treatment of $*y_{onset}$ as a consonant, and appears in **FIGURE 4.1** as the sporadic change $*i > j$.

4.1.2.2 Vowel innovations

Irarutu retained the three cardinal vowels *i, *a, *u that are reconstructed in both PMP and PCEMP, but the history of PMP *ə, which is reflected as PCEMP *e, *ə, and *o, is complicated because Irarutu lost PCEMP *ə, and reflexes of PCEMP *e and *o are erratic, sometimes agreeing with the PCEMP segment and other times conflicting with it. Due to the regularity of retained cardinal vowels (*a, i, u*) in stressed syllables, most innovations in the vowels must be seen as context-dependent; however, several PCEMP monophthongs even shifted, although sporadically, in stressed syllables.

	* 'σ ₁	* 'σ ₂	*' σ ₁	*' σ ₂	* 'σ	*'σ ₁	*'σ ₂	*'σ'
*ay>'e (3)	—	—	1	1				1
*ay>Ø (3)						3		
*a>Ø (21)	1	6			4	8	2	
*u>Ø (17)	1	1			5	10		
*i>Ø (12)		1			1	8	2	
*ə>Ø (5)					1	4		

Table 4.5 – Six recurrent vowel innovations. The first two changes concern diphthongs, the remaining four concern monophthongs.

Three generalized positions are pertinent for present purposes: final syllable, penultimate syllable, and ante- (as well as ante-ante-) penultimate syllable. Antepenults and ante-antepenults are grouped together as ‘prepenults’ (terminology courtesy of Blust 2009a). Innovations in the Irarutu vowel system show two (semi-) recurrent shifts, i.e., ‘guna assimilation’ and ‘labialized mid-vowels’, and several mergers, which fall into three categories: prepenultimate vowel deletion, apocope, and syncope. Sporadically, Irarutu also developed lax vowels *ɪ*, in a conditioned environment, and *ü*.

4.1.2.2.1 Recurrent vowel shifts

A change that is commonly called ‘guna assimilation’ is encountered in contexts where a vowel interacts with the semi-vowels *y and *w. From a historical perspective, there were arguably four falling diphthongs in final position in PMP: *-ay, *-aw, *-iw, *-uy. Only reflexes of *ay and *aw are observed in final position in Irarutu (see Table 4.1). For *ay, reflexes that show guna assimilation are recurrent, but the distribution of this segment needs elaboration. For *aw, these reflexes are merely semi-recurrent in the data set. From a phonetic point of view, these two complex segments exemplify falling diphthongs because, all else being equal, they have a particularly sonorous nucleus (*a) and a heavily colored, i.e., labialized (*w) or palatalized (*y) – hence low sonority – off-glide. As a phonological process, guna assimilation causes the nucleus of the diphthong to raise, in the direction of coloring from the off-glide, followed by deletion of the glide. The result is, the diphthongs that fall in sonority, *ay and *aw, become monophthongs, *e* and *o*, respectively.

PCEMP *-ay > *e*

*kanzupay > *sfe* ‘rat/mouse’

*qatay > *te* [tei~ti] ‘liver’ (supplementary item)

but, PCEMP *-ay > Ø (apocope)
 *matay > -mat ‘dead’
 *ma-Ruqanay > mran ‘male/man’

The factor that causes *ay > *e* guna assimilation to be considered recurrent arises where *a immediately precedes *y, or de-syllabified *i, functioning as the onset to an adjacent syllable. In this environment, and sometimes due to independent factors, a sequence of *a.i is thought to have become *a.y, which was in turn reanalyzed as a falling diphthong *ay, allowing guna assimilation to take place. On a side note, the word meaning ‘wood’, *e*, ultimately derives from a pre-PMP form reconstructed with *-iw: PAn *kaSiw > PMP kahiw > PMP *kayu. The reflex of PMP *wahiR (> PCEMP *waiR) shows that guna assimilation is not constrained to final position.

PCEMP *a / __i or __y > ay > IRH *e*
 *kayu (> kay > ay) > *e* ‘wood’
 *waiR (> wayR) > *wer* ‘water’

In spite of the fact that *aw > *o* is itself not a recurrent change, it falls under the canopy of guna assimilation as a recurrent process, and in strikingly similar contexts as the falling palatal diphthong.

PCEMP *a / __u > *o*
 *daun (> dawn) > *ro* ‘leaf’
 *i-kau (> i-kaw) > *o* ‘thou’
 but, PCEMP *-aw > Ø (apocope)
 *qaləjaw > *rre* ‘day’
 *panaw > *n-fa(n)* ‘go/walk’

4.1.2.2.2 Semi-recurrent and sporadic vowel shifts

There is one semi-recurrent and one sporadic vowel shift observed in **Table 4.1** that deserve comment because they occur in forms that are often cited. ‘High vowel laxing’ is recurrent, whereas ‘low vowel breaking’ is sporadic. The first shift occurred in vowels located in the penultimate syllable, which is the most difficult position to account for because of the general trend for stability of segments in this position due to co-incidence of stress; but, it must be addressed because of several instances of the recurrent merger called syncope, which also targeted a vowel in the penult.

The minimal pair *in* ‘drink’ : *m* ‘sleep’ is important because it shows a phonemic distinction between tense and lax high palatal vowels etymologically tied to PCEMP *i. The origin of *i* and *ɪ* is attributed to a phonetic split of *i in penults followed by a phonemic split that was conditioned by a tense/lax distinction in the final syllable. A lax vowel in the final syllable correlates with a lax vowel in the historical penult. The word *in* ‘drink’ is cognate with *inum ‘drink’, whereas *m* ‘sleep’ is cognate with *qinəp ‘lie down’, which was also subject to semantic shift. ‘Drink’ had a high tense vowel, *u in the final syllable, which – due to harmony – allowed the vowel in the penult to retain the phonetic feature [+tense] with no further changes. By contrast, ‘lie down’ had [-tense] *ə

in the ultima, causing the vowel *i/e in the penult (which is considered [+tense, -low]) to become [-tense]. Laxing of penultimate PMP *i, which was triggered by a lax vowel in the final syllable in Irarutu, therefore contrasts with optional lowering in PMP *qinep > PCEMP *qinəp/qenəp. After these phonological correlates were in place, the conditioning environments were subsequently lost through ‘word-final rime deletion’, i.e., final consonant deletion plus apocope. The phonetic split therefore resulted in a phonemic contrast that was context-dependent.

PCEMP *i_{penult} / CV_[-tense] > IRH *ɪ* (laxing)
 *qinəp > -*m* ‘sleep’

A parallel correlation is observed for context-dependent laxing of the labial vowel *u in instances where it precedes a syllable with a lax vowel, such as *a, but is complicated by labiopalatalization (*ü* represents [y], a lax labiopalatal high vowel), possibly conditioned by a coronal.

PCEMP *u > IRH *ü*
 *tudan > *mtür* ‘sit’
 but, PCEMP *u > IRH *u* [u ~ ʊ]
 *buaq > *fū* ‘fruit’

Although ‘laxing’ can generally be attributed to a specific context, i.e., a lax vowel in the final syllable, its occurrence is sporadic, shown by the word for ‘fruit’, above, which has a lax vowel in the final syllable, but lacks laxing in the penult.

In synchronic data from casual speech, the tense/lax distinction is frequently blurred.

PCEMP *i > IRH *i* [i ~ ɪ]
 *ini/ani > *adini* [adini ~ admi] ‘this’

The most noteworthy sporadic vowel shift is labeled ‘low vowel breaking’. It occurs in the oft-cited cognate *mata ‘eye’.

PCEMP *a_{final} > IRH *ie*
 *mata > *mtie* ‘eye’

This word has a high degree of variation, occurring in compounds as *mce*, cf. *mcerum* ‘face (lit. eye-house)’ (see Chapter 2 for more examples). Cowan (1953a), Anceaux (1961), and Smits and Voorhoeve (1992) report additional forms such as ‘*metia*’, ‘*matie*’, and ‘*mece*’. The compounded forms, and those cited by Cowan, Anceaux, and Smits and Voorhoeve suggest that the low vowel in the final syllable underwent phonetic palatalization due to perseverative pressure from the preceding [+coronal] *t. Corroborated with the fact that as a low vowel, which is highly sonorous, [a] has relatively great temporal quantity, the environment and the segment’s inherent qualities seem to have caused *a* to break into an on-glide *ja*. Subsequently, the nucleus raised in the direction of the color of the onset, producing *ie*, mirroring guna assimilation (but

there is an exceptional reflex of *mata in *wer+mta* ‘spring of water (lit. water eye)’, suggesting this compound could be a borrowing).

4.1.2.2.3 Semi-recurrent vowel mergers

Three recurrent vowel mergers are identified in **TABLE 4.5**: prepenultimate vowel deletion, syncope, and apocope.

4.1.2.2.3.1 Prepenultimate vowel deletion

The prepenult, being unstressed, is one position that was ubiquitously targeted for vowels to merge with Ø.

PCEMP *V_{prepen} > IRH Ø

*qabaRa > *fra-fu* ‘shoulder’

*taliŋa > *tgra* ‘ear’

*ma-Ruqanay > *mran* ‘male/man’

*taumataq > *matii* ‘person’

*ma-Ruqanay > *mran* ‘male/man’

There is a possibility that the sequence *au in prepenultimate syllables could have been affected by *guna* assimilation, but even if that were the case, the expected vowel *o* would also have been deleted.

4.1.2.2.3.2 Apocope

Apocope is observed in many coda-less PCEMP cognates that are identified in Irarutu. Furthermore, apocope deleted some final diphthongs (possibly after *guna* assimilation: *-ay > -e > Ø), i.e., *matay > *mat* ‘dead’, whereas *guna* assimilation, not apocope, is reflected in others, *ka(n)zupay > *sfe* ‘rat’. The unpredictable nature of this change again suggests lexical strata in the sense of Blust 1991.

PCEMP *-i > IRH Ø

*zuRi > *rur* ‘bone’

*diRi > *m-rir* ‘stand’

PCEMP *-u > IRH Ø

*susu > *sus* [sus ~ sus^u ~ sus^a] ‘breast’

*kutu > *ut* ‘louse’

PCEMP *-a > Ø

*sida > *ir* ‘they’

*dua > *ru* ‘two’ (alternate: *rifo*)

However, a number of forms that underwent final consonant deletion (FCD) also show apocope, and other forms that underwent apocope also show FCD, suggesting that FCD occurred simultaneously with apocope, otherwise it would be necessary to propose that FCD applied recursively, once prior to apocope for some words, and then again after apocope for other words.

Together, FCD and apocope resulted in the multisegmental change referred to as ‘linear’ deletion in **FIGURE 4.2**, of which there are two varieties, one deleted the word-final rime (*CV₁.CV₂(C) > CV₁C) and occurred in an overwhelming number of cognates (34/61), the other deleted the entire word-final syllable (*CV₁.CV₂ > CV₁), which usually consisted of only an open syllable, *-CV₂. There are only four examples of ‘final syllable deletion’.

PCEMP *-VC > IRH Ø (final rime deletion)

- *qinəp (> *qinə) > -*m* ‘sleep’
- *bunuq (> bunu) > *n-fun* ‘kill v.’
- *taŋis (> *taŋi) > *tag* ‘cry’
- *inum (> *inu) > *in* ‘drink’

PCEMP *-CV > IRH Ø (final syllable deletion)

- *buŋa > *fu* ‘flower’
- *panaw > *fa* ‘walk, go’
- *maRi > *ma* ‘come’
- *i-aku > *ja* ‘I’

4.1.2.2.3.2.3 Syncope *V_{penult} > Ø

Similar to the process of degemination, discussed in §4.1.2.2.4, below, syncope resulted in monosyllables from historical disyllables. Syncope has been noted for many Austronesian languages of Eastern Indonesia (Donohue and Grimes 2008, Blust 2009a). However, the special status of syncope, occurring distinctively for SHWNG languages after a reconstructed nasal (N_CVC), was noted as a general trait of the SHWNG language group as early as a century ago (Adriani and Kruijt 1914). Assuming that PCEMP had penultimate stress, stress shift necessarily preceded syncope, although the conditions for both are unclear. Blust (1978) proposed that syncope occurred between a nasal and a stop (frequently a voiceless stop), noting PEMP *mata > SHWNG *mta* ‘eye’ (IRH *mtie*), and *natu > SHWNG *ntu* ‘child/offspring’ (IRH *ntü*), and *banua > *bnu* ‘village’ (but, IRH *bin* ‘land’). However, syncope is more complex than this in Irarutu. More cognates reflect forms that violate Blust’s generalized syncope rule than support it.

PCEMP *a_{penult} > IRH Ø

- *malif > *brif* ‘laugh’
- *mata > *mtie* ‘eye’
- but, PCEMP *a_{penult} > IRH *a*
- *matay > *n-mat* ‘dead’
- *mataq > *m-mat* ‘green/unripe’
- *taumataq > *matü* ‘person’

Furthermore, syncope occurs in a much wider range of contexts than *N__C[-voice], such as such as adjacent to a rhotic, or after a fricative.

PCEMP *i_{penult} > IRH Ø

- *taliŋa > *tgra* ‘ear’ (cited as *təgəra* in the literature)

PCEMP *u_{penult} > IRH Ø
 *kulit > *rit* ‘skin’
 *kanzupay > *sfe* ‘rat/mouse’

Therefore, syncope in Irarutu cannot be predicted based on any type of environment, but, overall, syncope appears to be a byproduct of stress shift, that is, syncope could not delete a stressed vowel, only unstressed vowels. It caused some penults to delete or reduce to phonetic schwa (*fərafu~frafu* ‘shoulder’), although this may be a synchronic property of Irarutu, but other items are not effected at all (*buni > *funi* ‘hide’). The resulting stress alignment can be captured through generalization.

Syncope

*CV.σ > CV.'σ (stress shift) > C^V.σ (vowel reduction) > IRH C^[ə].σ ~ C.'σ

The interconnectedness between syncope and stress shift is seen in the reallocation of stress to vowels in previously unstressed final syllables, followed by relative stability of segments in historically secondary stressed syllables. This suggests that syncope preceded the majority of segmental shifts and mergers, because they are not observed in syncopated forms. The synchronic data supports this generalization.

One clue to the conditioning factors that spurred syncope is seen in a pair of potentially homophonous words, ‘three’ and ‘egg’. It appears to be the case that stress shift and syncope were used to differentiate these items; however it seems arbitrary for syncope to have occurred in ‘egg’ and not in ‘three’, which shows apocope, and not vice versa.

PCEMP *V_{penult} > IRH Ø
 *qatəluR > *tru* ‘egg’
 but, PCEMP *V_{penult} > IRH *o*
 *təlu > *tur* ~ *tor* ‘three’

When considered in relation to apocope, the data strongly suggests three layers of Austronesian lexical items: a syncope stratum, an apocopated stratum, and a stratum with no changes to word shape. Without this observation, the occurrence of syncope and apocope must be seen as arbitrary.

4.1.2.2.4 Sub-recurrent and sporadic vowel mergers

Three vowel mergers are semi-recurrent: degemination, *a > *u*, and targeted loss of the vowel from final syllables. There is also sporadic merger of non-high vowels with *o*. In two proto-forms that, due to consonant mergers, ended up with a sequence of identical vowels, this type of sequence reduced to a single vowel. This is not true degemination, since neither PCEMP nor Irarutu are proposed to have phonemic vowel length, but it is a useful label to refer to the general phenomenon. The process of degemination caused disyllabic proto-forms to become monosyllabic in Irarutu.

PCEMP $V^{\alpha}.V^{\alpha} > \text{IRH } V^{\alpha}$

*daqan (> **daan**) > *ran* ‘branch’

*wakar (> **waar**) > *war* ‘root’

‘Degemination’ is observed as a type of coalescence in between PMP and PCEMP, before the formation of Irarutu.

PMP $*V^1-V^1 > \text{PCEMP } *V^1$

*si ida (> si-ida) > *sida

The next recurrent change shows that *a in the final syllable merged with *u*, or *ü* after a coronal stop, which appears to have conditioned palatalization. The Irarutu word for ‘person’, *matü* (Kuri *tmatu*), could be explained by sporadic long-range metathesis, but this is obviously speculative. Of interest is the fact that one of Irarutu’s neighbors, Kamrau, is reported to have *eso* ‘one’, similar to Kuri (Peckham 1981a).

PCEMP $*a_{\text{final}} > \text{IRH } \ddot{u}/u$

*taumataq > *matü* ‘person’

*əsa > *esu* ‘one’

but, PCEMP $*a_{\text{final}} > \text{IRH } \emptyset$

*mataq > *m-mat* ‘green/unripe’

In two other items, the vowel in an unstressed final syllable was targeted by deletion, similar to apocope, but retained a coda in reconstructed and synchronic forms.

PCEMP $V_{\text{final}} > \text{IRH } \emptyset$

*laŋit > *ragt* ‘sky’

*zalan > *rarn* ‘path/road’

Two vowels sporadically merged with *o* in the penult, *ə, and *e. The changes of *ə > *o* and *e > *o* were discussed as instances of ‘labialized mid vowels’, above. There are no examples that verify *o > *o*. A notable property of these words is the presence of a rhotic in coda position.

PCEMP $*ə_{\text{penult}} > \text{IRH } o$

*təlu > *tor* ‘three’

PCEMP $*e_{\text{penult}} > \text{IRH } o$

*deŋəR > *fnogr* ‘to hear’

To conclude the discussion of ‘reliable’ cognates, it is notable that every generalization about reflexes of individual segments has one or more exceptions. There are no lexical items that can be predicted with any degree of certainty from a novel, i.e., selected at random, PCEMP form. Best guesses can be made, but these always necessitate some degree of reservation. In this regard, it could be said that there are no entirely regular reflexes of PCEMP forms in Irarutu. The only possible exception is **ini* ‘this’, but even

that term is unusual because it violates the regular process of apocope, and synchronically, the penultimate vowel tends to become lax.

4.1.3 Pseudo-cognates and faux cognates

In addition to the cognates identified in TABLE 4.1, there are twenty-one items that possess some degree of similarity with reconstructed PCEMP lexical items. Twelve of the words are labeled ‘pseudo-cognates’ (including secondary cognates) due to a fairly strong amount of correlation with reconstructed forms. Nine are designated as ‘faux cognates’ because there is some resemblance between these items and reconstructed forms, and existing resources identify them as possible cognates, but the similarity is equally due to either chance or borrowing.

4.1.3.1 Pseudo-cognates

The twelve Irarutu words identified in TABLE 4.6 appear to be good candidates as cognates but closer evaluation reveals the correlation is spurious due to the fact that they reflect a higher proportion of sporadic changes than is tolerable for inclusion in the preceding discussion. For this reason they are labeled ‘pseudo-cognates’. In other words, these forms each reflect a significant proportion of legitimate sound correspondences listed in TABLE 4.2, and ensuing discussion, but simultaneously reflect one or two qualitatively aberrant segmental correlations in regard to their PCEMP cognate that disqualify them as such.

No.	PMP	PCEMP	IRH	Gloss
<u>Pseudo-Cognates</u>				
65.	waqay	waqay	<u>f</u> a	‘leg, foot’
66.	kaRat	kaRat	m- <u>g</u> or	‘bite’ (m- is lexicalized)
67.	tu(m)buq	tumbuq	<u>s</u> ub	‘grow’
68.	ŋajan	ŋajan	<u>s</u> nan	‘name’ (IRH <i>nü</i> ‘name’; <i>snan</i> > ‘word/name’)
69.	talih	tali/waRəj	war <u>a</u>	‘rope’
70.	aliR/aluR/ saliR	aliR/saliR	sr- <u>s</u> wir	‘to flow’
71.		taje ni laŋit	taje	‘cloud’
72.	quzan	quzan	sy <u>e</u> m	‘rain’
73.	ma-nipis	ma-nipis	bnif <u>n</u>	‘thin (material)’ (*m/ __N > b is problematic)
74.	ma-labeR	labəR	f- <u>r</u> ifr	‘wide’
<u>Secondary Cognates</u>				
75.	bilin/puter	putər	putr	‘turn’
76.	taqun	taqun	taun	‘year’

TABLE 4.6 – Pseudo-cognates and secondary cognates. Twelve words from Appendix 3a are good candidates for comparison with PCEMP but contain one or two segments that deviate in unpredicted ways (suspect segments are underlined), or can be identified as a loanword on semantic grounds.

In general, pseudo-cognates possess systematic correlations that appear to represent borrowings. These words are likely to be true cognates in other languages, such as nearby SHWNG languages, e.g. Wandamen, but were borrowed into Irarutu. This is provable for recent loanwords from Indonesian that represent non-native concepts, e.g., IND *tahun* ‘year’: IRH *taun* ‘year’, and IND *putar* ‘turn’: IRH *putr* ‘turn’ (also expect PCEMP *p- > IRH f).

Consonants

- *w > f conflicts with retention of *w
- *-t > Ø “ retention of *-t
- *t- > s “ retention of *t; could be conditioned / __ V_[+high]
- *-q- > j “ ‘dorsal stop deletion’
- *-n > -m “ retention of *n and *m “
- *-s > -n is unmotivated, and conflicts with known reflexes of *s

Vowels

- *a > o conflicts with retention of *a
 - *a > i “ all other attested and putative reflexes of *a
 - *ə / __ j > a “ ‘labialized mid vowels’, and other reflexes of *ə
 - *a / __ l > w “ ‘rhoticization’ and its sub-recurrent exceptions
 - *a > e conflicts with apocope, retention of *a, and *-a > u
-

FIGURE 4.3 – Aberrant sound correspondences in ‘pseudo-cognates’, listed in order of occurrence in **Table 4.6**. All of these correspondences are sporadic. The term sub-recurrent denotes sporadic and semi-recurrent innovations.

Nearly all of the pseudo-cognates, and the divergent sound correspondences they reflect, could be included in the set of changes enumerated in **TABLE 4.2**, in particular, *snan* ‘say/word’ and *wara* ‘rope’ are important in this regard, but if they were included in the prior list, the recurrent changes would be over-ridden by exceptions, see **FIGURE 4.3**. For example, although *mmaje* ‘tongue’ is included as a true cognate, due to similarity with ‘breaking’ of *a_{final} σ in *mata > *mtie*, if *frifr* ‘wide’ is also accepted, the reflexes of *a become extremely unstable, being able to range over nearly every possible vowel quality without rhyme or reason. Identification of the aberrant sound correspondences in pseudo-cognates is useful for completeness’ sake, but they will not be discussed in detail.

4.1.3.2 Faux cognates

The remaining nine items to be presented, see **TABLE 4.7**, are labeled ‘faux cognates’ because the number of attested sound correspondences occur in a markedly low proportion of segments in relation to word length and each item unanimously requires proposing two – or more – creative innovations that conflict with the evaluative innovations discussed in §4.1.1–2, above. These items are identified here because available resources, such as the Austronesian Basic Vocabulary Database (Greenhill et al 2008), suggest they may be candidates, but such proposals must be eliminated. Suspect correspondences are identified below in **boldface**. Notes on nonconforming phenomena are added parenthetically.

No.	PMP	PCEMP	IRH	POc	Gloss
77.	qabu	qabu	<u>bob</u>	qapu	‘ash’ (IRH <i>tugwan</i> ; <i>bob</i> > ‘dust, smoke’)
78.	demdem	dəmdəm	<u>mr̥n</u>	ronrom	‘think’ (IRH <i>derbun</i> ; <i>mr̥n</i> > ‘know’)
79.	t-ina	t-ina	- <u>den</u>		‘mother’
80.	Rabun	Rabun	<u>mud</u>	Rapun	‘cloud’
81.	cemeD (?)	ma-qetəm/ ma-qitəm	<u>grmutn</u>	ma-qitom/ ma-qetom	‘dirty’
82.	ma-qitem	ma-qetəm	<u>grmutn</u>	ma-qetom	‘black’
83.	pa-nahik/ sakay	panaik/ sakay	<u>iet</u>	panek/ sake	‘climb’
84.	dalem	dələm	<u>garn</u>	ralom	‘in, inside’
85.	qazi/diaq	diaq	<u>ti</u>	(?)	‘no, not’

TABLE 4.7 – ‘Faux cognates’, nine Irarutu words that have two or more segments or changes morphological changes in relation to the proposed cognates.

- *qapuk ‘dust’ and *qabu ‘ash’ > **bob** ‘dust/smoke’ (semantic shift, lexical merger, reduplication, *a>o ; violates: fricativization)
- *dəmdəm ‘think’ > *mr̥n* ‘know’ (lexical split, *d>Ø, *ə>ʊ, *-m>-n),
- *dəmdəm ‘think’ > *derbun* ‘think’ (lexical split, *d>d, metathesis: *md>dm, *ə>u, *-m>-n)
- *t-ina > *-den* ‘mother’ (*t>d, *i>e)
- *Rabun > *mud* ‘cloud’ (*R>Ø, *b>m, *n>d)
- *ma-qetəm/ma-qitəm ‘black’ and *ma-qetəm ‘dirty’ > **grmutn** ‘black/dirty’ (innovative prefix *gr-*, *a>u, *-m>-n)
- *panaik > *iet* ‘climb/ascend’ (prepenult deletion, *n>Ø, breaking, *i>Ø, *k>t)
- *daləm > *garn* ‘in, inside’ (*d>g, *-m>-n)

4.1.3.3 Other cognates

There are three on-line resources that have made Irarutu lexical data from Grace’s 1955–6 fieldnotes, supplied courtesy of Dr. Robert A. Blust, widely available: Blust’s ongoing ‘Austronesian Comparative Dictionary (ACD)’ (www.trussel2.com/ACD/), which was initiated in 1995; Greenhill, Blust, and Gray’s 2008 ‘Austronesian Basic Vocabulary Database (ABVD)’ (www.language.psy.auckland.ac.nz/austronesian/); and the TransNewGuinea.org webpage (www.TransNewGuinea.org/language/irarutu). The ACD, which is the most careful about discerning cognates, only contains eight Irarutu words: 1) PMP *baRa>*fra* ‘hand’; 2) PMP *ida>*ire* ‘3P’ (NB: synchronic evidence shows -e represents non-phonemic [ə] in the Fruata dialect); 3) PEMP *wakaR-i>*kwakare* ‘root’, 4) PAN *ma-tunuh>*ma-tun* ‘to fry’, 5) PAN *diRi>*m-rir* ‘to stand’, 6) PMP *i-kahu>*o* ‘2S’, 7) PCEMP *kazupay>*səfe* ‘rat’ (ə is non-phonemic), and 8) PMP *tebuh>*tof* ‘sugarcane’. Lack of primary data introduced orthographic irregularities which are uncritically retained in the ABVD and the TransNewGuinea websites. Nevertheless, two notable example cognates are *-matotə* ‘to cut, hack’ < PCEMP *tətək < PMP *tektek* ‘to cut, hack’ (not confirmed in the research for this description), and *-mətu-*

tə ‘to pound, beat’ < PCEMP **tutuk*. The form for ‘cut, hack’ could bring the number of reflexes of *ə > IRH *o* up to two instances, which might therefore be considered a semi-recurrent shift. In the absence of identifying justifiable sound correspondences, other forms that are suggested to reflect direct inheritance must be analyzed as words with unknown provenance.

