A Grammatical Description of Hurukui, With Special Focus on Syntactic Gradience

by

Wesley Kuhron Jones

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in Linguistics

Dissertation Committee:

Don Daniels, Chair

Doris L. Payne, Core Member

Gabriela Pérez Báez, Core Member

Spike Gildea, Core Member

Vsevolod Kapatsinski, Core Member

Zhuo Jing-Schmidt, Institutional Representative

University of Oregon

Summer 2026

© 2026 Wesley Kuhron Jones

DISSERTATION ABSTRACT

Wesley Kuhron Jones

Doctor of Philosophy in Linguistics

Title: A Grammatical Description of Hurukui, With Special Focus on Syntactic Gradience

This dissertation describes the grammar of the Hurukui language of Papua New Guinea, in addition to a deep dive into syntactic gradience in the language's inventory of light verb constructions (LVCs). Part I consists of a sketch grammar, detailing the phonology, morphology, basic syntax, and some more complex syntax in Hurukui. Part II consists of the study of syntactic gradience. This study employs bottom-up quantitative methods to observe how the LVCs cluster into similar groups based on their properties.

Hurukui LVC lexical items exhibit a variety of phonological and morphosyntactic behaviors. These constructions often are not clearly either one word or two, with different criteria giving conflicting answers. The LVCs analyzed here provide a detailed case study of how multiple source constructions can give rise to complex predicates with different properties in synchrony. I use this case study to shed light on broader questions about the nature of linguistic categories, syntax, and cognition.

Acknowledgments

This dissertation is the largest project I have ever undertaken, and it is a miracle that I have completed it. I could never have come close to accomplishing such a thing without the assistance, in various ways, of many people in my graduate program and my personal life.

To Don Daniels, the best advisor I could have ever asked for. He read countless drafts of my papers and chapters that were littered with unresolved to-do items, jarring transitions, disorganized prose, nonsensical analyses, and obnoxious grandstanding, and despite all this, he consistently provided incredibly thorough, helpful feedback that has made me a better writer, thinker, and linguist. I have been proud to be his first advisee, and the experience has been unforgettable and will surely leave a lasting impression on my linguistic work, as well as my human traits, for the rest of my days. Without such a supportive, understanding, detail-oriented, organized, and clear-headed advisor, the task of completing a dissertation like this would have been impossible for me.

To the other members of my committee: Doris Payne, Gabriela Pérez Báez, Spike Gildea, Volya Kapatsinski, who provided not only professional and analytical help, but also personal advice, validation of the struggles I've felt during grad school and academia in general. Zhuo Jing-Schmidt, my institutional representative, also had helpful and insightful commentary on this thesis, and I thank her for it. All y'all's detailed and constructively critical comments on my work have helped me grow over the course of this Ph.D., and I will look back on them in the years to come for further insights into linguistic theory, fieldwork, literature engagement, argumentation, and surely other areas of wisdom that I can barely glimpse at this point in my life.

To various other linguists and linguistics students with whom I have had very interesting and productive discussions that have enriched my understanding and analysis of the topics addressed in this thesis, including but not limited to: Nathaniel Sims, Emma Mazzuchi, Sylvain Loiseau, Laura Arnold, Margaret Ransdell-Green, Gregory Anderson, Daniel Kraußе, Andrey Drinfeld, Eric Pederson, and Marianne Mithun.

To Dorothea Hoffmann and Willem de Reuse, who were very influential to me as a young linguist, and who encouraged me to pursue a Ph.D. focusing on language documentation and description. And to Frau Penrod, my high school German teacher, who recognized my love of and talent for languages, and who pushed me toward further development of my abilities by showing me such events as the University of Memphis's Foreign Language Fair and the North American Computational Linguistics Olympiad. These were incredibly formative experiences that solidified to me that I was meant to study linguistics.

To the Hurukui people, especially my host families: Miti Kapi and Ruben Joseph, Asel Giamaru and Yasakingo Yakam, Prisilla Bilaga† and Sebron Sende, Lynn Emmanuel and Patrick Suat, Yatap Winda, Nelson Emmanuel and Prisilla Nelson; and especially the consultants who work most closely with me: Asel Kapi, Pinias Titus, Lawrence Samuel, Nayamo Bilaga, Asel Giamaru, Kikingai Samson, Silas Emmanuel, Lynn Suat, Prisilla Bilaga†, and Yakam Asel. I would not have understood certain aspects of the grammar without the explicit explanation given by Hurukui people. Pinias Titus methodically explained the elevation-oriented deictic system to me by using props and moving to different places around them so that he could show me how the words used would change. Yakam Asel noticed that the marker *-ku* on deictics indicates location, while the deictic without *-ku* indicates direction. My host families taught me innumerable things, which I cannot list here, during casual conversations at home. Many consultants, as well as Hurukui people with whom I did not do language work, walked miles with me over treacherous ground, carrying my heavy bags, holding my hands because I cannot walk correctly on the trails, keeping me safe and well fed so that I could make it to my destinations and do the work I flew halfway across the world to do. You have all made me feel welcome and at home with you, and without all of your support and help none of this work would have been possible. *Ima moko isone piteko ekhu apa neka keku wore. Khaneito harawohe khai.*

To my mom, Cindy Jones, my dad, David Jones, and my sister, Evelyn Bailey Jones, with whom I have had long conversations over the phone and when visiting each other, and who have provided numerous hours of emotional support, encouragement, and wise advice. My family has supported me in my efforts and ambitions ever since it became clear in my youth that I was

destined to be obsessed with languages for the rest of my life. They have pointed me toward enriching activities and opportunities involving my interests and have always made it clear that what I wanted to do was worthwhile in their eyes because of how much I loved it. They have always lifted me up, expressed that they are proud of me, and encouraged me to follow these dreams and talents. Without that, I may never have majored in linguistics in the first place.

To those who assisted in my access to books and papers that would otherwise have been prohibitively expensive, I thank you for serving the cause of open and equitable access to knowledge.

To my good friends both within and outside of the University of Oregon linguistics program, including but not limited to Allison Taylor-Adams, Ludwig Adisiswoyo, Becky Paterson, Shayleen Macy, Xuan Guan, DeAndré Espree-Conaway, Dakota Whisler, Carla Consolini, and Shiloh Drake, who have been there for me while I have gone on at length about the woes and wonders of graduate school. Y'all's advice, listening, commiseration, celebration, and various other modes of support have been invaluable in making me certain that I was not alone in what I have gone through to get to this point, and in helping me get there as healthily and happily as I could.

To my cat, Mouse, my emotional support animal, the fluffiest softest cutest sweetest critter, for her nighttime cuddles, slow blinks, calming purrs, and reminders to take a break from stressful tasks to relax in the present moment, take naps in the sunlight, and eat snacks.

To Marxie, my dearest, who has provided copious words of encouragement and affection that helped melt away the stress of writing a dissertation. Spending time getting to know them deeply has been an enrichingly different flavor of life, a much-needed respite from the 24/7 mental noise of my brain worrying about academic tasks and deadlines. Their presence has reminded me of some of the things I value most in this life outside of what I have spent the last seven years most focused on. Their acceptance, curiosity, compassion, and love have been incredibly healing and rejuvenating, giving me mental and emotional energy without which I would have struggled much more to complete this project.

And to Nicola, *mekumlisun*, who has stood by my side and supported me throughout my graduate program. She moved across a continent with me seven years ago and soon will move across an ocean with me as I take the next step in my career as a linguist. She took over the cooking for the entire year while I was writing. She listened to me talk through various unorganized ideas on afternoon walks. She held space for my complaints and despair about this project, as well as my probably more-than-somewhat annoying grandiosity about the philosophy of categorization, among other ideas I was obsessed with during this time. She has helped me with understanding the rhetorical structure of literature reviews, promises to the reader, and other concepts about good writing that would otherwise completely elude my grasp. Certainly I have not done them justice here, and I continue to learn as we discuss more media together. She has also been incredibly emotionally supportive, cheering me on, validating my frustrations and struggles, reassuring me, and continually affirming her commitment to working together whatever the weather. Her steadfastness and reliability, together with her care and love, have been a solid, secure base from which I have been able to embark on my efforts. I owe a great deal of my accomplishments, and my happiness, to her presence in my life. *Sun mekemiur uraqamitilimia!*

This work was generously supported by the University of Oregon's Raymund Fellowship and NSF grants BCS-2048220 and BCS-2347759.

Table of Contents

1 Introduction: The Hurukui language, methods, and theoretical approach....23
1.1 Location and setting....23
1.2 Previous work....23
1.3 Language overview....23
1.3.1 Phylogeny....23
1.3.2 Typological profile and notable features....24
1.3.3 Dialects....24
1.3.4 Language status....25
1.4 Data collection methods and corpus....25
1.4.1 Field trips....26
1.4.2 Text types....26
1.4.3 Elicitation....26
1.5 Theoretical background of this description....26
1.5.1 Determining the meanings of forms....27
1.6 Overview of the thesis....30
2 Phonology....32
2.1 Consonants....32
2.1.1 Inventory and allophony....32
2.1.2 The status of the alveolar obstruents *t*, *nd*, and *s*....35
2.1.3 Minimal pairs for consonants....37
2.2 Vowels....40
2.2.1 Inventory and allophony....40
2.2.2 Vowel sequences....42
2.2.3 Minimal pairs for vowels and diphthongs....43
2.3 Syllable structure....44
2.4 Stress....46

2.5 Vowel reduction processes....49
2.5.1 Deletion....49
2.5.2 Monophthongization....52
2.6 Anticipatory semivowel insertion....53
2.7 Prosody....54
2.8 Orthography....55
3 Nominals....57
3.1 Nouns....58
3.1.1 Common nouns....59
3.1.2 Proper nouns....65
3.2 Pronouns....67
3.2.1 Personal pronouns....67
3.2.2 Emphatic pronouns....69
3.2.3 Vocative pronouns....71
3.2.4 Interrogative pronouns....72
3.3 "Whatchamacallit" words....73
4 Deictics....75
4.1 Deictic bases....76
4.1.1 Non-oriented bases....77
4.1.2 Participant-oriented bases....78
4.1.3 Elevation-oriented bases....78
4.1.4 Uses of deictic bases as independent words....81
4.2 The proximal prefix *aka-* ~ *ako-* ~ *ak-*....84
4.3 The demonstrative prefix *khai-* ~ *kha-* ~ *kh-*....85
4.4 The plural prefix *khe-*....89
4.5 The locative suffix *-ku* ~ *-ngu*....90
4.6 Deictic elements in 'do thus' verbs....92
5 Verbs....94

5.1 Preliminaries....95
5.1.1 Morphophonemic notation....95
5.1.2 Persons and numbers....96
5.1.3 The medial-final verb distinction....97
5.1.4 Roots, stems, and null morphemes....98
5.1.5 Realis and irrealis status....99
5.1.6 The morphological structure of the Hurukui verb....100
5.1.7 The alternating shapes of verbal morphemes....101
5.2 Verb stem formation....105
5.2.1 The repetitive....106
5.2.2 Object marking....107
5.2.2.1 Object suffixes....107
5.2.2.2 The reciprocal....110
5.2.2.3 The 'affect' verb....110
5.2.2.4 The 'give' verb....113
5.2.3 Adverbial stem formants....115
5.2.4 Aspect marking....117
5.2.4.1 The perfect....117
5.2.4.2 The distal past....120
5.2.4.3 The desiderative....122
5.2.4.4 The continuative....124
5.2.4.5 The unmarked aspect....125
5.2.5 Negation....126
5.2.6 The conditional....127
5.3 Inflection of final verbs....129
5.3.1 Realis marking....129
5.3.2 The future tense....133
5.3.3 The negative future in *-ome*....136

5.3.4 The imperative....138
5.3.5 The potential markers *-o* and *-Ka*....140
5.3.6 The negative habitual with possessive marking....141
5.4 Inflection of medial verbs....142
5.4.1 Realis same-subject....143
5.4.2 Irrealis same-subject....144
5.4.3 Serial marking....145
5.4.4 Different-subject....146
5.4.5 Simultaneous switch reference....148
5.4.6 Sequential switch reference....149
5.4.7 Switch reference combined with other TAM....152
5.4.7.1 The perfect plus switch reference....152
5.4.7.2 The distal past plus switch reference....153
5.4.7.3 The desiderative plus switch reference....154
5.4.7.4 The continuative plus switch reference....155
5.4.8 Switch reference combined with negation....155
5.5 Subordinating morphology....156
5.5.1 The purposive *-Ka*....156
5.5.2 The adjectivizer *-(i)khu*....157
6 Other word classes....159
6.1 Adjectives....159
6.1.1 Reduplication of adjectives....160
6.1.2 Intensive suffix *-ko* ~ *-ngu*....161
6.1.3 Adjectives used nominally....162
6.2 Quantifiers....163
6.2.1 Numerals....163
6.2.2 General quantifiers....164
6.2.3 Indefinite quantifiers....165

6.2.4 Interrogative quantifiers....166
6.3 Adverbs....166
6.4 Adpositions....168
6.5 Particles....168
6.5.1 Negative particles....169
6.5.2 Topic particles....170
6.5.3 The contrastive focus marker *kha*....171
6.5.4 Intensifying particles....174
6.5.5 Emphatic particles....177
6.5.6 Question particles....179
6.5.7 Polar answer particles....179
6.6 Connectives....180
6.7 Ideophones....181
7 Syntax of the clause....183
7.1 Phrases....183
7.1.1 Noun phrase....183
7.1.1.1 Noun-noun possession construction....186
7.1.1.2 Noun compounds....187
7.1.2 Adjective phrase....189
7.2 Nonverbal clauses....190
7.2.1 Copular clauses....190
7.2.2 Negative existential clauses....192
7.3 Verbal clauses....193
7.3.1 Core, oblique, and bare....194
7.3.2 The definite null test for covert 3SG objects....196
7.3.3 Argument and adjunct....198
7.3.4 Arguments....201
7.3.5 Oblique adjuncts....203

7.3.6 Bare adjuncts....204
7.3.7 Adjuncts of uncertain obliquity....205
7.3.8 Experiencer predicates....206
7.3.8.1 Stimulus-affect experiencer predicates....207
7.3.8.2 Possessive experiencer predicates....208
7.3.9 Placement of adverbial modifiers....209
7.3.10 Periphrastic passive....210
7.3.11 Light verb constructions....211
7.4 Interrogative clauses....214
7.4.1 Polar questions....214
7.4.2 Content questions....215
7.4.3 Intonation with question-final particles....215
8 Multi-verb constructions....217
8.1 Clause chains....217
8.1.1 Basic use of switch reference....218
8.1.2 Referential overlap....221
8.1.3 Relative tense....221
8.1.4 Polarity within chains....224
8.1.5 Mood within chains....226
8.1.6 Shared object marking....227
8.1.7 Medial nonverbal predicates....228
8.1.8 Interaction of experiencer predicates with switch-reference....229
8.1.9 Clause chains expressing manner....231
8.1.10 Clause chains expressing periphrastic aspect....232
8.1.10.1 Clause chains expressing durative....232
8.1.10.2 Clause chains expressing perfective....233
8.1.11 Clause chains expressing periphrastic causative....234
8.2 Clause juxtaposition....234

8.2.1 Comparative clause juxtaposition....235
8.2.2 Imperative clause juxtaposition....235
8.3 Serial verb constructions....235
8.4 Subordinate clauses....236
8.4.1 Complement clauses....236
8.4.1.1 Direct speech construction....238
8.4.1.2 Internal speech construction....238
8.4.1.3 Desiderative same-subject using internal speech construction....239
8.4.2 Relative clauses....241
8.4.2.1 Internally headed relative clauses....241
8.4.2.2 Headless relative clauses....242
8.4.3 Adverbial clauses....242
8.4.3.1 'While' clauses with continuative final verb....243
8.4.3.2 Conditional clauses with conditional marker....244
8.4.3.3 Conditional clauses with future final verb....244
8.4.3.4 Purposive clauses....245
9 Introduction: Gradience, categorization, and complex predicates....247
9.1 Motivation and point of departure of this study....247
9.2 Gradience in morphosyntax....251
9.3 Complex predicates and light verb constructions....255
10 Properties for diagnosing univerbation of light verb constructions....258
10.1 Terminology....258
10.2 Phonological and morphophonological properties....259
10.2.1 Stress positions and quantities....259
10.2.2 Vowel raising domain....260
10.2.3 Verb elision....260
10.3 Morphological properties....261
10.3.1 Possessor on initial....261

10.3.2 Object on verb....262
10.3.3 The possessor/object alternation....263
10.3.4 Initial reduplication....265
10.3.5 Verb reduplication....267
10.4 Syntactic properties....267
10.4.1 Modifier insertion....268
10.4.2 Time adverb insertion....269
10.4.3 Anaphora insertion....270
10.4.4 Focused argument insertion....270
10.4.5 PP insertion....272
10.4.6 Initial independence....274
10.4.7 Initial used clearly as noun....275
10.4.8 Initial modification by adjective when independent from verb....276
10.4.9 Initial used clearly as adjective....276
10.4.10 Topic/focus marking of initial....277
10.4.11 Verb alone in response (ellipsis)....279
10.5 Lexical and semantic properties....280
10.5.1 Variation in verb selection....280
10.5.2 Verb retaining some of its lexical semantics....281
11 Individual LVCs spanning the space....283
11.1 Unfused LVC: *moko e-* 'speak'....283
11.2 Fused LVC: *sáhi-* 'fix'....294
11.3 Partially fused LVC: *aipV e-* 'cry'....302
11.4 Partially fused LVC: *(h)ambukha i-* 'like'....314
11.5 Summary....323
12 Quantitative mapping of LVC space....327
12.1 Data encoding....327
12.1.1 Multicollinearity and logical dependency among features....327

12.1.2 One-hot encoding scheme....327
12.1.3 N/A values....329
12.1.4 Splitting probability among multiple indistinguishable analyses....330
12.2 Imputation....331
12.3 Principal component analysis....332
12.4 Methods using distance measurements....342
12.4.1 Weighted distance metric....342
12.4.2 Multi-dimensional scaling....343
12.5 Clusters within the space....345
12.6 Discussion....346
12.6.1 Discussion: Light verb constructions and similar constructions....347
12.6.1.1 Relationships of the LVCs within speakers' mental grammars....348
12.6.2 Discussion: Categorization and gradience....350

List of figures

Figure 4.1: Elevation-based deictic roots, excluding *we-* 'alongland'.....78
Figure 8.1. Relative time meanings.....223
Figure 12.1: VCs with PC1 (morphosyntactic univerbation) on the x-axis and PC2 (argument structure) on the y-axis.....339
Figure 12.2: VCs with PC1 (morphosyntactic univerbation) on the x-axis and PC3 (morphophonemic univerbation) on the y-axis.....340
Figure 12.3: VCs with PC2 (argument structure) on the x-axis and PC3 (morphophonemic univerbation) on the y-axis.....341
Figure 12.4: MDS projection of the VCs using the variance-explained-weighted distance metric 344

List of tables

Table 2.1: Consonant phonemes of Hurukui....32
Table 2.2: Velar palatalization before front vowels....34
Table 2.3: The velar-alveolar alternation....36
Table 2.4: Vowel phonemes of Hurukui....40
Table 2.5: Vowel sequences and diphthongs....42
Table 2.6: Examples of syllabification....44
Table 2.7: Syllabification of diphthongs....45
Table 2.8: Syllabification of non-diphthong vowel sequences....45
Table 2.9: Examples of regular stress assignment....47
Table 2.10: Examples of verb stems marked with floating final stress....47
Table 2.11: Stress assignment in univerbated forms....48
Table 2.12: Examples of stress assignment in proper names....48
Table 2.13: Word-final vowel deletion....49
Table 2.14: Word-final vowel deletion in univerbating phrases....50
Table 2.15: Word-initial vowel deletion....50
Table 2.16: Phonetic minimal pairs for lower and higher allophones of mid vowels....51
Table 2.17: Multiple options for vowel deletion in certain phrases....51
Table 2.18: Regular monophthongization *ai* > *a*....52
Table 2.19: Sporadic monophthongization *ai* > *a*....52
Table 2.20: Anticipatory insertion of [j]....53
Table 2.21: Anticipatory insertion of [w]....54
Table 2.22: Examples of the multi-word primary stress domain....54
Table 2.23: Utterance-final lengthening and shortening....55
Table 3.1: Possessive affixes on inalienable nouns....60
Table 3.2: Possessive paradigm of *opopo* 'food'....64
Table 3.3: Some ambialienable nouns....65