Nine more Irarutu words that reflect Austronesian etyma according to the criteria established in §4.1 occur outside this section but support the claims made here: PMP **nusa* > *nu* ‘island’, **baRaq* > *bar* ‘lung’, **taqi* > *ta ~ te* ‘feces’, **puki* > *fi* ‘vagina’, **tuRun* ‘descend, go down’ > *tur* (as in *wer ntur* ‘low tide’), **taRaq* > *tar* ‘cut, hack’, **punti* > *fud* ‘banana’, and possibly also **puket* > *fukr* ‘dragnet, net’, and **pusej* > *fwir* ‘center, navel’.

4.1.4 Other historical innovations

Voorhoeve (1989) describes several other historical innovations that he observed in Irarutu phonology and morphology: an increase from (what he thought was) the PAN five-vowel system to the contemporary Irarutu seven-vowel system, which is noteworthy because other SHWNG languages employ a five-vowel system; and, morphological simplification: i) loss of historically marked grammatical number in verbs and nouns, ii) loss of what he refers to as PSHWNG plural marking on nouns that were [+human], iii) loss of the inclusive/exclusive distinction, iv) fossilization of the PSHWNG possessive ligature **ne*, which he hypothesized takes affixes, and v) loss of the edible-inedible contrast on nouns (cf. Buli and Oc where this contrast is observed). The preceding description is neutral on points ii)–v), which are presented here to acknowledge Voorhoeve’s work, but are considered artifacts of his analysis.

However, there are several non-phonological innovations in the history of Irarutu that stand out in the discussion above, and are implied in the description in Chapter 3. A frequent historical change that is tied to the identification of cognates is relexification/semantic shift, such as PCEMP **qinəp/qenəp* ‘lie down’ > IRH *m* ‘sleep’. This point played a prominent role in the discussion of both reliable and faux cognates.

A more difficult factor has to do with morphology. So-called frozen morphology overlaps in several instances with productive synchronic morphology, reflecting a grammatical retention. For example, **ma-putiq* ‘stative-white’ > Irarutu *bfut* ‘white’ reflects prepenultimate vowel deletion, word-final rime deletion, fricativization, and unpacking, which strongly suggests that the stative marker was lexicalized in this word and is no longer treated as a predicate that can take stative morphology ***m-bfut* ‘to be white’. However, **ma-* ‘stative’ is reflected as Irarutu *m-*, and is productive in some words in the present-day language: *ja m-tim frrufn* ‘I closed the door (i.e., the door is closed)’ versus *frrufn timtim* ‘the door is closed’, but nonproductive in others *mrir* ‘to stand’ (*ja=m-rir* ‘I stand’, *o=m-rir* ‘you stand’, *i n-m-rir* ‘she stands’, but a form such as *rir-rir* does not exist). The non-productive counterpart occurs in a number of intransitive verbs having to do with physical states (*m-tür* ‘sit’, *m-rur* ‘sit cross legged’, *m-tugr* ‘sit with one leg out’, *m-sutr* ‘sit holding knees up’), adjectival predicates (*m-mat* ‘green/unripe’ and *m-genir* ‘salt(y)’), as well as transitive verbs such as *mgür* ‘buy’ (*ja*

mgür ‘I buy’, *o mgür* ‘you buy’, *i n-mgür* ‘she buys’). In these instances, initial *m-* is part of the root and makes no contribution to meaning (see also Chapter 3).

Two grammatical subsystems can be historically tied to Irarutu personal pronouns: possessive morphology (alienable prefixes and inalienable circumfixes) and agreement markers (for more details, see Chapter 3).

	PCEMP	Pronominal	Alienable	Inalienable	Agreement
1S	*i-aku	<i>ja</i>	<i>a-</i>	<i>-g</i>	<i>a-</i>
2S	*i-kau	<i>o</i>	<i>o-</i>	<i>-m</i>	<i>m-</i>
3S	*s-ia	<i>i</i>	<i>i-</i>	— (<i>-an</i>)	<i>n-</i>
1Px	*k-ami	<i>am</i>	<i>a-</i>	<i>-g</i>	<i>a-</i>
1Pn	*k-ita	<i>it</i>	<i>a-</i>	<i>-g</i>	<i>a-</i>
2P	*kamiu	<i>e</i>	<i>o-</i>	<i>-m</i>	<i>m-</i>
3P	*sida	<i>ir</i>	<i>i-</i>	—	<i>n-</i>

Table 4.8 – Paradigms related to reconstructed personal pronouns: free pronouns, possessives, and agreement markers.

On the topic of possession, several Papuan consultants reported a tendency in Papuan Malay to prefer a head-modifier construction such as *buku saya* ‘my book’ over the ‘standard’ Indonesian pattern of suffixal possession, e.g., *buku-ku* ‘my book’; however, the most natural construction is periphrastic: *saya punya buku* ‘I possess book’ (contracted to: *sa’ pu’ buku*). This trait can be seen as an areal feature due to diffusion (see also Klamer 2002). The Irarutu construction *ja ri buku* ‘I have book’, when used to mean ‘my book’ is structurally identical to periphrastic possession in Papuan Malay.

Another area which possibly shows that Irarutu is a syncretic language is the all-but-forgotten use of inflectional morphology, such as the Philippine-type voice system. There appear to be remnants of, for example, <in> ‘patient focus/nominalizer’ as <en> in a derived form of the word *fan* ‘go’, which is part of the proper family name *nafur-b<en>an*. This phenomenon needs more research and can only be mentioned in passing. Speakers are not able to identify the meaning of this morpheme, or use it in novel forms. Irarutu *fi* fulfills a similar double function – nominalizer/verbalizer – in contemporary Irarutu (see Chapter 3).

4.1.5 Evidence for lexical strata

The presentation of forms in **Table 4.1**, and the discussion above based on those forms, implicitly assumed that there are two, or more, strata of lexical items in Irarutu. One pair of words in particular, *tor* ‘three’ and *tru* ‘egg’, makes that assumption crucial for an understanding of Irarutu historical phonology. Due to the fact that segmental changes are attested across the categories that were identified in **Table 4.1**, it is especially challenging to verify the existence of the strata; however, four semi-independent phenomena allow them to be coaxed apart: stress shift, penultimate vowel deletion (i.e., syncope), final syllable vowel deletion (i.e., apocope), and reflexes of PMP schwa (*o* or *Ø*).

The necessity of considering the possibility that there are two lexical strata in Irarutu is imposed by the fact that the sequence *təlu, which occurs in unrelated words, has noticeably different reflexes in their modern-day counterparts, with no other reasonable explanation:

PMP *e > PCEMP *ə > IRH o
 *təlu > **təlu** > *tor* ‘three’
 PMP *e > PCEMP *ə > IRH e
 *qateluR/qiteluR > **qatəluR** > *tru* > ‘egg’

Apart from identical treatment of *t > t and *l > r in both words, and differing lengths (plus a coda in the word for ‘egg’), two separate patterns are apparent, each with a pair of concomitant phenomena. One pattern, attested in the word for ‘three’, shows apocope/final-syllable vowel deletion (*u > Ø), and a labialized mid vowel (*ə > o). The other pattern, attested in the word for ‘egg’, shows stress shift to the final syllable (*'σσ > σ'σ), and syncope (*ə_{penult} > Ø). Two more words that follow the ‘three’-type pattern are:

PMP *deŋer > *nogr* ‘hear’
 PMP *tebu > *tof* ‘sugar’

Not all words that can be assigned to one strata or the other contain schwa in the penult; and not all Austronesian etyma in Irarutu can be assigned to one of these strata, such as *qatep > *atif* ‘thatch’. Nevertheless, the patterns sketched for the ‘three’-type stratum, in which the segments in the penult are resistant to deletions but segments in the final syllable are particularly susceptible to deletions, and the ‘egg’-type stratum, in which segments in the penult are susceptible to deletions but segments in the final syllable are resistant to deletions, are reinforced by several additional etyma, some of which are not contained in the sample set from Appendix 3a. More ‘three’-type stratum items are:

PMP *kutu > *ut* ‘louse’
 PMP *salaq > *sar* ‘false’
 PMP *alaq > *-ar* ‘take’
 PMP *kita > *it* ‘1Pn’
 PMP *kami > *am* ‘1Px’
 and perhaps even
 PMP *ia > *i* ‘3S’
 PMP *aku > *a-* ‘1S’ (marker in RCs).

These words contrast with ‘egg’-type stratum items such as:

PMP *beli > *fri* ‘price’
 PMP *lipen > *rfo* ‘tooth’
 PMP *bulu > *fru* ‘feather’
 PMP *taliŋa > *tgra* ‘ear’
 PMP *kanzupay > *sfe* ‘rat’
 PMP *kulit > *rit* ‘skin’

PMP *qulu > *rü* ‘head’
PMP *qutin > *ti* ‘penis’
PEMP *natu > *ntü* ‘offspring’.

Granted, the evidence for lexical strata in Irarutu is less robust than for say, Rotuman (cf. Biggs 1965) or Tiruray (Blust 1991), partly due to the presence of similar segmental changes in both strata (such as *b > f), there are repeated instances where identical – or functionally identical – strings of segments in groups of semantically unrelated reconstructed forms have predictable reflexes in their modern-day Irarutu counterparts based on one or the other of the two lexical strata patterns. Two additional facts, the predominance of final syllable deletions in Irarutu (shown in **Table 4.1**) and the nature of the semantic domains in which each strata is found (numerals + pronouns + some nouns + one verb versus various nouns), suggest that the ‘three’-type stratum is probably older than the ‘egg’-type stratum. No further conclusions can be drawn at this time.

4.2 Genetic affiliation

This section investigates the relationship between Irarutu and historically prior Austronesian languages by comparing the historical phonology description above, and other innovations, with proposals from the literature. The discussion below identifies several forms of evidence that situate Irarutu within the Austronesian language family, but it also raises questions about the usual view that Irarutu belongs within the SHWNG subgroup. The role of non-Austronesian elements is briefly considered after the Austronesian hypothesis is presented.

4.2.1 The Austronesian component

Since Cowan (1953a), who used impressionistic evidence from pronouns, numerals, and some basic vocabulary, the general perception among linguists is that Irarutu belongs to the Austronesian language family (Anceaux 1961; Blust 1974, 1978; Matsumura and Matsumura 1991; Ross 1995; Tryon 1995; van den Berg and Matsumura 2008; Voorhoeve 1989, 1995). Grace (1955–6) designated Irarutu as an isolate within the An language family, but Anceaux (1961) classified it as an isolate within what has come to be known as South Halmahera – West New Guinea. As pointed out at the beginning of this chapter, lack of information has prevented a better analysis.

The grounds that have been used to establish higher level Austronesian subgroups such as Proto-Malayo-Polynesian (PMP), Proto-Central Eastern Malayo-Polynesian (PCEMP), and Proto-Eastern Malayo-Polynesian (PEMP) include phonological, morphosyntactic, lexical, and semantic evidence. Anceaux 1961 cites considerably more data than Cowan and classifies Irarutu as an isolated member of his ‘Cenderawasih Bay’ group but without concrete supporting evidence. Blust (1974, 1978) proposed that the Austronesian languages of Cenderawasih bay, including Irarutu, are most closely related to those of South Halmahera. He labelled the group South Halmahera – West New Guinea (SHWNG). Voorhoeve (1989) was the first to suggest a specific position for Irarutu within SHWNG based on sound correspondences. He proposed that Irarutu constitutes a primary branch of the SHWNG subgroup due to the failure of *t > s before *i*. On the

other hand, Ross (1995) postulates that Irarutu is affiliated with the South Halmahera branch. The present understanding is fragmentary at best.

A better understanding of various Irarutu dialects/closely related languages may further improve the classification. This is a crucial task for future research. As mentioned in Chapter 1, a significant problem concerns the classification of Nabi (a.k.a. Kuri), especially because community members by-and-large consider Kuri to be a dialect of Irarutu. On the other hand, according to the *Ethnologue*, Nabi is an unclassified CEMP language (Lewis et al 2013). However, that proposal is problematic due to linguistic data, folk belief, and the geographic location of Nabi in relation to other CEMP languages such as Koiwai, Sekar, Onin, and Uruangnirin, which are located at the opposite end of the Bomberai peninsula (see also Blust 2008a). It is helpful to briefly explain the historic context for classifying Austronesian languages in Western Melanesia/Eastern Indonesia before looking specifically at Irarutu.

Recognition of an Oceanic subgroup dates back to the 1930s (Dempwolff 1934–38). Existence of the Oceanic group raised two questions about the structure of the Austronesian language family: 1) where is the Western Boundary of the Oceanic languages? 2) what is the next higher subgroup that includes the Oceanic languages (i.e., how are the Oceanic languages related to other An languages?) Grace (1971) provides a widely accepted answer to the first question: Oceanic languages are found East of Cenderawasih Bay (see also Ross 1988). Blust responded to the second question and supported his position in several publications (1974, 1978, 1982, 1983/84, 2009a). Subsequently, the discussion of Irarutu’s genetic affiliation has been situated in the context of the commonly accepted “standard” Austronesian family tree model (Blust 2009a, also cited in Donohue and Grimes 2008), see **FIGURE 4.4**.

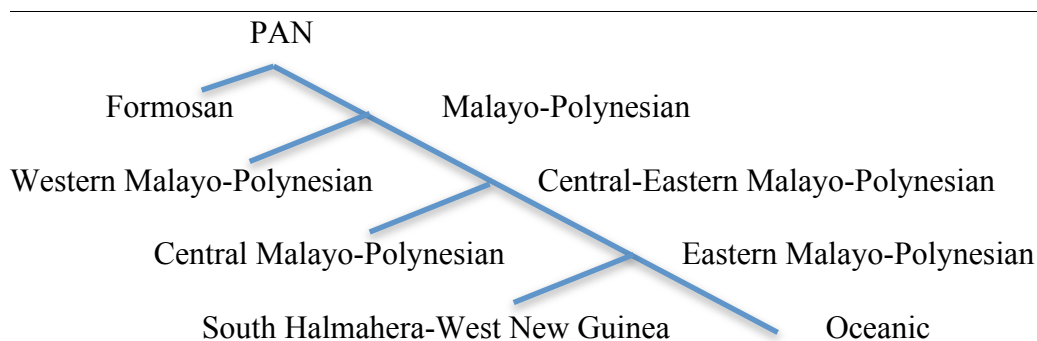


FIGURE 4.4 – “Standard” tree model of the Austronesian language family (Blust 2009a).

SHWNG languages are located below three major nodes: Malayo-Polynesian, Central Eastern Malayo-Polynesian, and Eastern Malayo-Polynesian; and, it is sister to the massive Oceanic subgroup (comprised of more than 460 languages). In **TABLES 4.5–8**, Irarutu data presented in parentheses cannot be verified as originating from Austronesian etyma.

4.2.1.1 Central-Eastern Malayo-Polynesian (CEMP)

The status of Central-Eastern Malayo-Polynesian as a valid subgroup has been questioned by Nothofer (1992), and by Donohue and Grimes (2008). At least one of its subcomponents, CMP is thought to have been a dialect chain (Blust 2009a), for which any given feature may belong to part of the chain but not necessarily the whole group. Due to its location in Eastern Indonesia, there is a good probability that Irarutu belongs under the CEMP node. Comparative evidence can verify the claim. In a recent publication on PCEMP (Blust 2009a), twenty-one innovations ranging over lexical, phonological, morphosyntactic, and semantic fields are evaluated for utility in subgrouping (TABLE 4.7). Since Donohue and Grimes (2008) give different weight to these innovations than Blust (2009a) does, the full set of twenty-one innovations is included, along with Irarutu data.

Title of innovation:	Irarutu data:
Lexical innovations	
1. Marsupial terms	(<i>jemug</i> ‘cuscus’)
2. Hawksbill turtle	
3. Other lexical innovations	
Phonological innovations	
4. Cluster reduction 1: $C_1V_1C_2C_1V_1C_2 > C_1V_1C_1V_1C_2$	
5. Cluster reduction 2: $CVC_iC_jVC > CVN_jC_jVC$	
6. PMP *uliq > *oliq ‘return’	(<i>f-ari</i>)
7. PMP *i-sai > *i-sei ‘who’	(<i>gata</i>)
8. PMP *ma-qitəm > *ma-qetəm ‘black’	(<i>grmutn</i>)
9. PMP *maRi > *mai ‘come’	ma
10. PMP *tudan > *todan ‘sit’	mtür
11. PMP *inum > *unum ‘drink’	in
12. PMP *inep > *enəp ‘sleep’	in
Morphosyntactic innovations	
13. Prefixal/proclitic agreement on verb	Yes
14. Alienable vs. inalienable possession	Yes
15. Frozen morphology:	
15a. PMP *həpat > *həpat, *pat, *pati ‘four’	(<i>gigti</i>)
15b. PMP *ma-huab > *mawab ‘yawn’	(<i>msgar</i>)
15c. PMP *ma-hiaq > *mayaq ‘shy’	(<i>rit der ti</i>)
Semantic innovations	
16. PMP *t-ina ‘mother’ > *t-ina ‘big’	-den ‘mother/big(rare)’
17. PMP *m-udehi ‘behind’ > *mudi ‘back (anat.)’	
18. PMP *ma-qitəm ‘black’ > *ma-qetəm ‘dirty’	(<i>grmutn</i> ‘black’; <i>sbitn</i> ‘dirty’)
19. PMP *tuqəla(n) ‘bone’ > Ø; *zuRi > ‘bone’	rur
20. PMP *buhək ‘hair’ > Ø; *daun ni qulu > ‘leaf of head’ feather’	drbun fru ‘hair (lit. head’s feather)’
21. PMP *daləm ‘inside’ > *daləm ‘mind, feelings’	(<i>gan</i>)

TABLE 4.9 - CEMP innovations (Blust 2009a). The use of ‘> Ø’ signifies ‘replaced by’. Change 5 reduces the number of possible consonant clusters, not the cluster itself. Changes 6–12 are lexically specified. Changes 13 and 14 imply cognate forms.

PCEMP change 1) ‘lexical innovations with special subgrouping import’ - PCEMP *kandoRa ‘cuscus, phalanger’, PCEMP *mansar/mansər ‘bandicoot’. Irarutu *jemug* is probably not related to *kandoRa. No data is available to compare with these items, but see PMP *ka-labaw ‘rat’ replaced by PCEMP *kanzupay > IRH *sfe* (*k > Ø, *a_{prepen} > Ø, *n > Ø, *z > s, syncope, *p > f, *guna* assimilation). However, the change *n > Ø could have been by chance. Change 2) PMP *peñu ‘green turtle’ was supplemented by PCEMP *keRa(nj) ‘hawksbill turtle’. No Irarutu data is relevant for comparison. Change 3) ‘other lexical innovations’ are elaborated in **TABLE 4.9** along with known Irarutu forms. Only a handful of valid comparisons are possible.

Just five of the forty innovations cited in **TABLE 4.10** show possible reflexes in Irarutu: *maya > *mmaje* ‘tongue’, *hiup > *uf* ‘blow’, *liqəR > *rü* ‘voice’, *malip > *brif* ‘laugh’, *ba > *fa* ‘go’. Most of the phonological correspondences between the PCEMP forms and Irarutu are regular. In *maya > *mmaje* ‘tongue’, *m > *m*, *y > *j*, and *-a > *e* (irregular). As mentioned above, consonant doubling, i.e., *m- > *mm-*, is occasional but irregular in the synchronic Irarutu language data and appears to be phonologically motivated (see Chapter 2).

In Irarutu, the word for ‘to blow’ is better attributed to PMP *upi > *uf*, with *p > *f*, and *-i > Ø, than PCEMP *hiup. In *liqəR > *rü*, rhoticization of *l, *q > Ø, and *-R > Ø are predictable; however, the vowel changes in *i and *ə that resulted in labiopalatal *ü* are unexplained (but syllables with rhotics seem to correspond to labiality, see §4.1.3). The correspondences of *m > *b*, *V_{penult} > Ø, *l > *r*, and *p > *f* describe *malip > *brif*. The correlation of PCEMP *ba to IRH *fa*, is possible because *b > *f*, but as noted previously, the occurrence of (optional) final -n in the Irarutu form *fa(n)* points to PMP *panaw.

Innovation:	Irarutu form:
PMP *dilaq ‘tongue’ > PCEMP *maya;	mmaje
PMP *surat ‘scratch, draw line or design’ > PCEMP *tusi	(ukir)
PMP *kapal ‘thick (of material)’ > PCEMP *telu;	(sarf)
PMP *hiup ‘to blow’ > PCEMP *upi ‘to blow’;	uf
*liqəR ‘voice’	rü
*malo ‘loincloth’	(ffur/tfaf)
*malip ‘laugh’	brif
*tosi ‘make a mark, scratch a line’	(tifn)
*tambu ‘taboo, forbidden’	(fnafejof)
*wan̄ka ‘canoe, boat’	(bir)
*qumun ‘earth oven’	(tugwan)
*abat ‘wound’	(jamt)
*au ‘dew’	(mud ruwer)
*ba ‘go’	fa(n)
*balaŋ ‘side, part’	(refid)
*belen ‘to swallow’	(tun)
*bina ‘shellfish sp.’	(tiga)
*birin̄ ‘to stone, throw a stone’	(ir nar kami dru)
*buu ‘to blow; conch shell trumpet’	(mabwin)
*ərit ‘scratch, scrape’	(figaris)
*i ‘numeral prefix’	--
*ima ‘pandanus w. leaves useful for plaiting’	(banro)
*isi ‘peel, strip off’	(dadn)
*bisik ‘to peel’	(nnis)
*bitak ‘mud’	(nbssi)
*bitu ‘sword grass: <i>Imperata cylindrica</i> ’	(sfetuf)
*bua ‘only, just’	(mtnot)
*bubu ‘sing; song’	(sis)
*budeq ‘sponge’	--
*buqal ‘levered up, uprooted’	(tafgign)
*butu ‘group, crowd, flock, school, bunch, cluster’	(me, jagf)
*matay ‘money, payment, medium of common exchange’	(sen)
*pali ‘side, half’	(refid)
*paniŋ ‘bait’	(fi nim wer)
*papaR ‘cheek’	(gewar)
*paRa- ‘reciprocal’	--
*qaue ‘exclamation of joy or sorrow’	(iteni ‘–joy’/afigi ‘–sorrow’)
*quRi ‘a fruit tree: <i>Spondias dulcis</i> ’	(ton ‘generic fruit tree’)
*wakir ‘k.o. root’	(war sfri ‘root above ground’)
*wari ‘sing; song’	(sis)

Table 4.10 – Forty ‘other PCEMP lexical innovations’, except where forms are noted as PMP.

Concerning the phonological innovations, change 4) ‘cluster reduction 1’: coda of the penult deletes, could be argued to have happened in *ma-dindiŋ > *dridn* ‘cold’, but based on the discussion of prenasalization in voiced obstruents, is more likely to be an instance of change 5) ‘cluster reduction 2’: coda of penult assimilates to onset of final syllable. However, change 5) was not intended by Blust for use as subgrouping evidence and

therefore is not considered further. Change 6) lower the penult in PMP *uliq ‘return, go back, restore’ > PCEMP *oliq ‘return’ > *grfari* ‘return’; is less convincing than *balik > IRH *grfari*, which reflects a greater proportion of attested correspondences. Change 7) irregular phonological raising of the penult from PMP *i-sai to PCEMP *i-sei ‘who’ is acknowledged as fairly weak evidence for subgrouping purposes because it may be convergent; furthermore, Irarutu replaced this form with *gata*, and so comparison is not possible. Change 8) lowering of the penult in PMP *ma-qitem ‘black’ > PCEMP *ma-qetəm ‘black’ cannot be compared with IRH *grmutn*. Change 9), deletion of *R in PMP *maRi ‘come’ > PCEMP *mai ‘come’ is supported by IRH *ma*, which reflects the irregular loss of medial PMP *R in this particular item. Change 10) lowering of penult in *todan ‘sit’ is ironic in the present context, because it is the only innovation that Donohue and Grimes 2008 supported, but it was discarded in Blust 2009a because of meager comparative data. Reflexes for both *tudān and *todan are cited in the literature on Irarutu; however, this is clearly due to an orthographic challenge posed by IRH labial vowels. Change 11) labialization of the penult in PMP *inum > PCEMP *unum ‘drink’ is not supported by IRH *in*; but, labialization is supported, albeit sporadically, by labio-palatals such as in *tatu > *tü* ‘true’ and *qulu > *rü* ‘head (source)’. Change 12), lowered penult in PCEMP *enəp ‘sleep’, marginally applies to IRH *m*, but it became lax, which could be interpreted as somewhat lower than *i*, but is not [-high]. Lowering in changes 6) and 12) support membership in CEMP, but differ in extent.

The first morphosyntactic change cited above is change 13) ‘prefixal/proclitic agreement on verb’, which, as Blust (2009a) notes, probably arose through several changes, therefore, it is weak as subgrouping evidence. Coincidentally, IRH does have prefixal agreement on the verb that cross-references an agent (more specifically the subject). Change 14) ‘alienable vs. inalienable possession’, relies on typological data, and it is also weak evidence. IRH does make this distinction. Despite its weak status, two publications, Voorhoeve 1989 and van den Berg and Matsumura 2008 cite this innovation to argue that IRH is a typical EMP language (see also Lichtenberk 1985). Change 15) ‘*h > Ø plus other change’, shown by frozen morphology in *həpat/pat/pati ‘four’, *mawab ‘yawn’, and *mayaq ‘shy’, is not reflected in Irarutu.

The semantic changes in PCEMP are better supported by Irarutu data. Change 16) is *t-ina ‘mother’ expands to mean ‘big’, but it is not certain that IRH *-den* is cognate. Supposing it is, Irarutu speakers use phrases such as *fa+ntü iden* (literally ‘foot offspring’s mother/biggest member’) ‘big toe’ where *fa ntu* (‘foot offspring’) means ‘toe’; however, it should be noted that the use of terms for concepts such as ‘big’ and ‘little’ in Irarutu are idiosyncratic. Change 17) *mudi ‘behind’ adds the meaning ‘back (of body)’. A corresponding term in IRH has not been discovered. Change 18) *maqetəm ‘black’ expands to mean ‘dirty’ is supported by the use of *grmutn* to mean ‘dusty (of clothes)’, but it is doubtful that these forms are cognate. Change 19) PMP *tuqelan/tuqelan ‘bone’ > Ø; PMP *zuRi ‘fish bone’ generalizes to PCEMP *zuRi ‘bone’ is witnessed by IRH *rur* ‘bone’. The phonological changes from PCEMP to IRH are regular, which further supports the semantic broadening of ‘fish bone’ to ‘bone’ prior to the development of Irarutu. Change 20) PMP *buhek ‘head hair’ > PCEMP *daun ni qulu ‘leaf of head’, i.e., a single lexical item is replaced by a periphrastic construction is

supported by IRH *rgwin fru* ~ *derbun fru* (‘head feather’) ‘head hair’ but also implies that *bulu ‘feather’ expanded to *fru* ‘bodily covering’. In support of Donohue and Grimes 2008, Irarutu data suggests that PMP *buhek was replaced more than once in the Austronesian languages of Eastern Indonesia because Irarutu does not reflect PCEMP *daun ni qulu, but rather a different epithet/periphrastic construction. This is a simpler explanation than postulating two changes (PMP *buhek > PCEMP *daun ni qulu and PCEMP *daun ni qulu > IRH *rgwin fru* ~ *derbun fru*); therefore, it is preferable, although it does not strictly rule out a serial change explanation. For change 21), Irarutu does not reflect *daləm ‘inside’ or ‘mind, feelings’.

A point related to diffusion is worth mention. In CMP, there are different reflexes for reconstructed diphthongs due to glide truncation or ‘guna assimilation’. Four proto-diphthongs were truncated: *-ay > -a, *-aw > -a, *-uy > -u, and *-iw > -i; by contrast, guna assimilation for the same protosegments resulted in -e, -o, -i, and -i, respectively. Reflexes of these diphthongs in Irarutu tend to follow the guna assimilation pattern, which is one trait that distinguishes Irarutu from most CMP languages.

To conclude, according to Donohue and Grimes (2008), only four of seventeen items for CEMP are supported by West Papuan languages: the presence of prefixal agreement, (in)alienable possession, the use of a descriptive phrase (but *daun ni qulu ‘leaf of head’ is replaced by *rguin fru* ‘head feather’) instead of a single lexeme for ‘head hair’, and the extension of *daləm to refer to ‘mind, feeling’; however, the evidence available on Irarutu is more substantial and does suggest that the language belongs under the CEMP node, with no substantial contradictions. For instance, *ma* reflects *mai ‘come’, with loss of PMP *R. Irarutu seems to have a reflex of *t-ina meaning ‘big’, as in *afagntu iden* ‘my big toe’, but more data is needed to substantiate this claim. Irarutu reflects *zuRi as *rur* ‘bone’. There are also reflexes of lexical innovations, such as PCEMP *maya > *mmaje* ‘tongue’; however, the picture is obscured by extensive replacement of protoforms, *i-sei ‘who’, *həpat ‘four’, *mawab ‘yawn’, *mayaq ‘shy’, *mudi ‘back’, and *daləm ‘mind, feeling’. Although Irarutu *in* ‘drink’ more closely resembles PMP and PCEMP *inum than PCEMP *unum, based on the full set of possible comparisons with proposed innovations, words in Irarutu that appear to be Austronesian etyma suggest the language does belong to the CEMP subgroup.

4.2.1.2 Eastern Malayo-Polynesian (EMP)

Like PCEMP, Proto-Eastern Malayo-Polynesian may have been a dialect chain, and it is expected that at least some of its features be present in Irarutu. Again, geographic location is a broad indicator of subgroup membership, but only shared innovations solidify the claim. Fifty-six innovations have been proposed for the Eastern Malayo-Polynesian subgroup (Blust 1978). These innovations, cited in **Table 4.11**, are mostly lexical, but there are four formal and four semantic innovations. It is pertinent to note that no phonological innovations have been identified for this subgroup.

Title of Change:

Irarutu form:

EMP formal innovations

- | | |
|---|-------------------|
| 1) PMP *qasu > PEMP *kasu ‘smoke’ | (<i>bob</i>) |
| 2) PMP *putul > PEMP *mutu ‘broken off, severed’ (*p>m)
(cf. PPN *mutu ‘cut off, ended’) | (<i>tubr</i>) |
| 3) PMP *besuR > PEMP *masuR/mosuR ‘satiated’ | (<i>risr</i>) |
| 4) PMP *duyuj > PEMP *dui ‘dugong’ (**-C) | (<i>funito</i>) |

EMP semantic innovations

- | | |
|--|---|
| 1) PMP *kakāt ‘peel’ > PEMP *kakāt-(i) ‘to peel (w/ teeth)’
(cf. PMP *kaRat, *getget ‘gnash teeth’) | (<i>mennis</i>) |
| 2) PMP *lumut ‘moss’ > PEMP *lumut ‘green’ | (<i>brumcie</i> ‘moss’, <i>mmat</i> ‘green’) |
| 3) PEMP *mimiR ‘urinate’;
(cf. PMP *iheq ‘urinate’)
PEMP *miRmiR ‘urinate’
(cf. PMP *miRmiR ‘spray’ and
POc *mimiR ‘spurt out, urinate’) | (<i>tbri</i> , <i>nbrufn</i>) |
| 4) PEMP *pa(ka)-salaq ‘punish (CAUS+wrong)’ | (<i>mtür tarir</i>) |

EMP Lexical Innovations

- | | |
|--|--|
| 1) *qanus-(i) ‘to spit’;
(cf. PMP *luzaq, POc *qanus ‘spittle’) | (<i>tfi</i> ‘to spit’) |
| 2) *ayawan ‘k.o. tree: ficus sp.’
(cf. PMP *nunuk ‘banyan’) | (<i>udaf</i> ‘banyan tree’) |
| 3) *boi/bui ‘smell, stench’
(cf. PMP *bahuq) | (<i>numir dirti</i> ‘bad smell’) |
| 4) *budan ‘white’
(cf. PMP *ma-putiq/buraq, POc *pulan ‘white’) | (<i>bfut</i> ‘white’ < PMP *ma-putiq) |
| 5) *butak ‘close the eyes’ | (<i>git mce</i> ‘close eyes’) |
| 6) *dadas-(i) ‘peel off’
(cf. POc *dadasi ‘scratch, peel, cut’) | |
| 7) *da(q)um ‘shade, shady’ | (<i>bwen</i> ‘shade’) |
| 8) *dege ‘cavity, small recess’
(cf. POc *ndeke ‘hollow, concavity...’) | (<i>riwar</i> ‘small cavity’,
<i>brufn</i> ‘recess’) |
| 9) *ibu/ubi ‘drinking vessel’
(cf. PPN *ipu ‘container for liquid’) | (<i>fitbi</i> , <i>bok</i>) ‘drinking vessel’) |
| 10) *iRiR ‘to fan’
(cf. POc *iRiR ‘to fan’) | (<i>mtfraf</i> ‘to fan’) |
| 11) *ka(dR)a ‘cockatoo; parrot’ | ra , (<i>akiko</i>) ‘parrot’ |
| 12) *kiñit ‘pinch’
(cf. POc *kiñit ‘pinch’) | (<i>nimt</i> ‘pinch’) |
| 13) *laman(a) ‘deep sea’
(cf. POc *laman(a) ‘sea, lake’) | (<i>riran</i> ‘sea’) |
| 14) *ma- ‘directional particle’ | — |
| 15) *mada- ‘depreciatory force (PREF)’ | — |
| 16) *mada ‘ripe, soft’
(cf. POc *mada ‘fermented, soft, ripe, wet’) | (<i>brmnir</i> ‘ripe’) |
| 17) *maiduj ‘evening dusk’ | (<i>rre nro</i> , <i>rre nmur</i> ‘dusk’) |
| 18) *maluRu ‘shade’ | (<i>bwen</i> ‘shade’) |

<p>(cf. PPN *malu ‘shade’)</p> <p>19) *matu ‘dry coconut’</p> <p>20) *mawa ‘enclosed space’</p> <p>21) *mo(dl)an ‘real, genuine, true’</p> <p>22) *momeq/momaw ‘crumbs, sweepings, rubbish’</p> <p>23) *mumo ‘whisper’</p> <p>24) *(n)a(dR)i ‘to wait’</p> <p>25) *natu ‘small/young’</p>	<p>(umagi ‘dry coconut’)</p> <p>—</p> <p>(tū ‘true’)</p> <p>(nsir ‘rubbish’)</p> <p>(nfi snnan ‘whisper’)</p> <p>(msgre ‘to wait’)</p> <p>ntū ‘smallest member of a set/offspring’</p>
<p>(cf. PMP *anak ‘child, POc *natu ‘child’)</p> <p>26) *(n)iwī ‘nest’</p> <p>27) *ŋu(dR)um ‘grunt, growl’</p> <p>28) *patote(k,q)/patotaw ‘outrigger boom’ (cf. POc *patoto ‘outrigger boom’)</p> <p>29) *beke ‘defecate’</p> <p>30) *sala ‘sharp-pointed object’ (cf. PPN *tala)</p> <p>31) *sakaRu ‘reef’ (cf. POc *sakaRu ‘reef’)</p> <p>32) *sepat ‘go past, go beyond’</p> <p>33) *sobu ‘go down, descend; dive’ (cf. POc *sompū ‘down’)</p> <p>34) *soRa/suRa ‘to help’ (cf. POc *soRa ‘to help’)</p> <p>35) *suda ‘comb’</p> <p>36) *supi ‘peel, pare’</p> <p>37) *tabus ‘taboo; sign of something interdicted’ (cf. POc *tampu ‘a ban, ritual restriction protected by supernatural sanction, marked by taboo sign’)</p> <p>*tambus ‘taboo’ (cf. PMP *palihī ‘taboo’)</p>	<p>(isan ‘house’)</p> <p>(sfut, narsrut ‘growl (dog)’)</p> <p>(nkrkar ‘grunt (pig)’)</p> <p>(bir fra ‘outrigger’)</p> <p>(fun nte ~ fun nta ‘defecate’)</p> <p>(krir ‘sharp-pointed object’)</p> <p>—</p> <p>nfat ‘go beyond’</p> <p>(msru ‘go down’)</p> <p>(nmur ‘descend’)</p> <p>(marfi mage ‘help’)</p> <p>(sisir, sir saft ‘comb’)</p> <p>(guf ‘peel by hand’)</p> <p>(fnafejof ‘prohibited’)</p>
<p>38) *ta(dR)i ‘steer a course (in navigating)’</p> <p>39) *taŋa ‘carrying container’ (cf. POc *taŋa ‘holder, bag’)</p> <p>40) *tatu ‘true’</p> <p>41) *tawan ‘a tree: pometia pinnata’</p> <p>42) *tinan ‘elder (of kinsmen)’</p> <p>43) *tobe(k,R) ‘throw down’</p> <p>44) *tobV ‘fishnet float and the wood from which it is made’</p> <p>45) *qutem/qutenj ‘fishnet float’</p> <p>46) *qutub ‘submerge to fill’ (cf. POc ‘qutup’ flood, draw water, fill with liquid, soak’)</p> <p>47) *wa(q)ip ‘scoop or container for water’</p> <p>48) *wataŋ ‘bring, carry, take’ (cf. PMP *alap ‘take’)</p>	<p>(birir ‘steer a canoe’)</p> <p>(nyatbi ‘carrying vessel’)</p> <p>tū, <i>fkue</i>, NBN <i>fetu</i> ‘true’</p> <p>—</p> <p>(nabrid ‘old person’)</p> <p>(mti ‘throw down’)</p> <p>—</p> <p>—</p> <p>(nfun ‘submerge to fill’)</p> <p>—</p> <p>(mar mama ‘take’, nan fa)</p> <p>(gud mama ‘bring’)</p>

TABLE 4.11 - EMP innovations (Blust 1995). Double asterisk indicates an expected form that is not substantiated, whereas ‘—’ indicates that no relevant data is available.

Based on the discussion in §4.1, the Irarutu data does not appear to be comparable with the four formal innovations proposed in TABLE 4.11. The two best candidates to show a correlation between PMP and IRH are: *putul > *tubr* ‘severed’; PMP *besuR > *risr* ‘satiated’. Both are unlikely.

It is a stretch to compare material in the domain of semantic innovation; however, there are two possible comparisons: the extension of *lumut from ‘moss’ to ‘green’, and *mi(R)miR from ‘spray’ to ‘urinate’. The change of *lumut > *mmat* ‘green’ can be ruled out because *mataq > *mmat* ‘green/unripe’ was proved on phonological grounds. The other comparison, *mimiR > *tbri* ‘urinate’ lacks reliable sound correspondences. These tenuous comparisons are therefore irrelevant for subgrouping purposes.

Only four Irarutu lexical items of forty-eight putative EMP lexical innovations are relevant for comparison: *ka(dR)a > *ra* ‘cockatoo’, *natu > *ntü* ‘offspring/small member of set’, *sepat > *nfat* ‘go past’, and *tatu > *tü* ‘true’. There is also one item, *bfut* ‘white’, that is retained from a pre-EMP source, PCEMP *ma-putiq.

Lexical innovation 11) *ka(dR)a > *ra* ‘parrot’ is found in Irarutu. It reflects stress shift, syncope, *k > Ø, and rhoticization, as well as justifiable semantic reference. Innovation 25) *natu > *ntü* ‘small’, with syncope, is also present, but the meaning has expanded to ‘smallest member of a set/offspring’, as in *fra ntü* (arm offspring) ‘finger’, and it is mostly used in compounds and idioms. The more frequent Irarutu term for ‘child/small’ is *mo*. Innovation 32) *sepat > *nfat* ‘go past’ reflects syncope, *p > f, and semi-recurrent *s > Ø (cf *si-ida > *ir* ‘they’). The final example, innovation 40) *tatu > *tü* ‘true’ appears to be cognate, but could also be derived from PMP *tuqu. Compared to *tatu, *tü* shows syncope, whereas *t- > Ø stands out, although it is not barred on phonological grounds, cf. FCD.

In contrast with these four examples, lexical innovation 4) *ma-putiq > *bfut* ‘white’, which was already shown to be cognate, is contrary to the proposed lexical replacement of PMP *ma-putiq/burak with EMP *budan ‘white’. In conclusion, less than ten percent (4/48) of the available comparisons for lexical innovation support a connection between Irarutu and EMP. This is hardly convincing for the purpose of genetic classification. Because much of the evidence for EMP is lexical, one problem with this kind of investigation is the difficulty in distinguishing innovation from retention (see Blust 1978). Although it is not a common type of explanation, the concept of ‘lemma’, which is a kind of metonym where an auxiliary term is taken to demonstrate a composition and is used in philology, lexicography, hermeneutics (Gadamer 1975), and psycholinguistic analysis of lexical priming, alleviates some of the inherent vagueness by illustrating that it is not only possible, but it is psychologically real, for various concepts to have more than one way to be referred to (see also Boretzky 1984). Nevertheless, it is hard to disprove that *budan and *ma-putiq coexisted in PMP with slightly different nuance in meaning, and that *budan, being a retention, supplanted PCEMP *ma-putiq within the EMP dialect chain, except in SHWNG languages such as Irarutu. For present purposes, the lexical material presented above is considered valid as subgrouping evidence, especially because EMP consists of SHWNG and Oc, and EMP is a daughter of Malayo-

Polynesian. These innovations can be assumed to be distinguishable from retentions based on probability. In other words, because a SHWNG language (i.e., Iraputu) has a form that looks like a shared innovation from EMP, based on forms in closely related SHWNG and OC languages, it is feasible to assume that the form is an innovation rather than a retention from PCEMP, despite the fact that retention cannot be ruled out. The correspondences in **Table 4.11** that agree with the proposed innovations outweigh the viable comparisons that disagree. As a result, the Austronesian lexical items in Iraputu that can be classified as EMP will be; but the somewhat marginal nature of evidence for this claim is acknowledged.

4.2.1.3 South Halamahera-West New Guinea (SHWNG)

As mentioned above, Oceanic languages are geographically located East of Cenderawasih Bay (Grace 1971). Therefore, Iraputu, which is located West of Cenderawasih Bay, in Eastern Indonesia, has a reasonable chance of belonging in the SHWNG subgroup, depending, of course, on linguistic evidence. Although there are thought to be less than three and a half dozen SHWNG languages, their interrelationships are not yet well understood. A modified family tree adapted from Gordon (2013; but see also Kamholz 2014, and forthcoming) is illustrated in **FIGURE 4.5**, below, with one major modification: Iraputu and Kuri are grouped together directly under the SHWNG node, following Voorhoeve (1989) because *t did not become Iraputu *s* before *i; however, based on the data presented so far, it could still be the case that Iraputu-Kuri subgroups outside of SHWNG but underwent extensive modification due to prolonged contact with SHWNG languages.

It is not the goal of this work to evaluate and repair the whole structure of the SHWNG subgroup, but some discussion of **FIGURE 4.5** will be helpful to explain the structure of relevant dimensions of the subgroup. The primary reason to group Iraputu and Nabi together is lexical similarity, which, based on roughly 1300 lexical items (collected by the author in 2010), is estimated to be about 65% lexically similar, but Gordon (2013) claims 90% lexical similarity without referring to any literature with specific information on Nabi. Furthermore, based on 2010 data, the grammars of these languages are also markedly similar. For our purposes, these languages can be considered on the verge of mutually intelligibility. They are also in mutual contact with several of the same Austronesian and non-Austronesian (Papuan) languages, such as Wandamen to the North and Mairasi (non-An), East of Iraputu and South of Nabi.

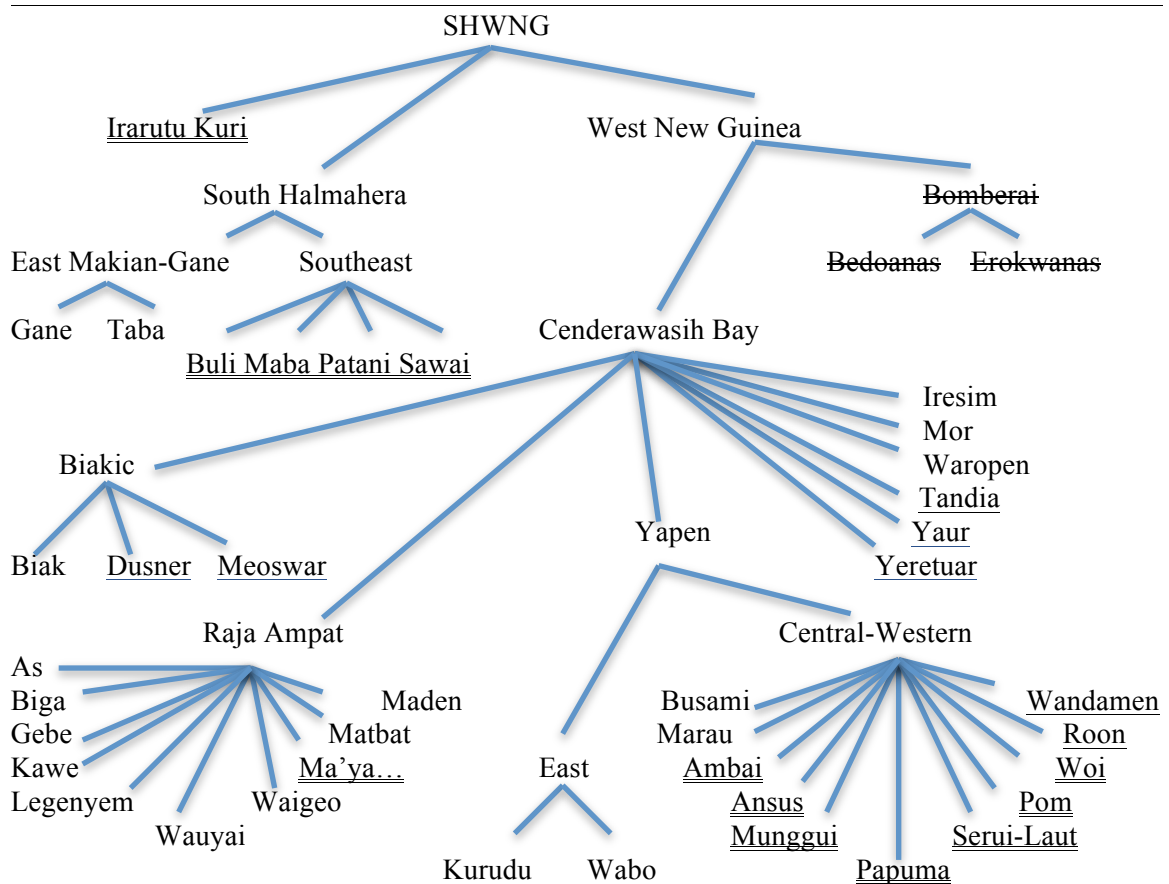


FIGURE 4.5 – SHWNG tree model, adapted from Lewis et al 2013. The primary change is the location of Irarutu as a primary branch of SHWNG, and the proposal that Irarutu and Kuri are a dialect chain. Other notation: the Southeast, South Halmahera languages Buli, Maba, Patani, and Sawai form a dialect chain; within Central-Western Yapen, Ambai, Ansus, Munggui, Papuma, Serui-Laut, Pom, and Woi form a dialect chain; and, Ma'ya is a dialect chain (Voorhoeve 1989). There is a sense of similarity for Papuans to associate Dusner, Wandamen, and Tandia with Meoswar, Roon, Yeretuar, and Yaur. According to Andreas Deda (p.c.), these seven languages probably belong in the Biakic group. ~~Bedoanas~~ and ~~Erokwanas~~ are place names, not languages.

According to several Papuans, some of whom are linguists, Tandia *is* Kuri as well as the name of a village on the back of the Bird's Neck (Andreas Deda p.c., but see also Deda et al. 2012). This is important because by all estimates, Tandia is the oldest Austronesian settlement in West Papua and was probably the original point of entry for Austronesian people into Papua. Dusner, an erroneous spelling of 'Usner', is located a short distance across Wandamen Bay to the East; it is moribund (Dalrymple and Mofu 2012). The available data on other nearby Austronesian languages such as Iresiam ('Iresim' in **FIGURE 4.5**) and Moor (Kamholz forthcoming), as well as Waropen (Held 1957), are notably distinct from Irarutu. Gordon (2013), presumably based on Ross (1995), classifies Irarutu as a South Halmahera language, but evidence shows Irarutu to be more similar to the West New Guinea languages. There are no signs of extensive contact

between Irarutu and the CEMP languages in Bomberai (Goodman 1998, Visser 1989), due to geographic separation and several intervening Trans-New Guinea languages. Assuming that various assertions about dialect chaining that are indicated in **FIGURE 4.5** are valid, there may be as few as thirty distinct SHWNG languages.

4.2.1.3.1 SHWNG Literature review

Before reviewing the Irarutu language data, it is beneficial to briefly review the history of scholarship on SHWNG languages. Adriani and Kruijt (1914) noticed the similarity between South Halmahera languages (Makian, Buli, Sawai, and Gane), ‘Kalana Fat’ languages, which are now known by the name ‘Raja Ampat’ (Waigeo, Salawati, Misool), and the Biakic language Numfor (Dutch ‘Noemfoor’). Four traits that Adriani and Kruijt comment on are: 1) loss of final vowels, 2) stress shift accompanied by syncope, 3) the morpheme *si* pluralizes nouns, and 4) a ‘reversed genitive’, where the modifier precedes the head, although this trait reflects an areal tendency that is even found in Papuan Malay. Esser (1938) includes four types of relevant languages in his classification (‘South Halmahera languages’, Numfor, Windesi, and Kowiai), whereas Dyen (1965, 1978) claims eight primary branches of An are located in Cenderawasih Bay in his classification (Buli, Minyafuin, Biga, As, Biak-Numfor, Wandamen, Yapen, and Waropen). Anceaux’s (1961) ‘Sarera group’ (Sarera Bay is an old name for Cenderawasih Bay) is based on 1) metathesis, both generally in the final syllable, and in the specific form *(ma)-dalem (e.g., Misool *malaman*) ‘inside/deep’ but *l > Sarera *n* elsewhere, 2) penultimate *e > Sarera *e* but expect *o*, 3) *d > Ø, and 4) sporadic palatalization, for example, the alveolar nasal in *t-ina ‘mother’ becomes palatal *ñ*.

Not only was Anceaux (1961) the first to classify Irarutu with languages of Cenderawasih Bay (WNG), he also provides what he calls a ‘quantitative’ analysis, where he segregates words that seem to correspond between languages in his data set from words that seem unrelated. He clearly asserts that Irarutu is closest to Mor (variant spelling of Moor), with 32% of the items appearing to be cognate with a high level of confidence, whereas Kurudu has an equal number of apparent cognates, but less confidence. Marau reflects the highest absolute percent of cognates, 35%, but at a poor level of confidence, followed by Mor and Dusner, the latter of which shares 21% vocabulary items with Irarutu. Dropping another degree of confidence, Ron also has 21% shared vocabulary with Irarutu. These percentages are strikingly low (but see Blust 1981, 2000). Based on his findings, Anceaux suggests that Irarutu belongs to a closely related but different subgroup than the Cenderawasih Bay languages, and may be an isolate within that subgroup. Anceaux supports this classification with the geographical position of Irarutu, which is distant from the South Halmahera languages.

4.2.1.3.2 SHWNG innovations

Blust (1978), using data from Anceaux (1961), but evidence from regular sound correspondences, also concludes that Irarutu belongs with the West New Guinea (Cenderawasih Bay) group, but assigns it no specific position within SHWNG. In several publications he identifies thirteen phonological innovations from PAn to PSHWNG organized into five categories: consonant shifts, vowel shifts, epenthesis, consonant mergers, and vowel mergers (Blust 1974, 1978, 1983/84; also summarized in Jackson

2008), see **Table 4.11**. The ‘distinctive’ changes, i.e., changes that are unusual outside the proposed group, 2), 6), 9) and 12), are **bold face**, whereas the ‘diagnostic’ changes, i.e., those that can be used to qualify a language as a member of the proposed group, 2), 4), 6), and 11) are underlined. Changes 2) and 6) are both distinctive and diagnostic. Voorhoeve (1989) cites distinctive change 9), but points out that *e in the final syllable became *e* or *Ø* in Irarutu. Subsequently, based on Voorhoeve’s description in Tryon (1995), Ross (1995) additionally appeals to Change 1).

Consonant shifts:

- 1) *p > *f*, *C, *t, *T > SH:*c*, or WNG:*k*; *b > SH:*p*, or WNG:*b*; *-d- > SH:*r*

Vowel shift:

- 2) *e_{penult} > *o***

Epenthesis:

- 3) *a > *ya*

Consonant mergers:

- 4) { *C, *c, *T, *t } / __i > *s*; and *-j-, *s > *s*

- 5) *k, *q, *‘, *H, *S, *x, *-j, *R, *-ŋ > *Ø*

- 6) *D, *d, *Z, *z, *l, *r > SH:*l*, or WNG:*r***

- 7) *n, *ñ, *ŋ > *n*

- 8) *-C_{coda}C_{onset}- > -C_{onset}-

Vowel mergers:

- 9) *e_{final}, *a > SH:*a*, or WNG:*e***

- 10) *i, *u_{penult}, *a_{final} > SH:*i* or WNG:*e*

- 11) Apocope 1: *-V > *Ø*

- 12) Syncope: *V_{penult} > *Ø***

- 13) Apocope 2: truncate final diphthongs

TABLE 4.12 – SHWNG innovations (Blust 1978). Diagnostic changes and **distinctive changes**.