Table 3.4: Possession of common nouns via NP juxtaposition....66
Table 3.5: Inclusory and appositive meaning involving proper nouns via NP juxtaposition....67
Table 3.6: Personal pronouns....68
Table 3.7: Emphatic pronouns....69
Table 4.1: Position class template for deictics....75
Table 4.2: Elevation-oriented deictic bases....81
Table 4.3: Correspondence between some deictics and their proximal counterparts....84
Table 5.1: Morphophonemic notation for the velar-alveolar alternation....95
Table 5.2: Position class template for most Hurukui verbs....101
Table 5.3: Named ablaut types....104
Table 5.4: Alternating shapes for some verbal morphemes....105
Table 5.5: The normal object marking paradigm....108
Table 5.6: Some verb roots and their forms with *-(i)ki* '23PL.OBJ'....109
Table 5.7: Object-marked fusional roots of 'affect'....111
Table 5.8: Object-marked fusional roots of 'fall down'....113
Table 5.9: Object-marked fusional roots of 'give'....114
Table 5.10: The realis perfect conjugation paradigm, with its realis completions....118
Table 5.11: Correspondence between forms of lexical *khai* 'be located' and the perfect....120
Table 5.12: The distal past conjugation paradigm, with its realis completions....121
Table 5.13: Some verb stems and their polite imperative forms....128
Table 5.14: The realis conjugation paradigm....130
Table 5.15: Allomorphy of the 1DU.R suffix *-rou ~ -reu*....132
Table 5.16: Verbs with irregular realis conjugations....132
Table 5.17: Combination of the continuative with the realis....133
Table 5.18: The future tense conjugation paradigm....134
Table 5.19: Regular and irregular future forms of *we-* 'go'....135
Table 5.20: Combination of the continuative with the future....136
Table 5.21: The negative future paradigm in *-ome*....137

Table 5.22: The imperative paradigm....138
Table 5.23: The negative habitual paradigm with possessive marking....142
Table 5.24: Realis same-subject conjugation paradigm....143
Table 5.25: Some verb stems and their SS.IRR forms....144
Table 5.26: Some verb stems and their serial forms....145
Table 5.27: Different-subject conjugation paradigm....147
Table 5.28: Some verb stems and their SS.SIM forms....148
Table 5.29: Simultaneous different-subject conjugation paradigm....149
Table 5.30: Same-subject sequential conjugation paradigm....150
Table 5.31: Different-subject sequential conjugation paradigm....151
Table 5.32: Inflecting combinations of perfect and switch reference suffixes....153
Table 5.33: Combinations of distal past and switch reference suffixes....154
Table 5.34: Desiderative different-subject conjugation paradigm....154
Table 5.35: Different subject negative conjugation paradigm....156
Table 6.1: Occurrence of *kha* 'FOC' as a suffix versus as a particle....171
Table 7.1: Headedness and intonation in noun compounds....188
Table 7.2: Arguments, oblique adjuncts, and bare adjuncts....200
Table 7.3: Examples of stimulus-affect EPs....207
Table 7.4: Examples of possessive EPs....208
Table 11.1: Properties of the LVC *moko e-* 'speak'....293
Table 11.2: Properties of the LVC *sáhi-* 'fix'....301
Table 11.3: Properties of the LVC *aipV e-* 'cry'....313
Table 11.4: Properties of the LVC *(h)ambukha i-* 'like'....322
Table 11.5: Properties of the LVCs detailed in Chapter 11....324
Table 12.1: Stress properties of selected LVCs, summary table format....328
Table 12.2: Stress properties of selected LVCs, one-hot encoding format....328
Table 12.3: Most heavily weighted features of the first three principal components....335
Table 12.4: Locations of the most extreme and middlemost VCs on the first three PCs....336

Table 12.5: Sample distances between VCs in variance-explained-weighted PCA space.....343

Abbreviations

1	first person
2	second person
3	third person
ADJZR	adjectivizer
AGT	agentive
AVC	auxiliary verb construction
COMIT	comitative
COMPL	completive
COND	conditional
CONT	continuative
DEM	demonstrative prefix
DES	desiderative
DISC.PRTC	discourse particle
DPST	distal past
DS / ds	different-subject
DU	dual
EMPH	emphatic
EXCL	exclusive
EXIST	existential
FOC	focus
FUT	future
GENER	generic form of inalienable nouns
HAB	habitual
HORT	hortative
IMP	imperative
INCL	inclusive
INDEF	indefinite
INSTR	instrumental
INTENS	intensive
INTR	intransitive
IRR	irrealis
LOC	locative
LV / lv	light verb
LVC	light verb construction
MDS	multi-dimensional scaling
MICE	multiple imputation by chained equations
n-p	none-primary stress pattern
NEG	negative
NP	noun phrase
NSG	non-singular
OBJ	object
p-n	primary-none stress pattern
p-p	primary-primary stress pattern
p-s	primary-secondary stress pattern
PC	principal component
PCA	principal component analysis
PFV	perfective
PL	plural
PNG	Papua New Guinea
POL	polite
POSS	possessive
POT	potential
PP	adpositional phrase
PRF	perfect
PROH	prohibitive
PROX	proximal
PUR	purposive
Q	question marker
R	realis
RECIP	reciprocal
RELZR	relativizer
REP	repetitive
s-p	secondary-primary stress pattern
SEQ	sequential
SER	serial
SG	singular
SIM	simultaneous
SR	switch reference
SS / ss	same-subject
SVC	serial verb construction
TAG.Q	tag question marker
TAM	tense/aspect/mood
TOPIC	topic
TR	transitive
VAC	verb-adjunct construction
VC	"verb construction", a generalization across VOCs, VACs, LVCs, etc.
VOC	vocative
VOC	verb-object construction

1 Introduction: The Hurukui language, methods, and theoretical approach

This chapter is structured as follows. §1.1 describes where Hurukui is spoken and the social setting. §1.2 lists previous work on the language, of which there is not much. §1.3 gives an overview of the nature of the language's features, dialects, and status/vitality. §1.4 describes the nature of the data on which this dissertation is based. §1.5 describes the functionalist, usage-based theoretical perspective from which I approach linguistic analysis and description, including my methods for finding out the meanings of forms based on what communicative functions they are used in. §1.6 gives an overview of the remainder of this dissertation.

1.1 Location and setting

Hurukui (ISO 639-3: [gsp]) is spoken by roughly 1,500 people in Morobe Province, Papua New Guinea. It has had significant contact with Iyo (its distant relative, in the Finisterre-Huon branch of Trans-New Guinea) and Adzera (Austronesian), but it is geographically separated from its closer relatives.

1.2 Previous work

In the 1970s, an SIL linguist named Kenneth McElhanon worked with a Hurukui schoolteacher named Sigkepe Sogum (whose given name in the modern orthography would be spelled *Singepe*). Together they produced a word list and a grammar sketch of 8 pages (McElhanon 1975, McElhanon & Sogum 1976). This work was done in the city of Lae, not on Hurukui land.

1.3 Language overview

1.3.1 Phylogeny

Hurukui belongs to the Rai Coast subgroup of the Madang branch of Trans-New Guinea (Z'graggen 1975, Pawley & Hammarström 2018). It is geographically isolated from all other Madang languages, located in the far southeast of the family's range. Further classification

remains elusive in the absence of detailed data on other Madang languages, especially those located the farthest to the east (in particular, the Yaganon group of Rai Coast).

1.3.2 Typological profile and notable features

Hurukui is a typical Madang language. The phonology is simple, with a system of five cardinal vowels and fifteen consonants. Syllables are almost all open, and there is a strong stress accent which regularly assigns stress to the penultimate syllable of a word. Unstressed vowel deletion is very common, especially at the ends of words. The morphology is mostly agglutinating, though fusion is common when person marking is involved (in verb suffixes as well as some verb roots that are fusional or suppletive for their object). The syntax is strongly verb-final, with complex verb morphology and little nominal morphology. It makes heavy use of clause chains involving switch reference, a coordination device common in Papuan languages (Roberts 1988; Roberts 1997: 102; deVries 2010; Fedden 2012; van Gijn 2016). Relative clauses are either internally headed or headless. The verbal lexicon is heavily enriched by light verb constructions, on which a large portion of this dissertation is focused.

1.3.3 Dialects

While a detailed study of sociolinguistic and dialectal variation in Hurukui remains to be done, some broad generalizations can be made. The eastern villages of Bawa, Rumpa, and Mamia all speak roughly the same dialect, which I call the Eastern dialect. The western villages of Wawiwi and Mapangomo speak more distinct varieties. Within Wawiwi, nominal vocabulary tends to vary by clan. Gusap has a large inventory of nominal vocabulary items that are different from all the other villages. Wawiwi and Gusap are also united in having some slightly different morphology than the eastern villages. There is noticeable variation in the vowel space among speakers, especially between the older and younger generations, although all speakers still have five vowel phonemes.

1.3.4 Language status

The language is used widely throughout the tribe, and nearly everyone is bilingual in Hurukui and Tok Pisin, to some extent, from birth. Children are more likely to speak Hurukui in the villages without primary schools, because primary school attendance often involves interaction with large numbers of children from outside the tribe, with whom Tok Pisin is spoken.

1.4 Data collection methods and corpus

The data source for Hurukui is mostly field data collected by the author, which is archived and publicly accessible in PARADISEC (Jones 2023). The parenthetical after each natural-speech example in this dissertation shows the source text and timestamp, e.g., "(PainimNaip 0:16-0:20)". I also draw Hurukui data from McElhanon (1975) and his work with Sigkepe Sogum, a Hurukui speaker (McElhanon & Sogum 1976). However, there is not much documentation there other than a word list and a grammar sketch of 8 pages.

The corpus consists of 316 natural-speech texts with a total length of over 25 hours, but this number is increasing as Hurukui speakers continue to record and transcribe new texts themselves while I am not in PNG. Most of this grammatical description is based on my first two field trips, which provided me with about 20 hours of natural-speech texts. I have had the time to analyze these texts more thoroughly than those from the two subsequent field trips. My third field trip, in Spring 2024, focused on elicitation of properties of light verb constructions and similar constructions. I conducted 42 elicitation sessions of about an hour each.

Every natural-language text has a video associated with it. I only recorded audio, not video, during elicitation sessions. Transcription sessions were not recorded, except for a few on my first field trip, in Summer 2021. I decided that the written information from transcription was enough for my purposes, since the audio did not provide nearly as much extra information as full elicitation sessions do, and so I stopped recording transcriptions to conserve battery power. As a result, transcription sessions yielded only handwritten notes in addition to the transcripts themselves.

1.4.1 Field trips

I conducted fieldwork for this dissertation for a total of seven months divided among four trips. The first was in 2021, when my fieldwork only lasted four weeks because of precautions and restrictions related to the COVID-19 pandemic. The second trip was in 2023 and lasted three months. Most of the natural-speech corpus comes from this trip. The third trip was in 2024 and lasted about six weeks. Most of the elicitation data about light verb constructions comes from this trip. The fourth and most recent trip was in 2025 and also lasted about six weeks.

1.4.2 Text types

The texts in the corpus are of a variety of genres. These include narrative monologues, conversations, interviews in which one speaker asks questions to another about a topic, procedural texts in which a speaker demonstrates how to do something while narrating (or having another speaker narrate), and others.

1.4.3 Elicitation

Much of the data used in this dissertation also comes from elicitation sessions. In these, I first attempted to prompt consultants to produce forms I was looking for or express meanings I was looking for, without using the forms myself. If they did not, I tried producing forms myself to see if they would repeat it and accept it. I did not count a form as acceptable if consultants responded that it was okay but did not repeat it themselves.

1.5 Theoretical background of this description

I approach grammatical description and analysis from a functional, usage-based standpoint, as detailed by such authors as Givón (2001) and DeLancey (2001), and Bybee & Beckner (2012). This approach incorporates knowledge about diachrony and typology while also remaining true to the idiosyncrasies of each language. Such a framework is called "panchronic", where diachrony matters just as much as synchrony (Kuteva 2001: 7). Accordingly, a large part of how I think about a language's synchronic state involves gradience between categories, which exists throughout linguistic systems because they are in flux over time (Traugott & Trousdale

2010). My description thus prioritizes accurate representation of categorical distinctions as well as gradience where appropriate.

My approach is heavily informed by Construction Grammar (Goldberg 1995; Croft 2001; Goldberg 2005), as well as Diachronic Construction Grammar (Fried 2009; Barðdal et al. 2015; Hilpert 2021), which deals with how constructions are created and change over time. Because the theory of Construction Grammar is monostratal (Goldberg 1995: 7), I posit no underlying forms that are transformed into surface forms, neither in phonology, nor morphology, nor syntax. Instead, I treat the surface form as the only level of language structure, and I posit internally reconstructed earlier forms to capture some generalizations about the behaviors of modern forms (Givón 2000). I consider this to be operationally similar (if not equivalent) to what is done by positing underlying forms, but theoretically more grounded in the reality of how language works in both the mind and across time. Additionally, while some researchers treat grammatical systems as having core phenomena and relegate gradience to things on the "periphery", usage-based and construction-based approaches instead treat all successful communication as equally valid within the linguistic system (McClelland & Bybee 2007).

1.5.1 Determining the meanings of forms

I often speak of what a form *means*. I think of a *meaning* as some idea or intention that a speaker is trying to express, which exists outside of the linguistic system, and for the expression of which the language is just one available medium. A meaning could be a function that an utterance has in terms of getting ideas into the listener's head, such as an image of a large pig that can be found in the nearby jungle. The interlocutors can compare these ideas with things in the real world, and they can determine how well the linguistic forms used fit with those ideas (i.e., how apt the forms were to express the ideas). This is semantic meaning: descriptions of the world. Another kind of meaning consists of communication that seeks to do something, to get someone to do something, because of social reasons or for goals in the world. An example would be the idea that the speaker wants the listener to go hunt the pig. This could be called social or pragmatic meaning, as it is not merely a description of the outside world as modeled in the speaker's mind, but it consists of information about the speaker's mind itself. The way a linguistic form is mapped

from the linguistic symbolic system to something else, e.g., some representation of the world is or something that needs to happen in the world, a thought, etc., is conventionalized, and I say that a form *means* (symbolizes) a *meaning* if this relation holds. All of this, broadly, is the concept of meaning.

Different languages do this mapping of forms to meanings in different ways, partitioning the conceptual space differently. This creates a problem in that transating text from one language to another does not preserve the available set of meanings that could be understood from the text itself, as recognized by Catford (1965: 35). The English and Tok Pisin translations used to create this grammatical description are only able to map to meanings in their own way, and they do so differently than Hurukui does. Each language treats certain groups of meanings the same in terms of what one says to encode them, but which sets these are varies across languages. There is a coarse-grainedness of linguistic encoding of portions of the meaning space, unless one spends an infinite amount of time describing things in ever finer detail.

Because of this fallibility of translation for getting at meaning, I use a variety of methods intended to elucidate more clearly the ideas themselves that speakers are trying to express. Sometimes I get the meaning from looking at what similar forms are used to mean in other situations. For instance, the verb root *ekha-* 'descend' occurs in *sono ekha-* (water descend-) 'bathe', *maha ekha-* (ground descend-) 'fall', and *ekha- yo-* (descend- come-) 'come down'. The lexeme can also be used to express a price going down. So this root is used in a lot of different situations and does not have one atomic meaning in conceptual space. In fact, no linguistic form has such a meaning. The set of meanings mapped to by a form is a cluster of conceptual points, occupying a subset of conceptual space. So I create approximations based on generalization over all those microscopic meanings, from knowing what context the forms were used in. Often I was there in the context or being told the story directly, and speakers would explain to me what was happening in the story as we transcribed it. I tried to understand what was actually going on in the story so I could know what kind of thing they were describing as happening in the world. Often the Tok Pisin and English translations were very rough at trying to encode the nuance of the Hurukui forms. I thus did not rely on translations to elucidate meaning, but speakers and I

created translations as ways of approximating what was being depicted or told to someone, in lieu of other representations such as direct depictions.

A domain in which I had to rely on a large amount of information other than translations is the deictics (Chapter 4), which involve information about closeness to the speaker and listener as well as geographic information involving elevation changes in the landscape. I have video information for all texts, showing gestures and the background. I know the location where each story was located (often whose house), and I have GPS coordinates for most places that I've stayed on Hurukui land (within approximately 12-15 yards' accuracy, as measured by an app on my smartphone). I have a topographic map of the area. I know where named places and things are located related to each other, e.g., uphill, downhill, upriver, downriver, etc. The deictics are used very frequently, for most mentions of places, directions, or paths (e.g., 'I went home uphill').

I made up a novel elicitation task while in PNG to get at the deictic meanings. I called it an *object placement task*. In that task, I arranged some items in someone's house in an unusual way. For instance, I put a chair upside-down on the floor, hung a bag on it, put a cucumber in the bag. Then I took a picture of it and dismantled it, the speakers not having seen what I had done. Then I gave them a task in which one person sees the picture and the other person is controlling the objects. The one looking at the picture has to the tell the other where to put the objects to recreate the picture. This task elicited a lot of use of deictics, and I collected five texts like this.

During transcription of those texts, a lot of my understanding of the deixis comes from a speaker with good metalinguistic awareness. We were transcribing these texts, and he was the only one who had showed up at that time. I told him I didn't understand the difference between all the different directional words, so he explained it all to me. He sat me down at the table where we were transcribing, and then he put things on the table in certain places. He explained to me what each of us would say about where the object is located, given where it was and given where we two were relative to the nearby landforms. He did the same thing from perspectives on all four sides of the table because this changed all the deictics depending where we were relative to each other, the object, and the landscape. I wrote all of this information down and drew pictures in my notebook, along with some notation about what he had said in each position. A lot of my

understanding of the deictic system comes from those texts and those interactions. These methods show my approach to understanding what the Hurukui forms mean, especially in cases such as deixis and TAM where the English and Tok Pisin translations do not do justice to the richness of the Hurukui semantics due to differences in how the three languages carve up conceptual space into sets expressed by similar forms.

1.6 Overview of the thesis

This dissertation is divided into two parts. Part I, the sketch grammar, contains Chapters 2 through 8. Chapter 2 is about the phonology. Chapter 3 is about the word class of nominals. Chapter 4 is about the word class of deictics. Chapter 5 is about the word class of verbs. Chapter 6 is about all other word classes, including adjectives, adverbs, ideophones, and the wastebasket label of particles. Chapter 7 is about the syntax of phrases and individual clauses. Chapter 8 is about multi-verb constructions, including clause chains, auxiliary verb constructions, and serial verb constructions.

Part II, the study of syntactic gradience, focuses on the variety of phonological and morphosyntactic behaviors in light verb constructions (LVCs) in Hurukui. These constructions often are not clearly either one word or two, with different criteria giving conflicting answers. I use this case study to shed light on broader questions about the nature of linguistic categories, syntax, and cognition. Chapter 9 opens the discussion with a review of literature on the topics of gradience in grammar and complex predicates. Chapter 10 then defines the properties I used to characterize Hurukui LVCs. Chapter 11 illustrates the values these properties take for four hand-picked LVCs that have different degrees of univerbation. Chapter 12 uses quantitative methods to show how the LVCs vary within the high-dimensional space defined by the properties, and interprets what the results can tell us about more broadly applicable questions about language.

Part I: Hurukui Sketch Grammar

2 Phonology

This chapter describes the phonology of Hurukui. §2.1 discusses consonants. §2.2 discusses vowels, vowel sequences, and diphthongs. §2.3 discusses syllable structure. §2.4 deals with stress, which is prominent in the language and factors into other phenomena such as reduction and univerbation. §2.5 discusses some processes of vowel reduction. §2.6 describes a coarticulatory process of anticipatory semivowel insertion. §2.7 deals with prosody, and §2.8 briefly shows the practical orthography and how it differs from the transcriptions used elsewhere in this dissertation.

Asterisks are used to indicate internally reconstructed forms if preceded by the phrase "Pre-Hurukui", and they indicate unacceptable forms otherwise.

2.1 Consonants

2.1.1 Inventory and allophony

The consonant inventory of Hurukui is shown in Table 2.1. For each phoneme in the table, IPA is shown first (e.g., /pʰ/), and orthography is shown in italics (e.g., *ph*).