There are four subcategories for the consonant shifts identified in Change 1). 1a) *p > *f* called ‘fricativization’ above, i.e., is attested. 1b) *C, *t, *T > SH *c* or WNG *k* is not supported. 1c) *b > SH *p* or WNG *b* is partially supported, but with phonetic prenasalization and two other reflexes, *f* and *m*. 1d) *-d- > SH *r* is supported, but appears to be better justified under change 6), ‘rhoticization’. PCEMP did not have *C or *T.

PCEMP *p > PSHWNG and IRH *f*

*panaw > *fa* ‘to walk’

*kanzupay > *sfe* ‘rat’

PCEMP *t > IRH *t* (expect SH *c* or WNG *k*)

*kulit > *rit* ‘skin’

*matay > *n-mat* ‘(it is) dead’

*tunu ‘burn’ > *tün* ‘cook’

PCEMP *-d- > SH and IRH *r* (but see Change 6)

*ma-dindin > *dridn* ‘cold’

PCEMP *b > WNG and IRH *b*, but also > IRH *f* and *m* (see §4.1)

*bitil > *bitr* ‘hungry’

*babinay ‘woman’ > *bfin* ‘female’ (rare)

The distinctive vowel shift proposed in Change 2), i.e., PAn/PMP **e*_{penult} > PSHWNG **o* is supported by three citable forms. One reconstructed form with the sequence **eCe* that has been identified in Irarutu lost the vowel in the final syllable due to ‘targeted deletion’. In an effort to classify Irarutu, this change was cited by both Blust (1978) and Voorhoeve (1989). It is reasonably strong evidence (see also the discussion of syncope and lexical strata in §4.1 regarding the word for ‘three’).

PMP **e*_{penult} > PSHWNG *o* > IRH *o*

PMP **dejeR* > IRH *f-nogr* ‘hear’ (Voorhoeve (1989) cites ‘*nonggar*’)

PMP **telu* > PCEMP **təlu* > IRH *tor* ‘three’

PAn **tebuS(u)* (Voorhoeve 1989) ‘sugarcane’ > IRH *tof* ‘sugar/sugarcane’

The epenthetic innovation, Change 3) **a*- > *ya*-, which also occurs in many CMP and Oceanic languages, is obscured by contextual factors, but there are two relevant pieces of information that suggest it did occur in Irarutu. It is possible that the PCEMP word initial morpheme **i*-, which occurs with speech act participant pronouns, may have vanished sometime between PCEMP and PSHWNG, but the facts related to glide fortition, discussed in §4.1 seem to suggest it did not. Supposing it was deleted before the development of PSHWNG, a cyclic paradox is created: **i*-aku > **aku* (syncretism, although it could simply be procope) > *a* (word final syllable deletion) > *ya* (y-epenthesis) > *ja* (glide fortition) ‘I’. This change is not well substantiated in Irarutu.

PCEMP *(*i*-)*a* > PSHWNG *ya*

**i*-aku > *ja* [ja] ‘I’

paralleled by, PCEMP **y*_{onset} > IRH *j* [j]

**maya* > *mmaje* ‘tongue’

The first consonant merger, Change 4) where PMP **c* and **t* before **i*, medial **-j-*, and **s* merged as PSHWNG **s* is a ‘diagnostic change’ that was utilized by Blust (1978) and Voorhoeve (1989) for classifying Irarutu. The merger of earlier **t* with **s* was conditioned: preceding the high front vowel **i*. This is not supported by any known data from Irarutu. The merger of medial **-j-* with **s* is observed in one reliable form, **pajay* > *fas*, and one ‘pseudo-cognate’, **ŋajan* ‘name’ > *snan* ‘word’. Therefore, one must infer that this portion of the merger was sporadic. On the other hand, the general retention of **s* as Irarutu *s*, is supported by recurrent data. Overall, Change 4) is not well-supported by Irarutu data. The diverse reflexes of **s* depend on the history of the personal pronouns, and could be attributed to a lexical strata (see §4.1.5).

PCEMP **t*/___*i* > IRH *t* (expect PSHWNG *s*)

**bitil* > *bitr* ‘hungry’

**ma-putih* > *bfut* ‘white’

PAN ‘penis’ (Voorhoeve 1989) > *ti* ‘penis’

PEMP **tini* (Voorhoeve 1989) > *tini* ‘body’

PCEMP **-j-* > PSHWNG *s*

PAN **pajay* > *fas* ‘rice’

**ŋajan* ‘name’ > *snan* ‘name/word’ (pseudo-cognate)

PCEMP *s > SHWNG and IRH s, but also > Ø
***susu** ‘breast’ > *sus*

Voorhoeve (1989) astutely deduced that the lack of attestation for *t / i > IRH s suggests that if Irarutu belongs in the SHWNG subgroup, it must have branched-off very early on, before SH languages, which do reflect the innovation, ‘migrated’ West (more on this in §4.2.1.3.3, below), as represented in the family tree model in **FIGURE 4.4**.

The second category of merger in the consonants, Change 5), is with Ø, in other words it represents deletion. Most instances of *k > Ø, except a few examples of initial *k, which became voiced under anticipatory voicing assimilation, and hence phonetically prenasalized. Only one example of *-j > Ø has been identified in the data, so it is difficult to endorse this merger because of its rarity. *R is generally reflected as IRH r, but it was deleted in at least one word. The only data on *-ŋ shows that it became IRH n, not Ø. Voorhoeve regarded change 5) as being peripheral, but for the most part it is well illustrated by Irarutu data.

PMP *k > SHWNG and IRH Ø

***manuk** > *man* ‘bird’

***kutu** > *ut* ‘louse’

but also, PCEMP *k > g (see §4.1)

***kali** > *gar* ‘dig’

PCEMP *q > PSHWNG and IRH Ø

***qatelur** > *tru* ‘egg’

***ma-Ruqanay** > *mran* ‘male’

PCEMP *-j > PSHWNG and IRH Ø (sporadic?)

***warəj** > *wara* ‘rope’

PCEMP *R > PSHWNG and IRH Ø (in one lexical exception)

***maRi** > *ma* ‘come’

but, PCEMP *R > IRH r

***ma-Ruqanay** > *mran* ‘male’

***zuRi** > *rur* ‘bone’

PCEMP *-ŋ > IRH n (sporadic, expect PSHWNG Ø)

***ma-dindiŋ** > *dridn* ‘cold’

The third consonant merger, Change 6), PMP *d, *l, *z > WNG r, is the most well-attested innovation in Irarutu. It was labeled ‘rhoticization’ in §4.1, and shows a very strong allegiance of Fruata dialect with WNG languages. For classificatory purposes, this change is considered both distinctive and diagnostic. It was an essential component in the arguments by both Blust (1978) and Voorhoeve (1989). There are only two known exceptions to ‘rhoticization’ of *z, both showing PCEMP *z > IRH s, but one of these is observed in a ‘pseudo-cognate’, ***quzan** > *siem* ‘rain’, which reflects a sporadic change, *n > m.

PCEMP *d > WNG and IRH *r*
 *ma-**d**indĩ > *dridn* ‘cold’
 ***d**aun > *ro* ‘leaf’
 ***d**ua > *ru* ‘two’
 PCEMP *l > WNG and IRH *r*
 *la**ŋ**it > *ragt* ‘sky’
 *za**l**an > *rarn* ‘road’
 PCEMP *z > WNG and IRH *r*
 *zu**R**i > *rur* ‘bone’
 *za**l**an > *rarn* ‘road’
 but, PCEMP *z > IRH *s*
 *kan**z**upay > *sfe* ‘rat’

By contrast, the fourth consonant merger, Change 7), *n, *ñ, *ŋ > n, is supported by a number of forms, as well as the synchronic segment inventory of Irarutu, which lacks phonemic palatal and velar nasals. There are no known examples of cognates that reflect *ñ > n. Unfortunately, this merger is not significant for the task of classifying Irarutu.

PCEMP *n > SHWNG and IRH *n*
 *ta**n**əm > *tan* ‘to plant’
 PEMP ***n**atu ‘child’ > *ntü* ‘offspring/smallest member of a set’
 PCEMP *ŋ > SHWNG and IRH *n*
 *ma-din**ŋ**ĩ > *dridn* ‘cold’
 and possibly,
 ***ŋ**ajan ‘name’ > *snan* ‘word’
 but, PCEMP *ŋ > IRH *g*
 *la**ŋ**it > *ragt* ‘sky’
 *de**ŋ**er > *nogr* ‘hear’

The fifth, and final, consonant merger, Change 8), depends on context. It is represented abstractly as *CC > C, but was originally proposed as *NC > C. This is true in two cognates in Irarutu, but under the stipulation that the resulting segment was voiced and subsequently became phonetically prenasalized ^NC, a quality that subsequently affected all voiced segments in all positions in Irarutu. One of the items, the pseudo-cognate *sub* ‘grow’ shows a sporadic, context-dependent change, *t > s/ __u, but the reduction in the consonant cluster matches the change in the other example, and is therefore included in the discussion here. Although this change is semi-recurrent, it bears little weight in the classification of Irarutu.

PCEMP *-C_{coda}C_{onset}- > SHWNG and IRH -C_{onset}-
 *ma-din**ŋ**ĩ > *dridn* [rⁿdn] ‘cold’
 *tum**buq** > *sub* [su^mb] ‘grow’

Change 9), *e_{final} and *a > SH *a* or WNG *e*, represents a distinctive, bi-partite vowel merger that should be able to be used to help differentiate SH from WNG languages. It was cited by Voorhoeve (1989). Robust Irarutu data supports *a > *a*, which is typical of

South Halmahera languages, but – assuming the validity of reconstructed penultimate stress – as a retention in stressed syllables, its association with SH languages loses significance. Furthermore, unstressed syllables that are reconstructed with *a lose this segment; and more importantly, due to the nature of the data, this superficial association can effectively be refuted. First, in the majority of instances, unstressed PCEMP *a was deleted, or semi-recurrently became *u*. Second, the discussion of guna assimilation in §4.1, above, showed that *ay became Irarutu *e*. Third, semi-recurrent context-dependent ‘breaking’ also entails a change of *a > *e*, albeit a diphthong [je]. Lastly, the reflex of PMP *e [ə] in the final syllable, which became PCEMP *ə, is reflected as IRH Ø due to apocope, save for a semi-recurrent shift to *o*. Therefore, despite the fact that PCEMP *a > IRH *a* superficially seems to support classifying Irarutu with SH languages, it only does so in some contexts, therefore that correlation does not represent a reliable change of merger of final *e with either SH *a* or WNG *e*.

PCEMP *a > SH and IRH 'a

*tanem > *tan* ‘to plant’

*panaw > *fa(n)* ‘go/walk’

*zalan > *rarn* ‘path/road’

*qabaRa > *fra-fu* ‘shoulder (lit. boney protrusion of the shoulder)’; PMP *buku > *fu*

but, PCEMP *a_{unstressed} > IRH Ø

*panaw > *fa(n)* ‘go/walk’

*zalan > *rarn* ‘path/road’

*qabaRa > *fra-fu* ‘shoulder (lit. boney protrusion of the shoulder)’

or, PCEMP *a > IRH *u* (semi-recurrent)

*asa > *esu* ‘one’

*taumataq > *matü* ‘person’

and, PCEMP *ay > IRH *e* (*ay > ey > *e*)

*kanzupay > *sfe* ‘rat’

*kayu > *e* ‘wood’

*waiR (> *wayr*) > *wer* ‘water’

and, PCEMP *a_{final} > *ie* (semi-recurrent/sporadic)

*mata > *mtie* ‘eye’

PMP *e_{final} > PCEMP *ə_{final} > IRH Ø (expect SH *a* or WNG *e*)

*denger > *dengər* > *f-nogr* ‘hear’

*inēp > *inəp* > IRH *m* ‘to sleep’

but, PCEMP *e_{final} > IRH *o* (sporadic)

*lipən > *rfo* ‘tooth’

The second vowel merger, Change 10): *i, *u_{penult}, *a_{final} > SH *i* or WNG *e*, is interpreted in a similar manner as Change 9). Stressed *i > Irarutu *i* is a retention, with the exception of context-dependent laxing (*i > ɪ preceding a syllable with a lax vowel). It is worth recapitulating the fact that synchronically *i* ranges over [i ~ ɪ ~ e] in casual speech. Stressed penults containing *u became *u* or *ü*, another retention/conditioned innovation, but one that contradicts merger; but, *u > Ø in unstressed, syncopated items. There are no straightforward instances of *a > *e* in the final syllable, but again, data from guna assimilation and the semi-recurrent/sporadic process of breaking suggest it did merge

with *e*. Therefore, the retention PCEMP *i > IRH *i* cannot be justified as evidence to group Irarutu with SH languages, whereas the reflex of final *a* mildly implies merger with *e*, and hence points to grouping IRH with WNG languages. Overall, the data suggests that Change 10) did not take place in Irarutu, but nevertheless, this innovation is neither distinctive nor diagnostic, so it is weak evidence for the claim that Irarutu belongs with the WNG languages.

PCEMP *i > SH and IRH *i* (coincidence)

*ma-dindin > *dridn* ‘cold’

*inum > *in* ‘drink’

PEMP *tini > *tñi* ‘body’

PCEMP *u_{penult} > IRH *u* (expect SH *i* or WNG *e*)

*bunuq > *fun* ‘kill’

*tunu > *tñn* ‘cook’

but, PCEMP *u_{penult-unstressed} > IRH Ø

*kanzupay > *sfe* ‘rat’

PCEMP *a_{final-unstressed} > IRH Ø (expect SH *i* or WNG *e*)

*mataq > *m-mat* ‘green’

*matay > *n-mat* ‘dead’

but, PCEMP *-ay > PSHWNG and IRH *e* (guna assimilation)

*kanzupay > *sfe* ‘rat’

and, PCEMP *a_{final} > IRH *je* (breaking)

*mata > *mtie* ‘eye’

The next merger, identified as Change 11) ‘Apocope 1’, i.e., final vowels were lost, is a diagnostic innovation for the SHWNG subgroup, and is reported on in Adriani and Kruijt 1914. More than half of the words identified as An cognates in **Table 4.1**, many also attesting to final consonant deletion, show this innovation. It is interesting to note that in some words, the vowel in the final syllable was subject to ‘targeted deletion’, whereas syncope prevented its application in a smaller set of items.

PCEMP *-V > PSHWNG and IRH Ø

*kutū > *ut* ‘louse’

*mataq > *m-mat* (assuming prior loss of *q)

may be extended to, PCEMP *V_{final σ} > IRH Ø (as a targeted deletion)

*lanjit > *ragt* ‘sky’

but, PCEMP *V_{final σ} > IRH 'V (syncopated forms)

*kulit > *rit* ‘skin’

The last of the distinctive innovations, Change 12) ‘Syncope’, is referred to specifically by Blust (1978), and in passing by Voorhoeve (1989), who considers it relatively strong evidence that Irarutu belongs to the SHWNG group. However, syncope in Irarutu was discussed at some length in §4.1, where it was shown that for this language, it is a context-free, possibly even strata-specific, change in Irarutu. It is only supported by eight Irarutu words in **Table 4.1**, a little more than 10% of the cognate data. In addition,

because *a > ie, is a highly marked form, *mtie* stands apart from the rest of the comparable data.

PCEMP *V_{penult} > PSHWNG and IRH Ø

*bulu ‘feather’ > fru ‘hair (i.e., body covering)’

*mata > *mtie* ‘eye’

*qatelur > *tru* ‘egg’

PEMP *natu ‘child’; IRH *ntü* ‘small, young’

The last innovation cited in **Table 4.12**, ‘Apocope 2’, i.e., truncate final diphthongs, is not reflected in Irarutu without raising in the diphthongal nucleus, that is, guna assimilation (see discussion in §4.1), but it is observed in several CMP languages. This distribution suggests quite strongly that Irarutus had little to no contact with CMP languages, such as those that are located at the opposite corners of the Bomberai peninsula.

4.2.1.3.3 A Note on SHWNG as a dialect chain

In addition to the historical correspondences discussed above, Blust (1978) noticed that SHWNG appears to have been a dialect chain. He observed that *k is retained in the SH languages at the Western edge of the subgroup, whereas *-V is retained at the eastern end of WNG. Based on these observations he proposed that *k > Ø (Change 5), and *-V > Ø (Change 11), started at East and West extremities and spread through SHWNG in opposite directions, affecting all languages except those in the extreme periphery, which reflect only one of the changes. In regards to Irarutu, PCEMP *manuk > *man* ‘bird’, suggests that both changes occurred, either simultaneously, or Change 5) before Change 11). This observation can be used to account for several instances of ‘linear deletion’, which, due to its central location, is compatible with the dialect chain hypothesis. It can be noted that Kuri retains *manuk > *manuk* ‘bird’, which suggests that the language was either not subjected to this change (Kuri is described as socially more isolated than Irarutu, see Chapter 1), or more likely that ‘linear deletion’ in Irarutu was independent of similar changes that took place in SHWNG languages (cf. Biak *man* ‘bird’).

4.2.1.4 Summary: the position of Irarutu

Heretofore, there has been no general agreement regarding the position of Irarutu in the Austronesian language family. Anceaux (1961) proposed that it is an isolate that does not group closely with his ‘Sarera Bay’ group (WNG); Grace (1955–6) speculated that it is a CMP language; Blust proposed that it is an isolate within the WNG branch of the SHWNG subgroup; Voorhoeve (1989) proposed that it was an early off-shoot from PSHWNG; Ross (1995) classified it with the SH languages; and most recently, van den Berg and Matsumura (2008) refer to it under the general heading of SHWNG. Jackson (2008) pointed out the need for better descriptions of the synchronic phonology and grammar of Irarutu, and reliance on primary data (cf. Klamer 2007), in order to justify any firm statement about the genetic affiliation of the language.

Two diagnostic and distinctive changes function as initial grounds for evaluation. Change 2) is a vowel shift of penultimate PMP *e > o. The general change of penultimate PMP *e [ə] (> PCEMP ə) > SHWNG o is supported by ‘labialized mid

vowels' such as **telu* > *tor* 'three', and **tətək* > *tot* 'cut, hack' (cited from the literature), but there are notable exceptions (see §4.1.1.4). Change 6), a consonant merger labeled 'rhoticization' can be amended to PCEMP **d*, **z*, **l*, **R* > *r* (no examples of **r* can be brought to bear on the evaluation). This innovation is overall well supported, except that **z* > *r* is attested in only one form (and is also counterbalanced by **z* > *s*). The exceptions to **d* > *r* are explained by environmental conditioning, following a homorganic nasal. All known instances of **l* and **R* > *r* except in final position, under FCD. Based on these two changes, sporadic shift of **e* > *o* and fairly strong 'rhoticization', Irarutu can be said to more closely resemble the WNG languages than the SH languages; however, the evidence is not convincing enough to claim that Irarutu is a SHWNG language. For example, 'rhoticization' could be due to borrowing or areal diffusion (Matsumura and Matsumura report that East Arguni dialect has independently developed *l*), and **e* > *o* belongs to a specific lexical strata.

Two diagnostic changes are helpful as secondary criteria. Change 4), **s*, **c* and **t* merging as *s* before **i* alongside medial **j* and **s* merging as *s*. The conditioned portion of this merger is unattested in Irarutu data (see also Voorhoeve 1989). In one word that has two instances of **s*, it is retained. This could be a borrowing. There is one reliable cognate which shows that medial **j* did become *s*, **pajey* > *fas* 'rice', and one pseudo-cognate, where the same change also occurs, but this is putative, **ŋajan* > *snan* 'name/word'. The vowel merger, Change 12) 'apocope', is supported by about half of the Irarutu data in §4.1 (36/67 cognates). This is the best evidence that Irarutu is a SHWNG language, but it has also been reported as an areal feature of languages in the Bomberai peninsula (Anceaux 1958). Based on 'diagnostic' changes, the evidence is a little stronger that Irarutu is a SHWNG language, but there is still considerable room for speculation.

The third test is presence of two distinctive innovations, both of which are vowel mergers. The details of Change 9) are fairly complicated but support grouping with WNG, rather than SH, languages. Unfortunately, as a categorical innovation, the data fails to support a merger of **a* and final **e*. This change is therefore not convincing evidence for status as a SHWNG language. The second distinctive merger, Change 12) 'syncope' is supported by about a sixth of the items in section §4.1 and is therefore recurrent. However, again, this process is known as an areal trait in the Bomberai peninsula (Anceaux 1958), and is also witnessed in some CMP languages (Blust 2009a and Donohue and Grimes 2008). Furthermore, syncope is overshadowed by apocope and therefore can be attributed to a specific lexical strata rather than to a general process.

The remaining seven changes are neither diagnostic nor distinctive. Therefore they are not persuasive forms of evidence that Irarutu is, or is not, a SHWNG language.

To account for the heavily mixed correlations with all of the criteria for classifying Irarutu as a member in the SHWNG subgroup, and the pervasiveness of exceptions to recurrent, context-free sound correspondences (§4.1), two contrasting positions must be compared. On the one hand, if Irarutu is asserted to be a SHWNG language, to avoid reversals (e.g., PMP **e* [ə] > PCEMP **e* > IRH *o*) and irregularity in reflexes (PCEMP **ə*

> *i, e, o, a*) in reconstructed forms, it must be proposed that Irarutu borrowed words from an unknown source. On the other hand, and more realistically, it can be proposed that Irarutu represents a primary branch of CEMP because it reflects pre-PEMP innovations, in which case Irarutu speakers may have borrowed a significant number of items from some nearby SHWNG language, or languages. This is more consistent with the data presented above, where the two recurring statements in the discussion have been 1) Irarutu appears to be more similar to WNG languages than to SH languages, which supports the conclusions of Anceaux, Blust, and Voorhoeve, and 2) conflicting sound correspondences appear to be due to lexical strata, which makes the second stance most easily reconcilable with Grace's proposition.

Although Irarutu has been classified as an Austronesian language, a careful look at the data raises the possibility that there are actually three options for classifying Irarutu. That is, Irarutu could be Austronesian, non-Austronesian (potentially Trans-New Guinea), or a contact language variety (creole or mixed language). The best reason to classify Irarutu as an Austronesian language is that the only reliable source for etymologies in the language points to the Austronesian language family, including formal innovations, semantic innovations, and lexical innovations. The reason to classify Irarutu as a non-Austronesian language, i.e., to propose that it is, for example, a Trans-New-Guinea language, is that more than three out of five words have etymologies of unknown provenance (which is quite likely due to borrowing), or on the other hand, that it is some type of contact language variety, because the grammar is relatively simple (Anceaux 1958), which is typical to a certain extent in language contact varieties.

More important than cognate count or typological features, it is important to consider the socio-cultural setting of the language. **Table 4.13** contrasts two widely acknowledged types of settings that lead to two systemically different contact language varieties: creoles and bilingual mixed languages (Appel and Muysken 1987; Baker 2000; Bickerton 1975; Mufwene 2001; Sebba 1997; and Thomason 2001, 2003).

<u>Pidgins/Creoles</u>	<u>Bilingual Mixed Languages</u>
usually >3 languages in contact	2 languages in contact
intergroup: lingua franca	intragroup: not lingua franca
imperfect learning	bilingual
grammatical simplification	little to no grammatical simplification
most words from 1 language	mix of words from source languages

TABLE 4.13 – Two generalized social settings and linguistic correlates attributed to the origins of two different types of contact language varieties.

As is typical for Papuan languages versus An languages, when Irarutu is compared typologically to its neighbors, Irarutu is described as being simpler ('easier to learn') than neighboring languages (Anceaux 1958). In relation to Mairasi (Trans-New Guinea phylum), both languages have prenasalized stops. Barau (TNG), to the west of Irarutu,

shares with Irarutu a strong bi-labial fricative, deletion of final vowels in casual speech, vowel length in syllables preceding a stressed syllable, and lexicalized stress in polysyllabic words (cf. T. Matsumura 1991). Tanah Merah (a.k.a Sebyar), also located northwest of Irarutu, shares with Irarutu long words with many open syllables that reduce to closed syllables and many consonant clusters in casual speech (see Chapter 2). Tanah Merah/Sebyar (TNG) verbs are morphologically complex. Irarutu verb complexes and verb phrases can also be fairly complex (see Chapter 3).

An areal feature of Papua is the rather extensive use of periphrastic constructions for body part terms, such as Irarutu *ruguin fru* ‘head’s feather’ for ‘hair’, *mcerum* ‘eye’s house’ for ‘face’, *tgra ro* ‘ear’s leaf’ for the external part of the ear, as opposed to ‘ear hole’ (this example is also found in Numbami, according to Joel Bradshaw, p.c. 2014). Melanesian languages are also famous for not using decimal-based counting systems (Briley 1977, Lincoln 2012), but have widely divergent systems, ranging from binary systems to exotic tally systems. Irarutu has a quinary-based system (see Chapter 3). Furthermore, Irarutu and neighboring languages are described as having a mixture of Iroquois- and Hawaiian-type kin-term systems (Matsumura 1992 and 1997, N. Peckham 1981b, 1983). These factors suggest that language learning across the history of Irarutu speaking people was generally not ‘imperfect’ in the way one would expect for pidgins. Furthermore, multilingualism is the norm in Melanesia (Foley 1986, 1998), and several ‘mixed’ languages (cf. Capell 1976, Wurm 1982), which fall under the canopy term ‘contact language varieties’ (Baker 2000, Janda and Joseph 2003), have been reported in the area. In this context, ‘contact’ is used to refer to the social interaction between peoples of different (micro-)cultures, who speak different languages but have the need to negotiate meaning to communicate. Importantly, in the process of acquiring a second or third language, and for bi- or multi-lingual children, the relevant linguistic codes are processed in each mind of each speaker (Ricoeur 1981), which often leads to interference and influence between the patterns and systems of the various languages as each individual tackles the task of working out the boundaries of each system. Subsequently, these speakers interact with others who are in the same context, faced with similar acquisition and socialization tasks, so that publicly negotiated meaning (and generalized patterns of interference and confluence between the various codes) may involve code switching, code-mixing, and borrowings.

4.2.2 Words of unknown provenance and notes on borrowings

This section discusses ninety-seven items from the 200-word list that cannot be shown to resemble any known Austronesian source, as well as several instances of attested borrowing/forms shared with nearby Austronesian languages through borrowing.

4.2.2.1 Words of unknown provenance

Table 4.14, below, lists words from Appendix 3a that have no known connection with Austronesian etyma. The nine ‘faux cognates’ from **Table 4.11** probably also belong in this category. The status of the thirteen ‘pseudo-cognates’ from **Table 4.10** is indeterminate.

No.	IRH	Gloss			
2.	trir	'left side'	104.	mtir	'fat, oil'
13.	fgye	'back'	105.	sy	'tail'
14.	fta	'belly'	106.	sawat	'snake'
16.	fta-ntü	'guts'	107.	seje	'(earth)worm'
23.	wams	'blood'	109.	bnisr	'mosquito'
25.	drbon	'neck'	111.	sum	'fish'
26.	drbun-fru	'hair'	112.	rumir	'rotten'
27.	wegur	'nose'	118.	samwin	'grass'
28.	sagwei	'breathe'	119.	mifr	'earth/soil'
29.	—	'smell'	120.	kami	'stone'
30.	rimta	'mouth'	121.	enyefu	'sand'
35.	—	'vomit'	124.	wer-fun	'sea, saltwater'
36.	wer-rimta	'spit'	125.	terir	'salt'
38.	gor	'chew'	126.	weto	'lake'
42.	—	'suck'	127.	witu	'forest'
46.	si	'see'	129.	syeba	'moon'
47.	—	'yawn'	130.	tür	'star'
49.	—	'lie down'	132.	fi-fras	'fog; mist'
56.	mo	'child'	134.	kararu	'thunder'
57.	waman	'husband'	135.	nffe	'lightning'
60.	-jie	'father'	138.	win	'warm/hot'
61.	san	'house'	140.	mmua	'dry'
63.	nü	'name'	141.	srwer	'wet'
66.	srit	'to tie'	142.	rimn	'heavy'
67.	nut	'sew'	144.	sfrer	'burn'
68.	raini	'needle'	146.	tugwan	'ash'
69.	fruft	'to hunt'	147.	grmutn	'black'
70.	ntun	'shoot (arrow)'	149.	krruer/ wams	'red' 'blood/red'
71.	ntaf	'stab'	150.	taf	'yellow'
72.	dbe	'hit; beat'	152.	kokon/kosi	'small'
73.	—	'steal'	153.	bidu	'big'
76.	rga	'living, alive'	154.	fut	'short (height)'
77.	skur	'scratch (an itch)'	155.	mrro	'long (objects)'
78.	ntifn	'cut (wood)'	157.	sarf	'thick'
80.	nagt	'split'	158.	sesut	'narrow'
81.	mdedan	'sharp'	159.	frifr	'wide'
82.	pum	'dull, blunt'	160.	ndridn	'sick'
83.	wedud	'work (in garden)'	161.	(rit dir ti)	'shy; ashamed'
85.	—	'choose'	162.	nabrid	'old (people)'
88.	nufr	'squeeze'	163.	bunat	'new'
89.	nud	'hold (in fist)'	164.	dir	'good'
91.	mgür	'buy'	165.	fit	'bad'
92.	sgafn	'to open'	167.	gisie	'night'
94.	dru	'to throw'	170.	[NP + nia]	'when'
95.	mtit	'to fall'	171.	—	'hide'
96.	fun	'dog' (E. Indonesia area)	173.	ge	'at'
100.	frifr	'wing'	176.	—	'below'
101.	sur	'to fly'	178.	—	'that'
			179.	nbabr	'near'

180.	nero	‘far’	191.	timebe	‘all’
181.	ge neno	‘where?’	192.	-ir	‘and’
186.	e	‘you (pl.)’	193.	—	‘if’
188.	nia	‘what?’	194.	fnano	‘how?’
189.	gata	‘who’	196.	nif	‘count’
190.	—	‘other’	200.	gigti	‘four’

TABLE 4.14 – Ninety-seven Irarutu words without known etymology, numbered according to Appendix 3a. Irarutu terms for 12 items, identified with ‘—’ need further research.

At least one word, *fun* ‘dog’ is found in Eastern Indonesia, but is not related to PCEMP. Four compounds and one phrase are composed of one Austronesian etymon and one or more roots of unknown provenance, i.e., *fta ntü* ‘guts’, *drbun fru* ‘hair’, *wer rimta* ‘spit’, *wer fun* ‘saltwater’, and *rit dir ti* ‘ashamed’ (Austronesian etyma are **boldfaced**).

The literature suggests that Mairasi to the southeast of Irarutu and Tanah Merah (a.k.a. Sebyar) to the northwest of Irarutu could be sources for the vocabulary that is not identifiable with an Austronesian source (Voorhoeve 1989, who calls these types of items ‘non-Austronesian’). For example, Voorhoeve speculated that *sot* (Voorhoeve’s orthography) ‘woman/female’, could be from a form reflected as Kamrau *yo:da*; Sabakor *yawoya*; Kamoro *kaoka*; Asmat *toot*, *cowoc* ‘woman’, but did not cite – or propose – a reconstructed form. All of these languages belong to the Trans-New-Guinea phylum, but see Pawley (1998), Ross (2005); Wurm (1975a, 1975b), and Wurm et al (1975) for more on historical reconstruction and the validity of this language family. Several of the consultants for the present research agreed that there are words from Mairasi in Irarutu, but no examples were cited.

	An	Other	Totals	An/Totals
Nouns	35	48	83	41% An
Verbs	21	30	51	41% An
Question words		5	5	0% An
Personal pronouns	6	1	7	85% An
Adverbs			--	--
(Adjectives)	4	24	28	14% An
Adpositions	1	3	4	.25% An
Coordinating Conj		1?	1	0% An
Numerals	3	1	4	75% An
Demonstratives	1	3	4	25% An
Negator	1?		1	100% An?
Totals	72	116	188	38%

TABLE 4.15 – Distribution of words with Austronesian versus unknown etymologies in ten syntactic categories.

To evaluate if words from specific parts-of-speech can be traced to an Austronesian source, a survey of distributions across various syntactic categories is summarized in

TABLE 4.15. Austronesian etyma in Irarutu are basically evenly distributed across the ten categories, and constitute about 40% of each. The concept of ‘basic vocabulary’ (in the sense of Blust 1991), that is, terms that are believed to be shared by all languages due to their fundamental nature (but see Ricoeur 1981) and are borrowed much less frequently than non-basic vocabulary, is limited to nouns, verbs, and adjectives, so it fails to explain the attested distributions.

4.2.2.2 Specific borrowings

Assuming the validity of the claim that Irarutu is indeed an Austronesian language, there are two possible ways to explain the presence of nearly one hundred items in **TABLE 4.14**: inventions (cf. Blust 1981) or borrowings. An adequate survey of neighboring languages as possible sources for borrowed vocabulary is not within the scope of this work, but some comments are in order.

Wandamen (An), Irarutu’s northeast neighbor is a likely source for some of the ‘pseudo-cognates’ and ‘faux cognates’. Unfortunately, none of the items cited up to this point have been confirmed as borrowings from Wandamen; but, Wandamen *ayai* ‘my father’ – distinct from *tama-mui* ‘your father’ and *tama-ni* ‘his father’ (forms from Anceaux 1961) – appears to have a stem *-ya-i* that matches in a fairly straight-forward way with Irarutu *a-jie* ‘my father’, and based on two semi-recurrent processes, glide fortition and vowel breaking, reflects a historically prior (possibly PEMP) form **ya* ‘father’. Flaming (1983a) points out that Wandamen and Irarutu share the word *arai* ‘same-sex sibling’. The ABVD cites correlation between Irarutu *samwine* ‘grass’ and Wandamen/Windesi *samuen* ‘grass’. Another item that shows the influence of Wandamen on its neighbors is the paraphrase *dian kariria* ‘crocodile (*lit.* bad fish)’, which correlates with the simplified Kuri (dialect of Irarutu) form *diankara* ‘crocodile’ (data cited from Smits and Voorhoeve 1992), whereas the Fruata dialect has *sieri~ceri* ‘crocodile’.

A little further afield, Irarutu has words that look strikingly similar to several Austronesian languages in Eastern Cenderawasih bay, such as Serui-Laut, Anus, Marau, Pom, Papuma, and Woi *kami* ‘stone’ (cited in Anceaux 1961), which is identical with Irarutu *kami* ‘stone’. By contrast, Kuri has *obaf* ‘stone’, which appears to reflect PCEMP **batu*. Less convincing are similarities with nearby Central Malayo-Polynesian languages such as Sekar *manif* ‘laugh’ and Irarutu *brif* ‘laugh’, or Koiwai (Walker 1990) and Kisar *rur* ‘bone’ with Irarutu *rur* ‘bone’. Voorhoeve (1989) suggests that some degree of contact with Koiwai and other CMP languages is probable, but no signs of this are apparent in lexical items. The similarities Irarutu has to CMP languages are probably due to chance or shared retentions from PCEMP (cf. Koiwai *rur* ‘bone’).

One additional form, *wams* ‘blood/kin’, is worth a brief comment because it was noted by Tryon (1995), who mistakenly attributed it to Sanskrit through Indonesian *bangsa* ‘kin/clan’. However, this is somewhat problematic because the sound mappings between the Irarutu form and the Indonesian are less transparent than those between Irarutu and Sanskrit. De Casparis (1997) cites Sanskrit **vaṃśa* ‘tribe, nation’ > Indonesian *bangsa* ‘nation, group’. Despite the semantic shift that it would be necessary to assume (SKT ‘tribe, nation’ > IRH ‘blood (relation), kin group’), and a lack of evidence that Papua was ever Indianized, the sound correspondences between Sanskrit **vaṃśa* and Irarutu *wams*,

are worth keeping in mind because it provides a source for this element, possibly by way of an old trade language used in some historic trade network (see §4.3), thereby eliminating the need to posit it as invention. A connection between putative early loan words, such as SKT **vaṃśa* > IRH *wams* ‘blood/kin’, fairly recent loan words, such as IND *putar* > *putr* ‘turn’, plus the faux cognates identified above, suggests that aberrant sound correspondences paired with shared semantics indicate borrowed vocabulary. The largest challenge for future research in Irarutu historical linguistics is to identify the sources of the remaining items.

4.3 Discussion

A difficulty in classifying Irarutu is caused by its geographical location between CMP languages, such as Koiwai to the South, as well as Onin, Sekar, and Uruangnirin to the West, and SHWNG languages, such as Wandamen, to the Northeast and East. This section synthesizes linguistic information with archeological and genetic information to explain Irarutu’s traditional location. The linguistic information includes lexical material, grammatical properties, oral literature (Miedema 1997), and the social and cultural context of the language. The archeological information (Bellwood 1998, Hope and Haberle 2005, Spriggs 1998) includes what we know about migration as a structural process based on Anthony (1990), and geographical information. Genetic information plays only a peripheral role in this discussion.

Anthony (1990) proposed five more-or-less distinct properties for long-distance migration: 1) leap-frogging, whereby great distances are crossed and large areas are bypassed; 2) well-defined routes with specific destinations and a highly restricted point of origin; 3) return voyages, especially when conditions at the goal and source become similar; 4) migration frequency, which implies that migration constitutes a self-propagating cultural tradition; and, 5) demographic skewing towards men.

It is important to reiterate that the Irarutu community is located in three general geographic environments: 1) on the North and South coasts of the Bomberai peninsula where there are swampy, coastal, and riverine areas, which are suitable for cultivating sago and fishing, 2) in the interior of the Bomberai peninsula, where there are foothills and small valleys, which are suitable for gardening, hunting and pig raising, and 3) on the nearby shores of Cenderawasih Bay where there are coastal lowland areas, which are well-suited for gardening, fishing, and cultivating tree crops (M. Walker 1990). In the present context of increasing globalization, each of these areas remains relatively isolated, but Irarutus have a history of extensive interaction with neighboring communities.

It is along the shores of Cenderawasih Bay that conditions are most appealing for human settlement at any point in time. The geographic features up to Manokwari and over the top of the Bird’s head are considerably less hospitable. This property aligns well with Anthony’s concept of ‘leap-frogging’, and suggests a route with a destination. Blust (1993) suggests that both PCEMP and PEMP took a considerable time to build-up before debouching, and that PCMP spread rapidly from the Northern Moluccas through Eastern Indonesia’s Lesser Sundas. It is therefore expected that at the time when Austronesians

expanded into New Guinea, their origin was somewhere around Northern Halmahera. It is also worth noting that Irarutu has at least one term, *kka* ‘bathe/swim’ that is shared with Oceanic languages, cf. POc *kakaRu, and a few other SHWNG languages, but is unknown outside EMP. The distribution of this form, and corroborating non-linguistic evidence, seems to suggest that PEMP was somewhere located in Cenderawasih Bay.

As for a specific route into New Guinea, the origin for migration into the island of New Guinea must be attributed to the location of PCEMP, which could have been nearly anywhere in the Northern Molukkas, but seems likely to have been near, if not on, Halmahera. Furthermore, the conjecture about rate of expansion entails migration as a cultural tradition. It may not be valid to generalize from Micronesian culture to PCEMP culture, but most sea-voyaging and navigation within Micronesian societies (cf. Marck 1986) is carried out by men. The most challenging aspect of Anthony’s structural migration theory to verify for the Austronesian settlement of New Guinea is back-migration in the early stages of their expansion out of the small islands of what is now Eastern Indonesia, a point to which we return shortly.

According to a variety of sources (Lynch, Ross, and Crowley 2002; Miedema, Odé, and Dam 1998; Pawley, Attenborough, Golson and Hide 2005), Austronesian-speaking people navigated around the Bird’s Head into the serene Cenderawasih Bay around 3500 years before the present era (BPE). Based on inference using the information that is available from community members’ local and traditional knowledge, combined with general information about the history of SHWNG languages and geographical features, the point of entry for the Austronesian speaking people probably was the area now known as Tandia (Cenderawasih coast, East of Kuri).

An interesting feature of New Guinea, and possibly Melanesia as a whole, is that cultural features have persevered with relatively minor and only superficial modification for several millennia. Assuming that the basic socio-political structure of Western Melanesia has also remained relatively intact, it seems probable that the Austronesians promoted amicable relations with the pre-existing people and co-habited with one or more unrelated ethnic groups. Foster (1973) comments that the Austronesians mixed extensively with the ‘Papuan’, a claim which is supported in genetic studies such as Kim et al. 2012, but maintained their own languages. At least for the Irarutus, a lifestyle of co-habitation and intermingling with pre-existing populations is supported by two cultural traits. Firstly, aggression, in a Western sense, has negative connotations in Irarutu. For example, ‘strong, brave, and warrior’ as well as ‘bad and evil’ are all part of the meaning of *fit*, which suggests that non-warlike behavior drives social structure. Secondly, between the two types of ‘big-man’ proposed by Sahlins (1963), one who excels at battle and one who excels in commerce, Irarutu favors the economic big-man, although the warlord big-man does play a role – especially in the centuries old cultural phenomenon ‘honggi’, in which a stronger group takes tribute in the form of humans and materials from weaker groups to increase the ‘fertility’ of the stronger group (see Haenen 1998, Koentjaraningrat 1994, Mansoben and Walker 1990, and Goodman 1998 for more on the subject). These two examples support the idea that early Irarutus probably tried to blend-in with their neighbors and seek peaceful resolution of conflict rather than battle.

Foster (1973) also concluded that the only endemic food crops on the island of New Guinea were bananas and sugarcane. Other crops, such as taro, cassava, and coconut originate from other sources, and taro and cassava were probably introduced by people. It is hypothesized here that the early Austronesians traded with indigenous peoples in exchange for permission to use raw materials, and to co-habit the land. A by-product of co-habitation was admixture of Austronesians with the people of New Guinea; but, due to differing language ideologies, the various languages persisted, perhaps through alternate forms of language use such as ‘dual lingualism’ (Lincoln 1976).

Blust (1978) was successful in uniting the SH languages with the WNG languages and indicated that the greater diversity of SHWNG languages in Cenderawasih Bay is explained by the principle of greatest diversity, i.e., the area that has the most diversity has been in place the longest. The South Halmahera languages are relatively homogeneous (see **Figure 4.5**).

Concerning the movement of Irarutus from the Cenderawasih Bay side of the Bird’s Neck region to their present location, Voorhoeve thought that several of the changes he noticed (presented in §4.1.4, above) could be attributed to contact with non-Austronesian languages as Irarutu moved into the Bomberai Peninsula. He comments that Irarutu’s geographical location suggests that the language entered from the Northeast and wedged itself between Tanah Merah and Mairasi (also indicated in Wurm and Hattori 1981). This hypothesis is corroborated by several forms of meta-linguistic and non-linguistic evidence, such as traditional beliefs, songs, and cultural traits, all of which point to the Nabi River (in the northeast area of the present language region) as their area of origin. Several Irarutu language consultants recounted that Irarutus migrated to their present location in search of means for subsistence.

Voorhoeve further speculates that Mer and Tanah Merah were located historically where Irarutu is today. This is supported by the genetic affiliation of Mer, Mairasi, and Tanah Merah (Sebyar). No forms of evidence point to early contact with Mor (Mer) or Baham. Sebyar arrived not more than a few centuries ago from the opposite bank of the McCluer Gulf. To the south, there is evidence of contact between Irarutu and Kamrau as well as Sabakor near the Asmat area (Southeast of **MAP 2**, see §1.1.2). For example, Kamrau and Sabakor share parallels in mythology and social organization with Irarutu, but this is probably due to mutual influence rather than common retention. Sabakor mythology also suggests that Sabakor people used to live further North, in the area that Mairasi is presently located. Coincidentally, Peckham (1982b) points out that *Mairasi* in the Mairasi language means ‘original’, a similarity to *Irarutu* ‘true language’.

From a linguistic standpoint, and more than any single historically derived feature, the existence of at least two lexical strata, the younger of which more closely typifies SHWNG languages (cf. syncope), suggests that Irarutu subgroups above SHWNG. Donohue and Grimes’ (2008) portrait of the boundary between WMP languages and so-called CMP languages, in which overlapping distributions of speakers resulted in diffusion of features due to contact, appears to help explain at least part of the history of the Irarutu language. Various waves of migrants settling in the Cenderawasih Bay area

alternated in carrying prestige. During each period of time when Irarutu carried less prestige than some WNG language in which Irarutu speakers were in contact with, Irarutu became susceptible to borrowing lexical items from the higher prestige language, see §4.2.2. When compared with the distribution and status of surrounding Austronesian languages that are located on the edges of the Bomberai peninsula, such as the trade language Onin, which is thought to have been a powerful force in the ‘raja’ system of social structure (see for example Goodman 1998), Koiwai to the South of Arguni Bay, and Wandamen, which currently carries social prestige, the pressure for Irarutu to borrow words from different sources is statistically reinforced (see also Thompson 2001).

4.4 Summary

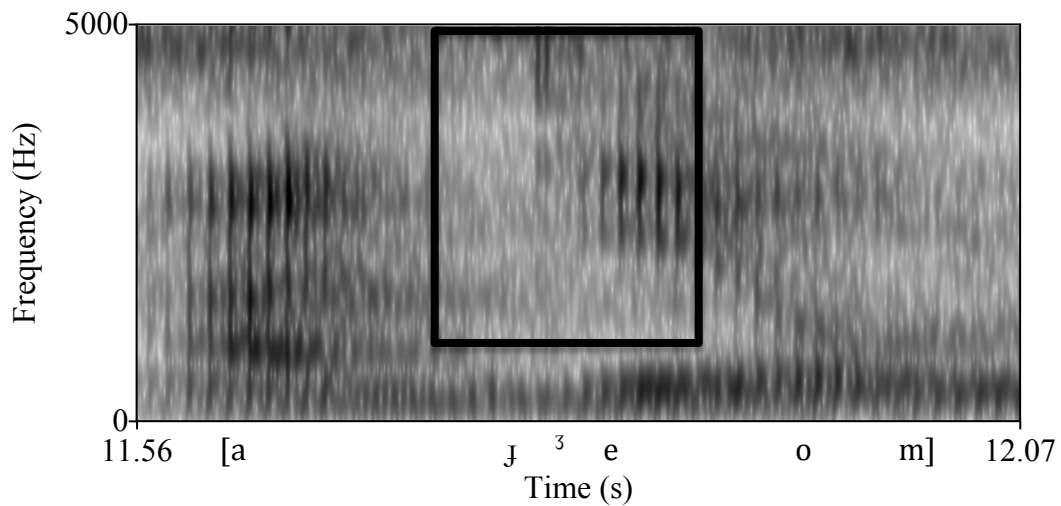
The discussion above utilized lexical comparison combined with the synchronic description in Chapters 2–3 to aid and contextualize historical inquiry. In the first section, a substantial sample set of words was associated with specific Austronesian etyma. Evidence from two competing, but conflicting changes in reconstructed disyllables, i.e., syncope and apocope, in corroboration with stress shift and reflexes of *ə, indicates the existence of two lexical strata within the Austronesian component in Irarutu. In the second section, a range of innovations were evaluated in relation to various criteria used to determine membership in several Austronesian subgroups (MP, CEMP, EMP, and SHWNG), which, taken as a whole, suggest Irarutu belongs to a proto-language that predates SHWNG (in accord with Grace 1955–6). In the third section, the linguistic data was corroborated with archeological and genetic information to briefly explore the history of the Irarutu people. The main directions for future research on the history of Irarutu are: survey neighboring languages for shared lexical items and develop a more detailed understanding of the grammar of these languages as possible sources for ‘transfer’. These tasks will pave the road into a better understanding of Irarutu dialectology and West Papuan languages in general.

The background information, phonological and grammatical descriptions above, supplemented by historical inquiry, lay a solid foundation for subsequent research in a variety of topics in the Irarutu language and the development of conservation materials for it.

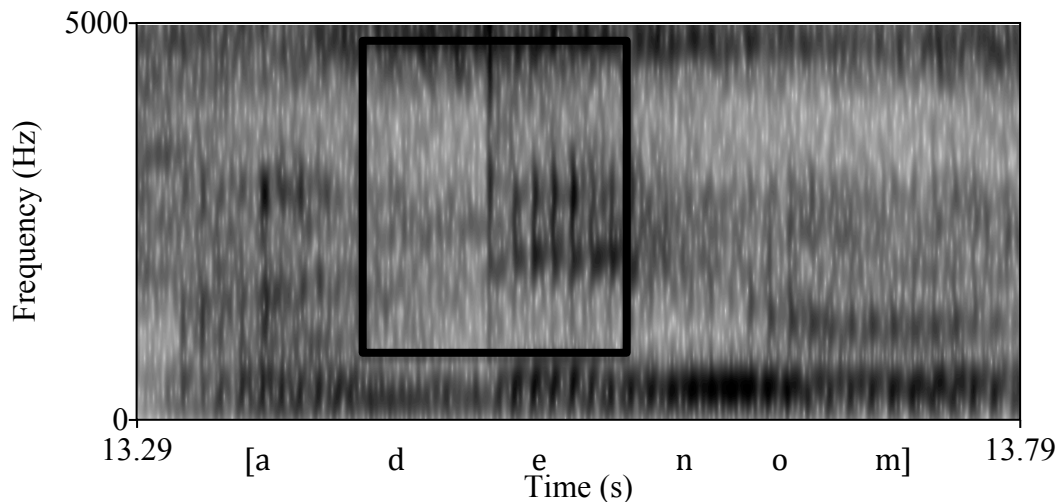
Appendix 1a – Spectrograms of Irarutu speech

The following pair of spectrograms 1–2) shows the [ʒ]-like fricated release portion of a palatal stop /j/ in comparison with the release of an alveolar voiced stop /d/. The items are pronounced using the ‘Kaimana’ (South Arguni) dialect that does not have prenasalized voiced stops.

- 1) [ajeom] ‘my father...’ from *aje o mtag* ‘my father, you cry’



- 2) [adenom] ‘my mother...’ from *aden o mtag* ‘my mother, you cry’



Appendix 1b - Irarutu Feature Chart

	y	ɪ	i	e	a	o	u	w	r	m	n	f	s	h	b	d	ʝ	g	p	t	c	k
syll	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
approx	–	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–
cons	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+
son	+	+	+	+	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–	–	–	–
cont	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–
nas	–	–	–	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–
tense	–	–	+	+	–	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
hi	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
low	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
lab	+	–	–	–	–	+	+	+	–	+	–	+	–	–	+	–	–	–	+	–	–	–
pal	+	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	+	–
cor	–	–	–	–	–	–	–	–	+	–	+	–	+	–	–	+	–	–	–	+	–	–
dors	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	+	–	–	–	+
voiced	+	+	+	+	+	+	+	+	+	+	+	–	–	–	+	+	+	+	–	–	–	–

Appendix 2 – Example texts

The following two examples of short, scripted texts are included here to supplement the description of Irarutu syntax and to provide a basis for future research into the structure of Irarutu discourse as well as to boost literacy in Irarutu. These examples were selected because they represent central themes for a number of documentation events/recording sessions. Slightly different versions of each example of discourse were documented repeatedly (see Appendix 4). Furthermore, each of the examples presented below was scripted by their authors after being created spontaneously. The texts provide a first step towards developing Irarutu pedagogical materials, which is why Indonesian translations (the underlined tier) are also included. The orthography is a mixture of native speaker conventions and those used in the syntax Chapter (Ch. 3), above.

The first example text, ‘A Brief husband and wife conversation’, is a typical domestic exchange centered around drinking coffee. It provides a glimpse into Irarutu gender norms and social relationships maintained through language. The second example text, ‘Preparing taro’, presents instructions on how to make pounded taro-with-cassava, which, in addition to papeda (see Chapter 1), is a main starch for the Irarutu people. It is typically made in large batches using traditional implements: *ndartu* ‘a large wooden bowl made from the trunk of a rubber tree for pounding taro’, and *titaru* ‘a long wooden stick, roughly the size of a baseball bat, made for pounding taro’. Individual servings of this food are wrapped in a banana leaf packet, which allows it to remain edible for several days without refrigeration.

Example text 1 – ‘A Brief husband and wife conversation (Percakapan suami-istri sehari-hari)’

by Markus Sefire (M) and Aligonda Nimbafu (A)
Associated documentation data folder: JJ-003
File: 130321-004 [3’29”]

M: *re-fefa ndir*
day-leaning good
‘good afternoon’
selamat sore

M: *o=ro mbun-nya fi mo-jie=e?*
2=TOP 2;do-something ACT IRR-complete=pol
‘what are you doing?’
ko bikin apa?