Table 2.1: Consonant phonemes of Hurukui

	bilabial	alveolar	palatal	velar	glottal
nasal	/m/ *m*	/n/ *n*			
prenasalized stop	/mb/ *mb*	/nd/ *nd*		/ŋg/ *ng*	
plain stop	/p/ *p*	/t/ *t*		/k/ *k*	
aspirated stop or fricative	/pʰ/ *ph*	/s/ *s*		/kʰ/ *kh*	/h/ *h*
approximant or flap	/w/ *w*	/ɾ/ *r*	/j/ *y*		

The phonemes *m*, *n*, *mb*, *nd*, *p*, *t*, *s*, and *h* have no noticeable allophony. The prenasalized stops *mb*, *nd*, and *ng* are always prenasalized, never voiced without prenasalization (which is common word-initially in other Madang languages).

The aspirated stops *ph* and *kh* are often realized as fricatives [ɸ] and [x]. This lenition is optional word-initially and obligatory word-medially. For *kh*, this lenition change is nearly complete, but not entirely; some speakers still have intervocalic [kʰ] in careful speech.

phai	[pʰá.i] [ɸá.i]	'knife'
epha	[é̱ɸa]	'hear'
khai	[kʰá.i] [xá.i]	'it is located'
ekha	[é̱xa]	'bitter'

Word-initial fricative realization is also obligatory if the next consonant is the same phoneme, as in *phaphi* [ɸaɸi] 'cover' and *khakhi* [xaxi] 'salt'.

The plain stops *p*, *t*, and *k* may become somewhat aspirated word-medially. This is because the aspirated stops *ph* and *kh* are almost always realized as fricatives word-medially, so speakers know that a word-medial [pʰ]/[kʰ] is actually *p/k* rather than *ph/kh*. This is especially common if the plain stop is the second consonant in the word, following an initial aspirated stop.

khakongu	[kʰakó̱ŋgu] [kʰakʰó̱ŋgu] [xakʰó̱ŋgu] *[xaxó̱ŋgu]	'that one right there'
khaipi	[kʰájpi] [kʰájpʰi] [xájpʰi] *[xájɸi]	'they are located'

The phoneme *k* is sometimes realized as [ɡ], [ɣ], or [ɯ̟] intervocalically. The conditions determining which of these allophones are possible in which environments require further investigation, and there is a lot of inter-speaker and intra-speaker variation in this regard.

Hurukui	[huɾukú.i] [huɾugú.i] [huɾuɣú.i] *[huɾuɯ̟ú.i]	'Hurukui'
phúkekòmu	[ɸúkɐ̟kɔ̟̀mu] [ɸúgɐ̟gɔ̟̀mu] [ɸúɣɐ̟ɣɔ̟̀mu] [ɸúɯ̟ɐ̟ɯ̟ɔ̟̀mu]	'it will break'

The velar stops are often postvelar, perhaps even uvular, but this does not contrast with velar articulation. For instance, *khai* 'it is located' is sometimes [k̟ʰá.i ~ qʰá.i], and *nga* 'where?' is sometimes [ŋa].

The velar stops *k*, *kh*, and *ng* are palatalized before the front vowels *i* and *e*. Some examples are shown in Table 2.2.

Table 2.2: Velar palatalization before front vowels

phonological form	**surface pronunciation**	**meaning**
***k**eku*	[kʲɐ̟́ku]	'LOC, someone's place'
***kh**ingaye*	[kʰʲiŋgáje̟]	'green'
***ng**e*	[ŋgʲe̟]	'look'

The semivowel *w* has the fortified allophone [β] and occasionally the even stronger [b], occurring before the front vowels *i* and *e*.

Wawiwi	[waβíβi] [waβíbi]	'Wawiwi' (a village)
weka	[wé̞ka] [βé̞ka]	'over there (alongland)'

The semivowel *y* has the fortified allophone [ɟ], optionally occurring word-initially. It can also occur word-medially in enunciation.

yoho	[jó̞ho̞] [ɟó̞ho̞]	'fire'
eyo	[é̞jo̞] [é̞ɟo̞]	'my friend'

The liquid *r* is almost always a flap [ɾ], but for some speakers it can occasionally be a trill [r] in careful speech. This is more common in older speakers.

While the fricative *h* has no allophones, it is optionally elided.

humo / *umo*	[húmo̞] [úmo̞]	'also'
Haumu / *Aumu*	[háwmu] [áwmu]	(a female vocative name)

2.1.2 The status of the alveolar obstruents *t*, *nd*, and *s*

The alveolar obstruents *t* and *nd* have a highly restricted distribution, and they are involved in allomorphy that brings their phonemic status into question. These sounds occur almost exclusively after *i*. The alveolar fricative *s* is also mostly found in this environment, but it is found more often in other environments than *t* and *nd* are, so its status as a phoneme is more clear. Word-initial examples of *s* include *sapuru* 'aibika' and *sono* 'water', the later being a Wanderwort among nearby languages. Word-medial examples include *uso* 'rope'.

A number of verb suffixes have allomorphy in which their initial consonant is alveolar after *i* and velar otherwise. In these suffixes, the alveolars {*t*, *nd*, *s*} correspond to the velars {*k*, *ng*, *kh*}. I refer to this as the *velar-alveolar alternation*. Table 2.3 illustrates this allomorphy with a few different suffixes.

Table 2.3: The velar-alveolar alternation

suffix	**meaning**	**alveolar allomorph**		**velar allomorph**	
-ka ~ -ta	'1SG.POT'	*i-**ta***	'let me eat'	*hoka-**ka***	'let me cut'
		*pi-**ta***	'let me stab'	*e-**ka***	'let me take'
		*isi-**ta***	'let me sleep'	*yo-**ka***	'let me come'
-kepi ~ -tepi	'3PL.FUT'	*i-**tepi***	'they will eat'	*hoka-**kepi***	'they will cut'
		*pi-**tepi***	'they will stab'	*e-**kepi***	'they will take'
		*isi-**tepi***	'they will sleep'	*yo-**kepi***	'they will come'
-nge ~ -nde	'COND, POL.IMP'	*i-**nde***	'please eat'	*hoka-**nge***	'please cut'
		*pi-**nde***	'please stab'	*e-**nge***	'please take'
		*isi-**nde***	'please sleep'	*yo-**nge***	'please come'
-kha ~ -sa	'SS.IRR'	*i-**sa***	'eat and (SS.IRR)'	*hoka-**kha***	'cut and (SS.IRR)'
		*pi-**sa***	'stab and (SS.IRR)'	*e-**kha***	'take and (SS.IRR)'
		*isi-**sa***	'sleep and (SS.IRR)'	*yo-**kha***	'come and (SS.IRR)'

These distributional facts are the results of a past sound change, in which alveolar stops became velar everywhere except after Pre-Hurukui *i (or a diphthong ending in Pre-Hurukui *j):

Pre-Hurukui	*t, t^h, nd	remained	t, t^h, nd	/	i _ , Vj _
		>	k, k^h, ŋg	/	otherwise

Additionally, the Pre-Hurukui stop *t^h subsequently became *s*, creating the *kh ~ s* alternation. Velars could still occur in all environments, including after Pre-Hurukui *i, because they were only the output of this backing change, not the input to a fronting change. There may have also been a Pre-Hurukui sound *s that did not undergo backing, which could account for the less restricted distribution of modern *s*.

The stops *t* and *nd* are very rarely found word-initially or after other vowels root-internally. For *t*, nearly all such cases used to have an *i* preceding the *t*, and this *i* was elided. For instance, the Pre-Hurukui derived stem **i-tape-* (eat-COMPL-) 'eat it all' has lost its initial *i* for all speakers and is now just *tape-*. Several of the emphatic pronouns (§3.2.2) are based on a root *ita-* 'self', but the initial *i* is optionally elided in some of them (e.g., *itana* ~ *tana* 'yourself') and obligatorily elided in others (e.g., *tonokapo* 'themselves'). The only native word that clearly has an initial *t* and cannot be traced back to having elided an initial *i* is *tata* 'father'.

For *nd*, there are only two word-initial cases: the vocative pronouns (§3.2.3). These are *ndapoki* 'you two (vocative)' and *ndapono* 'y'all (vocative)', in which the numerals *apoki* 'two' and *apono* 'three, a few' are prefixed with a form *nd-* of unknown origin. Only one root has *nd* after a vowel other than *i*: *andeke* 'sore'.

Despite these restricted distributions of *t*, *nd*, and *s*, and their near-complementarity with the corresponding velars (especially for *t/k* and *nd/ng*), they should be analyzed as phonemes. In §2.1.3, I show minimal and near-minimal pairs for the alveolars versus the velars, in addition to various other pairs of similar phonemes.

2.1.3 Minimal pairs for consonants

k/kh

kaki	'centipede'	*khakhi*	'salt'
koko	'with'	*khokho*	'great-grandparent/child'
eka	'let me take it'	*ekha*	'bitter'
kono	'mountain, peak'	*khomo*	'flea'

k/ng

kore	'that person says'	*ngore*	'that person sees'
aka	'bite'	*anga*	'day, time, dawn, light'

p/ph

pai	'new'	*phai*	'knife'

epa	'I took it and (DS)'	*epha*	'hear'

p/mb

pa	'stab'	*mba*	'grandfather, grandson'
pite	'big'	*mbite*	'corn'
opo	'taro, thing'	*hombo*	'pitpit'
apa	'I put it and (DS)'	*Amba*	(a male vocative name)

t/nd

tapore	'that person ate it all'	*ndapoki*	'you two (vocative)'
khaita	'let me stay'	*khainda*	'which?, where?'

kh/h

kha	'and'	*ha*	'eat'
khokho	'great-grandfather, great-grandson'	*hokho*	'plate'
ekha	'bitter'	*eha*	'dog'

kh/s

isikhu	'I slept and (SS.R)'	*sisu*	'mosquito'
akha	'put it and (SS.IRR)'	*asakha*	'only, needlessly'

ng/nd

ngapa	'from where?'	*ndapoki*	'you two (vocative)'
hunginga	'claw, fingernail'	*minda*	'song, religion'

k/t

kono	'mountain'	*tonokapo*	'they (emph.)'
ikepi	'they took them'	*itepi*	'they will eat'
akare	'it bit that person'	*atare*	'that person put us'

m/mb

mungaito	'our child'	*mbungai*	'red'
muhi	'gecko'	*mbúhì*	'carry it and (SS.IRR)'
ama	'you put it'	*Amba*	(a male vocative name)

n/nd

ine	'that person ate and (DS)'	*inde*	'please eat'

Word-initial occurrence of *nd* is rare, and there are no known minimal pairs of initial *n/nd*.

w/p

wore	'that person went'	*pore*	'INTENS'
awa	'that person did not put it'	*apa*	'I put it and (DS)'

r/t

rene	'it burned and (DS)'	*tenekapo*	'yourselves'
pira	'that person's head'	*pita*	'let me stab it'

Word-initial occurrence of *t* is rare, and there are no known minimal pairs of initial *r/t*.

k/h

keme	'how much?, how many?'	*Heme*	(name of a mountain)
eka	'let me take it'	*eha*	'dog'

s/h

saporo	'metal axe' (loanword)	*hapoka*	'cut a small object'
uso	'rope'	*uho*	'firewood'

Ø/h

epa	'I took it and (DS)'	*hepa*	'heavy trap for small animals'

ahe	'one (singular indefinite)'	*hahe*	'some (plural indefinite)'
mui	'cucumber'	*muhi*	'gecko'

2.2 Vowels

2.2.1 Inventory and allophony

Hurukui has a cardinal five-vowel system.

Table 2.4: Vowel phonemes of Hurukui

	front	**central**	**back**
high	/i/ *i*		/u/ *u*
mid	/e/ *e*		/o/ *o*
low		/a/ *a*	

The vowels *a*, *i*, and *u* do not have much noticeable allophony. They can be more tense or more lax, but this does not matter much in the language, so I disregard it here.

The mid vowels *e* and *o* each have a lower allophone and a higher allophone. I label the lower ones [e̯] and [o̯], and the higher ones [e̬] and [o̬] (I do not otherwise use these diacritics on vowels, so their meaning will not be ambiguous). My reason for labeling them this way is that there is significant inter-speaker variation in how the two allophones are realized phonetically. The lower allophones [e̯, o̯] can be as low as low-mid [ɛ, ɔ] for some speakers and as high as high-mid [e, o] for others. The higher allophones [e̬, o̬] can be as low as high-mid [e, o] for some speakers and as high as pure high [i, u] for others. I am thus treating [e̯, e̬] as "sub-phonemes" of *e* and [o̯, o̬] as "sub-phonemes" of *o*, in the sense that they are groups of phones that pattern together but do not themselves have phonemic status.

The lower allophones [e̯, o̯] are the default.

keme	[kḛ́mḛ]	'how much?, how many?'
sono	[sǫ́nǫ]	'water'
eko	[ḛ́kǫ]	'there near you'
owe	[ǫ́βḛ]	'clay pot'
itepa	[itḛ́pa]	'bandicoot'
ngora	[ŋgǫ́ɾa]	'I saw'

Raising to [ẹ, ọ] occurs when the next vowel is high [i, u] or higher-allophone mid [ẹ, ọ].

*h**e**ki*	[hḛ́ki]	'yam'
*k**o**pi*	[kǫ́pi]	'they said'
*k**e**pu*	[kḛ́pu]	'yesterday'
*mb**o**khu*	[mbǫ́xu]	'grave'

This regressive raising extends back along a chain of consecutive mid vowels.

*p**e**r**e**i*	[pẹɾẹ́.i]	'y'all/they (DU) stabbed'
*k**o**k**e**pi*	[kọkḛ́pi]	'y'all/they (PL) will say'
***e**m**e**nu*	[ẹmḛ́nu]	'junglefowl'
***o**r**o**mu*	[ọɾǫ́mu]	'shell money'

If a mid vowel *e* or *o* is followed by a semivowel *y* or *w* in the onset of the next syllable, this does **not** by itself trigger raising. For instance, the following words have mid vowels which are followed by semivowels but remain in their lower allophones, because the next syllable's vowel is not high or raised-mid.

- *eya* [ḛ́ja] 'stinging nettle'
- *òmewa* [ǫ̀mḛ́wa] 'it didn't light up'
- *apoyo* [apǫ́jǫ] 'my child'
- *owe* [ǫ́βḛ] 'pot'

For understanding Hurukui phonology, the phonetics of how these lower and higher allophones are realized does not matter much. However, in Part II of this dissertation, I will discuss how the ongoing merger of the higher allophones [ę, ǫ] with high [i, u] is contributing to changes in the light verb system (discussed in Part II).

2.2.2 Vowel sequences

Almost all sequences of non-identical vowels are attested. The sequence *oi* is merged with *ui* for all speakers, except across certain morpheme boundaries (§2.3). I write *ui* rather than *oi* for morpheme-internal tokens, but there is variation as to whether it is phonetically [ǫj] or [uj]. The sequence *ea* is only found in one single word, while *oa* is not found in any single word[1].

Table 2.5 shows examples of the attested vowel sequences.

Table 2.5: Vowel sequences and diphthongs

V1 \ V2	*a*	*e*	*i*	*o*	*u*
a		*men**ae*** 'true'	*p**ai*** 'new'	*pu**ao*** 'my nephew'	*h**au**phe* 'betel nut'
e	*m**ea*** 'tongue'		*khar**ei*** 'y'all/they (DU) are here'	***eo*** 'yes'	*ir**eu*** 'we (DU) ate'
i	*hum**ia*** 'sun'	*am**ie*** 'many'		*h**io**ho* 'woman'	*kh**iu*** 'leftover essence of a visitor'
o		*h**oe*** 'snake'	(merged with *ui*)		*khar**ou*** 'python'
u	*hak**ua*** 'landslide'	*hik**ue*** 'bed, mat'	*angam**ui*** 'dregs'	*n**uo**ra* 'I gave to them (DU)'	

1 By 'single word' I mean wordforms consisting of a single root plus zero or more affixes. This excludes recently univerbated forms in which the *e/o* came from the first word and the *a* came from the second word, such as *wo are* 'that person planted it' > *wóàre*.

2.2.3 Minimal pairs for vowels and diphthongs

i/e

ima	'I (1SG)'	*ema*	'you did'
pira	'that person's head'	*pera*	'I stabbed'
iki	'pain'	*eke*	'take'

e/ai

ena	'leaf'	*aina*	'your neck'
weni	'over here (alongland)'	*waini*	'up here (upland)'
pe	'no'	*pai*	'new'

ei/ai

arei	'y'all/they (DU) put'	*arai*	'journey, wandering'
kharei	'y'all/they (DU) are located'	*Sarai*	(name of a river)

ai/ae

pai	'new'	*apae*	'animal nest'
Sarai	(name of a river)	*harae*	'together'

e/a

eme	'that person's eye'	*ama*	'you put'
weka	'alongland'	*waka*	'upland'
hepe	'rain' (Arapipi dialect)	*hepa*	'heavy trap for small animals'

u/o

uma	'die'	*oma*	'you took'
ikhunu	'unripe'	*ekhono*	'now, today'
mbukhu	'wet'	*mbokhu*	'grave'
hombu	'heart of the *soko* vine'	*hombo*	'pitpit'

o/au

omu	'breast'	*Aumu*	(a female vocative name)
nayo	'my elder brother'	*nayau*	'sorcery'

ou/au

kharou	'python'	*khikhirau*	'mud'

au/ao

ewau	'we (DU) didn't take'	*ewao*	'I won't take'

o/a

ora	'I took'	*ara*	'I put'
okora	'I tied'	*hokara*	'I cut'
moho	'headwaters'	*maha*	'ground, land'

2.3 Syllable structure

Syllables are all open under normal circumstances. No consonant clusters are allowed. The syllable structure is thus a very simple (C)V. Some examples of syllabified words following this pattern are shown in Table 2.6.

Table 2.6: Examples of syllabification

o.ra	'I took'	*kho.mbo*	'moon, month'
hi.mbi.nu	'yellow pandanus'	*wo.ti.e*	'red pandanus'
ho.ka.ko.pu	'we (PL) will cut'	*no.ro.ka.ma*	'you think'

The vowel sequences *ai*, *au*, and *ui* are tautosyllabic (diphthongs) when they occur in non-word-final position. In this environment, I transcribe them as [aj], [aw], and [uj]. Word-finally,

they are heterosyllabic. In this environment, I transcribe them as [a.i], [a.u], and [u.i]. This is illustrated in Table 2.7.

Table 2.7: Syllabification of diphthongs

	non-word-finally		**word-finally**	
ai	*ha.r**ai**.ru*	'cordyline'	*kha.w**a.i***	'y'all/they (DU) are not located'
	*kh**ai**.ngu*	'that one'	*mbu.ng**a.i***	'red'
au	*ha.ng**au**.ngu*	'spider'	*kha.w**a.u***	'we (DU) are not located'
	*n**au**.mba*	'epiphyte'	*khi.khi.r**a.u***	'mud'
ui	***ui**.ta*	'down there'	*a.ra.k**u.i***	'taproot'
	*phi.m**ui**.to*	'our navel'	*khu.kh**u.i***	'hidden'

Other vowel sequences are treated as heterosyllabic in all environments. This is illustrated in Table 2.8 for a few of them.

Table 2.8: Syllabification of non-diphthong vowel sequences

	non-word-finally		**word-finally**	
ia	*mb**i.a**.ra*	'long'	*hu.m**i.a***	'sun'
oe	*mb**o.e**.ko*	'plant.sp'	*i.p**o.e***	'that person's father'
ua	*n**u.a**.ra*	'I hit them (DU)'	*ko.p**u.a***	'we (DU) said, and (DS) ...'

An exception to this is when an inalienable noun root (§3.1.1) ending in *e* receives the suffix *-ito* '1PL.POSS'. This creates a tautosyllabic diphthong *ei* word-medially, e.g., *ipe* [í.pe̱] 'that person's name' versus *ipeito* [i.péj.to̱] 'our names'. A small number of proper nouns (§3.1.2) also have this diphthong, such as *Heiwa* 'Adzera tribe', which can be pronounced [héj.wa] or [hé̱.wa].

Similarly, a sequence *oi*, distinct from *ui*, is retained for speakers who make the distinction between [o̱] and [u]. It occurs when inalienable noun roots ending in *o* are suffixed with *-ito* '1PL.POSS'. The resulting *oi* straddling the morpheme boundary is pronounced [o̱j], not [*uj]. For instance, *ipo-ito* 'our father' is pronounced [ipó̱jto̱], not [*ipújto̱], while *numu-ito* 'our body' is clearly [numújto̱], not [*no̱mó̱jto̱], for these speakers.

In these rare tokens of *ei* and *oi*, the semivowel offglides **do** trigger mid vowel raising. Thus, the mid vowels in the sequences *ei*, *eu*, *oi*, and *ou* are always raised, whether the sequences are tautosyllabic or heterosyllabic. As a result, presence or absence of mid vowel raising cannot bear on the question of whether it is better to analyze the merged *oi/ui* as *oi* or as *ui*, since both of these would result in raising of preceding mid vowels (via the raised *o* in the case of *oi*, and via the *u* in the case of *ui*). Further phonetic and variationist analysis is a direction for future research, and eventually may be able to bear on this question more clearly for lexical items in which morphophonemic vowel changes do not exist to help differentiate *e* from *i* and *o* from *u* (especially words that take no affixes, such as *Hurukui*, *kui* 'bag', and *nuimi* 'tomorrow').

Syllable structures other than (C)V are found in the following kinds of words that can be used in full sentences:

- proper names, e.g., *Gu.**sap***, *Ya.**kam***
- ideophones, e.g., ***khim khum*** 'sprouting'
- loanwords from Tok Pisin, e.g., ***baim*** 'buy', *bai.si.**kol*** 'bicycle'

as well as in the following kinds of words that veer more toward the paralinguistic:

- interjections, e.g., ***ih*** [ih] 'oh! (surprise, shock)'
- sounds made at animals, e.g., ***oh*** [ʔo̲h] (sound for calling to dog)
- onomatopoeia, e.g., [kɔ̋k̄] (sound imitating bird of paradise)

2.4 Stress

Hurukui has two degrees of stress: primary and secondary. Primary stress is realized by higher pitch, higher intensity, and a bit of length. Secondary stress has lower pitch but still has higher intensity and a bit of length. When I mark stress in this dissertation, I mark primary stress with acute accent (*á é í ó ú*) and secondary stress with grave accent (*à è ì ò ù*).

In most words, the penult bears primary stress, while all other syllables are unstressed. This is illustrated in Table 2.9, using forms from Tables 2.6, and 2.7, and 2.8.

Table 2.9: Examples of regular stress assignment

***ó**.ra*	'I took'	***khó**.mbo*	'moon, month'
*hi.**mbí**.nu*	'yellow pandanus'	*wo.**tí**.e*	'red pandanus'
*ho.ka.**kó**.pu*	'we (PL) will cut'	*no.ro.**ká**.ma*	'you think'
*ha.**rái**.ru*	'cordyline'	***náu**.mba*	'epiphyte'
*mbu.**ngá**.i*	'red'	*khi.khi.**rá**.u*	'mud'
*mbi.**á**.ra*	'long'	*nu.**á**.ra*	'I hit them (DU)'
hu.mí.a	'sun'	*ko.p**ú**.a*	'we (DU) said, and (DS) ...'

One type of exception occurs when multisyllabic verb stems are marked with a morpheme that is realized as floating final stress. This can mean same-subject irrealis (§5.4.2) or imperative with object marking (§5.3.4). In all such cases, the stress-shifted form coexists with a synonymous zero-marked form, in which there is no suffix nor any stress shift. Table 2.10 illustrates some of these forms.

Table 2.10: Examples of verb stems marked with floating final stress

enóka ~ enoká	'take them (PL)'	*éki ~ ikí*	'take y'all/them (PL)'
éya ~ eyá	'take me'	*húmba ~ humbá*	'fill'
úki ~ ukí	'cook'	*hapóka ~ hapoká*	'cut a small object'

Another type of exception occurs in univerbated forms, which is also where secondary stress appears. During the univerbation process, one of the words retains its primary stress, while the primary stresses of the others are demoted to secondary. The factors determining which word retains primary stress will require further research. Some examples are illustrated in Table 2.11, with the reconstructed multi-word form alongside the modern univerbated form.

Table 2.11: Stress assignment in univerbated forms

separate words (Pre-Hurukui)	**univerbated (modern)**	**meaning**
**náma áre*	*námàre*	'that person sat down'
**pé wóre*	*péwòre*	'it ran out, it went away'
**hiná kháka*	*hinàkháka*	'keep watch'
**mbúkhu ímu khái*	*mbúkhìmukhài*	'it is wet'

Further exceptions to regular penultimate stress assignment occur in some of the same kinds of words that violate the language's canonical syllable structure (§2.3): proper names and ideophones.

In proper names, the exceptions are sporadic and likely came about due to how names from other languages were adapted into Hurukui. This is not surprising, because these names often break the rules of Hurukui phonology in various other ways as well.

Table 2.12: Examples of stress assignment in proper names

Bilaga	[mbiɾága]	(a male given name)	*Yakam*	[jakʰám] [jaxám]	(a male given name)
Daniel	[dɛ̱́njɛ̱ɾ]	(a male given name)	*Nganining*	[ŋaniníŋ]	(a male given name)
Yasakingo	[jasakíŋo̱]	(a female given name)	*Pamela*	[pámɛ̱la]	(a female given name)
Sawao	[sawá.o̱]	(a female given name)	*Khayahá*	[kʰajahá]	(a female vocative name)
Gusap	[gusáp]	(a village)	*Hingia*	[híŋgi.a]	(a village)

Ideophones, in contrast to proper names, mostly obey Hurukui phonology. The main way they break the rules is that they can end with consonants. In such cases, the word bears final stress, as if there had been a final vowel that was elided, leaving the stress in regular position on the formerly penultimate vowel. The word *sisos* [sisó̱s] 'squeezing' is a good example of this.

2.5 Vowel reduction processes

2.5.1 Deletion

Word-final vowel deletion is pervasive in Hurukui. When the final vowel in a word is unstressed, and the word is not the last word in the utterance, the vowel is often deleted. This feature is also found in the nearby Iyo language (Minter 2009: 13), which is a distant relative in the Finisterre-Huon branch of Trans-New Guinea and with which Hurukui has had significant contact. When I mark vowel deletion in this dissertation, I mark deleted vowels with breve accent (*ă ĕ ĭ ŏ ŭ*). Table 2.13 shows some examples.

Table 2.13: Word-final vowel deletion

phonological form	**surface pronunciation**	**meaning**
imă pahe	[ímpáhe̯]	'my house'
opopŏ musŭ maisa	[o̯pò̯pmùsmájsa]	'small things'
sonŏ humbă e yo	[só̯nhúmbé̯jó̯]	'fill up the water and bring it!'

Some pairs of words are univerbating both grammatically and phonologically, such that no other material is allowed to intervene between them. In many of these cases, the deletion of the first word's final vowel has become obligatory, and the identity of the deleted vowel is either unknown or only knowable via internal reconstruction. Table 2.14 shows some phrases of this kind.

Table 2.14: Word-final vowel deletion in univerbating phrases

phonological form	**surface pronunciation**	**meaning**
hengV$_{[-hi]}$ *ora*	[hé̱ŋgo̱ɾa]	'I jump over'
hungV ora	[húŋgo̱ɾa]	'I pull'
omV$_{[-hi]}$ *ore*	[ó̱mo̱ɾe̱]	'it lights up'
perV$_{[+hi]}$ *ara*	[pé̱ɾàɾa]	'I tie'
orV$_{[-hi]}$ *kora*	[ó̱ɾko̱ɾa]	'I cover up'
wekă era	[wé̱ke̱ɾa]	'I want to go'
wekă kore	[wé̱kːo̱ɾe̱]	'that person wants to go'

Unstressed word-initial *i* is also often deleted, especially when it carries low functional load. This occurs often before *t*, because the mere presence of the alveolar is a strong signal that an *i* precedes it (§2.1.2). It also occurs with the common verb root *i-* 'be, become, do'. Table 2.15 shows some examples of these sorts.

Table 2.15: Word-initial vowel deletion

phonological form	**surface pronunciation**	**meaning**
wekă imara *wekă ĭmara*	[wé̱kimàɾa] [wé̱kmàɾa]	'I want to go'
maisă ĭtokhV$_{[+hi]}$ *ĭtomu*	[májstọ̈xtọ̈mu]	'it will rot quickly'
Pre-Hurukui **itapora* > *tapora*	[tapó̱ɾa]	'I ate it all'
na itana *na ĭtana*	[na.itána] [natána]	'yourself'
yoho ĭteme	[jò̱htḛ́me̱]	'ashes'

Because final vowel deletion often removes a vowel that conditions the lower or higher allophones of mid vowels (§2.2.1), there are some surface realizations that are phonetically minimal pairs between these allophones.

Table 2.16: Phonetic minimal pairs for lower and higher allophones of mid vowels

phonological form	**surface pronunciation**	**meaning**
emenă ngorane	[e̞mè̝nŋgo̞ráne̞]	'your eyeball'
emenŭ ngorane	[ĕ̞mè̝nŋgo̞ráne̞]	'junglefowl egg'
pera	[pé̞ɾa]	'I stabbed'
perĭ a	[pĕ̞́ɾà]	'tie and (SS.IRR)'

Vowel deletion has more options in some situations in which the second word of a pair is a verb whose root consists of only a single vowel and carries a low functional load. Table 2.17 shows some common phrases like this. The options for deletion are (1) the last vowel in the first word, (2) the first vowel in the second word, or (3) both of these. The table shows an asterisk in disallowed cells.

Table 2.17: Multiple options for vowel deletion in certain phrases

form	**last vowel in first word deleted**	**first vowel in second word deleted**	**both vowels deleted**	**meaning**
moko emara	[mó̞ke̞màɾa]	[mó̞ko̞màɾa]	[mó̞kmàɾa]	'I am talking'
ruru imare	[ɾúɾimàɾe̞]	[ɾúɾumàɾe̞]	[ɾúɾmàɾe̞]	'it is getting soft'
Pre-Hurukui **o̞po itopu* > *opitopu ~ opotopu*	[ò̞pitó̞pu]	[ò̞po̞tó̞pu]	*	'our food'

Final vowels can also be devoiced instead of fully deleted, which often leads to the plain stops *p*, *t*, and *k* immediately preceding them being realized as [pʰ, tʰ, kʰ]. An example is *wopi* [wó̞pi ~ wó̞pʰi̥] 'y'all/they (PL) went'.

In casual speech, root-internal vowels between voiceless obstruents can be devoiced, leaving coarticulation on adjacent consonants as the primary cue to the phonological vowel. An example is *wakhupena* [waxupé̞na ~ waxu̥pé̞na ~ waxʷpé̞na] 'your blood'.

2.5.2 Monophthongization

Diphthongs sometimes lose their offglides. The most common of these changes is *ai* > *a*, which can happen when affixation moves the stress off of the *ai*. This is shown in Table 2.18 with some forms of the verb root *khai-* 'be located', as well as some deictics that contain a prefix which seems derived from the same verb (§4.3).

Table 2.18: Regular monophthongization *ai* > *a*

parsed form	**form**	**pronunciation**	**meaning**
khai-pi	***kháipi***	[kʰájpi]	'y'all/they (PL) are located'
khai-wa-pi	***kha**wápi*	[kʰawápi]	'y'all/they (PL) did not stay'
khai-ma	***kháima***	[kʰájma]	'you are located'
khai-ma-re	***kha**máre*	[kʰamáɾe̱]	'that person lives in a place'
khai-no	***kháino***	[kʰájno̱]	'this'
khai-no-kha	***kha**nókha*	[kʰanó̱xa]	'this (FOC)'
khai-wa-ka	***kha**wáka*	[kʰawáka]	'that (upland, far)'

For some roots, however, this change occurs more sporadically. This is shown in Table 2.19 with the verb root *yai-* 'build'. Some of the forms where *ai* is unstressed require monophthongization, some of them optionally allow it, and some of them prohibit it.

Table 2.19: Sporadic monophthongization *ai* > *a*

parsed form	**form**	**pronunciation**	**meaning**
yai-pi	***yáipi***	[jájpi]	'y'all/they (PL) built'
yai-wapho	***ya**wápho*	[jawáɸo̱]	'don't build'
yai-ma-re	***ya**máre* ***yai**máre*	[jamáɾe̱] [jajmáɾe̱]	'that person is building'
yai-[komu~tomu]	***yai**tómu*	[jajtṓmu]	'that person will build'

The diphthong *ai* can also become *e*, including in stressed syllables. This happens root-internally in casual speech for some words and proper names. One example is *araiku ~ areku*

'already'. This change is sporadic enough that it may simply be lexically-selected sound change rather than part of the language's phonology.

The diphthong *ui* can become *u* in certain situations. Similarly to the change *ai* > *a*, the change *ui* > *u* often happens in unstressed syllables that were created when affixation moved the stress elsewhere. Additionally, similarly to the change *ai* > *e*, the change *ui* > *u* also occurs sporadically in lexical items, including in stressed syllables. For instance, the word *uita ~ uta* 'down there' has the locative-suffixed form *utaku*, while **uitaku* is disallowed.

The diphthong *au* does not monophthongize, nor do the rarer tautosyllabic *ei* and *oi* which occur across certain morpheme boundaries (§§2.2.2, 2.3).

2.6 Anticipatory semivowel insertion

When the vowel *e* precedes a palatalized velar stop, a [j] glide is inserted between the *e* and the stop. This does not apply when the consonant is *kh*, however, because it is intervocalically a fricative (§2.1.1). Some examples are shown in Table 2.20

Table 2.20: Anticipatory insertion of [j]

phonological form	**insertion?**	**surface pronunciation**	**meaning**
reke	yes	[ɾe̞jkʲe̞]	'that person's belly'
nengi	yes	[ne̞jŋgʲi]	'village, outside'
humekhe	no	[hume̞xe̞]	'container'

Similarly, when the vowel *a* precedes a consonant followed by *u*, a [w] glide can be inserted between the *a* and the consonant. This is a more sporadic change, however. This is shown in Table 2.21.

Table 2.21: Anticipatory insertion of [w]

phonological form (* = Pre-Hurukui)	**insertion?**	**surface pronunciation**	**meaning**
**hahu > hauhu*	yes	[háwhu]	'game, play, happy'
**makhu > maukhu*	yes	[máwxu]	'laughter'
**i-wa-mu > iwaumu*	yes	[iwáwmu]	'that person will not eat'
**i-wa-pu > iwaupu*	yes	[iwáwpu]	'we (PL) will not eat'
**i-ma-mu > imaumu*	yes	[imáwmu]	'that person will continue eating'
i-wa-pu	no	[iwápu]	'we (PL) did not eat'
i-tanu	no	[itánu]	'you will eat'
hoka-mu	no	[hǫ̱kámu]	'that person cut and (SS)'
nge-na-pu	no	[ŋgę̱nápu]	'we (PL) saw you'

2.7 Prosody

There is a prosodic domain larger than the phonological word, in which only one primary stress is assigned and other word-level primary stresses are demoted to secondary or unstressed entirely. Later in this grammar I discuss how this is used to mark functions such as headedness in noun compounds (§7.1.1.2). Some examples of this primary-stress domain are shown in Table 2.22.

Table 2.22: Examples of the multi-word primary stress domain

phonological form	**surface pronunciation**	**meaning**
haráě para	[haɾápaɾa]	'together'
ngú khàipi	[ŋgúxàjpi]	'y'all/they (PL) are in that place'
ngú khàhu	[ŋgúxàhu]	'with that person'
ìmă úmo	[ìmúmǫ̱]	'me too'
ímă ùmo	[ímùmǫ̱]	'my tree'
ngù khái	[ŋgùxá.i]	'it exists'
ngú khài	[ŋgúxà.i]	'it is in that place'

Because the difference between primary stress and secondary stress is mostly a difference in pitch, intonational effects often sound like changes to stress quantity and position. Later in this grammar I describe how intonation is used for marking questions (§§7.4.1, 7.4.3) among other functions.

At the end of an utterance, one often finds either lengthening of the final vowel or a phonetic glottal stop cutting it short. Table 2.23 shows a few examples of this.

Table 2.23: Utterance-final lengthening and shortening

phonological form	**surface pronunciation**	**meaning**
eo	[e̯o̯ː]	'yes'
pe	[pe̯ː] [pe̯ʔ]	'no, it does not exist'
mbòkhǒ pá	[mbò̯xpáʔ]	'cover it up'

2.8 Orthography

The current official orthography was decided on at a multi-day orthography summit in August 2024 in Bibuai village. This was the first such meeting, resulting in the first official orthography for Hurukui. Some loose ends remain because further analysis and speaker input is needed. The orthography is not quite the same as the transcription conventions used in examples in this dissertation, because the dissertation needs to show phenomena such as secondary stress, vowel elision, explicitly marked regular primary stress, and other things that are not used in the community-facing orthography.

The consonant and vowel symbols are all the same as those used for the segmental phonemes in this dissertation. For consonants, these are listed in Table 2.1 in §2.1.1. For vowels, these are listed in Table 2.4 in §2.2.1.

The only diacritic is the acute accent, used to mark primary stress in irregular position. For consonant-final words, the orthography treats the regular stress position as the last vowel, before the last consonant (as if the word had a final vowel that has been lost, which is often the

case). The orthography does not mark secondary stress at all. Thus, words with final stress on a closed syllable such as *sisos* [sisˌós] 'squeezing' do not have any diacritics.

Final vowel elision is not shown in the orthography. Final vowels that can be spoken must be written. Thus, one writes *ima pahe* 'my house', not **im pahe*, even though the latter reflects a pronunciation that is overwhelmingly more common. However, if the final vowel has been completely lost and cannot be recovered even in careful speech, it is not written.

In a small number of words, the sequences [kuV] and [ŋguV] have become [kwV] and [ŋgwV] when the V is stressed. The orthography writes these with *kw* and *ngw*, but there are no labialized velar phonemes in Hurukui. An example is *angwane* 'vein'.

3 Nominals

Nominals are words which can head noun phrases, which can be the arguments of verbs or postpositions and can serve as nonverbal predicates. Nominals in Hurukui can mostly be divided into nouns and pronouns. There is also a small class of nominals that I call "whatchamacallit" words, used when the speaker cannot come up with the name of something or someone. Deictics are also often used as nominals, but they have additional morphology and syntactic behavior. For this reason, I treat deictics in Chapter 4.

Nouns can be divided into *common nouns* and *proper nouns*. Common nouns generally denote words for types of people, places or things, but not the names of specific instances of those types. Within common nouns, there are *alienable nouns*, which take no morphology, and *inalienable nouns*, which require a suffix indicating their possessor. These possessive suffixes distinguish first/second/third person and singular/dual/plural number, but not clusivity. Inalienable nouns denoting humans also mark dual and plural number of their referents. Some nouns show behaviors of both alienable and inalienable nouns, and I refer to these as *ambialienable nouns*. Proper nouns generally are the names of specific people or places, and they take no morphology. Common nouns and proper nouns differ somewhat in their syntactic distribution.

Pronouns can be divided into *personal pronouns*, which are the typical anaphora corresponding to translations like 'I' and 'you'; *emphatic pronouns*, which are also anaphoric but with a semantic element of emphasis; *vocative pronouns*, used to call out to two or more people without using their names; and *interrogative pronouns*, used to ask 'who?' and 'what?'. The personal and emphatic pronouns distinguish first/second/third person and singular/dual/plural number, and the first-person nonsingulars also distinguish inclusive and exclusive. Many of the personal pronouns are identical in form to the corresponding possessive suffixes found on inalienable common nouns. There are only two vocative pronouns, one for calling to two people and one for calling to more than two. The interrogative pronouns do not mark any person or number categories.

3.1 Nouns

Nouns are nominals that are non-deictic and non-anaphoric. Here I show some examples of the syntactic distribution of noun phrases. For simplicity, all of the highlighted noun phrases consist only of a single noun. (1)-(5) show nouns acting as arguments to verbs, subjects in (1)-(3) and objects in (4)-(5). In (6), a noun is used as a nonverbal predicate (predicate nominal). In (7), a noun is used as the argument of a postposition.

In (1), the noun *eha* 'dog' is the subject of the verbs *akai-* 'follow' and *we-* 'go'.

(1)

Eha	*mbo*	*khaporo*	*akai-mu*	*wo-re.*
dog	animal	wallaby	follow-3SG.SS.R	go-3SG.R

'The dog chased the wallaby.' (Hunt10 0:22-0:26)

In (2), the noun *khukhuyo* 'my nose' is the subject of an existential predicate in the positive polarity (which uses a conjugated verb *khai-* 'be located, exist').

(2)

Khukhu-yo	*khai-Ø.*
nose-1SG.POSS	exist-3SG.R

'I have a nose.' (Bodi 0:38-0:41)

In (3), the noun *hikue* 'blanket, mat' is the subject of an existential predicate in the negative polarity (which uses a non-inflecting particle *pe*, §6.5.1).