A: *ja mbsyu m-tür fene*
1S just STAT-sit only
‘I’m just sitting’
saya duduk saja

A: *ja fun-nya nyati*
 1S do-something nothing
 ‘I’m not doing anything’
tidak ada apa

M: *na-jiro m-ro mbengguen wer+win nene termus andi=ro*
 COND-EXPL 2-go 2;look water+hot PREP thermos DET=TOP
 ‘If that’s the case, go look if there is hot water in the thermos.’
Kalau begitu, pergi lihat kalau ada air panas dalam termos itu

A: *termus andi=mo wer+win=ti=o*
 thermos DET=DISJ water.hot=NEG=POL
 ‘There’s no hot water in the thermos.’
Tidak ada air panas dalam termos di sana

M: *na-jiro m-tün wer+win mbunat nirngge ngga-nggin kopi fande*
 COND-EXPL 2-cook water+hot new because 1;DES-SAP;drink coffee just
 ‘If that’s the case, heat some fresh hot water because I want to drink coffee.’
Kalau begitu, masak air panas karena saya mau minum kopi saja.

Example text 2 – ‘*Wanggt ssi* (Pounding taro Proses keladi tumbuk)’
 by Englebert Kofiaga

Associated documentation folder: JJ-003
 File: not recorded (but see 130328-002 [3’10’], 130328-004 [4’48’], etc.)
 Notes: summarizes contents of ‘making taro-with-cassava’; written in Iraputu with Indonesian translation

1: *guf wanggt and-i*
 peel taro DEF-DET
 ‘peel the taro’
kupas keladi/kuliti keladi

2: *mbr wanggt and-i mbr n-fa kurni*
 2;CAUSE taro DEF-DET 2;ACT 3-go pot
 ‘put the taro in a pot’
masukan keladi di dalam belanga

3: *mbu m-tün and-i*
 PROG STAT-cook DEF-DET
 ‘begin to cook it’
mulai masak

4: *and mbu n-krkar ro-ngge mse n-ar and-i met ro-ngge m-nggrffar and*
 DEF PROG 3-boil COP-COMP until 3-give DEF-DET little COP-COMP 2-turn DEF
 ‘if it has boiled for a long time, stir it’
kalau sudah mendidi agak lama dibolak-balik

5: *n-fa wanggt and-i n-ir sirnyef uf ro-ngge m-ssi sirnyef and-i*
 3-go taro DEF-DET 3-with cassava REL COP-COMP STAT-pound cassava DEF-DET
 ‘if it is already cooked, take it out of the heat’ (put the taro with the cassava to use)
kalau sudah masak dikeluarkan dari api

mbi ro-f se-mbu su wanggt
 2-ACT go-CMP ?-PROG CPL taro
 ‘begin pounding the taro’
dan mulai ditumbuk

6a: *n-fa and mbu n-m-besi rongge mbu su and nene fina se-mbu me-sesi.*
 3-go DET PROG 3-stat-pound COP. PROG CPL DET PREP fire ?-PROG STAT-meet
 ‘if it is already cooked, take it out of the heat’
kalau sudah masak/matang di keluarkan dari api dan mulai ditumbuk.

me-sesi warajef andi mbi-rot se-mbu su wanggt and-i
 STAT-pound cassava DEF-DET 2;ACT-?mix ?-PROG CMPLV taro DET-DEM
 ‘pound the cassava til it is smooth and follow with the taro’
kalau itu kasbi dan keladi berarti, kasbi ditumbuk lebih dulu sampai lumat baru menyusul keladi

6b: *me-sesi wanggt and-i n-ir sirnyef and-i namse and na-n-fi esu*
 STAT-pound taro DET-DEM 3-with cassava DET-DEM until DET INF-3-ACT one
 ‘pound the taro with the cassava until it becomes one (incorporated)’
tumbuk keladi campur kasbi tadi sampai kedua-duanya betul-betul menyatu

7: *n-fa and ndu serangg n rongge su-ma me-saft wesusun ruwe and-i*
 3-go DET PROG smooth COP CPL-join STAT-lift taro.water water DET-DEM

sefu mafr
 so.that easy
 ‘if it is not smooth, mix in a little hot water’
kalau belum lumat, dicampur sedikit dengang air panas bekas rebusan tadi

8: *n-fa and mbu n-fi jadi esu ro-ngge ande-j=ro*
 3-go DET PROG 3-ACT become one COP-COMP DET-DEM=TOP

ngge mbu siap ngge na
 COMP PROG ready 3;DES eat
 ‘if it is already mixed thoroughly, it is ready to be served and eaten.’
kalau sudah betul-betul adonan atau campuran tadi (kasbi dan keladi red) sudah

menyatu artinya siap juga di sajikan (di makan)

- *- *a-matu Jack, and-i ro catatan met fnote,*
1 SPOS-person J., DEF-DET COP note little merely,
'my friend Jack, this is a little note (how to make taro-with-cassava)'
teman Jek, ini cuma sekedar catatan kaki

sefu o mbi ingat-ingat / sefu o ma-m-run
so.that 2S 2;act remember / so.that 2S 2-2-know
'to help you remember'
untuk diingat-ingat saja

mo ja ra-run ngge tese o mbis.
CONJ 1S 1-know COMP surely 2S 2;can
'but I know that you can do it anyway'
tetapi prosesnya Jek ketahui saya yakin pasti sukses

APPENDIX 3a – Comparative 200-word lists

Comparative 200-word lists for PMP, PCEMP, IRH, and POc with glosses. This is an updated and slightly modified and updated version of the wordlist published by Blust (2008a), with the addition of Irarutu language data for the sake of comparison. As much as possible, cognates are listed on the same line. The plus symbol ‘+’ indicates a word in Irarutu is thought to be cognate with a proto-language, while those in italics are thought to be non-cognate. Forms with neither ‘+’ nor italics may be secondary cognates. The tilde ‘~’ indicates reduplication, whereas the hyphen ‘-’ indicates a morphological boundary.

No.	PMP	PCEMP	IRH	POc	Gloss (Notes)
1.	(qa)lima	lima	fra	lima	‘hand’
2.	ka-wiRi	ka-wiRi	trir	ma-wiRi	‘left side’
3.	ka-wanan/ ma-taqu	ka-wanan/ mataqu	—/ tu	wanan/ mataqu	‘right side’
4.	qaqay	waqay	fa+	waqe	‘leg, foot’
5.	lakaw/ panaw	lakaw/ panaw	—/ fa+	kalo/ pano	‘walk, go’ (PEMP *ba)
6.	zalan	zalan	rarn	salan	‘road’
7.	um-aRi	mai/ maRi	ma+	mai/ maRi	‘come’
8.	bilin/ puter	—/ putər	—/ putr	liu	‘turn’
9.	laɲuy/ naɲuy	—/ naɲuy	kka	kakaRu (?)	‘swim’ (PEMP *kakaRu?)
10.	cemeD (?)	ma-qetəm/ ma-qitəm	grmutn	ma-qitom/ ma-qetom	‘dirty’
11.	qabuk/ qapuk	—/ qapuk	bob	—/ qapuk	‘dust’
12.	kulit	kulit	rit+	kulit	‘skin’
13.	likuj	mudi	fgie	muri	‘back’
14.	tian	tian	fta	tian	‘belly’
15.	tuqela(n,ŋ)	zuRi	rur+	suRi	‘bone’
16.	tinaqi	tinaqi	fta ntu	tinaqi	‘guts’
17.	qatay	qatay	ti	qate	‘liver’
18.	susu	susu	sus+	susu	‘breast’
19.	qabaRa	qabaRa	fra fu+	(qa)paRa	‘shoulder’ (<i>fra</i> > ‘arm’)
20.	taqu	taqu	mrn	taqu	‘know’
21.	demdem	dəmdəm	derbun, mrn+	ronrom	‘think’ (<i>mrn</i> > ‘think’)
22.	ma-takut	ma-takut	mtat+	ma-takut	‘afraid’
23.	daRaq	daRaq	wams	nraRaq	‘blood’
24.	qulu	qulu	rū+	qulu	‘head’ (IRH: ‘source’)
25.	liqəR	liqəR	drbon	Ruqa/liqəR	‘neck’
26.	buhek	buk/ daun ni qulu	—/ drbun fru	—/ nraun ni qulu	‘hair’
27.	ijun	ijisuK	wegur	—	‘nose’
28.	ma-ñawa	ma-ñawa	sagwei	mañawa	‘breathe’
29.	hajek	hajək	—	ajok	‘smell’
30.	baqbaq	babaq	rimta	papaq/qawa	‘mouth’

31.	ipen/ nipen	ipən/ nipən/ lipən	—/ —/ rfo+	—/ nipon	‘tooth’
32.	dilaq	maya	maje+	maya	‘tongue’ (IRH m~maje)
33.	tawa	tawa/ malip	—/ brif+	—/ malip	‘laugh’
34.	tañis	tañis	tag+	tañis	‘cry’
35.	um-utaq	mutaq	—	mu-mutaq/ luaq	‘vomit’
36.	luzaq	(?)	wer rimta	qanusi	‘spit’
37.	kaen	kan	ga+	kani	‘eat’
38.	mamaq	mamaq	m-gor	mama	‘chew’
39.	tanek/ zakan	tanək/ zakan	tün	—/ sakan	‘cook’
40.	inum	inum	in+	inum	‘drink’
41.	kaRat	kaRat	-gor+	kaRat	‘bite’ (<i>m-gor</i> ‘bite’)
42.	sepsep	səsəp	—	sosop	‘suck’
43.	taliŋa	taliŋa	tgra+	taliŋa	‘ear’
44.	deŋeR	deŋəR	-nogr+	roŋoR	‘hear’ (<i>f-nogr</i> ‘hear’)
45.	mata	mata	mtie+	mata	‘eye’
46.	kita	kita	si	kita	‘see’
47.	ma-huab	mawap	—	mawap	‘yawn’
48.	tiduR/ tuduR	tiduR/ tuduR	in	tiruR/ turuR	‘sleep’
49.	qinep	qenəp/qinəp	in+	qenop/qinop	‘lie down’ (<i>in</i> > ‘sleep’)
50.	mipi/ nipi/ h-um-ipi	—	n-in si	—	‘dream’
51.	ma-tudan	tudan/ todan	mtür+	nopo	‘sit’
52.	ma-diRi/ tuqud	diRi/ tuqud	mrir+	—/ tuqur	‘stand’
53.	tau/ taumataq	tau/ taumataq	—/ matu+	tau/ matu+	‘person’
54.	laki/ ma-Ruqanay	laki/ ma-Ruqanay	mran+	maRuqane	‘male/man’
55.	bahi/ b-in-ahi/ ba-b-in-ahi	bai/ b-in-ay/ ba-b-in-ay	—/ sot, bfin+	—/ pine/ papine	‘woman’ (<i>bfin</i> > ‘wife’)
56.	anak	anak	mo, ntu+	natu	‘child’ (<i>ntu</i> ‘offspring’)
57.	bana/ qasawa	—/ qasawa	—/ wa-	—/ qasawa	‘husband’ (<i>wa-man</i> ‘husband’)
58.	qasawa	qasawa	wa-	qasawa	‘wife’ (<i>wa-bfin</i> ‘wife’)
59.	t-ina	t-ina	-den+	tina	‘mother’
60.	t-ama	t-ama	-jie	tama	‘father’
61.	Rumaq	Rumaq	san, rum+	Rumaq	‘house’ (<i>rum</i> > ‘village’)
62.	qatep	qatəp	atif+	qatop	‘thatch’ (<i>atif-ro</i> ‘thatch’)
63.	ŋajan	ŋajan	nü, snan+	ŋajan	‘name’ (<i>snan</i> > ‘word’)
64.	kaRi/ tutur	—/ tutur	fier	kunu(?)	‘say’
65.	talih	tali/waRəj	wara+	tali/waRoj	‘rope’

66.	hiket	hiket	srit	(?)	‘to tie’
67.	tahiq/ zaqit	—/ zaqit	nut	—/ saqit	‘sew’
68.	zaRum	zaRum	raini	saRum	‘needle’
69.	qanup	(?)	fruft	(?)	‘to hunt’
70.	panaq	panaq	ntün	panaq	‘shoot (arrow)’
71.	suksuk	susuk	ntaf	soka	‘stab’
72.	palu	palu	dbe	palu	‘hit; beat’
73.	takaw	takaw/ pa-nakaw	—	—/ panako	‘steal’
74.	bunuq	bunuq	-fun+	punuq	‘kill’ (<i>n-fun</i> ‘kill’)
75.	matay	matay	-mat+	mate	‘die (dead)’ (<i>n-mat</i> ‘dead’)
76.	ma-qudip	ma-qudip	rga	maqurip	‘living, alive’
77.	kaRaw	kaRaw	skur	karo	‘scratch (an itch)’
78.	taRaq/ tektek	taRaq/ tətək	ntifn	taRaq	‘cut (wood)’
79.	kahiw	kayu	e-+	kayu	‘wood’ (<i>e-ntu</i> ‘wood’)
80.	belaq/ silaq	bəlaq	nagt	polaq	‘split’
81.	ma-tazem/ ma-tazim	—/ ma-tazim	mdedan	—/ ma-tasim	‘sharp’
82.	pundul/ pudul/ dumpul	—/ —/ dumpul	pum	—/ purul	‘dull, blunt’
83.	quma	quma	wedud	quma	‘work (in garden)’
84.	tanem	tanəm	-tan+	tanom	‘to plant’ (<i>n-tan</i> ‘to plant’)
85.	piliq	piliq	—	piliq	‘choose’
86.	tu(m)buq	tumbuq	sub+	tumpuq	‘grow’
87.	baReq	baRəq	br-bar+	paroq	‘swell’
88.	pereq/ peRes Rames	—/ pəRəs/ Raməs	nufr	—/ poRos	‘squeeze’
89.	gemgem	gəgəm	-ud	kokom	‘hold (in fist)’
90.	kali	kali/keli	gar+	kali/keli	‘dig’
91.	beli	bəli	mgür	poli	‘buy’
92.	buka	buka	sgafn	puka	‘to open’
93.	bayu/ tuktuk	bayu/ tutuk	—/ -tut+	—/ tutuk	‘to pound’ (<i>m-tut</i> ‘to pound’)
94.	tudaq	tudaq	dru	(?)	‘to throw’
95.	nabuq	ka-nabuq/ ma-nabuq	mtit	(?)	‘to fall’
96.	asu	asu	fun	(?)	‘dog’
97.	manuk	manuk	man+	manuk	‘bird’
98.	qateluR/ qiteluR	qatəluR	tru+	qatoluR	‘egg’
99.	bulu	bulu	fru+	pulu	‘feather’
100.	kapak/ panij	kapak/ panij	frifr	—/ panij	‘wing’
101.	Rebek	Rəbək	sur	Ropok	‘to fly’
102.	labaw	labaw/ kanzupay	—/ sfe+	lapo/ kansupe	‘rat’ (<i>sfe</i> ‘mouse’)

103.	hesi/isi	isi	fi-je	pinsiko	‘meat, flesh’
104.	himaR/ menak/ mipak	—/ menak/ mipak	mtir —	—/ monak	‘fat, oil’
105.	ikuR	ikuR	su	siku	‘tail’
106.	nipay/ ulaR	nipay/ ulaR	sawat	mwata	‘snake’
107.	kalati/ qali-wati	kalati	seje	(?)	‘(earth)worm’
108.	kutu	kutu	ut+	kutu	‘head louse’
109.	namuk/ lamuk	namuk/ lamuk	bnisr	namuk/ lamuk	‘mosquito’
110.	lawaq	lawaq	rara, baru	lawaq	‘spider’
111.	hikan	hikan	sum	ikan	‘fish’
112.	buRuk/ busuk	buRuk/ busuk	rumir	mapu	‘rotten’
113.	daqan	daqan/ saṇan	ran+	raqan	‘branch’
114.	dahun	daun	-ro +	ndraun	‘leaf’ (<i>e-ro</i> ‘leaf’)
115.	uRat/ wakaR	—/ wakaR	—/ war+	—/ wakaR(a)	‘root’
116.	buṇa	buṇa	fū+	puṇa	‘flower’
117.	buaq	buaq	fu+	puaq	‘fruit’
118.	baliji/ udu	baliji/ udu	samwin	paliji	‘grass’
119.	tanaq/ taneq	tanaq/ tanəq	mifr	tanaq/ tanoq	‘earth/soil’ (cf. IRH san ‘house’)
120.	batu	batu	kami	patu	‘stone’
121.	qenay	qənay	errie-fu, enjefu	qone	‘sand’
122.	danum/	danum/	—/	danum/	‘water’
123.	wahiR/ aliR/ aluR/ saliR	waiR/ aliR/ saliR	wer + srswir	waiR/ aliR/ saliR	‘to flow’
124.	tasik	tasik	werfun	tasik	‘sea, saltwater’
125.	qasiRa/ timur	qasiRa/ tasik	terir	—/ tasik	‘salt’
126.	danaw	danaw	weto	rano	‘lake’
127.	halas	halas	witu	(?)	‘forest’
128.	laṇit	laṇit	ragt+	laṇit	‘sky’
129.	bulan	bulan	seba	pulan	‘moon’
130.	bituqen	bituqən	tur	pituqon	‘star’
131.	Rabun	Rabun/ taqe ni laṇit	mud/ taje	Rapun/ taqe ni laṇit	‘cloud’
132.	kabut	kabut	fi-fras	kaput	‘fog; mist’
133.	quzan	quzan	sem+	qusan	‘rain’
134.	gurgur/ kudug	gugur/ kudug	kararu	—/ kuruk	‘thunder’
135.	kilat/ qusilaq	kilat/ qusilaq	nffe	—/ qusilaq	‘lightning’

136.	hañin	hañin	nof	añin	‘wind’
137.	hiup	upi	uf+	upi	‘to blow’
138.	ma-panas	ma-panas	win	ma-panas	‘warm/hot’
139.	ma-diñdiñ	ma-dindiñ	dridn+	ma-dindiñ	‘cold’
140.	ma-Raṇaw	ma-Raṇaw	mmua	ma-Raño	‘dry’
141.	ma-baseq	basəq	srwer	(?)	‘wet’
142.	ma-beReqat	ma-bəRat	rimn	mapat	‘heavy’
143.	hapuy	hapuy	fina	api	‘fire’
144.	tunu	tunu	sfrrer, tün+	tunu	‘burn’ (<i>tyn</i> > ‘cook’)
145.	qebel/	—/	bob	—/	‘smoke’
	qasu	qasu		qasu	
146.	qabu	qabu	tugwan	qapu	‘ash’
147.	ma-qitem	ma-qetəm	grmutn	ma-qetom	‘black’
148.	ma-putiq/	ma-putiq	bfut+	—/	‘white’
	burak	burak		purak	
149.	ma-iRaq	meRaq	krruer, wams	(me-)meRaq	‘red’ (<i>wams</i> ‘blood/red’)
150.	ma-kunij	kunij	taf	año	‘yellow’
151.	mataq	mataq	-mat+	kesa (?)	‘green’ (<i>m-mat</i> ‘green’)
152.	dikiq/	dikiq/	kokon	rikiq	‘small’
	kedi	kedi			
153.	ma-Raya	Raya	bidi	lapa	‘big’
154.	ma-babaq	babaq	fut	papaq (?)	‘short (height)’
155.	anaduq	anaduq	mrro	anaruq	‘long (objects)’
156.	ma-nipis	ma-nipis	bnfn	manipis	‘thin (material)’
157.	ma-kapal	kapal/	sarf	—/	‘thick’
		telu		ma-tolu	
158.	kepit	kəpit	sesut	kopit	‘narrow’
159.	ma-labeR	labəR	-rifr+	(?)	‘wide’ (<i>f-rifr</i> ‘wide’)
160.	ma-sakit	ma-sakit	ndridn	masakit	‘sick’
161.	ma-hiaq	mayaq	rit dir ti	mayaq	‘shy; ashamed’
162.	ma-tuqah	ma-tuqa	nabrid	matuqa	‘old (people)’
163.	baqeRu	baqəRu	bunat	baqəRu/	‘new’
				bəqəRu	
164.	ma-pia	ma-pia/diaq	dir	ma-pia	‘good’
165.	zaqat	zaqat	fit	saqat	‘bad’
166.	ma-bener	bənər/	—/	mponor	‘correct, true’
		qutu	tü+		
167.	beRñi/	—/	gisie	—/	‘night’
	beRñin	bəñin		mponji	
168.	qalejaw	qaləjaw	re	qaso	‘day’ (IRH <i>r~re</i>)
169.	taqun	taqun	taun	taqun	‘year’
170.	p-ijan	p-ijan	—	ña-ijan	‘when’
171.	buni	buni	-funi+	puni	‘hide’ (<i>br-funi</i> ‘hide’)
172.	pa-nahik/	panaik/	iet	panek/	
	sakay	sakay		sake	‘climb’
173.	i/	i/	ge	i	‘at’
	di	di			
174.	dalem	daləm	garn	ralom	‘in, inside’
175.	a taqas/	atas/	ffu+	atas/	‘above’
	i babaw	i babaw		i papo	
176.	i babaq	i babaq	—	i papaq	‘below’

177.	i ni/ a ni	ini	-ini+	ni/ ani	‘this’
178.	i na	ina	—	ina	‘that’
179.	hazani	hazani/ raŋi	nbabr	—/ raŋi/tata	‘near’
180.	ma-zauq	zauq	nero	sauq	‘far’
181.	i nu	(?)	ge neno	pea	‘where?’
182.	i-aku	i-aku	ja+	aku	‘I’
183.	i-kahu	i-kau	o+	iko	‘you’
184.	si-ia	s-ia	i+	ia	‘he/she’
185.	i-k-ita/ kami	k-ita/ k-ami	it+/ am+	kita/ kami	‘we (incl.)’ ‘we (excl.)’
186.	i-kamu/ kamiu	—/ kamiu	e	—/ kamiu	‘you (pl.)’
187.	si ida	sida	ir+	sira	‘they’
188.	apa	apa/ sapa	nia	—/ sapa	‘what?’
189.	sai	sai/ sei	gata	sai/ sei	‘who’
190.	duma/ liqan	—/ liqan	—	—/ liqan	‘other’
191.	amin	(?)	timbe	(?)	‘all’
192.	ka/ ma	ka/ ma	-ir	ka/ ma	‘and’
193.	ka/ nu	ka/ ma	—	ka/ ma	‘if’
194.	kuja/ kua	—/ kua	fnano	kua	‘how?’
195.	qazi/ diaq	—/ diaq	—/ =ti	(?)	‘no, not’
196.	bilan/iap	—/ ihap	—/ nif	(?)	‘count’
197.	esa/ isa/ tasa	əsa/ isa/ tasa	esu+	—/ isa/ tasa	‘one’
198.	duha	dua	ru+	rua	‘two’
199.	telu	təlu	tor+	tolu	‘three’
200.	epat	(ə)pat/ pati/ pani	gigti	pat/ pati/ pani	‘four’

APPENDIX 3b – List of sound correspondences

Sound correspondences for cognates identified in **Table 4.1** and discussed in **Table 4.2** and **Figure 4.1** in Chapter 4. Coda segments are double indent, nucleus is single indent, and onsets are not indented. To facilitate comparison with **Figure 4.1**, correspondences are grouped according to general word-shape changes and subgrouped according to final, penultimate, antepenultimate, and preantepenultimate syllable. An environment that contains parentheses indicates that the segment is lost in the Irarutu reflex.

A. Words with main pattern of deletions in historically unaccented syllables

final syllables

- *-p > Ø / (ə)_#
- *-k > Ø / (u)_#
- *-q > Ø / (a)_#, (u)_#, (ə)_#, (a)_#, (a)_#
- *-s > Ø / (i)_#
- *-n > Ø / (a)_#, (u)_#, a_#
- *-m > Ø / (u)_#, (ə)_#

- *i > Ø / (R)_#, R_#, η(s), (R)_#, l_#, m_#
- *u > Ø / s_#, n_(m)#, n_(q)#, ay_#, n_(k)#, a_(n)#, l_#, (q)_#
- *ə > Ø / n_(p)#, n_(m)#, R_(q)#
- *a > Ø / (η)_#, u_#, d_(n)#, n_#, m_(q)#, t_(q)#, t_#, d_#, t_#, u_#
- *ay > Ø / t_#
- *aw > Ø / (n)_#, (j)_#

- *-t- > t / a_ay, u_(u), a_(a), i_a
- *-d- > r / u_(a), i_(a)
- *-j- > Ø / ə_aw
- *-q- > Ø / u_(u)#
- *-s- > s / u_(u)
- *-m- > m / u_(a), a_(i)
- *-n- > n / i_(ə), i_(u), u_(u), a_(ə), a_(u)
- *-n- > n~Ø / a_(aw)#
- *-η- > Ø / u_(a)#
- *-η- > g / a_(i)
- *-R- > r / u_(i), i_i, a_(ə)
- *-R- > Ø / a_(i)#
- *-l- > r / a_(i), ə_(u)

penultimate syllable

- *'i > 'i / #_n(u), d_R, k_t, s_d
- *'i > 'i / q_.nə
- *'u > 'u / d_a, z_R, s_s, R_m, b_n, (k)_t, b_(a)
- *'u > 'ü / t_d, t_n, b_(η)
- *'ə > 'o / t_l
- *'ə > 'e / l_j

*'a > 'a / p_(n), m_(R), t_η, m_t, t_n, b_R, k_l, m_n, m_t, k_m, k_(n)
 *'a > 'o / d_(u)
 *'ay > 'e / (k)_u

*p- > f / #_a
 *b- > b / #_a
 *b- > f / #_u, #_u, #_u
 *t- > t / #_a, #_u, #_a, #_ə, *#_u
 *k- > g / #_a, #_a
 *k- > Ø / #_ay, #_u, #_i, #_a
 *q- > Ø / #_i
 *s- > s / #_u
 *s- > Ø / #_i
 *z- > r / #_u
 *m- > m / #_a, #_a, #_a, #_a
 *R- > r / #_u
 *-l- > r / a_ə

antepenultimate

*a > Ø / (q)_l

*q > Ø / #_(a)

B. Words with minor patterns of deletions in unaccented syllables

final syllable

*-p > f / ə_#
 *-t > t / i_#, (u)_#
 *-q > Ø / a_#, (i)_#
 *-n > n / (a)_#
 *-η > n / (i)_#
 *-R > r / (ə)_#, (a)_#, (i)_#
 *i > Ø / t_q, d_n, η_t, p_#, a_R
 *i > i / n_#, n_#
 *u > Ø / n_#, k_#, (k)_t
 *ə > i / t_p
 *ə > Ø / η_R
 *a > Ø / (q)_n, (k)_R
 *a > u / s_#
 *a > ü / t_q
 *a > e / y_#
 *a > o / k_(u)#
 *a > Ø / i_#, l_n
 *ay > Ø / n_#, n_#

*-p- > f / i_i, u_i
 *-t- > t / a_a, u_i, i_#, a_ə
 *-d- > d / i(n)_i
 *-k- > k / a_u
 *-k- > Ø / i_a, *a_(a)
 *-s- > s / i_a
 *-ŋ- > g / a_i, e_ə
 *-n- > n / i_i, a_ay, i_ay, u_(u)
 *-r- > r / a_ə
 *-l- > r / a_a
 *-y- > j / a_a

penultimate syllable (stressed)

*-p- > f / a_u
 *-b- > f / a_i
 *-d- > r / a_i
 *-k- > Ø / i_a
 *-q- > Ø / u_a
 *-n- > n / a_i
 *-n- > Ø / i_.(d) (accompanies *-d- > d / i(n)_i)
 *-m- > m / u_a

*'i > i / #_n, s_a, n_p, d_nd, b_n
 *'i > 'e / #_s
 *'u > 'u / #_p
 *'u > 'ü / t_n²³
 *'e > o / d_ŋ
 *'a > a / m_y, m_t, w_r, i_k, q_n, l_ŋ, *d_(q), w_(k)
 *'a > e / w_(i)

*t- > t / #_u
 *d- > n / #_e
 *d- > r / *#_a
 *q- > Ø / #_a
 *s- > Ø / #_i
 *m- > m / #_a (cf. *-m-)
 *l- > r / #_a
 *w- > w / #_a, #_a, #_a

antepenult

*i > j / #_a
 *i > Ø / #_k
 *u > Ø / a_m, R_q
 *a > Ø / m_n, m_p, m_d, b_b

*b- > b / #_a
 *m- > b / #_a, #_a
 *m- > Ø / #_a
 *-R- > r / a_u

ante-antepenult

*a > Ø / t_u, m_R

*t- > Ø / #_a
 *m- > m / #_a

C. Stress shift PCOMP *'σ.σ > IRH σ'σ

final syllable change

*-p > f / i_#
 *-t > t / i_#
 *-n > Ø / ə_#
 *-r > Ø / 'u_#

*u > 'u / q_#, l_r, l_#
 *u > 'ü / l_#, t_#
 *a > 'a / R_#, l_#(η_#-metath)
 *i > 'i / l_t, l_p
 *ə > o / p_n
 *a > 'ie / t_#
 *ay > 'e / p_#

*-t- > t / a_a, a_u
 *-p- > f / u_ay, i_ə
 *-b- > f / a_aw
 *-q- > Ø / a_u
 *-R- > r / a_a
 *-l- > r / u_i, u_u, a_i, ə_u, u_u
 *-ŋ- > g / i_a

historically stressed penult

*-b- > f / a_a
 *-t- > t / a_a, a_ə
 *-z- > s / n_u
 *(-)l- > r / a_i, #_i

*'i > Ø / l_η, l_p
 *'u > Ø / k_l, b_l, q_l, z_p
 *'ə > Ø / t_l
 *'a > Ø / b_R, m_l, m_t, n_t, b_b

*b- > f / #_u, #_a

*k- > Ø / #_u

*q- > Ø / #_u

*m- > b / #_a

*m- > m / #_a

*n- > n / #_a

antepenult

*-n.- > Ø / a_z

*a > Ø / m_t, q_b, t_l, q_t, k_n

*t > t / #_a

*k > Ø / #_a

*q- > Ø / #_a, #_a

*m- > Ø / #_a

Appendix 4 – List of associated data items

The following list of digital materials forms a significant portion for the basis of the grammatical description provided in the body of this work. The materials are archived at Kaipuleohone, the University of Hawai'i digital archive (www.kaipuleohone.org). The Irarutu documentation materials are catalogued under the 'Jason Jackson collection' and include some word-lists and other data files on several additional West Papuan languages. The research-tool language is Indonesian. Email the kaipuleohone.org administrator or the author (jjackson@kuboa.com) for permission to access restricted files.