(3)

Kaya	***hikue***	*pe.*
long.ago	blanket	NEG.EXIST

'Long ago there were no blankets.' (Blanket 1:21-1:25)

In (4), the noun *pahe* 'house' is the object of the verb *yai-* 'build'.

(4)

Ima	***pahe***	*yái-tè-ra.*
1SG	house	build-DES-1SG.R

'I want to build a house.' (Haus 1:43-1:47)

In (5), the noun *imenoka* 'their mother' is the object of the verb *a-* 'kill.3SG'.

(5)

Opo	*khangu-kha*	***ime-noka***	*á-mukhài-Ø.*
thing	this-FOC	mother-3PL.POSS	kill.3SG-3SG.PRF.R-3SG.R

'This thing killed their mother.' (Raunwara 2:32-2:38)

In (6), the noun *hangeyaka* 'large stick used for making trap' is a predicate nominal.

(6) *Khaino* *ngu* ***hangeyaka***.
this 3SG trap.stick
'This is a large stick for making traps.' (Trap 2:13-2:16)

In (7), the noun *Nungui* (a clan name) is the argument of the postposition *para* 'from'.

(7) *Khongoromu* ***Nungui*** *para* *yó-nekhài-Ø.*
ancestor Nungui.clan from come-3SG.DPST-3SG.R
'My ancestors came from Nungui clan.' (Huakono 0:03-0:06)

Nouns in Hurukui can be divided into several subclasses. First, nouns may be common nouns or proper nouns. Common nouns may further be alienable or inalienable, and some behave as members of both of these classes.

3.1.1 Common nouns

Common nouns are the largest class of nominals. It is an open class allowing loanwords and neologisms. It can be divided into *alienable nouns* and *inalienable nouns*. The semantic notion of alienability is defined such that an entity is called *alienable* if it is "inherently relational, inseparable from [its possessor], not normally acquired or lost" (Mithun 1999: 251), and *inalienable* otherwise. I use these terms only as names that give a rough sense of what kinds of nouns are in the Hurukui categories of alienable and inalienable nouns. The criterion I use for labeling a noun as alienable or inalienable is strictly morphological. If it can take suffixes indicating its possessor, then it is inalienable; otherwise, it is alienable. This does not perfectly correspond to the cross-linguistic semantic notion of alienable possession, but it overlaps significantly enough that I see these labels as helpful approximations (in line with my general approach to definitions given the ubiquity of gradience in the world and language, as discussed in Chapter 9).

Alienable nouns are those common nouns that take no affixes. As such, there is not much to describe about them here. They can denote various objects, people, kinship, spatial relations, feelings, and other semantic categories. Examples include *pahe* 'house', *moko* 'speech', *kopohe* 'child', *tata* 'father', *khokhophe* 'side of valley', and *ua* 'illness'.

Inalienable nouns are those common nouns that must take affixes indicating their possessor. Inalienable nouns denoting humans can also take suffixes for their number. Most inalienable nouns denote body parts (*pira-* 'head', *kepane-* 'chest'), growths attached to the body (*mapa-* 'sore'), excrement (*khikho-* 'urine', *hipa-* 'feces'), kinship (*ime-* 'mother', *aphi-* 'taboo in-law'), or spatial relations (*khane-* 'inside', *makara- ~ makare-* 'middle, between'). Some of them denote more abstract things, such as feelings (*ura-* 'fear', *hambukha-* 'liking') or states (*akho-* 'waiting'). All of these semantic categories also contain alienable nouns, so it is not possible to be sure from the meaning whether a noun will be inalienable.

Table 3.1 lists the possessive affixes. For the pronouns *nia* '2DU', *nua* '3DL', *neka* '2PL', and *noka* '3PL', as well as the suffix *-na* '2SG.POSS', the possessive suffixes are identical to the pronoun. It is clear, however, that these are suffixes rather than a postposed pronoun because the word is a single stress domain (§2.4).

Table 3.1: Possessive affixes on inalienable nouns

possessor	**affixes**
1SG	*-(y)o*
2SG	*-na* *ni-*
3SG	*-Ø* *-(y)e*
1DU	*-yo*
2DU	*-nia*
3DU	*-nua*
1PL	*-ito*
2PL	*-neka*
3PL	*-noka*

The suffixes immediately follow the root. For instance, the forms of the inalienable noun root *oma-* 'hand, arm' include:

- *oma-**na*** 'your hand/arm'

- *oma-**ito*** 'our hands/arms'
- *oma-**noka*** 'their hands/arms'

Some of the suffixes have allomorphy. The 1SG.POSS suffix *-(y)o* takes the form *-yo* after *o* or *u*, and it takes the form *-o* elsewhere. Examples of the *-o* allomorph include:

- *oma-* 'hand, arm' > *oma-**o*** 'my hand/arm'
- *eme-* 'eye' > *eme-**o*** 'my eye'
- *kuni-* 'foot, leg' > *kuni-**o*** 'my foot/leg'

Examples of the *-yo* allomorph include:

- *epo-* 'ear' > *epo-**yo*** 'my ear'
- *khukhu-* 'nose' > *khukhu-**yo*** 'my nose'

The meaning of 3SG possessive can be marked by two different suffixes, (1) *-Ø* and (2) *-(y)e*. The choice between these is lexically selected by the noun root. The zero suffix has no allomorphy. Examples of words taking it include:

- *oma-* 'hand, arm' > *oma-**Ø*** 'that person's hand/arm'
- *epo-* 'ear' > *epo-**Ø*** 'that person's ear'

The other 3SG.POSS suffix, *-(y)e*, takes the form *-ye* after *e*, and it takes the form *-e* elsewhere. Examples of the *-e* allomorph include:

- *mba-* 'grandfather/-son' > *mba-**e*** 'that person's grandfather/-son'
- *kuni-* 'foot, leg' > *kuni-**e*** 'that person's foot/leg'
- *mokono-* 'sister' > *mokono-**e*** 'that person's sister'
- *phimu-* 'navel' > *phimu-**e*** 'that person's navel'

Examples of the *-ye* allomorph include:

- *ime-* 'mother' > *ime-**ye*** 'that person's mother'
- *neke-* 'cross cousin' > *neke-**ye*** 'that person's cross cousin'

There are some inalienable noun roots which end in *a*, lexically select for 3SG.POSS suffix *-(y)e*, and have their final *a* removed when the suffix is attached. In these cases, the 3SG.POSS suffix takes the *-e* allomorph.

- *pokhona-* 'back' > *pokhon-e* 'that person's back'
- *munga-* 'child' > *mung-e* 'that person's child'
- *pira-* 'head' > *pir-e* 'that person's head' (one possible variant)

But not all *a*-final inalienable noun roots do this.

- *oma-* 'hand, arm' > *om**a**-Ø* 'that person's hand/arm'
- *mapa-* 'sore' > *map**a**-Ø* 'that person's sore'
- *mba-* 'grandfather/-son' > *mb**a**-e* 'his grandfather/-son'
- *khura-* 'by oneself' > *khur**a**-e* 'by themself'

The 2SG possessive meaning is marked by one of two different affixes. These are a suffix *-na* and a prefix *ni-*. These affixes have a complementary distribution which is lexically selected: each inalienable noun can take only one of them. Despite this, I hesitate to call them allomorphs, because of their different placement relative to the root, which clearly shows historically different origins of the two. The suffix *-na* is much more common and follows a pattern of suffixing the pronoun's form, also seen in second and third person nonsingular (Table 3.1). In contrast, the prefix *ni-* is one of the only prefixes in the language and is only used with a few lexemes. For most of these, the roots in other possessive inflections are significantly altered or suppletive.

- ***ni**-muna* 'your wife' (other roots: *hioho*, *akhure*)
- ***ni**-munga* 'your child' (other roots: *apoyo*, *anuke*)
- ***ni**-oho* 'your big brother (to a man)' (other roots: *papa*, *na-*)
- ***ni**-okha* 'your little sibling' (other root: *usi-* ~ *ukhi-*, probably cognate with *-okha*)
- ***ni**-pua* 'your uncle/nephew' (regular root: *pua-*)

Inalienable nouns also have a form that I call *generic*, glossed 'GENER'. This form refers to the inalienable noun's denotatum in general, without specifying who it belongs to. The generic form can be marked in two ways: suffixed or zero-marked. This choice is lexically selected, with some nouns requiring one or the other, and some nouns allowing either one.

Suffixed generics are marked using *-noka*, which also means '3PL.POSS'.

- *hipa-**noka*** 'feces', cf. *hipa-Ø* 'that person's feces'

Zero-marked generics are marked with a zero suffix, even if their 3SG.POSS form takes *-(y)e*.

- *ona-**Ø*** 'bone', cf. *ona-Ø* 'that person's bone'
- *omu-**Ø*** 'breast, breastmilk', cf. *omu-e* 'her breast'

Some generics can be marked with either zero or *-noka*. For these nouns, the zero-marked generic denotes the item in general, which may or may not be possessed by anyone. The suffixed generic denotes an item that belongs to someone, but it is not specified who that possessor is. The noun *phephe* 'grandmother, female ancestor' exemplifies this:

- *phephe-**Ø*** 'female ancestor(s)' vs *phephe-**noka*** '(a particular) old woman',
  cf. *phephe-ye* 'that person's grandmother'

For some noun roots, this results in a divergence of meaning because some of the lexeme's denotata are inextricably connected to a person while others are not. The noun *kokope* 'ghost, picture, shadow' exemplifies this, as ghosts are inextricably connected to a person while pictures are not (they may not have any people in them).

- *kokope-**Ø*** 'picture, video' vs *kokope-**noka*** 'ghost'

The noun *opopo* 'food' has an anomalous possessive paradigm. It is the result of the univerbation of erstwhile relative clauses (§8.4.2) meaning 'the thing/food that {possessor} will

eat'. The source forms and modern forms are listed in Table 3.2. This paradigm preserves the root *opo*, now mainly meaning 'taro', in its older meaning of 'food, thing', which is now mainly expressed by the reduplicated *opopo*. The syncretism of second and third persons in nonsingular numbers is inherited from the way verbs are conjugated (§5.1.2).

Table 3.2: Possessive paradigm of *opopo* 'food'

possessor	**source relative clause (* = Pre-Hurukui)**	**modern possessed form**[2]
1SG	**opo i-o*	*opio*
2SG	**opo i-tanu*	*opitanu*
3SG	**opo i-tomu*	*opitomu ~ opotomu*
1DU	**opo i-tohu*	*opitohu ~ opotohu*
23DU	**opo i-tekhu*	*opitekhu*
1PL	**opo i-topu*	*opitopu ~ opotopu*
23PL	**opo i-tepi*	*opitepi ~ opotepi*

Inalienable nouns which refer to humans take suffixes indicating nonsingular number: *-mo* 'DU' and *-po* 'PL'. These follow the possessive suffix.

- *ni-munga-**po*** (2SG.POSS-child-PL) 'your children (PL)'
- *mung-e-**po*** (child-3SG.POSS-PL) 'that person's children (PL)'
- *e-ye-**mo*** (friend-3SG.POSS-DU) 'that person's two friends'

When *-mo* 'DU' or *-po* 'PL' is present on a noun marked for 1SG possessor, the 1SG.POSS suffix must take the *-yo* allomorph, even if it otherwise takes the *-o* allomorph.

- *mba-**yo**-po* 'my grandfather/sons (PL)', cf. *mba-**o*** 'my grandfather/son'

2 The tokens of *o* in the root portion of these forms, *op(o)-*, are all the higher allophone [o̩], even though they would have been the lower allophone [o̞] before univerbation. This is unlike more recent univerbations, in which the mid vowel raising process (§2.2.1) has not yet spread far enough left to cross erstwhile word boundaries.

Some inalienable nouns may optionally be used in a possession construction without bearing a suffix marking the possessor. I call these *ambialienable nouns*. All of them denote kinship, and they likely come from vocative forms that are being recruited into the system of inalienable nouns. Table 3.3 shows some examples.

Table 3.3: Some ambialienable nouns

meaning	**alienable form**	**inalienable form**
'my aunt'	*ima meme*	*(ima) meme-o*
'that person's aunt'	*ngu meme*	*(ngu) meme-ye*
'my older sister'	*ima nene*	*(ima) nene-o*
'your older sister'	*na nene*	*(na) nene-na*

3.1.2 Proper nouns

Proper nouns are an open class of nouns that denote the names of people or places. They are capitalized in the orthography. They can be the names of:

- people: *Yakham* [jaxám] 'Yakam', *Sakingo* [sakíŋo] 'Sakingo', *Sebron* [sɐ̭bɾó̜n] 'Sebron'
    - this also includes *vocative names*, a small closed class used to call out to people or refer to them indirectly
        - for men: *Amba*, *Amoko*, *Anai*, *Anao*, *Awe*, *Awehi* (exhaustive list)
        - for women: *Amo*, *Aumu ~ Haumu*, *Awae*, *Khayahá* (exhaustive list)
- clans: *Ngarumbaya* 'Ngarumbaya clan', *Pisou* 'Pisou clan'
- tribes and some languages: *Hurukui* 'Hurukui', *Heiwa* 'Adzera', *Herawa* 'Iyo'
- places: *Wawiwi* 'Wawiwi village', *Haruka* 'Haruka mountain', *Kurama* 'Bibuai river'

Proper nouns, like alienable common nouns, have no morphology. They largely share the syntactic distribution of common nouns, but there are a few differences. The rest of this section will compare their distributions.

The indefinite quantifiers *ahe* 'one, INDEF.SG' and *hahe* 'some, INDEF.PL' (§6.2.3) are often found modifying common nouns, but are completely unattested modifying proper nouns.

Proper nouns, like common nouns, can be modified by adnominal deictics, but this is mostly done with place names. In (8), the place name *Top Town* (an English name because it is a neighborhood in Lae) is modified by the adnominal deictic *yaka* 'up there'.

(8) *Anga* *ahe* *we-nge* *ngu,* *ane,* ***Top_Town*** ***ya-ka,***
time INDEF go-COND TOPIC whatchamacallit Top.Town up-there

khaingu *we-kanu* *nge,* *ngu, ...*
there go-2SG.FUT COND TOPIC
'If you go sometime, what's it called, Top Town up there, if you go there, then ...' (Hos 3:00-3:14)

Proper nouns denoting people or groups of people are rarely found with adnominal deictics, and only with ones have the demonstrative prefix (§4.3). The latter is seen in (9), where the name *Sakingo* is modified by *khayaka* 'up there'.

(9) *Ngingara,* *Betty,* *meme* *Sakingo* ***kha-ya-ka,***
Ngingara.name Betty.name aunt Sakingo.name DEM-up-there

noka *ngu* *itae* *pahe* *ro-ka* *háye-pi.*
3PL TOPIC 3SG.EMPH village downland-there come.to.exist-3PL.R
'Ngingara, Betty, and aunty Sakingo up there, they were born in the village.' (Marit 2:05-2:15)

Common nouns can be possessed, but proper nouns cannot. For common nouns, this is done with a construction [NP_{possr} NP_{possd}] (§7.1.1.1). Some of these are listed in Table 3.4.

Table 3.4: Possession of common nouns via NP juxtaposition

meaning	**alienability**	**possessed form**	**gloss**
'my pig'	alienable	*ima ihe*	1SG pig
'my house'	alienable	*ima pahe*	1SG house
'my husband'	inalienable	*(ima) rame-o*	(1SG) husband-1SG.POSS
'my head'	inalienable	*(ima) pira-o*	(1SG) head-1SG.POSS

The same form [NP_1 NP_2] is found where the second NP is a proper noun. However, it instead has inclusory meaning ('NP_1 together with NP_2') or appositive meaning (expressing that

NP_1 and NP_2 have the same referent). An additional meaning resembles possession, expressing that NP_1 is from the place/tribe NP_2, but this is the reverse order from the [NP_{possr} NP_{possd}] found with common noun possessors. Some examples of these meanings are shown in Table 3.5.

Table 3.5: Inclusory and appositive meaning involving proper nouns via NP juxtaposition

meaning	**meaning type**	**form**	**gloss**
'Yatu and I'	inclusory	*hia Yatu*	1DU.EXCL Yatu
'grandfather Gambe'	appositive	*mba Gambe*	grandfather Gambe
'we two Hurukui people'	appositive	*naya Hurukui*	1DU.INCL Hurukui.tribe
'a woman from Mamia'	possessive-like	*haikhu Mamia*	woman Mamia.village

3.2 Pronouns

Hurukui pronouns can be divided into personal pronouns, emphatic pronouns, vocative pronouns, and interrogative pronouns. Personal and interrogative pronouns can be marked with the contrastive focus suffix *-kha* 'FOC' (§6.5.3).

3.2.1 Personal pronouns

Personal pronouns distinguish three numbers: singular, dual, and plural. The first-person non-singular pronouns distinguish clusivity. These are listed in Table 3.6.

Table 3.6: Personal pronouns

gloss	pronoun
1SG	*ima*
2SG	*na*
3SG	*ngu*
1DU.EXCL	*hia*
1DU.INCL	*naya*
2DU	*nia*
3DU	*nua*
1PL.EXCL	*hita*
1PL.INCL	*naita*
2PL	*neka*
3PL	*noka*

In (10)-(12), pronouns are used in some basic syntactic roles: subject of a verb, object of a verb, and argument of a postposition.

In (10), the pronoun *neka* '2PL' is the subject.

(10) ***Neka*** *ínake* *khaníngè-pi?*
2PL why do.this-2PL.R
'Why did y'all do this?' (KilimPkn 2:18-2:22)

In (11), the pronoun *ngu* '3SG' is the object.

(11) ***Ngu*** *apira-kepi.*
3SG take.off-3PL.FUT
'They will take it (clothing) off.' (Bilas 3:59-4:01)

In (12), the pronoun *ima* '1SG' is the argument of the postposition *khahu* 'COMIT'.

(12) ***Ima*** *khahu* *khai-wa-Ø.*
1SG COMIT be.located-NEG-3SG.R
'He didn't stay with me.' (Nayamo 1:22-1:25)

Personal pronouns can take the suffix *-kha* 'FOC' which marks contrastive focus. I do not consider the resulting forms a separate set of pronouns from the personal pronouns, but inflected

forms of them. This focus marker is analogous to the separate-word particle *kha* 'FOC' (§6.5.3) that is found with most other types of nominals. With personal pronouns, though, the movement of stress to the new penultimate position (§2.4) and the mandatory pronunciation of *kh* as [x] (§2.1.1) show that this marker is a suffix.

In (13), the pronoun *ima* '1SG' is marked for contrastive focus because the speaker thinks he can do a better job starting the fire than the person who is currently trying it.

(13) *Umi~umi-nde* *ngu* ***ima-kha*** *ukumí-tara* *o.*
REP~die-COND TOPIC 1SG-FOC hold-1SG.FUT EMPH
'If it (fire) keeps dying, then I'll start it (by rubbing bamboo).' (Kaikai 4:00-4:03)

3.2.2 Emphatic pronouns

The emphatic pronouns are derived from a Pre-Hurukui inalienable noun root **(h)ita-* 'self', but the forms have been changed somewhat and become less analyzeable. These are listed in Table 3.7.

Table 3.7: Emphatic pronouns

gloss	**source construction (* = Pre-Hurukui)**	**pronoun**
1SG.EMPH	**(h)ita-o*	*itao ~ tao*
2SG.EMPH	**(h)ita-na*	*itana ~ tana*
3SG.EMPH	**(h)ita-e*	*hitae ~ itae ~ tae*
1DU.EXCL.EMPH	**(h)ita-(yo?)-mo*	*hitamo*
1DU.INCL.EMPH	**na (h)ita-(yo?)-mo*	*natamo*
2DU.EMPH	**(h)ita-nia-mo*	*tinamo*
3DU.EMPH	**(h)ita-nua-mo*	*tunamo*
1PL.EXCL.EMPH	**(h)ita-(ito?)-po*	*hitapo*
1PL.INCL.EMPH	**na (h)ita-(ito?)-po*	*natapo*
2PL.EMPH	**(h)ita-neka-po*	*tenekapo*
3PL.EMPH	**(h)ita-noka-po*	*tonokapo*

These pronouns are used in various situations in which one wants to put emphasis on an anaphor in addition to referring to it. The exact nature of this emphasis will require more research, but it often has elements of reflexivity, focus, surprise, and/or sarcasm.

In (14), the pronoun *tana* '2SG.EMPH' is used to add emphasis. It is also marked with the contrastive focus particle *kha* to contrast the listener from the travelers being talked about, but this is not the same as the emphasis being expressed. The emphasis here seems to be because speaker B thinks speaker A has asked a silly question (the journey in question being about 170 kilometers) and responds sarcastically.

(14) A: *Maha* *para* *aráya-mài-pi?*
ground LOC travel-HAB-3PL.R

O *kunie_apono* *khe~kheko* *para* *aráya-mài-pi?*
or car PL~that LOC travel-HAB-3PL.R

B: ***Tana*** *kha* *maha* *para* *aráya-kànu* *o?*
2SG.EMPH FOC ground LOC travel-2SG.FUT or

A: 'Do they travel on foot or in those cars?' (Lae 2:11-2:18)
B: 'Would *you* want to travel on foot?'

In (15), the pronoun *hitapo* '1PL.EXCL.EMPH' is the argument of the postposition *khahu* 'COMIT'. The emphasis meaning is present because the speaker is describing how children who complete secondary school are still, unexpectedly, returning to the village and living the same way as uneducated villagers. It thus carries an element of surprise or incredulity.

(15) *Hameme* ***hitapo*** *khahu* *harawohe* *hí-pikhài-pi.*
again 1PL.EXCL.EMPH COMIT one become-3PL.PRF.R-3PL.R
'They have become like us again.' (Aser 1:29-1:33)

The emphatic pronouns can be used as an argument or adjunct of a verb to give reflexive or reciprocal meaning. In (16), the pronoun *itae* '3SG.EMPH' has reflexive meaning in the semantic role of recipient.

(16) ***Itae*** *moko* *e-ma-re.*
3SG.EMPH speech take-CONT-3SG.R
'He's talking to himself.' (overheard)

In (17), the pronoun *tunamo* '3DU.EMPH' is used with reciprocal meaning. In addition, there is a reciprocal suffix on the verb that reinforces this meaning.

(17) *Hamame* ***tunamo*** *pukhúmbi-nì-khu ...*
the.other.way 3DU.EMPH collide-RECIP-3DU.SS.R
'They crashed into each other, and ...' (WP1 3:31-3:36)

The emphatic pronouns can also be used as possessives, to emphasize that the possessum belongs to a particular possessor rather than someone else. In (18), the pronoun *hitapo* '1PL.EXCL.EMPH' is used as a possessor, to emphasize that the speaker is using lexemes from their own smaller dialect rather than the more common one. This use thus has a bit of contrastive focus meaning.

(18) *Eo, ngu* ***hitapo*** *moko e-khu ngáe-ma-ra.*
yes 3SG 1PL.EXCL.EMPH language take-1SG.SS.R tell-CONT-1SG.R
'Yes, I'm telling it in our own language.' (Abus 17:40-17:45)

The emphatic pronouns are also often used in contexts where there is not a clear emphasis being made. Some of these uses even include both a personal pronoun and the corresponding emphatic pronoun together. Why this is done, rather than only the personal pronouns being used in these situations, will require further research. This kind of use is shown in (19) and (20).

(19) *Humura keya ngu* ***itae*** *kaku.*
frog frog.sp 3SG 3SG.EMPH different
'The *keya* frog is different.' (Abus 6:57-7:00)

(20) ***Tonokapo*** *ake ekha-pínakhài-Ø.*
3PL.EMPH song go.down-3PL.DPST-3SG.R
'They sang with drums.' (Histry2 5:30-5:34)

3.2.3 Vocative pronouns

The vocative pronouns are two special forms *ndapoki* '2DU.VOC' and *ndapono* '2PL.VOC'. They are most commonly used to call out to groups of people. They are derived from the numerals *apoki* 'two' and *apono* 'three', but the latter vocative pronoun can also refer to more than three. The origin of the element *nd-* found in these pronouns is obscure.

In (21), *ndapoki* '2DU.VOC' is used to call out to two people.

(21) *Oo* ***ndapoki!*** *Humia* *nge* *araiku* *we-ma-re.*
oh 2DU.VOC sun TOPIC already go-CONT-3SG.R
'Hey you two! It's already dusk.' (BusHaus 6:58-7:02)

There is also a referential use of these pronouns with third-person meaning, in which *ndapoki* means 3DU and *ndapono* means 3PL. Why these are used instead of the personal pronouns *nua* '3DU' and *noka* '3PL' in these situations requires further research. This usage seems to have a meaning like 'those people over there' in a place that the speaker is indicating. In (22), *ndapono* is used to mean 'these people' referring to the speaker's children, who were mentioned a bit earlier in the story.

(22) *Kohe,* ***ndapono*** *kheko* *para* *ni_nge_ngu* *háyi-pinakhài-Ø.*
okay 3PL that from TOPIC come.into.being-3PL.DPST-3SG.R
'And so these children were born from that situation.' (Marit 1:28-1:32)

3.2.4 Interrogative pronouns

Hurukui has a small closed class of interrogative pronouns: *neku* 'who' and *ino* 'what'. In (23), the pronoun *neku* 'who' is a predicate nominal.

(23) *Na* *ipe-na* ***neku?***
2SG name-2SG.POSS who
'What is your name?' (overheard)

In (24), *neku* 'who' is used as a modifier on the noun *hamunge* 'man'.

(24) *"Hamunge* ***neku*** *arakua* *wa-ka* *khai-Ø* *khuní?"* *ko-mai-pu.*
man who jungle upland-there live-3SG.R EMPH.Q say-HAB-1PL.R
'"What man lives up there in the jungle?", we ask.' (Dwarf 2:23-2:26)

In (25), *ino* 'what' is the object.

(25) *E,* *eha* *oro-ka* ***ino*** *a-mai-pi?*
hey dog downland-there what kill.3SG-CONT-3PL.R
'Hey, what are those dogs killing?' (Hunt5 1:44-1:48)

In (26), *ino* 'what' is the argument of a postposition.

(26) ***Ino*** *para* *we-mánakhài-Ø?*
what LOC go-2SG.DPST-3SG.R
'By what means of transportation did you go there?' (Kimbe 0:23-0:26)

Some interrogative pronouns can take the focus suffix *-kha*. The pronoun *neku* 'who' fuses with it irregularly to become *nékŭkhă*, pronounced [nẹ̯x]. This form is nearly obligatory when *neku* is in the semantic role of agent. This is seen in (27).

(27) ***Néku-kha*** *hamame* *anongo* *khaingu* *para* *we-komu?*
who-FOC back place that LOC go-3SG.FUT
'Who will go back to that place?' (PapaPait 126)

With *ino* 'what', the focus-marked form *ino-kha* can also carry instrumental meaning, as shown in (28). This is another meaning of both the suffix *-kha* and the corresponding separate-word particle *kha* 'FOC' used on other kinds of nominals.

(28) ***Inó-khă*** *sàhi-mài-pi?*
what-INSTR make-HAB-3PL.R
'What is it made of?' (overheard)

3.3 "Whatchamacallit" words

The words *ane* and *akeko*, both glossed 'whatchamacallit', are used in place of something whose name one cannot currently think of. The items replaced are usually nouns, so I treat these words as nominals.

In (29), *ane* 'whatchamacallit' replaces whatever noun would refer to the type of seed/fruit that the man is picking.

(29) ***Ane*** *amburine* *ekhono* *riki-mu ...*
whatchamacallit seed now pick-3SG.SS.R
'He picks the what's-its-name fruit and ...' (WP1 3:44-3:51)

In (30), *ane* 'whatchamacallit' is used to replace a kinship term referring to a woman named Sawang. It thus bears inalienable noun morphology, the generic possessor suffix *-noka* (§3.1).

(30) *We-pu* *pore,* ***ane-noka,***
go-1PL.SS.R all.the.way whatchamacallit-GENER

Sawang *ara* *kha-ma-re* *ya-ka, ...*
Sawang.name place be.located-HAB-3SG.R up-there
'We went all the way up to where what's-her-name, Sawang, lives, ...' (Dokta 2:10-2:17)

The word *akeko* 'whatchamacallit' is used similarly. It is likely derived from the homophonous deictic (§4.2) meaning 'right there near you'. In (31), it is used to replace a Tok Pisin loanword *kantri* 'country', which has no Hurukui equivalent, then the speaker decides to just use the Tok Pisin word.

(31) *Australia* *na* *America* *noka-kha* *China* *khahu* *anáni-pi*
Australia and America 3PL-FOC China COMIT fight-3PL.SS.R

akeko *khaino* *e-pínakhai-Ø,* *kantri* *khaino,*
whatchamacallit this take-3PL.DPST-3SG.R country this

Papua_Niu_Gini *kantri* *khaino* *e-pínakhai-Ø.*
Papua.New.Guinea country this take-3PL.DPST-3SG.R
'It was Australia and America who fought with China and took over this whatchamacallit, this country of Papua New Guinea.' (Komplen 1:54-2:08)

4 Deictics

Hurukui has an elaborate system of deictics, which is a class of words that carry certain spatial information. These have a variety of syntactic uses, so the word class is defined morphologically. The semantic categories expressed include nearness to speaker, nearness to listener, and elevation-based information. The Hurukui-speaking area is mountainous and contains many river gorges, and similarly rich systems of elevation-based deixis are also found in nearby languages such as Iyo (Minter 2009).

Deictics have quite varied syntactic distribution. A common use of deictics is as noun phrases. Like nominals (Chapter 3), they can serve as arguments or adjuncts of verbs, they can be nonverbal predicates, and they can be the arguments of postpositions. Some non-uniformity in this syntactic distribution is observed due to pragmatic factors. For instance, few deictics occur as the agentive subject of a verbal predicate, because agents are more likely to be referred to using their names or other nouns. Deictics can also be adjuncts to verbs. This use is often structurally ambiguous, analyzeable as either nominal or adverbial (see §7.3.6 on bare adjuncts and §7.3.9 on adverbial modifiers).When deictics are used in this way, they are often found in the roles of location, source, or goal. Deictics can also be adnominal, occurring after the modified noun.

Another important function of deictics is as sentence-final discourse particles. These look similar to postverbal adverbial uses, but their meanings are somewhat different, more pragmatic than simply spatial. The choice of deictic value for these seems straightforward in many cases. For instance, when someone is showing or handing you something, 'here' is used, and when someone is telling you about something more relevant to you or a direct response to something you said, 'near you' is used.

Deictics in Hurukui have a significant amount of morphology, and as a result there is a large number of possible wordforms. Their morphological template is shown in Table 4.1.

Table 4.1: Position class template for deictics

Slot -3	**Slot -2**	**Slot -1**	**Slot 0**	**Slot +1**
plural	demonstrative	proximal	base	locative

There is a significant amount of overlap in the syntactic distributions of deictics with and without certain affixes. For instance, deictics both with and without the locative suffix are found in many of the same roles. Additionally, despite the common use of the demonstrative prefix to make a pronominal word meaning the person/thing located somewhere, the corresponding deictic that lacks this prefix can often also be used in the same way. Detailed analysis of how these different forms are used, especially regarding teasing apart syntactic functions such as modification versus reference, will have to await future research.

This chapter is organized starting from the deictic base, which is the smallest string of deictic morphemes that can be an independent word, and goes leftward through the prefix slots followed by the suffix slot. Each section shows various syntactic uses of deictics containing the morphemes that have been introduced so far. §4.1 introduces the deictic bases themselves. §4.2 describes the proximal prefix. §4.3 describes the demonstrative prefix. §4.4 describes the plural prefix, which is only found on deictics with the demonstrative prefix. §4.5 describes the locative suffix. §4.6 discusses a phenomenon whose morphological source is more obscure: deictic-like elements found in the initials of light verb constructions meaning 'be thus, do thus'.

4.1 Deictic bases

Deictic bases are the smallest strings of morphemes that can form an independent deictic word, and to which further affixes can attach. Deictic bases are a small closed class and can be divided into three different groups: *non-oriented*, *participant-oriented*, and *elevation-oriented*. The non-oriented roots can refer to any direction or location, without relating it to participants or geographic features. The participant-oriented roots express direction and location with respect to the speaker and listener. The elevation-oriented roots express direction and location with respect to nearby landforms such as mountains, valleys, rivers, sloped ground, and flat ground. These three subclasses also differ in co-occurrence with the various affixes, so they are separable on both semantic and morphological grounds.

For the non-oriented and participant-oriented deictics, the bases themselves are free and can stand as words without any affixes. In contrast, elevation-oriented deictic bases are composed

of two bound morphemes, the first expressing information about elevation and the second expressing information about distance. Because neither of these can occur without the other, I do not privilege either as being called the "root"[3]. This also means that elevation-oriented deictics always also carry some participant-oriented information.

§4.1.1 describes the non-oriented bases. §4.1.2 describes the participant-oriented bases. §4.1.3 describes the elevation-oriented bases and how they are formed from the combination of the two bound affix sets. §4.1.4 shows the syntactic uses of deictic bases as independent words.

4.1.1 Non-oriented bases

There are three non-oriented bases. The first is the general referential *ngu* 'that place', which can refer to any direction or location. It is homophonous with the personal pronoun *ngu* '3SG', and it is possible that these are actually different senses of the same lexeme with a broader meaning (i.e., simply 'that'). In (32), *ngu* 'that place' is used to refer to a place that had been introduced earlier in the story.

(32)

Nenge	*hi-mu*	***ngu***	*a-mu*	*kha ...*
village	make-3SG.SS.R	that.place	put-3SG.SS.R	and

'He founded a village there, and ...' (Graun 12:08-12:11)

The other two non-oriented roots are the interrogatives *nga* 'where' and *ngapa* 'what place'. The interrogative *nga* 'where' is completely general, asking for any location. It can be answered with nouns, deictics, place names, etc.: anything that answers the question 'where'. (33) shows a very basic use.

(33)

Nga	*wo-re?*
where	go-3SG.R

'Where did he go?' (VD1 2:05-2:07)

The interrogative *ngapa* 'what place' is much rarer. Like *nga* 'where', it can ask for any location, but it must be answered with a specific place name. This is shown in (34). The interlocutor responds in detail by listing the different places.

3 There are some conspicuous resemblances between elevation morphemes and verbs: *ya-* 'uphill, up' resembles *ya(i)-* 'ascend', and *we-* 'alongland' resembles *we-* 'go'. These may be coincidences, or they may be clues about the historical development of this deictic system. For now, I treat these morphemes as unrelated.

(34) ***Ngapa*** *náma-pinakhài-Ø* *kho?*
what.place settle-3PL.DPST-3SG.R I.wonder
'What are the names of the places where they settled?' (Graun 3:18-3:21)

4.1.2 Participant-oriented bases

The participant-oriented deictic roots are *no* 'here' and *eko* 'near you', which do not encode any elevation-oriented information. They are free roots and thus are also bases. Simple uses are shown in (35) and (36).

(35) ***No*** *yo-Ø!*
here come-2SG.IMP
'Come here!' (TP1.Tr 52)

(36) ***Eko*** *ino* *umo* *hapoka-ma-re?*
near.you what tree cut-CONT-3SG.R
'What kind of tree is he cutting (there, by you)?' (VD4 1:14-1:17)

4.1.3 Elevation-oriented bases

Anything that is far from both speaker and listener, and/or whose displacement in the upward or downward direction is relevant, must use the elevation-oriented deictics. Each of these contains exactly one of five bound roots. Figure 4.1 gives a schematic of the elevation-oriented deictic categories.

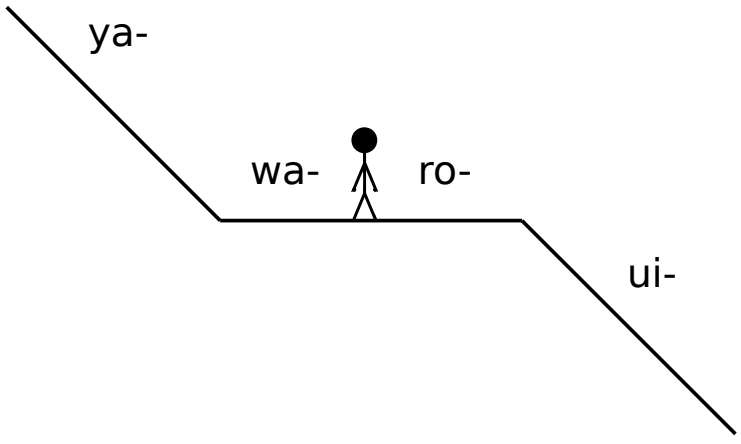


Figure 4.1: Elevation-based deictic roots, excluding *we-* 'alongland'

The concepts of 'up' and 'uphill' are colexified into one root *ya-* 'up', and the concepts of 'down' and 'downhill' are colexified into another root *ui- ~ u-* 'down'. In addition to these two, there are three roots expressing horizontal direction. One, *wa- ~ wai-* 'upland', means direction along a flat surface in the general direction of higher ground. Another, *ro- ~ oro-* 'downland', means direction along a flat surface in the general direction of lower ground. Finally, the root *we-* 'alongland', means horizontal direction towards neither higher ground nor lower ground, but sideways relative to the upland-downland axis. In Figure 4.1, *we-* 'alongland' is the direction perpendicular to the page (both inward and outward from the reader's perspective).

In (37), the root *ya-* 'up' is used to refer to an uphill direction, because the house is uphill from where the story is being told.

(37)

Ima	*yo-khu*	*pahe*	***ya**-ka*	*wo-ra.*
1SG	come-1SG.SS.R	house	up-there	go-1SG.R

'I came and went home.' (Asde 3:32-3:36)

In (38), the root *we-* 'alongland' is used because the speaker knows (even though this took place indoors) that the hill outside slopes up in front of him and down behind him, and the water container is to his side.

(38)

Sono	*humekhe*	*ahe*	***we**-ka*	*wi-sa*	*e-Ø*	*yo-Ø.*
water	container	INDEF	alongland-there	go-SS.IRR	take-SER	come-2SG.IMP

'Go get one of those water containers over there.' (ObJP1 0:15-0:20)

The elevation-oriented deictic roots cannot stand on their own. In order to form a base, they must be immediately followed by one of three *distality* suffixes: *-ni* 'here', *-ka ~ -ta* 'there', or *-ke ~ -te* 'far'. The first, *-ni* 'here', indicates somewhere closer to the speaker. The second, *-ka ~ -ta* 'there', indicates somewhere farther from the speaker (either near the listener or far away from both). The next few examples will illustrate these two distality suffixes before moving on to the third.

In (39), *yani* 'up here' is used because the speaker is talking about some decorations that he has put on his head. They are closer to him than to the listener.

(39) *Opopo* *khuha* *khai-ngu* *nge_ngu* *e-khu* *yo-khu*
thing all DEM-that.place TOPIC take-1SG.SS.R come-1SG.SS.R

ya-ni *keku* *a-pa* *póke-mukhai-Ø.*
up-here LOC put-1SG.DS be.attached-3SG.PRF.R-3SG.R
'I've brought all of this stuff and attached it up here.' (Samting 3:59-4:04)

In (40), *uni* 'down here' is used by speaker A, and *uita* 'down there' is used by speaker B. They refer to the same direction, directly downward from speaker A.

(40) A: *E-Ø* ***u-ni*** *a-o?*
take-SER down-here put-1SG.POT

B: *Maha,* *maha* ***ui-ta*** *a-ke.*
ground ground down-there put-2SG.IMP
A: 'Should I put it down?' (ObjP1 1:24-1:29)
B: 'The ground, put it on the ground.'

Based on how these two suffixes are used in natural situations, *-ka ~ -ta* 'there' is the default choice, while *-ni* 'here' is mostly reserved for locations significantly closer to the speaker. As a result, I often gloss *-ka ~ -ta* as 'there', meaning distance from the speaker rather without necessarily including closeness to the listener.

In (41), *weka* 'over there' is used to talk about the city of Kimbe, which is on another island at a distance of about 455 km. It does not matter which participant the city is closer to, because it is so far away and out of view. As such, both speakers use the suffix *-ka* 'there' for it.

(41) A: *Khimbe* ***we-ka*** *ni* *inóng~inòngi-mukhài-Ø?*
Kimbe.city alongland-there FOC PL~be.how-3SG.PRF.R-3SG.R

B: *Khimbe* ***we-ka*** *nge ...*
Kimbe.city alongland-there TOPIC
A: 'What are things like in Kimbe?' (Kimbe 1:28-1:36)
B: 'In Kimbe, ...'

The third suffix, *-ke ~ -te* 'far', is rarely used. Even for places that are objectively far away compared to distances encountered in everyday life (e.g., the city of Lae at a distance of about 170 km), it is much more common to use *-ka ~ -ta* 'there'. The suffix *-ke ~ -te* 'far' is only used when one wants to emphasize that a place is "**way** up/down/over there".