Folder	Description	Notes
JJ-001	<u>Austronesian languages</u> Irarutu: 200-word list, Irarutu: short narrative; Serewen: 200-word list; Ambai: 200-word list, Ambai: short narrative; Biak: 200-word list 1, Biak: 200-word list 2. <u>Non-Austronesian language</u> Mpur: 200-word list.	Metadata for several speakers; Audio recordings; Audio-visual recordings; Location: UNIPA campus.
JJ-002	Irarutu: 1698-word list. Irarutu: spontaneous sentences. Nabi/Kuri: 1479-word list; Nabi/Kuri: poem.	Metadata for several speakers; Audio recordings; Audio-visual recordings; Photographs; Xcel file; Location: Manokwari and Aipiri.
JJ-003	Irarutu: vocabulary elicitation, sentence elicitation, spontaneous discourse, scripted text, folk songs.	Metadata for several speakers; Audio recordings; Photographs; Location: Manokwari and Bintuni
JJ-004	Irarutu: elicited vocabulary and sentences.	Metadata; Audio recordings; Telephone conversations.

Works Cited

- Adriani, N. and A. C. Kruijt. 1914. *De Bare'e-sprekende Toradja's an midden-Celebes*. Vol. 3. VKNA 46. Batavia: Landsdrukkerij.
- Anceaux, J. C. 1958. Languages of the Bomberai Peninsula, outline of a linguistic map. *Nieuw-Guinea Studiën* 2: 109-120.
- . 1961. *The linguistic situation in the islands of Yapen, Kurudu, Nau and Miosnum, New Guinea*. Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkonden 35. The Hague: Nijhoff.
- Andrews, Avery. 1985. The major functions of the noun phrase. In *Language typology and syntactic description*, vol. 1, ed. by Timothy Shopen, 62–154.
- Anthony, David W. 1990. Migration in archeology: the baby and the bath water. *American Anthropologist* 92:895–914.
- Appel, René and Pieter Muysken. 1987. *Language contact and bilingualism*. London: Edward Arnold.
- Badan Koordinasi Penanaman Modal Daerah. 1995. *Investment project profile in Irian Jaya*. Jayapura: Regional Investment Coordinating Board of Irian Jaya.
- Baker, Mark C. 2003. *Lexical categories*. Cambridge: Cambridge University Press.
- . 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Baker, Philip. 2000. Theories of creolization and the degree and nature of restructuring. In *Degrees of restructuring in creole languages*, ed. by Ingrid Neumann-Holzschuh and Edgar W. Schneider, 41–61. Amsterdam: Benjamins.
- Barr, Donald F. and Sharon G. Barr. 1978a. *Index of Irian Jaya languages*. Pre-publication draft. Universitas Cenderawasih and Summer Institute of Linguistics.
- . 1978b. *Mairasi Language Survey Report*, ms.
- Bellwood, Peter. 1998. From Bird's Head to bird's eye view: long term structures and trends in Indo-Pacific prehistory. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 951–975.
- Bennett, J. Fraser. 1994. Iambicity in Thai. *Studies in the Linguistic Sciences* 24(1/2): 39–57.
- Bickel, Balthasar. 2011. Grammatical relations typology. In *Oxford handbook of linguistic typology*, ed. by Jae Jung Song, 399–444. Oxford: Oxford University Press.
- Bickerton, Derek. 1975. *Dynamics of a creole system*. Cambridge: Cambridge University Press.
- Biggs, Bruce. 1965. Direct and indirect inheritance in Rotuman. *Lingua* 14:383–415.
- Bink, G. L. 1891. Listje van telwoorden en enige zelfstandige naamwoorden enz. van Wandammen. *Tijdschrift voor Indische Taal-, Land-, en Volkenkunde* 34.62–71.
- Blust, Robert A. 1974. Eastern Austronesian: A note. *Working Papers in Linguistics* 6(4):101–107. Honolulu: Department of Linguistics, University of Hawai'i.
- . 1978. Eastern Malayo-Polynesian: a subgrouping argument. In Stephen Wurm and Lois Carrington, 181–234.
- . 1981. Variation in retention rate among Austronesian languages. (Paper presented at the Third International Conference on Austronesian Linguistics. Denpasar, Bali, January 1–11.)
- . 1982. The linguistic value of the Wallace line. *Bijdragen tot de Taal- Land- en Volkenkunde* 138:231–250.

- . 1983/1984. More on the position of the languages of Eastern Indonesia. *Oceanic Linguistics* 22/23:1–28.
- . 1991. On speech strata in Tiruray. In *Papers in Austronesian Linguistics*, ed. by Malcolm D. Ross, no. 1:1–52. Canberra: Pacific Linguistics (PL A-82).
- . 1993. Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics* 32:241–293.
- . 1995. Austronesian comparative dictionary. [[http:// www.trussel2.com/ACD/](http://www.trussel2.com/ACD/)]
- . 2000. Why lexicostatistics doesn't work: the 'universal constant' hypothesis and the Austronesian languages. In *Time depth in historical linguistics*, vol. 2, ed. by Colin Renfrew, April McMahon, and Larry Trask, 311–331. Cambridge: The McDonald Institute for Archaeological Research.
- . 2004. Austronesian nasal substitution: A survey. *Oceanic Linguistics* 43:73–148.
- . 2005. Must sound change be linguistically motivated? *Diachronica* 22(2):219–269.
- . 2008a. Is there a Bima-Sumba subgroup? *Oceanic Linguistics* 47:46–114.
- . 2008b. Complex sound changes. (Presentation at Austronesian Formal Linguistics Association 15).
- . 2009a. The position of the languages of eastern Indonesia: a reply to Donohue and Grimes. *Oceanic Linguistics* 48:36–77.
- . 2009b. *The Austronesian languages*. Canberra: Pacific Linguistics.
- Blust, Robert A. and Stephen J. Trussel. ongoing. Austronesian comparative dictionary. [[http:// www.trussel2.com/ACD/](http://www.trussel2.com/ACD/)]
- Boretzky, Norbert. 1984. The Indo-Europeanist model of sound change and genetic affinity, and change in exotic languages. *Diachronica* 1(1):1–51.
- Bowden, John. 1997. The meanings of directionals in Taba. In *Referring to space: studies in Austronesian and Papuan languages*, ed. by Gunter Senft, 251–268. Oxford: Clarendon Press.
- . 2001. *Taba: Description of a South Halmahera language*. Canberra: Pacific Linguistics (PL 521).
- Bradshaw, Joel. 1993. Subject relationships within serial verb constructions in Numbami and Jabêm. *Oceanic Linguistics* 32:133–161.
- Briley, Joyce. 1977. Some counting systems of Irian Jaya. *Irian: Bulletin of Irian Jaya development* 6(3):28–32.
- Broschart, Jürgen. 2000. The Tongan category of preverbials. In *Approaches to the typology of word classes*, ed. by Petra M. Vogel and Bernard Comrie, 351–369. Berlin: Mouton de Gruyter.
- Capell, Arthur. 1976. Austronesian and Papuan 'mixed' language: general remarks. In *New Guinea area languages and language study 2: Austronesian languages*, ed. by Stephen A. Wurm, 527–579. Canberra: Pacific Linguistics (PL C-39).
- Carr, Philip. 1993. *Phonology*. London: Macmillan.
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and Topic*, ed. by Charles N. Li, 25–55. New York: Academic Press.
- . 1994. *Discourse, consciousness, and time*. Chicago: The University of Chicago Press.
- Chomsky, Noam. 1995. *Minimalist Program*. Cambridge: MIT Press.

- Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. Boston: MIT Press.
- Comrie, Bernard and Michael Cysouw. 2009. New Guinea through the eyes of WALS. *Language and Linguistics in Melanesia* 30:65–95. [<http://wals.info/>]
- Comrie, Bernard, and Norval Smith. 1977. *Lingua* descriptive studies: questionnaire. *Lingua* 42:1-72.
- Cowan, H. K. J. 1953a. *Voorlopige resultaten van een ambtelijk taalonderzoek in Nieuw-Guinea*. 's-Gravenhage: Martinus Nijhoff.
- . 1953b. De Austronesisch-Papoea'se taalgrens in de onderafdeeling Hoolandia (Nieuw-Guinea). *Tijdschrift Nieuw-Guinea* 13.133-43, 161-76, 201-6.
- . 1955. Notes on Windesi Grammar. *Oceania* 26.42–58.
- Croft, William. 1991. *Syntactic categories and grammatical relations: the cognitive organization of information*. Chicago: University of Chicago Press.
- Crowley, Terry. 1987. Serial verbs in Paamese. *Studies in Language* 11(1):35–84.
- . 2002. *Serial verbs in Oceanic: A descriptive typology*. Oxford: Oxford University Press.
- Dalrymple, Mary, and Suriel Mofu. 2011. *Dusner dictionary*. [<http://dusner.clp.ox.ac.uk/>]
- . 2012. *Dusner*. Berlin: Lincom.
- De Casparis, J. G., ed. 1997. Sanskrit loan-words in Indonesian: an annotated checklist of words from Sanskrit in Indonesian and traditional Malay. *NUSA* 41: 1–49.
- De Clerq, F. S. A. 1890. *Bijdragen tot de kennis der Residentie Ternate*. Leiden: Brill.
- Deda, Andreas. 2012. Bahasa Tandia tinggal fosil. ms.
- Deda, Andreas, Musa Ayorbaba, Sugeng D. Riyanto, Alfred Umpain, Elais T. Auri, Eklevina C. Nenepat, Erin Matius, John P. Waroi, Rafael Natkime, Risnawati, Welma Kawari, Nehemia Wallianggen, and Nelly D. Sinaga. 2011a. *Dusner: bahasa dan budaya Papua yang tersisih* [Dusner: A lost Papuan language and culture]. Undergraduate student project, Universitas Negeri Papua di Papua.
- . 2011b. *Dusner-Indonesian-English dictionary*. Undergraduate student project, Universitas Negeri Papua.
- Dempwolff, Otto. 1934–1938. 3 vols. *Vergleichende Lautlehre des austronesischen Wortschatzes*. Zeitschrift für Eingeborenen-Sprachen, Supplement 1. *Induktiver Aufbau einer indonesischen Ursprache* (1934), Supplement 2. *Deduktive Anwendung des Urindonesischen auf austronesische Einzelsprachen* (1937), Supplement 3. *Austronesisches Wörterverzeichnis* (1938). Berlin: Reimer.
- Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- . 2011-12. *Basic Linguistic Theory* vols. I-III. Oxford: Oxford University Press.
- Donegan, Patricia. 1986. *On the Natural Phonology of Vowels*. PhD diss, Columbus, OH: Department of Linguistics, Ohio State University.
- . 1993. On the phonetic basis of phonological change. In *Historical linguistics: problems and perspectives*, ed. by Charles Jones, 98–130. London: Longmans.
- Donegan, Patricia and David Stampe. 1978. The syllable in phonological and prosodic structure. In *Syllables and segments*, ed. by Alan Bell and Joan B. Hooper. Amsterdam: North Holland.
- . 1983. Rhythm and the holistic organization of language structure. In *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*, ed. by

- John F. Richardson, Mitchell Marks, and Amy Chukeman, 337–353. Chicago: Chicago Linguistic Society.
- Donohue, Mark. 2006. Coronals and velars: support for Blust. *Oceanic Linguistics* 45:229–241.
- Donohue, Mark and Charles E. Grimes. 2008. Yet more on the position of the languages of eastern Indonesia and East Timor. *Oceanic Linguistics* 47(1):114–158.
- Dyen, Isidore. 1965. *A lexicostatistical classification of the Austronesian languages*. Baltimore: International Journal of American Linguistics Memoir 19.
- . 1978. The position of the languages of eastern Indonesia. In Stephen Wurm and Lois Carrington, 235–54.
- Edmondson, Jerold A. and Yang Quan. 1989. Word-initial preconsonants and the history of Kam-Sui resonant initials and tones. In *Comparative Kadai: Linguistic studies beyond Tai*, ed. by Jerold Edmondson and David Solnit, 143–166. Arlington, Texas: University of Texas at Arlington.
- Esser, S. J. 1938. Talen. Sheet 9 of *Atlas van Tropisch Nederland*. Batavia: Topographische Dienst in Nederlandsch-Indië.
- Fautngil, Christ, Frans Rumbrawer, and Bartolomeus Kainakaimu. 1994. *Sintaksis Bahasa Biak*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa Departemen Pendidikan dan Kebudayaan.
- Fields, Phil. n.d. *Bahasa Irarutu: Fruata Report*, ms.
- Flaming, Rachel. 1981. *Wandamen Verbs and Verb phrases*, ms.
- . 1983a. Wandamen kinship terminology. In *Gods, heroes, kinsmen: ethnographic studies from Irian Jaya, Indonesia*, ed. by William R. Merrifield, Marilyn Gregerson and Daniel C. Ajamiseba, 245–53. Dallas, TX: International Museum of Cultures. (also published in *Irian* 1983 11(1): 1–17)
- . 1983b. Cohesion in Wandamen Narrative. In *Miscellaneous studies of Indonesian and other languages in Indonesia*, part 7, ed. by Soenjono Dardjowidjojo and John W. M. Verhaar, 41–49. Jakarta: Atma Jaya University, Badan Penyelenggara NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia 15.
- Florey, Margaret. 2004. Countering purism: Confronting the emergence of new varieties in a training program for community language workers. In *Language Documentation and Description*, vol. 2, ed. by Peter K. Austin, 9–27. London: The Hans Rausing Endangered Languages Project.
- Foley, William A. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- . 1998. Toward understanding Papuan languages. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 503–518.
- . 2003. Genre, register and language documentation in literate and preliterate communities. In *Language Documentation and Description*, vol. 1, ed. by Peter K. Austin, 85–98. London: The Hans Rausing Endangered Languages Project.
- Foster, Peter. 1973. The origin and introduction of basic food crops of Irian Jaya. *Irian* 2(3):49–54.
- Fox, Charles E. 1978. *Arosi Dictionary*. Revised edition with English-Arosi Index prepared by Mary Craft. Canberra: Pacific Linguistics (PL C-57).
- Gadamer, Hans-Georg. 1975. *Truth and Method*. London: Continuum.

- Galis, K. W. 1955. Talen en dialecten van Nederlands Nieuw-Guinea. *Tijdschrift Nieuw-Guinea* 16(4–6).
- Gil, David and John Bowden. Intercontinental Dictionary Series (IDS) at JFS. Leipzig. [<http://lingweb.eva.mpg.de/ids>]
- Givón, Talmy. 1990. *Syntax: A Functional-Typological Approach*. 2 vol.s. Amsterdam: John Benjamins. (2nd ed. 1994)
- Goodman, Tom. 1998. The *sosolot* exchange network in Eastern Indonesia during the seventeenth and eighteenth centuries. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 421–454.
- Grace, George W. 1955–6. Field notebooks. (New Guinea language data) [<http://digicoll.manoa.hawaii.edu/grace/Pages/PDFlist.html>]
- . 1971. Notes on the phonological history of the Austronesian languages of the Sarmi Coast. *Oceanic Linguistics* 10:11–37 (also published in *Irian: Bulletin of West Irian Development* 1(3):21–53).
- Granadillo, Tania and Heidi Orcutt-Gachiri. 2011. *Ethnographic Contributions to the Study of Endangered Languages*. Tuscon: The University of Arizona Press.
- Greenhill, Simon J., Robert A. Blust, Russell D. Gray. 2008. The Austronesian Basic Vocabulary Database: from bioinformatics to lexomics. *Evolutionary Bioinformatics*, 271–283. [<http://www.language.psy.auckland.ac.nz/austronesian/>]
- Gumperz, John J. 1968. Linguistics 3: The speech community. In *International Encyclopedia of the Social Sciences*, ed. by David L. Sills and Robert K. Merton, 381–386
- Guérin, Valerie. 2007. Discovering Mavea. PhD diss., Honolulu: Department of Linguistics, University of Hawai‘i.
- Hadi, A.C. Sungkana. 1999. Irian Jaya Libraries and the preservation of Indigenous Languages. *Irian* 22(1):48–53.
- Haenen, Paul. 1998. History, exchange, and myth in the southeastern Bird’s Head of Irian Jaya. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 235–256.
- Hajek, J. and J. Bowden. 2002. Taba and Roma: clusters and geminates in two Austronesian languages. *Proceedings of the XIVth International Congress of Phonetic Sciences, San Francisco*:1033–1036.
- Hale, Ken, Michael Krauss, Lucile J. Watahomigie, Akira Y. Yamamoto, Colette Craig, LaVerne Masayeva Jean, and Nora C. England. 1992. Endangered languages. *Language* 68(1):1–42.
- Harris, Alice C. and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge Studies in Linguistics 74. Cambridge: Cambridge University Press.
- Hawkins, John A. 1983. *Word Order Universals*. New York: Academic Press.
- Hayes, Bruce. 2009. *Introductory Phonology*. Malden: Wiley-Blackwell.
- Held, G. J. 1942a. Grammatica van het Waropensch-Nederlandsch Noord Nieuw-Guinea. *Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen* 77(1).
- . 1942b. *Woordenlijst van het Waropensch*. Bandoeng: A. C. Nix.
- . 1956. Waropensche Teksten. *Verhandelingen van het Koninklijk Insitituut voor Taal-, Land- en Volkenkunde* 20.

- . 1957. *The Papuas of Waropen*. Translation Series No. 2. The Hague: Koninklijk Instituut voor Taal- Land- en Volkenkunde.
- Henning, Jean C. and Theodore A. Henning (compilers). 1991. *Sane pai ve pir kavavo nana kavo Wondama = Perbendaharaan kata bahasa Wandamen = Wandamen vocabulary*. Publikasi khusus bahasa-bahasa daerah, seri B, 8. Jayapura: Universitas Cenderawasih and Summer Institute of Linguistics.
- Himmelman, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36:161-195.
- . 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In *The Austronesian languages of Asia and Madagascar*, ed. by Alexander Adelaar and Nikolaus P. Himmelman, 110–81. London: Routledge.
- . 2012. Linguistic data types and the interface between language documentation and description. *Language Documentation and Conservation* 6:187–207.
- Himmelman, Nikolaus P. and D. Robert Ladd. 2008. Prosodic Description: an introduction for fieldworkers. *Language Documentation and Conservation* 2(2):244–274.
- Hockett, Charles F. 1955. A manual of phonology. *International Journal of American Linguistics* 21(4). Part 1: Indiana University publications in anthropology and linguistics. Memoir 11. Baltimore: Waverly Press.
- . 1958. *A Course in Modern Linguistics*. New York: Macmillan
- Hope, Geoffrey S., and Simon G. Haberle. 2005. The history of the human landscape of New Guinea. In Andrew Pawley, Robert Attenborough, Jack Golson and Robin Hide, 541–554. Canberra: ANU.
- Hopper, Paul and Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2):251–299.
- Huffman, Franklin. 1972. The Boundary between the monosyllable and the disyllable in Cambodian. *Lingua* 29:54–66.
- Huizinga, F. 1998. Relations between Tidore and the north coast of New Guinea in the nineteenth century. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 385–419.
- Hymes, Dell. 1968. Linguistics 1: The field. In *International Encyclopedia of the Social Sciences*, ed. by David L. Sills and Robert K. Merton, 351–371.
- Jackson, Jason. 2004. Complementary coals. (Presentation for the University of Hawai‘i at Mānoa Linguistics Department Tuesday Seminar).
- . 2008. *South Halmahera-West New Guinea Austronesian areal linguistics*, ms.
- Janda, Richard D. and Brian D. Joseph, eds. 2003. *The handbook of historical linguistics*. Malden: Blackwell Publishing.
- Kamma, Freerk Christiaans. 1947. De verhouding tussen Tidore en de Papoeese eilanden in legende en historie. *Indonesie* 1:361–70; 536–59.
- . 1994. *Ajaib di Mata Kita*, vol. 3. Disadur oleh Dr. Th. van den End. Jakarta: PT BPK Gunung Mulia.
- . n.d. *Enkele aanwijzingen voor het Wandammen*. Het Utrechts Archief, nummer toegang 6737, inventarisnummer 1102-2.
- Kamholz, David. 2014. South Halmahera–West New Guinea: The history of Oceanic's closest relative. (Presentation at Linguistic Society of America Annual Meeting. Minneapolis, MN.)
- . forthcoming. *Austronesians in Papua: Diversification and change in South*

- Halmahera–West New Guinea*. PhD. diss. Berkeley: Department of Linguistics, University of California at Berkeley.
- Karubaba, Sara Yulita. 2000. *A case of personal pronoun in Ambai language*. Jayapura: Department of Language and Art Faculty of Teacher Training and Educations Cenderawasih University.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge: The MIT Press.
- Keenan, Edward L. 1985. Relative clauses. In *Language typology and syntactic description*, vol. 2, ed. by Timothy Shopen, 141–170.
- Keenan, Edward L. and Bernard Comrie. 1977. NP accessibility and universal grammar. *Linguistic Inquiry* 8:63-100.
- Kern, H. 1885. Over de verhouding van het Nufoorsch tot de Maleis-Polynesische talen. *Verspreide Geschriften* 6.35-76.
- Kijne, Isaac Samuel. n.d.(a). [Notebook containing Umar wordlist and Tandia grammatical notes]. Archived at Koninklijk Instituut voor Taal-, Land-, en Volkenkunde (Leiden), aanvraagnummer D Or. 421, inventarisnummer 33.
- . n.d.(b). [Notebook containing Waropen, Roon, and Dusner wordlist and grammatical notes]. Archived at Koninklijk Instituut voor Taal- Land- en Volkenkunde (Leiden), aanvraagnummer D Or. 421, inventarisnummer 13.
- Kim, Sung K., Christopher R. Gignoux, Jeffery D. Wall, Annette Lum-Jones, Hansong Wang, Christopher A. Haiman, Gary K. Chen, Brian E. Henderson, Laurence N. Kolonel, Loic Le Marchand, Daniel O. Stram, Richa Saxena, and Iona Cheng. 2012. Population genetic structure and origins of native Hawaiians in the multi-ethnic cohort study. *PLoS ONE* 7(11):e47881. [doi:10.1371/journal.pone.0047881]
- Klamer, Margaretha A. F. 2002. Typical Features of Austronesian Languages in Central/Eastern Indonesia. *Oceanic Linguistics* 41(2): 363-383.
- Klamer, Marian. 2007. The use of language data in comparative research: a note on Blust (2008) and Onvlee (1984). *Oceanic Linguistics* 48(1):250–263.
- Klamer, Marian, Ger Reesink, and Miriam van Staden. 2007. East Nusantara as a linguistic area. In *From linguistic areas to areal linguistics*, ed. by Pieter C. Muysken, 95–150. Amsterdam: John Benjamins.
- Koentjaraningrat, ed. 1994. *Irian Jaya membangun masyarakat majemuk*. Jakarta: Djambatan.
- Ladefoged, Peter. 1992. Another view of endangered languages. *Language* 68(4):809–811.
- . 2001. *A Course in Modern Phonetics*. New York, Harcourt-Brace Jovanovich.
- Lass, Roger. 1984. *Phonology: an introduction to basic concepts*. Cambridge: Cambridge University Press.
- Laycock, Donald C. 1974. A little Mor. In Stephen Wurm and Lois Carrington, 285–316.
- . 1977. Me and You versus the rest. *Irian: Bulletin of Irian Jaya development* 6(3):35–39.
- Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig, eds. 2013. *Ethnologue: Languages of the World*. 17th ed. (web edition). Dallas, Texas: Summer Institute of Linguistics, Inc. [http://www.ethnologue.com]
- Lichtenberk, Frantisek. 1985. Possessive constructions in Oceanic languages and in Proto-Oceanic. In *Austronesian linguistics at the 15th Pacific Science Congress*, ed. by Andrew Pawley and Lois Carrington, 93–140. Canberra: Pacific Linguistics.