In (42), *roke* 'way down there (downland)' is used to refer to the large inland valley where some fighting occurred during World War II.

(42) *Anáni-pinakhài-Ø* *makara* ***ro-ke***.
fight-3PL.DPST-3SG.R middle downland-far
'They fought way down there (in the large valley).' (PapaPait 5:54-5:58)

There is thus a five-by-three grid of elevation-oriented deictic bases, listed in Table 4.2, sorted from most-upward to most-downward.

Table 4.2: Elevation-oriented deictic bases

distality / elevation	**here**	**there**	**far**
up	*ya-ni*	*ya-ka*	*ya-ke*
upland	*wa-ni ~ wai-ni*	*wa-ka*	*wai-te*
alongland	*we-ni*	*we-ka*	(unattested)
downland	*ro-ni ~ oro-ni*	*ro-ka ~ oro-ka*	*ro-ke*
down	*ui-ni ~ u-ni*	*ui-ta ~ u-ta*	*ui-te*

4.1.4 Uses of deictic bases as independent words

Deictic bases can function as nonverbal predicates that express where the subject is. Many location predicates, however, use the verb *khai* 'be located' along with the deictic. In (43), *nga* 'where' is a predicate expressing location.

(43) *Ima* *ngiha* ***nga?***
1SG walking.stick where
'Where is my walking stick?' (Kofi 8:54-8:59)

Deictic bases can also function as nominals, including as arguments of verbs and arguments of postpositions. In (44), *waka* 'up there (upland)' refers to an object and is used as the subject of a nonverbal predicate.

(44) *"**Wa-ka*** *ino?"* *ko-khu* *nge-khu* *sáhe-ra.*
upland-there what say-1SG.SS.R see-1SG.SS.R do.well-1SG.R
'I wondered "What is that up there?" and I looked at it well.' (Hunt7 2:48-2:51)

In (45), *eko* 'near you' refers to a person and is used as the subject of the verb *ko-* 'say'.

(45) ***Eko*** *ino* *ko-re?*
near.you what say-3SG.R
'What did that person say?' (Kofi 8:25-8:27)

In (46), *eko* 'near you' refers to an animal and is the object of the verb *ruma-* 'shoot'.

(46) *We-khu* *nge-khu,* *o* ***eko*** *ngu* *rumá-pakhai-Ø.*
go-1SG.SS.R see-1SG.SS.R oh near.you TOPIC shoot-1SG.DS-3SG.R
'I went and saw that I had shot it (pig).' (Hunt11 1:57-2:01)

In (47), *nga* 'where' is the argument of the postposition *para* 'LOC'.

(47) *Ne,* *imisi* ***nga*** *para* *we-make-re?*
so road where LOC go-HAB-3SG.R
'So, where does the road go?' (Trap 1:04-1:07)

Bare adjunct uses of deictics also seem very adverbial, and more research is needed into this area of Hurukui syntax. In (48), *no* 'here' refers to a place and is used as an adjunct or adverbial modifying the verb *nukhi-* 'leave'. It has a semantic role of source.

(48) ***No*** *nukhí* *imisi* *kere* *wé-ka* *wopu* *nge_ngu ...*
here leave.SER road along alongland-there go:1PL.FUT TOPIC
'We'll leave here and follow the road there, and ...' (Lae 2:51-2:55)

In (49), *no* 'here' is used as an adjunct or adverbial modiying the verb *khai-* 'be located, stay, live', in the role of location.

(49) *Hitapo* ***no*** *kha-mai-pu*
1PL.EXCL.EMPH here live-HAB-1PL.R

na *tana* *khura-na* *ui-ta* *kha-ma-ma.*
2SG 2SG.EMPH by.self-2SG.POSS down-there live-HAB-2SG.R
'We live here, and you live by yourself down there.' (Nayamo 0:20-0:23)

In (50), *uini* 'down here' is used in the role of location in the frame of the verb *náma-* 'sit down, settle'.

(50) *Misin* *yo-pi* ***ui-ni*** *námà-pi* *ni_ngé_ngu ...*
mission come-3PL.SS.R down-here settle-3PL.R TOPIC
'The missionaries came and settled down here, and ...' (Man3 4:19-4:23)

In (51), *nga* 'where' is used in the role of goal.

(51) *Neke-noka* *eko* ***nga*** *wo-re?*
cross.cousin-GENER near.you where go-3SG.R
'Where did that guy (cousin) go?' (Kaikai 11:45-11:47)

In (34), reproduced here as (52), *ngapa* 'what place' is used in the role of location or goal.

(52) ***Ngapa*** *náma-pinakhài-Ø* *kho?*
what.place settle-3PL.DPST-3SG.R I.wonder
'What are the names of the places where they settled?' (Graun 3:18-3:21)

Deictic bases can also be used pronominally in other constructions. In (53), *ngapa* 'what place' is pronominal because it precedes the noun *hamunge* 'man' in a possession construction (§7.1.1.1). Adnominal modifiers must follow the noun, so *ngapa* 'what place' is not adnominal here. The phrase *ngapa hamunge* 'man from where' is an instance of a construction expressing tribe or village someone is from, which has this word order. Another example is *Hurukui hamunge* 'Hurukui man'.

(53) *Hamunge* *khaingu* ***ngapa*** *hamunge?*
man this what.place man
'Where is this guy from?' (VD5 1:05-1:08)

In (54), *ngapa* 'what place' is used because the speaker has forgotten a toponym. Whether this use is adnominal (modifying *awiru* 'valley') or appositive requires further investigation.

(54) A: *Hita* *mba-ito-po* *ngu* *awiru* [hesitates]
1PL.EXCL grandfather-1PL.POSS-PL TOPIC valley

ngapa *kho?*
what.place I.wonder

B: *Bintia.* A: *Bintia.*
place.name place.name

A: *Bintia* *akai-pu* *yo-pu.*
place.name follow-1PL.SS.R come-1PL.R
A: 'Our ancestors, the valley of... where again?' (Graun 1:16-1:32)
B: 'Bintia.'
A: 'Bintia. We followed Bintia Valley here.'

Another important use of deictic bases is as sentence-final discourse particles. In (55), *eko* 'near you' is used in this way to point out something to the listener that they did not realize but is

closely connected or relevant to them. In this context, the mother had intended to speak Tok Pisin to a non-Hurukui person but spoke Hurukui instead, without realizing it.

(55) *Akaime* *khai-ngu* *moko* *wohe* *e-ma-ma* ***eko***.
mother DEM-3SG language native take-CONT-2SG.R near.you
'Mom, you're speaking Hurukui.' (Kofi 5:46-5:48)

4.2 The proximal prefix *aka-* ~ *ako-* ~ *ak-*

Deictics also mark an extra distance category, *proximal*, using the prefix *aka-* ~ *ako-* ~ *ak-* 'PROX'. It adds a meaning of nearness. For instance, the base *we-ka* 'over there alongland' has proximal counterpart ***aka****-we-ka* 'right there alongland'. The proximal prefix takes the form *ak-* before a vowel, *ako-* before a consonant when the next vowel is rounded (except with the root *ro-* ~ *oro-* 'downland'), and *aka-* otherwise. Table 4.3 shows the correspondence between some deictic bases and their proximal counterparts.

Table 4.3: Correspondence between some deictics and their proximal counterparts

meaning	**base**	**proximal**
here	*no*	***ako****-no*
near you	*eko*	***ak****-eko*
downland + there	*ro-ka* ~ *oro-ka*	***aka****-ro-ka*
down / downhill + there	*ui-ta* ~ *u-ta*	***ak****-ui-ta* ~ ***ak****-u-ta*

In (56), *akuita* 'right down there' is used to express that Susuruai is nearby.

(56) *Eha* *anute_maisa* *kh-ako-ngu* *ngu* *itae*
dog small.SG DEM-PROX-that.place 3SG 3SG.EMPH

Susuruai ***ak-ui-ta*** *khai-ne* *e-khu* *we-ma-ra.*
place.name PROX-down-there be.located-3SG.DS take-1SG.SS.R go-CONT-1SG.R
'This small dog had been down at Susuruai and I was taking it away.' (Guria5 6:19-6:24)

There are no proximal counterparts to the non-oriented roots *ngu* 'that place', *nga* 'where', or *ngapa* 'what place'. There are also no proximal deictics in which the distality suffix is *-ni* 'here'.

The proximal and 'far' affixes **can** co-occur, although this is rare. The forms *aka-wai-te* 'PROX-upland-far' and *aka-ro-ke* 'PROX-downland-far' are attested in the corpus. What exactly these mean, as opposed to forms like *wai-te* 'upland-far' and *aka-wa-ka* 'PROX-upland-there', will require further investigation, especially regarding the topography of the specific places mentioned. Insights from further analysis may warrant changes to the glosses of some of these morphemes.

In (57), *akawaite* 'PROX-upland-far' is used adverbially with the second verb *yo-* 'come', in the role of goal.

(57) *Ngu_para* *yo-khu* ***aka-wai-te*** *yo-khu*
and.then come-1SG.SS.R PROX-upland-far come-1SG.SS.R

ngahaino *ahe* *ngu* *khai-Ø.*
waterfall INDEF that.place exist-3SG.R
'I came to this place a little ways up, and there was a waterfall there.' (Mamia2 4:20-4:25)

In (58), *akaroka* 'right down there (downland)' is used in the role of location, adverbially with the verb *khai-* 'be located'.

(58) *Itae* *we-pi* ***aka-ro-ka*** *khai-pi.*
3SG.EMPH go-3PL.SS.R PROX-downland-there be.located-3PL.R
'They went and lived right down there (downland).' (Graun 5:51-5:56)

In (59), *akono* 'right here' is adnominal, modifying *nangora* 'greens'.

(59) *Nangora* ***ako-no*** *ngu,* *kepu* *kanapa* *we-wa-pa* *nake*
greens PROX-here TOPIC yesterday garden go-NEG-1SG.R because

hai *sapuru* *ona-Ø* *harawohe.*
only aibika bone-3SG.POSS one
'These greens are only one bunch of aibika, because I didn't go to the garden yesterday.' (Marita 5:11-5:17)

4.3 The demonstrative prefix *khai- ~ kha- ~ kh-*

The prefix *khai- ~ kha- ~ kh-* 'DEM' makes the word more likely to refer to a person or thing that is located in a place, rather than the place itself. For this reason I call it the *demonstrative* prefix, as it makes words that better translate as 'this' and 'that' rather than 'here' or

'there'. Deictics marked with this prefix also tend to be used in a semantic role like agent, patient, or theme. However, these are only tendencies; deictics with this prefix can also refer to places and be used in semantic roles of location, goal, and source, much like their counterparts without the demonstrative prefix.

The prefix, *khai- ~ kha- ~ kh-* 'DEM', appears to be derived from the verb *khai* 'be located'. Its allomorphs are:

- *khai-* when stressed (i.e., in penultimate syllable) and before a consonant
- *kha-* when unstressed (i.e., not in penultimate syllable) and before a consonant
- *kh-* before a vowel

This morphology seems to be derived from a headless relative clause structure (§8.4.2.2) in which the verb *khai* 'be located' occurred together with a deictic to refer to something which was located in that direction. For example, *yaka* means 'up there', and the Pre-Hurukui source clause **khai yaka* meant '(the thing which) is up there' univerbated to become *khayaka* 'that up there'.

The deictic base *nga* 'where' undergoes the velar-alveolar alternation (§2.1.2) when prefixed with *khai-* 'DEM', creating *khai-**nd**a* 'which one?'

Deictics with the demonstrative prefix can occur in many syntactic roles. Like their counterparts without the prefix, they can be used as nominals, including as nonverbal predicates, arguments of verbs, and arguments of postpositions.

In (60), *khaino* 'this' is a nonverbal predicate, expressing identification. The speaker is talking about a particular kind of water, and he is showing it to the audience.

(60) *Sono* *numu-e* ***khai-no***.
water body-3SG.POSS DEM-here
'This is the water.' (Planim 4:31-4:34)

In (61), *khaino* 'this' is the subject of a nonverbal predicate.

(61) ***Khai-no*** *ino?*
DEM-here what
'What is this?' (SickGirl 8:13-8:16)

In (62), *khakono* 'this right here' is the subject of a verbal predicate.

(62) ***Kh-ako-no*** *épŭ_a-mukhài-Ø?*
DEM-PROX-here suffice-3SG.PRF.R-3SG.R
'Is this okay right here?' (Kaikai 4:46-4:49)

In (63), *khaingu* 'that one' is the object of the verb *nge-* 'see'.

(63) ***Khai-ngu*** *nge-pu* *khane-ito* *iki* *ni_ngé_ngu* *khai-Ø.*
DEM-that.place see-1PL.SS.R inside-1PL.POSS pain TOPIC exist-3SG.R
'We see this, and we feel upset.' (Komplen 1:44-1:49)

In (64), *kheko* 'that near you' is the argument of the postposition *para* 'LOC'. Why a deictic plus a locative postposition is used, rather than a deictic with the locative suffix (§4.5), will require further research.

(64) ***Kh-eko*** *para* *ui-ta* *wi-sa*
DEM-near.you LOC down-there go-SS.IRR

pahe *khiri* *ui-ta* *wo-kohu* *hi.*
village old down-there go-1DU.FUT EMPH
'Let's go down this way and go to the old village.' (Hunt10 0:36-0:41)

Deictics with the demonstrative prefix can also be used as bare adjuncts or adverbials in roles relating to location and motion. In (65), a very similar construction to that in (49) is seen (where *no* 'here' was used in a location role with the verb *khai-* 'be located'), but *khaino* 'this' is used instead.

(65) ***Khai-no*** *kha-mai-pi.*
DEM-here live-HAB-3PL.R
'They lived here.' (Misin 0:32-0:36)

In (66), *khainda* 'where, which one' is used in the role of location.

(66) ***Khai-nda*** *we-kopu?*
DEM-where go-1PL.FUT
'Which way will we go?' (Kofi 4:45-4:47)

In (67), *khuita* 'that down there' is an adjunct to the verb *we-* 'go', in the role of goal.

(67) ***Kh-ui-ta*** *wi-Ø* *kha-kepi.*
DEM-down-there go-SER stay-3PL.FUT
'They will go stand down below.' (GiamanPik 0:42-0:45)

More clearly adverbial are the postverbal uses of these deictics, since nominal adjuncts cannot occur in this position. In (68), *khakaroka* 'that right down there (downland)' is used adverbially, modifying the verb *khai-* 'be located, live'.

(68) *Okh-e* *ngu* *khai-Ø* ***kh-aka-ro-ka.***
younger.sibling-3SG.POSS 3SG live-3SG.R DEM-PROX-downland-there
'His younger sibling lives right down there.' (Nadzab 5:35-5:39)

Deictics with the demonstrative prefix can also be used adnominally, as determiners. In (69), *khaino* 'this' is used adnominally. Participant-oriented deixis is used because she is showing a picture to her interlocutor.

(69) *Ara* ***khai-no*** *nge* *ara* *hawera.*
place DEM-here TOPIC place empty
'This is an empty place.' (Chameleon 2:21-2:25)

In (70), *khuita* 'that down there' is used adnominally. The deictic value of 'down' is used because the cucumber is on the floor.

(70) *Mui* ***kh-ui-ta*** *hi-mana*
cucumber DEM-down-there make-2SG.DS

ngu *khane-Ø* *we-ka* *wi-nde-Ø.*
3SG inside-3SG.POSS alongland-there go-POL.IMP-3SG.IMP
'Put the cucumber inside it, please.' (ObjP5 0:15-0:20)

In (71), *khaingu* 'that one' is used adnominally, modifying the noun *ngohu* 'trap'.

(71) *Ngohu* ***khai-ngu*** *nge-ka* *wo-ra* *nge.*
trap DEM-that.place see-PUR go-1SG.R TOPIC
'I went to see this trap.' (Trap3 1:14-1:17)

In (72), *khaino* 'this' modifies the noun *sono* 'water'.

(72) *Sono* ***khai-no*** *horórŭ_a-kòmu.*
water DEM-here spill.INTR-3SG.FUT
'This water will spill.' (Kofi 1:35-1:37)

In (73), *khuita* 'that down there' modifies the noun *uho* 'firewood'.

(73) *Uho* ***kh-ui-ta*** *e-Ø* *a-khokhe-khi!*
firewood DEM-down-your.side take-SER put-fast-2DU.IMP
'You two get that firewood and put it!' (Kaikai 10:00-10:03)

Deictics with the demonstrative prefix can also function as sentence-final discourse particles. In (74), *kheko* 'that near you' is used to express that the speaker is showing something to the listener, or the utterance is sort of directed at them in a more pointed way than a normal informative sentence. The speaker wants the listener to empathize with the ashamed chameleon.

(74) *Moko* *ě-pia* *ngu* *ako*
speech take-3DU.DS 3SG poor

orápě_ì-ne *khai-Ø* ***kh-eko***.
be.ashamed.3SG-3SG.DS stay-3SG.R DEM-near.you
'The two of them talk to him and he (chameleon) feels ashamed and stays there (I'm showing you).' (Chameleon 4:10-4:13)

In (75), *kheko* 'that near you' is used as a sentence-final discourse particle, in a very similar way to *eko* 'near you' in (55).

(75) *Na-kha* *moko* *wana-ma-ma* ***kh-eko***.
2SG-FOC speech do.poorly-CONT-2SG.R DEM-near.you
'You're speaking poorly.' (Kofi 6:19-6:22)

4.4 The plural prefix *khe-*

Deictics with the demonstrative prefix can mark the number of their referents, while those without it cannot. Singular number is unmarked, while plural (here, meaning two or more) is obligatorily marked. The main strategy for marking plurality in these deictics is the prefix *khe-* 'PL' before the demonstrative prefix *khai- ~ kha- ~ kh-*. For instance, *khaweka* 'that over there' has plural *khekhaweka* 'those over there'.

This structure seems to derive from headless relative clauses, in which the location verb *khai* was doubled to indicate multiple objects (in this example, Pre-Hurukui **khai khai weka* > *khekhaweka*). This is in line with Hurukui's use of full reduplication to mark pluractionality in verbs (§5.2.1) and ideophones (§6.7) and nonsingular number in adjectives (§6.1.1).

The pluralized deictics are used in the same syntactic roles as their singular counterparts. However, sentence-final discourse particle uses of the pluralized deictics are not attested.

In (76), *khekhaingu* 'those ones' is used as a noun phrase, the object of the verb *e-* 'take'.

(76) ***Khe-khai-ngu*** *e-pi* *kha* *hamunge* *noka-pi.*
PL-DEM-that.place take-3PL.SS.R and man kill.3PL-3PL.R
'They took these things (weapons) and killed men.' (PapaLaif 1:34-1:39)

They can also be used adnominally. In (77), *khekhaingu* 'those ones' is used in this way.

(77) *Ekhono* *uphue* ***khe-khai-ngu*** *yo-make-re.*
now problem PL-DEM-that.place come-HAB-3SG.R
'Now these problems are still coming.' (Hevi 1:57-2:01)

In (78), *khekhaingu* 'those ones' is used adnominally to modify *mehu* 'tree bark'.

(78) *Mehu* ***khe-khai-ngu*** *nge*
tree.bark PL-DEM-that.place TOPIC

phai *kha* *hapoka-pa* *hikh-hukh* *i-ne ...*
knife INSTR cut.small-1SG.DS REP~short become-3SG.DS
'I cut this bark into short pieces with a knife, and ...' (SkinDiwai 3:26-3:30)

It should be noted that deictics with the demonstrative prefix can also be pluralized by full reduplication. In (79), *khaino* 'this' is adnominal and is fully reduplicated to mark non-singular number.

(79) *Opopo* *numu-e* *nge_ngu* *opopo* ***khai-no~khaino***
food body-3SG.POSS TOPIC thing DEM-here~PL

e-khu *yo-ma-ra.*
take-1SG.SS.R come-CONT-1SG.R
'For food, I've brought these things.' (Fluffy 6:10-6:15)

4.5 The locative suffix *-ku ~ -ngu*

The suffix *-ku ~ -ngu* 'LOC' forces a deictic word to refer to a place, not a person or thing, and to be used in a semantic role like location, goal, or source. They cannot be used in argument roles like subject and object. For these reasons, I call the suffix *-ku ~ -ngu* a *locative* marker. As noted in previous sections, however, deictics that lack it can also often be used in these same syntactic roles. This requires further research.