- Lincoln, Peter Craig. 1976. *Describing Banoni, an Austronesian language of South Bougainville*. PhD. diss., Honolulu: Department of Linguistics, University of Hawai'i at Mānoa.
- . 2012. Counting. In *Language in Hawai'i and the Pacific*. University of Hawai'i at Mānoa Linguistics 100 Course Reader, ed. by Hiroko Sato and Jake Terrell, 45–56. Honolulu: University of Hawai'i.
- Lynch, John, Malcolm Ross, and Terry Crowley. 2002. *The Oceanic Languages*. Surrey: Curzon Press.
- Maan, G. 1940. *Boelisch-Nederlandsche woordenlijst met Nederlandsch–Boelisch register*. Deel LXXIV. Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen. Bandoeng: A. C. Nix.
- . 1951. *Proeve van een Bulische Spraakkunst*. 's-Gravenhage: Martinus Nijhoff.
- Malkiel, Yakov. 1968. Linguistics 2: Historical linguistics. In *International Encyclopedia of the Social Sciences*, ed. by David L. Sills and Robert K. Merton, 371–380.
- Mandopma, F. 1992. *Sistem Pengetahuan tradisional masyarakat Irarutu di desa Fruata kecamatan Babo kabupaten Manokwari*, ms.
- Mandowen, Willy. 1999. Taboo as sociolinguistic shadow among Biak speakers. *Irian* 22(1):32–40.
- Mansoben, Johsz and Malcolm Walker. 1990. Indigenous political structure and leadership patterns in Irian Jaya. *Irian* 18:17–24.
- Marck, Jeffery C. 1986. Micronesian dialects and the overnight voyage. *Journal of the Polynesian Society* 95:253–258.
- Matisoff, James. 1990. Bulging monosyllables: Areal tendencies in Southeast Asian diachrony. *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society*, 16–19 February 1990, ed. by Kira Hall, Jean-Pierre Koenig, Michael Meacham, Sondra Reinman, and Laurel A. Sutton, 543–559. Berkeley, CA.: Berkeley Linguistics Society.
- Matsumura, Michiko. 1992. Irarutu kinship and marriage. *Irian Bulletin of Irian Jaya* 20:1–16.
- . 1997. Irarutu kinship and marriage. In *Kinship and social organization in Irian Jaya: a glimpse of seven systems*, ed. by Marilyn Gregerson and Joyce Sterner, 221–242. Jayapura and Dallas: Universitas Cenderawasih and Summer Institute of Linguistics.
- Matsumura, Takashi. 1984. *Survey report on Irarutu language group*, ms.
- . 1985a. *Allocation survey for Irarutu villages in Arguni Bay*, ms.
- . 1985b. *Laporan sigkat pelaksanaan survei di Teluk Arguni*, ms.
- . 1991. Irarutu phonology. *Workpapers in Indonesian Languages and Cultures* 10:37–74.
- . 1999. *Lukas nfier frgrgir snan frfier dir weni: Injil Lukas*. Lembaga Alkitab Indonesia.
- Matsumura, Takashi and Michiko Matsumura. 1991. A preliminary grammar sketch of the Irarutu language. *Workpapers in Indonesian Languages and Cultures* 10:75–110.
- Mayer, John. 2001. *Code-switching in Samoan: t-style and k-style*. Unpublished PhD. diss. Honolulu: Department of Linguistics, University of Hawai'i.

- McCawley, James. 1981. *Everything that linguists have always wanted to know about logic but were ashamed to ask*. (2nd edition 1993.) Chicago: University of Chicago Press.
- Meyer, A. B. 1874a. Über die Mafoor'sche und einige andere Papuasprachen auf Neu Guinea. *Sitzungsberichte der Wiener Akademie der Wissenschaften, Philosophisch-Historische Classe* 77:299–356.
- . 1874b. Prove der Mafoor'schen Sprache. *Sitzungsberichte der Wiener Akademie der Wissenschaften, Philosophisch-Historische Classe* 78.
- Miedema, Jelle (compiler). 1997. *Texts from the oral tradition in the Southern Bird's Head peninsula of Irian Jaya: Inanwatan-Berau, Arandi-Bintuni, and Hinterland*. Irian Jaya Source Materials no. 15 series B no. 7. Leiden/Jakarta: Projects Division of the Department of Languages and Cultures of South-East Asian and Oceania (DSALCUL)/Irian Jaya Studies: A Programme for Interdisciplinary Research (ISIR).
- Miedema, Jelle, Cecilia Odé, and Rien A. C. Dam, eds. 1998. *Perspectives on the Bird's Head of Irian Jaya, Indonesia: Proceedings of the conference, Leiden, 13-17 October 1997*. Amsterdam and Atlanta: Rodopi.
- Mofu, Surriel. 2005. *Biak Morphosyntax*. M.Phil thesis, University of Oxford.
- . 2008. *Biak Morphosyntax*. D. Phil. Thesis. University of Oxford.
- Mufwene, Salikoko S. 2001. The Founder Principle in the development of creoles. In *The ecology of language evolution*, 25–80, 211–215. Cambridge: Cambridge University Press.
- Nathan, David. 2006. Thick interfaces: mobilizing language documentation with multimedia. In *Essentials of language documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 363–379. Berlin: Mouton de Gruyter.
- Noonan, Michael. 1985. Complementation. In *Language typology and syntactic description*, ed. by Timothy Shopen, 42–140.
- Nothofer, Bernd. 1992. CEMP and CMP—two Austronesian subgroups? A critical evaluation of Blust (1990) and Grimes (1991). (Paper presented at the Maluku Studies Conference, Honolulu, Hawai'i, July 30–August 2, 1992.)
- O'Grady, William. 2013. *The Syntax Files*. Honolulu: University of Hawai'i at Mānoa.
- O'Grady, William and John Archibald eds. 2011. *Contemporary linguistic analysis*. 7th ed. Toronto: Pearson.
- Ongkodharma, Nitya and Y. A. Nelwan. 1983. Kepercayaan orang Wandamen. In *Gods, heroes, kinsmen: Ethnographic studies from Irian Jaya, Indonesia*, ed. by William R. Merrifield, Marilyn Gregerson and Daniel C. Ajamiseba, 69–113. Dallas, TX: International Museum of Cultures. (also published in *Irian* 13:87–127)
- Ongkodharma, Nitya and Naomi Saggars. 1977. *Survey report, Wasior area*, ms.
- Ongkodharma, Nitya, Naomi Saggars, and Rachel Flaming. n.d. *Wandamen phonology*, ms.
- Ottow, W. 1862. Woordenlijst der te Doreh en omstreken gesproken wordende Myfoorsche (Noemfoors, Mafoors, Nufoors etc.) taal. *Nieuw-Guinea* 1862:201–33. Amsterdam: Fred Muller.
- Patz, Elisabeth. 1978. The case marking and role coding system of Numfoor Biak. *Oceanic Linguistics* 17(2):141–161.
- Pawley, Andrew. 1998. The Trans-New Guinea Phylum hypothesis: A reassessment. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 655–690.

- Pawley, Andrew, Robert Attenborough, Jack Golson and Robin Hide, eds. 2005. *Papuan Pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples*. Canberra: Pacific Linguistics.
- Payne, Thomas E. 1997. *Describing morpho-syntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Peckham, Lloyd. 1978. *Mairasi Grammar*, ms.
- . 1982a. Mairasi verb morphology. *Workpapers in Indonesian Linguistics* 1:75–96. Universitas Cenderawasih and Summer Institute of Linguistics.
- . 1982b. All kinds of people: A Mairasi view of the spirit world. *Irian* 10(2):35–91.
- . 1991. Mairasi Phonology. *Workpapers in Indonesian Languages and Cultures* 10:111–145.
- Peckham, Nancy. 1981a. Day and night songs in Mairasi restival music. *Irian Bulletin of Irian Jaya*. 9(1):55–65.
- . 1981b. My elders, my helpers. *Irian Bulletin of Irian Jaya*. 9(3):43–65.
- . 1983. Mairasi kinship and marriage. In *Gods, heroes, kinsmen: ethnographic studies from Irian Jaya, Indonesia*, ed. by William R. Merrifield, Marilyn Gregerson and Daniel C. Ajamiseba, 255–270. Dallas, TX: International Museum of Cultures.
- Peckham, Lloyd and Nancy Peckham. n.d. *Mairasi Phonology*, ms.
- Perlmutter, David M. and Paul D. Postal. 1983. Some proposed laws of basic clause structure. *Studies in Relational Grammar* 1:83–128.
- Piggott, Glyne L. 2004. The representation of extrametricality. *McGill Working Papers in Linguistics* 18(2): 45–88.
- Pike, Kenneth. 1947. Grammatical prerequisites to phonemic analysis. *Word* 3:155–172.
- Price, David S., and Mark Donohue. 2009. *Report on the Ansus survey, West Yapen island, Papua, Indonesia*. SIL Electronic SurveyReport 2009-001. Summer Institute of Linguistics International. [<http://www-01.sil.org/SILESR/2009/silesr2009-001.pdf>]
- Ramar, Origenes, Nitya Ongkodharma and Rachel Flaming. 1983. *Veka vavo nana kavo Wondama = Percakapan-percakapan dalam bahasa Wandamen = Wandaman conversations*. Publikasi khusus bahasa-bahasa daerah A-6. Jayapura: Universitas Cenderawasih and Summer Institute of Linguistics.
- Ratman, Nana. 1998. Geology of the Bird's Head, Irian Jaya, Indonesia. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 719–755.
- Reesink, Ger. 1998. The Bird's Head as Sprachbund. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 603–642.
- . 2005. West Papuan languages: roots and development. In Andrew Pawley, Robert Attenborough, Jack Golson and Robin Hide, 185–218.
- Rehg, Kenneth. 2012. The representation of long vowels and diphthongs in three central Pacific languages: Hawaiian, Tongan, and Fijian. (Presentation for the University of Hawai'i at Mānoa Department of Linguistics Tuesday Seminar).
- Remijsen, Albert C.L. 2002. *Word-prosodic systems of Raja Ampat languages*. Utrecht: LOT.
- . 2010. Nouns and verbs in Magey Matbat. In *Typological and areal analyses: Contributions from East Nusantara*, ed. by Michael Ewing and Marian Klammer, 281–311. Canberra: Australian National University (Pacific Linguistics 618). [http://www.ling.ed.ac.uk/~bert/mageymatbat_grammar.pdf]

- . n.d.(a) *The tone contrast of Magey Matbat*
[http://www.ling.ed.ac.uk/~bert/ngtpaper_sub.pdf]
- Ricoeur, Paul. 1981. *Hermeneutics and the Human Sciences*. Cambridge: Cambridge University Press.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, 281–337. The Netherlands: Springer.
- Robinson, Laura. 2008. *Dupaningan: Grammar, texts, and vocabulary*. PhD diss., Honolulu: Department of Linguistics, University of Hawai‘i.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1):75–116.
- Ross, Malcolm. 1988. *Proto-Oceanic and the Austronesian languages of western Melanesia*. Canberra: Australian National University (Pacific Linguistics C-98).
- . 1995. Some current issues in Austronesian linguistics. In *Comparative Austronesian dictionary*, part 1: fascicle 1, ed. by Darrell Tryon, 45–120. Berlin: Mouton de Gruyter.
- . 2002. The history and transitivity of western Austronesian voice and voice-marking. In *The history and typology of western Austronesian voice systems*, ed. by Fay Wouk and Malcolm Ross, 17–62. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University (Pacific Linguistics).
- . 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In Andrew Pawley, Robert Attenborough, Jack Golson, and Robin Hide, 45–120.
- Rusman. 1998. Youth, education and employment in Irian Jaya. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 365–381.
- Sadock, Jerold M. and Arnold M. Zwicky. 1985. Speech act distinctions in syntax. In *Language typology and syntactic description*, vol. 1, ed. by Timothy Shopen, 155–196.
- Saggers, Naomi. 1979. *A sketch grammar of Wandamen*. M.A. thesis (Mimeo.)
- Sahlins, Marshall D. 1963. Poor man, rich man, big-man, Chief: Political Types in Melanesian and Polynesia. *Comparative Studies in Society and History* 5(3):285–303.
- Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace & World, Inc.
- Saragih, Yansen. 2006. *Subject–Verb Agreement in Ansum Language*, ms.
- Sawaki, Yusuf. 2006. *Penyebahan bahasa dan budaya di Kabupaten Raja Ampat suatu kajian sociolinguistik dan etnografi/antropologi*. Manokwari: Konsortium Pembuatan Atlas Sumber Daya Kabupaten Raja Ampat, ms.
- . forthcoming. *Description of Woi*. PhD. diss., Canberra: Department of Linguistics, Australia National University.
- Schachter, Paul 1985. Parts-of-speech systems. In *Language typology and syntactic description*, vol. 1, ed. by Timothy Shopen, 1–61.
- Sebba, Mark. 1997. *Contact languages: Pidgins and creoles*. Hampshire: Macmillan.
- Selkirk, Elisabeth. 1984. On the major class features and syllable theory. In *Language sound structure: studies in phonology*, ed. by Mark Aronoff & Richard T. Oehrle, 107–136. Cambridge: MIT Press.
- Shopen, Timothy, ed. 1985. *Language typology and syntactic description*. 3 Vols. Vol. 1: Clause structure. Vol. 2: Complex clauses. Vol. 3: Syntactic categories. Cambridge: Cambridge University Press.
- Sievers, Eduard. 1901. *Grundzüge der Phonetik*. 5th ed. Leipzig: Breitkopf and Härtel.

- Sills, David L. and Robert K. Merton eds. 1968 *International Encyclopedia of the Social Sciences*. New York: Macmillan.
- Silzer, Peter J. 1982a. *Ambai: some notes on syntactic universals*, ms.
- . 1982b. Kepentingan linguistik diakronis bagi penelitian sinkronis dalam bahasa Ambai di Irian Jaya. *Workpapers in Indonesian linguistics* 1:6–13. Universitas Cenderawasih and Summer Institute of Linguistics.
- . 1983. *Ambai: an Austronesian language of Irian Jaya, Indonesia*. Canberra: Australia National University.
- Silzer, Peter and Daniel C. Ajamiseba. 1981. Laporan hasil penelitian bahasa-bahasa Austronesis di pulau Yapen yang diadakan pada bulan Pebruari 1981 oleh projek Universitas Cenderawasih and Summer Institute of Linguistics. Dallas, TX: Universitas Cenderawasih and Summer Institute of Linguistics.
- Silzer, Peter J. and Heljä Heikkinen. 1984. Index of Irian Jaya languages. *Irian Bulletin of Irian Jaya* 12.
- Silzer, Peter J. and Heljä Heikkinen Clouse. 1991. *Index of Irian Jaya Languages*. Special publication of Irian: Bulletin of Irian Jaya. Jayapura: Program Kerjasama Universitas Cenderawasih dan Summer Institute of Linguistics.
- Silzer, Peter J. and Sheryl Silzer. 1985. *Ambai dictionary*, ms.
- Silzer, Sheryl. 1979. Some figures and ground functions in Ambai narrative. *Irian* 8(1):33–52.
- Slump, F. 1924–38. Grammatica en woordenlijst van de Seroei-Laut taal. Het Utrechts Archief, nummer toegang 2274, inventarisnummer 1102-1.
- Smits, L. and C. L. Voorhoeve, eds. 1992. *Irian Jaya Source Materials No. 4–5 Series B No. 1–2 The J.C. Anceaux Collection of Wordlists of Irian Jaya Languages, A: Austronesian Languages*. 2 parts. Leiden/Jakarta: Projects Division of the Department of Languages and Cultures of South-East Asia and Oceania (DSALCUL)/Irian Jaya Studies: A Programme for Interdisciplinary Research (ISIR).
- . 1994. *Irian Jaya Source Materials No. 9–11 Series B No. 3–5 The J.C. Anceaux Collection of Wordlists of Irian Jaya Languages, B: Non-Austronesian Languages*. 3 parts. Leiden/Jakarta: Projects Division of the Department of Languages and Cultures of South-East Asia and Oceania (DSALCUL)/Irian Jaya Studies: A Programme for Interdisciplinary Research (ISIR).
- Soeparno. 1975. *Kamus bahasa Biak–Indonesia*. Report of the Results of Lexicographic Research, Project for the Development of the National and Areal Languages of Indonesia (ILDEP). Jayapura: Universitas Cenderawasih.
- . 1976. Masalah kata ganti dalam penyusunan kamus Biak–Indonesia. *Majalah Universitas Cenderawasih* 22(13):21–30.
- . 1977a. *Kamus bahasa Biak–Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.
- . 1977b. Fungsi dan pola perulangan sempurna dalam bahasa Biak. *Majalah Universitas Cenderawasih* 24(14):18–27.
- Spriggs, Matthew. 1998. The archaeology of the Bird’s Head in its Pacific and Southeast Asian context. In Jelle Miedema, Cecilia Odé, and Rien A. C. Dam, 931–939.
- Steinhauer, Hein. 1985. Number in Biak: counterevidence to two alleged language universals. *Bijdragen tot de Taal-, Land- en Volkenkunde* 141(4):462–85.

- . 1986. Number in Biak: counterevidence to two alleged language universals (a summary). In *FOCAL I: papers from the Fourth International Conference on Austronesian Linguistics*, ed. by Paul Geraghty, Lois Carrington, and Stephen Wurm, 171–3. Canberra: Pacific Linguistics (PL C-93).
- Teljeur, Dik. 1982. Short word-lists from South Halmahera, Kayoa, Makian, Ternate, Tidore and Bacan. In *The Makian languages and their neighbors*, ed., by C. L. Voorhoeve 129–148. Canberra: Pacific Linguistics (PL D-46).
- . 1983. Spatial orientation among the Gimán of South Halmahera. *Buletin Leknas* 2(2): 347–479.
- . 1990. *The symbolic system of the Gimán of South Halmahera*. Dordrecht: Koninklijk Instituut voor Taal- Land- en Volkenkunde.
- . 1994. Life-cycle rituals among the Gimán of South Halmahera. In *Halmahera and beyond: social science research in the Moluccas*, ed. by Leontine E. Visser, 165–194. Leiden: KITLV Press
- Thieberger, Nicholas. 2004. *Topics in the grammar and documentation of South Efate, an Oceanic language of Central Vanuatu*. PhD diss., Melbourne: Department of Linguistics and Applied Linguistics, University of Melbourne.
- Thomason, Sarah. G. 2001. Contact languages II: Other mixed languages. In *Language contact: An introduction* 196–221. Washington DC: Georgetown University Press.
- . 2003. Contact as a Source of Language Change. In *The handbook of historical linguistics*, ed. by Richard D. Janda and Brian D. Joseph. Malden: Blackwell Publishing.
- Tryon, Darrell ed. 1995. *Comparative Austronesian dictionary: An introduction to Austronesian studies*. 4 parts. Berlin: Mouton de Gruyter.
- van den Berg, Rene and Takashi Matsumura. 2008. Possession in Irarutu. *Oceanic Linguistics* 47(1):213–222.
- van den Heuvel, Wilco. 2006. *Biak, description of an Austronesian language of Papua*. The Netherlands: LOT Janskerhof 13 3512 BL. Utrecht. (Doctoral diss., Netherlands Graduate School of Linguistics.)
- van der Crab, P. 1862. *De Moluksche eiland*. Batavia: Lange.
- van der Leeden, A. C. 1980. *Preliminary notes on the language of Maya*, ms.
- . 1993. *Ma'ya*. Language study 97. Leiden: Lembaga Ilmu Pengetahuan Indonesia–Rijksuniversiteit Leiden.
- . 1995. *Ma'ya a language study a phonology*. Leiden: Lembaga Ilmu Pengetahuan Indonesia–Rijksuniversiteit Leiden.
- van Hasselt, F.J.F. 1902a. Gebruik van vermomde taal door de Nufooren. *Tijdschrift voor Indische Taal-, Land-, en Volkenkunde* 45.276-80.
- . 1902b. Nog iets over vermomde taal in het Nufoorsch en over Nufoorsche raadsels. *Tijdschrift voor Indische Taal-, Land-, en Volkenkunde* 45. 557-62.
- . 1905. *Spraakkunst der Nufoorsche taal*. Den Haag: M. Nijhoff.
- . 1936. Het Noemfoorsch als eenheidstaal op het Noordwestelijk deel van Nieuw Guinea. *Tijdschrift Nieuw-Guinea* 1.114–7.
- . 1947. *Biak–Dutch Dictionary*. Nieuw-Guinea Studie-Comité. Koninklijk Instituut voor de Taal- Land- en Volkenkunde.
- van Hasselt, J.L. 1868. *Allereerste beginselen der Papoesch-Mafoorsche taal*. Utrecht.
- . 1876. *Beknopte Spraakkunst der Noefoorsche taal*. Utrecht.

- . 1876/1893. *Hollandsch–Noefoorsch en Noefoorsch–Hollandsch woordenboek*. Utrecht.
- van Hasselt, J.L en F.J.F van Hasselt. 1947. *Noemfoorsch Woordenboek*. Amsterdam: N.V. Drukkerij Uitgeverij J.H. de Bussy.
- Visser, Leontine E. 1989. The Kamrau Bay area: between Mimika and Maluku. *Irian Bulletin of Irian Jaya* 17:65–76
- von Herder, Johann Gottfried. 1772. Treatise on the origin of language. In *Herder philosophical writings*, trans. and ed. by Michael N. Forster [2002], 65–164. Cambridge: Cambridge University Press (originally published as ‘Ursprung der Sprache’, available in *Herders Sämmtliche Werke*, ed. by Bernhard Ludwig Suphan. Band 5 [1891]. Berlin: Weidmannsche Buchhandlung.)
- Voorhoeve, Clements L. 1975a. *Languages of Irian Jaya: checklist, preliminary classification, language maps, wordlists*. Canberra: Pacific Linguistics (PL B-31).
- . 1975b. A hundred years of Papuan linguistic research: western New Guinea area. In *New Guinea area languages*, ed. by Stephen Wurm, 117–42.
- . 1989. Notes on Irarutu. *Irian, Bulletin of Irian Jaya* 16:106–119.
- . 1995. Irarutu: introduction and wordlist. In *Comparative Austronesian dictionary*, part 1: fascicle 1, ed. by Darrell Tryon, 651–657.
- Walker, Malcolm. 1990. Irian Jaya cultures: an overview. *Irian* 18:1–16.
- Walker, Roland. 1983a. *Fakfak Survey Report*, ms.
- . 1983b. *Waropen Survey Report*, ms.
- . 1990. *Koiwai–English Wordlist*, ms.
- . 1993. Language shift in Europe and Irian Jaya, Indonesia: Toward the heart of the matter. In *Case studies in minority languages*, ed. by Kees de Bot, 71–87. Amsterdam: Association International de Linguistique Appliquée.
- Walker, Roland and Michael Werner. 1978. *Bomberai Survey Report*, ms.
- Wanma, T. n.d. *Eerste beginselen van het Biaks*, ms.
- Warami, Hugo. n.d. *Sintaksis Bahasa Waropen*. Manokwari: Universitas Negeri Indonesia di Papua di Cenderawasih, ms. (printed by its author in 2006)
- Whisler, Jacqui. n.d. *A grammar of Sawai*. Pattimura University and Summer Institute of Linguistics, ms.
- Whisler, Ronald. 1992. Phonology of Sawai. In *Phonological studies in four languages of Maluku*, ed. by Donald A. Burquest and Wyn D. Laidig, 7–32. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- . n.d. *Sawai discourse notes*. Pattimura University and Summer Institute of Linguistics, ms.
- Whisler, Ronald, and Jacqui Whisler. 1995. Sawai: Introduction and wordlist. In *Comparative Austronesian dictionary*, part 1: fascicle 1, ed. by Darrell Tryon, 659–65.
- Wilson, William H. 2012. Hawaiian language revitalization. In *Language in Hawai‘i and the Pacific*. University of Hawai‘i at Mānoa Linguistics 100 Course Reader, ed. by Hiroko Sato and Jake Terrell, 118–29. Honolulu: University of Hawai‘i.
- Wurm, Stephen. 1975a. The application of the comparative method to Papuan languages: General and Highlands. In *New Guinea Area Languages*, ed. by Stephen A. Wurm, 237–261.
- . 1975b. Personal Pronouns. In *New Guinea Area Languages*, ed. by Stephen A. Wurm, 191–217.

- . 1982. Austronesian and non-Austronesian (Papuan) languages in contact: some notes. In *Gava` Studies in Austronesian languages and cultures*, ed. by Rainer Carle, Martina Heinschke, Peter Pink, Christel Rost, and Karen Stadlander, 88–109. Berlin: Dietrich Reimer Verlag.
- . 2000. *Ways and methods for maintaining and re-invigorating endangered languages*. (Paper presented at Conference on Endangered Languages of the Pacific Rim, Nov 24–25, 2000. Kyoto Conference Hall, Kyoto, Japan.)
- . 2003. The language situation and language endangerment in the greater Pacific area. In *Language death and language maintenance: Theoretical, practical and descriptive approaches*. Current Issues in Linguistic Theory, vol. 240, ed. by Mark Janse and Sijmen Tol, 15–48. Amsterdam: John Benjamins.
- . 2007. Australasia and the Pacific. In *Encyclopedia of the world's endangered languages*, ed. by Christopher Moseley, 425–578. London: Routledge.
- Wurm, Stephen, ed. 1975. *New Guinea Area Languages*. Vol. 1: Papuan Languages and the New Guinea Linguistic Scene. Canberra: Pacific Linguistics (PL C-38).
- Wurm, Stephen A. and Lois Carrington, eds. 1978. *Second international conference on Austronesian linguistics: Proceedings*. Canberra: Pacific Linguistics (PL C-61).
- Wurm, Stephen A. and Shiro Hattori. 1981. *Language atlas of the Pacific area*. Part 1: *New Guinea area, Oceania, Australia*. Canberra: Australian Academy of Humanities in collaboration with the Japan Academy.
- Wurm, Stephen A., C. L. Voorhoeve, and Kenneth A. McElhanon. 1975. The Trans New Guinea Phylum in general. In *New Guinea area languages*, ed. by Stephen Wurm, 299–322.
- Zec, Draga. 1994. *Sonority constraints on prosodic structure*. New York: Garland Publishing.
- Zorc, R. David. 1983. Proto-Austronesian accent revisited. *Philippine Journal of Linguistics* 14(1):1–24.
- Zwicky, Arnold M. 1985. Clitics and particles. *Language* 61(2):283–305.