Only the participant-oriented and elevation-oriented deictics can take the locative suffix. The non-oriented deictics cannot. The suffix takes the allomorph *-ngu* after the suffix *-ni* 'here', and *-ku* otherwise.

In (80), *uningu* 'down here (LOC)' is the argument of the postposition *para* 'from'.

(80) ***U-ni-ngu*** *para* *we-pi* *ya-ka* *we-pi* *kha ...*
down-here-LOC from go-3PL.SS.R up-there go-3PL.SS.R and
'From down here, they went up there, and ...' (Man3 4:19-4:26)

In (81), *noku* 'here (LOC)' bears the locative suffix and is used in the role of goal with the verb *a-* 'put'.

(81) *Ikhu* *Anutu* *kha* *sáhi-mu* ***no-ku*** *a-ne* *kha* *ni_ngé ...*
native God FOC create-3SG.SS.R here-LOC put-3SG.DS and TOPIC
'God created the native people and put them here, and ...' (Mahapa 1:25-1:29)

In (82), the deictic *u(i)taku* 'down there (LOC)' is used as the argument of the postposition *keku* 'LOC'. That is, both the deictic and the postposition explicitly mark locative meaning. The reason for this will require further research.

(82) *Rumpa* *nge* ***u-ta-ku*** *keku* *hi-pina*
village.name TOPIC down-there-LOC LOC make-3PL.DS

háyi-nekhai-Ø.
arise-3SG.DPST-3SG.R
'They founded Rumpa down there.' (Graun 13:01-13:04)

In (83), *u(i)taku* 'down there (LOC)' is used as an adverbial or adjunct in the role of location.

(83) *Hindaura* *yo-pu* *Bibuai* ***ui-ta-ku*** *isí-pukhài-pu.*
afternoon come-1PL.SS.R village.name down-there-LOC sleep-1PL.PRF.R-1PL.R
'In the afternoon we came and slept in Bibuai.' (Raun 1:56-2:00)

In (84), *wakaku* 'at that place (upland)' is used as an adverbial or adjunct in the semantic role of goal.

(84) *Nare-khu* *e-khu* *yo-khu* ***wa-ka-ku*** *a-khu ...*
carry-1SG.SS.R take-1SG.SS.R come-1SG.SS.R upland-there-LOC put-1SG.SS.R
'I carried it here and left it up there, and ...' (BusHaus 1:26-1:29)

In (85), *khuitaku* 'that down there' is used adnominally on *ngunu* 'lake'.

(85) *Ngunu* ***kh-ui-ta-ku*** *ekha-mu* *we-ne ...*
lake DEM-down-there-LOC descend-3SG.SS.R go-3SG.DS
'It (lizard) went down into that lake down there, and ...' (PutimBebi 2:20-2:23)

Locative-marked deictics are also often used as sentence-final discourse particles. In (86), *khanoku* 'this (LOC)' is used to direct A's attention to what is happening right in front of him, because A told B to do something that B was already doing.

(86) A: *Newehi,* *na* *mburumu* *khakhuna-ke!*
cross.cousin 2SG sweet.potato peel-2SG.IMP

B: *Mburumu* *khakhuna-ma-ra* ***kha-no-ku****.*
sweet.potato peel-CONT-1SG.R DEM-here-LOC
A: 'Cousin, peel the sweet potatoes!' (Kaikai 5:30-5:36)
B: 'I *am* peeling the sweet potatoes (here, you see).'

There is a deictic form *ngungu*, but it does not involve the locative allomorph *-ngu*. Instead, it is a reduplication of *ngu* 'that place' used to indicate repeated motion towards a goal. This is shown in (87).

(87) *Hita* ***ngu~ngu*** *ekha-pu* *óngi-pu* *kha ...*
1PL.EXCL that.place~PL fall-1PL.SS.R rest-1PL.SS.R and
'We fell all over the place, and rested, and ...' (Skiu 1:04-1:08)

4.6 Deictic elements in 'do thus' verbs

In Hurukui's extensive system of light verb constructions (LVCs), discussed in Part II of this dissertation, some deictic-like forms are used, but these differ a bit from the independent-word deictics on which this section focuses. These include the synonymous *kha(i)nongV*, *kha(i)ningV*, *kha(i)nokV*, and *kha(i)nikV*, which combine with the light verb root *i-* 'be, do' to make complex stems meaning 'be thus, do thus'. These elements all appear to be derived from the deictic *khaino* 'this'. One of these stems is illustrated in (88).

(88) *E-pi* *khánà-pi* ***khanóngĭ****-mai-pi.*
take-3PL.SS.R wrap-3PL.SS.R do.thus-HAB-3PL.R
'They wrap it like this.' (Sik 1:00-1:04)

There is also an element *inongV*, which combines with *i-* 'be, do' to make a complex stem meaning 'be how?, do what?'. This element appears to be derived from the interrogative pronoun *ino* 'what' (§3.2.4). One usage of this stem is shown in (89).

(89) *Mui* *e-Ø* ***inóngi**-o?*
cucumber take-SER do.what-1SG.POT
'What should I do with the cucumber?' (ObjP5 0:35-0:37)

Many of the other deictics with the demonstrative prefix *khai- ~ kha- ~ kh-* can undergo similar derivation, but these are significantly more rare. In (90), the deictic *kheko* 'that near you' is used as a deictic element in a 'do thus' verb, to refer to the way the listeners are behaving.

(90) *Eme-Ø* *ahe* *í-kha-kòmu* *nge,*
eye-GENER another become-PRF.IRR-3SG.FUT COND

neka *ekhono* ***khekóngi**-tèpi?*
2PL now do.that.near.you-2PL.FUT
'If it had turned out otherwise, would y'all be doing that now?' (Robert 3:01-3:05)

Other conceivable stems derived in the same way, such as *khawekángi-* 'do like that over there' and *khuitángi-* 'do like that down there', are so far only attested in elicitation.

These deictic elements are highly fused with the light verb root, so their final vowel has been lost. The element *-(i)kV ~ -(i)ngV* which follows *khaino* 'this' or *ino* 'what' is possibly related to the locative *-ku ~ -ngu*. Thus, the form *kha(i)nokV* possibly corresponds to the deictic *kha(i)noku* 'this (LOC)'. However, there are no extant deictics or pronouns corresponding to the other forms: **kha(i)ningu*, **kha(i)nongu*, **kha(i)niku*, and **inongu*. Whatever process created these elements is no longer productive.

5 Verbs

The verb is the most morphologically complex word class in Hurukui. The class is most reliably defined morphologically rather than syntactically or semantically, but syntactic generalizations can still be made. Verbs head verb phrases, which can stand alone as the only constituent in a clause that predicates. They take agreement suffixes showing the person and number of their subjects.

(91) ***Ume-re***.
die-3SG.R
'It (fire) died.' (Kaikai 5:54-5:56)

(92) ***Khaka-Ø***.
stay-2SG.IMP
'Stay!' (Samting 2:16-2:18)

This chapter focuses on verbs whose roots are morphologically simple. The Hurukui verb lexicon is heavily enriched by light verb constructions in various states of univerbation. These are introduced instead in Chapter 7 and investigated in detail in Part II.

The verb marks three persons and three numbers (singular, dual, plural), with some pervasive patterns of syncretism (§5.1.2). Matrix-clause verbs can be divided into two classes, medial and final, based on whether they contain switch reference morphology, which is used for subject tracking in clause chains (§5.1.3). Various verb roots and suffixes undergo processes of ablaut (as well as some less predictable alternations and some suppletion), which are conditioned both lexically and by the subsequent morpheme (§5.1.4). The TAM categories marked on verbs are divided into two types of reality status: realis and irrealis, based on which switch reference morphemes are used in clause chains of that TAM category (§5.1.5).

Hurukui verb inflection and derivation is mostly agglutinating and suffixing. Exceptions include object marking fused with the root in a few verbs (§§5.2.2.3 and 5.2.2.4). There is also a reduplicative process marking repetitive meaning (§5.2.1).

This chapter is structured as follows. §5.1 gives some conceptual preliminaries necessary for understanding the structures and functions involved in Hurukui verbs. §5.2 discusses the

formation of stems, which are strings of morphemes ready to receive inflection. §5.3 discusses inflectional morphology for final verbs. §5.4 discusses inflectional morphology for medial verbs. §5.5 describes subordinating morphology.

5.1 Preliminaries

To aid understanding of the Hurukui verb system, some preliminary concepts are introduced in this section. In §5.1.1 I briefly define some morphophonemic notation which I use to make the forms of morphemes more concise to write. In §5.1.2 I describe the basic person and number categories and syncretism patterns. In §5.1.3 I describe the important morphological distinction between medial verbs and final verbs. In §5.1.4 I describe the distinction between roots and stems, and I discuss the need for positing phonologically null morphemes. In §5.1.5 I describe the reality status system (realis vs irrealis). In §5.1.6 I give a position class template showing the slots into which verbal affixes can be sorted. In §5.1.7 I describe the phenomenon of *shapes*, which are allomorphic alternants found in many verbal morphemes (but distinct from more predictable morphophonemics such as allomorphy that occurs due to the velar-alveolar alternation), and the conditions under which they appear.

5.1.1 Morphophonemic notation

To make some allomorphy more concise to write, I use a few morphophonemic notations consisting of capital letters. These are all related to the velar-alveolar alternation (§2.1.2). Table 5.1 shows the symbols and their realizations.

Table 5.1: Morphophonemic notation for the velar-alveolar alternation

notation	**realization after *i***	**realization elsewhere**
K	*t*	*k*
KH	*s*	*kh*
NG	*nd*	*ng*

For instance, the future suffixes all begin with *K*, e.g., *-́Kara* '1SG.FUT' and *-Kanu* '2SG.FUT'. These are realized with initial *t* after *i*[4] and with initial *k* elsewhere. Thus, the verb *pi-* 'stab' has 1SG.FUT form *pí-tara* and 2SG.FUT form *pi-tanu*, while the verb *humba-* 'fill (tr.)' has 1SG.FUT form *humbá-kara* and 2SG.FUT form *humba-kanu*.

Some suffixes begin with a velar stop *k*, *kh*, or *ng*, but do not undergo this alternation, so they are written with lowercase letters. For instance, the 1SG.SS.R suffix *-khu* (§5.4.1), when following *i*, is always *-khu*, never **-su*.

Another notational convention that I use is a hyphen with an acute accent over it at the beginning of some suffixes, e.g., *-́Ki* 'DES'. This indicates that the syllable before the suffix is stressed, even if it is not penultimate in the word. I do the same with a grave accent when showing a complex of multiple suffixes which assigns secondary stress to the preceding syllable, as in *-̀mukháí-wa-Ø* '-3SG.PRF.R-NEG-3SG.R'.

5.1.2 Persons and numbers

Hurukui verbs agree with their subject in person and number. The persons marked are first, second, and third. The numbers marked are singular, dual, and plural. These are very similar to the distinctions made in the personal pronouns (§3.2.1), but there are some key differences.

The first difference between verb agreement and pronouns is that verb agreement does not mark clusivity in the first-person non-singulars. While pronouns distinguish *hia* '1DU.EXCL' and *naya* '1DU.INCL', verbs mark these values the same: 1DU. Similarly, while pronouns distinguish *hita* '1PL.EXCL' and *naita* '1PL.INCL', verbs mark these values the same: 1PL. This pattern is found in all verb marking that contains person and number information, including subject markers on final verbs (§5.3), switch reference markers on medial verbs (§5.4), and object marking (§5.2.2). Within object marking, the pattern is found not only in the object suffixes (§5.2.2.1), but also in the 'affect' and 'give' verbs, which have separate fusional roots for each object category (§§5.2.2.3 and 5.2.2.4).

4 If this *i* is elided, it still conditions the alveolar-initial allomorph, but if it is part of a diphthong *ai* that is reduced to *a*, then the velar-initial allomorph will be used.

The second difference between verb agreement and pronouns is that verb agreement often syncretizes second and third person in the non-singular numbers. While pronouns distinguish *nia* '2DU' and *nua* '3DU', many person-indexing markers on verbs conflate these two (e.g., *-rei* '23DU.R'). Similarly, while pronouns distinguish *neka* '2PL' and *noka* '3PL', many person-indexing markers on verbs conflate these two (e.g., *-pi* '23PL.R'). This pattern is found in all subject markers on final verbs (§5.3) and switch reference markers on medial verbs (§5.4). Within object marking (§5.2.2), it is found in only one of the object suffixes (the anomalous *-iki* '23PL.OBJ', §5.2.2.1) as well as the paradigms of the fusional 'affect' and 'give' verbs (§§5.2.2.3 and 5.2.2.4). Additionally, both the 'give' verbs and the imperative paradigm conflate 2SG with 3SG, e.g., *ne-* 'give.23SG' and *-ke* '23SG.IMP'.

In my glosses, I use the shortest applicable abbreviation unless I am specifically drawing attention to the syncretism in the example. The abbreviations used are '1DU' and '1PL' for the forms showing syncretism for clusivity. For the syncretic 23DU and 23PL forms, I gloss '2DU' or '3DU' and '2PL' or '3PL', depending on the context.

5.1.3 The medial-final verb distinction

Hurukui makes frequent use of a morphosyntactic structure called clause chaining. This is described in more detail in Chapter 8, but here I provide a short overview of the concept that is needed for understanding Hurukui verb morphology.

In Hurukui, as in many Papuan languages, sentences can consist of long chains of multiple clauses in coordination. In such a chain, the last clause is called final, and its verb is also called final. If there is only one clause in a sentence, then it is a final clause and its verb is a final verb. Thus, a final verb is a verb that can end a sentence or be the only verb in an utterance. All the preceding clauses in the chain are called medial, and their verbs are also called medial. Thus, a medial verb is a verb that occurs at the end of a non-subordinate clause but cannot occur at the end of a sentence[5].

5 There are complications to this distinction, such as pragmatics that allow a medial verb to occur in the last clause of certain chains, but those are not of importance for understanding the distinction between medial morphology and final morphology. Such complications will be discussed in Chapter 8.

This distinction between medial and final verbs is crucial to Hurukui verbal morphology. These two types of verbs take completely different morphology in the inflection slot (Slot 6). The inflectional morphology found on final verbs is treated in §5.3, and that on medial verbs is treated in §5.4.

5.1.4 Roots, stems, and null morphemes

A verb root is a single morpheme, the smallest thing to which verbal inflectional morphology can attach. (93) and (94) show some very simple forms involving the root *yo-* 'come'. In (93), it is marked with a zero suffix for second-person singular imperative. In (94), it is marked with a suffix for 3SG subject in the realis mood.

(93) ***Yo**-Ø!*
come-2SG.IMP
'Come!'

(94) ***Yo**-re.*
come-3SG.R
'That person came.'

Inflectional morphology can also attach to longer, multi-morphemic forms. These are called stems. A stem is a string of morphemes that is ready to receive verbal inflection, and it consists of a root plus optional other suffixes. An example of a stem that is longer than a root is *nge-ya-* (see-1SG.OBJ). (95) and (96) show it with the same two suffixes as (93) and (94).

(95) ***Nge-yá**-Ø!*
see-1SG.OBJ-2SG.IMP
'Look at me!'

(96) ***Nge-ya**-re.*
see-1SG.OBJ-3SG.R
'That person saw me.'

In this description, I treat phonologically null forms as things that can trigger morphophonemic alternations and other effects on nearby morphemes. I consider this not to be reflective of what is actually going on when speakers learn and use language, due to my usage-based theoretical persuasion. However, I consider it the easiest and most efficient way to impart to the reader an understanding of these phenomena in Hurukui verb morphology. This resembles the concept of an *effective theory* in the philosophy of science: it is limited by design and meant as a useful approximation, systematizing what is relevant and irrelevant for the purposes at hand, rather than a full explanation of the true mechanisms (Wells 2012: 1).

I will thus talk in terms of null/zero morphemes conditioning changes and occupying slots in paradigms. For instance, the negative realis paradigm contains the suffix -Ø '3SG.R', while other persons use overt markers. This zero has a specific meaning: a negative realis verb with no overt person-marking suffix must be interpreted as having 3SG subject, rather than being unspecified for person. I will also treat roots/stems that must be followed by a paradigm containing a zero morpheme as being bound rather than free. For instance, the fusional root *opa-* 'give to me' can be uttered on its own, but this must be interpreted as a command to a singular listener; that is, it must be parsed as *opa-Ø* (give.1SG-2SG.IMP). For this reason, it is not free in the same way as a morpheme like the deictic *no* 'here', which can be marked for more specific meanings such as locative (with *-ku*, §4.5) or proximal (with *ako-*, §4.2), but remains unspecified for those semantic features in the absence of the relevant affixes[6].

5.1.5 Realis and irrealis status

For some morphosyntactic purposes, Hurukui makes a binary division of TAM into two categories: *realis* and *irrealis*. These are together called the dimension of *reality status*, or just *status* for short. To speak very roughly, realis is for predicates that have already happened or are happening, i.e., events that are "real", whereas irrealis is for everything else. However, there are plenty of gray areas between these semantic prototypes, both within individual languages and cross-linguistically (Jones in prep.).

In Hurukui, the reality status of a clause chain is determined by which TAM category is used on its final verb. The Hurukui inflection slot (Slot 6) has a single realis paradigm (§5.3.1), and all other paradigms in that slot are treated as irrealis. This includes imperative (§5.3.4), conditional (5.2.6), and future (5.3.2). The status of a chain is reflected by which same-subject markers are used on its medial verbs (§§5.4.1 and 5.4.2).

To illustrate, consider the markers *-mu* '3SG.SS.R' and *-KHa* 'SS.IRR'. In (97), the final verb *khai* is marked as realis, so the medial verb *ekhamu* uses the realis same-subject marker *-mu*. In

6 This distinction is not unlike the conceptualization of verbal arguments versus adjuncts in which arguments are required by the verb's frame while adjuncts are optional extra additions.

contrast, in (98), the final verb *yo* is marked as imperative, so the medial verb *ekhakha* uses the irrealis same-subject marker *-KHa*.

(97) *Sono* *ekha-**mu*** *ngu* *khai-**Ø**.*
water go.down-3SG.SS.R that.place be.located-3SG.R
'It (cassowary) bathed and was there.' (Muruk2 2:21-2:25)

(98) *Sono* *ekha-**kha*** *yo-**Ø!***
water go.down-SS.IRR come-2SG.IMP
'Bathe and come!' (elicited)

The desiderative (§5.2.4.3) presents a more complicated case, because the verbs that are under its scope are treated as irrealis, but the desiderative verb itself and any other verbs with which it is chained but are not under its scope could be either realis or irrealis. In (99), the verbs *wisa*, *nukhisa*, and *we-* are all desiderative (wanting to go, leave, and go), shown by the SS.IRR suffix *-KHa* (here in its allomorph of *-sa*) on the first two of these. However, the final TAM marking on *wékimare*, after the desiderative marker, is realis. This is because the person really did want to do these things, so the state of wanting is real even though the things they wanted to do are not.

(99) *Wi-**sa*** *sono* *nukhi-**sa*** *pore* *arakua* *ya-ka* *wé-ki-ma-**re**.*
go-SS.IRR water leave-SS.IRR INTENS jungle up-there go-DES-CONT-3SG.R
'He wanted to leave the river and go all the way up into the jungle.' (Mamia3 2:43-2:46)

The purposive (§5.5.1), which creates subordinate clauses, behaves similarly to the desiderative in this way. All the verbs in the subordinate chain are irrealis, but the verbs in the matrix chain can have either status.

5.1.6 The morphological structure of the Hurukui verb

Table 5.2 shows a broad position class template for the Hurukui verb. It is a simplification to give an overview of what typical verbs look like, and complications will be described in the relevant sections. The slots are macroscopic, some of them containing multiple sub-slots. For

instance, multiple aspect markers (Slot 3) can co-occur with a strict ordering, and some of the classes of inflectional morphemes (Slot 6) can be followed by further suffixes.

Table 5.2: Position class template for most Hurukui verbs

Slot 0	**Slot 1**	**Slot 2**	**Slot 3**	**Slot 4**	**Slot 5**	**Slot 6**
root	object	adverbial stem formant	aspect	polarity	conditional	inflection
fusional object-marked root						

I use the term *stem* to mean a string of morphemes that is ready to be inflected with something in Slot 6. I thus first treat stem-forming processes, before moving on to inflection. In the discussion of stem formation, I largely ignore the question of whether certain affixes are more inflectional or more derivational. While some of them are clearly derivational, such as the adverbial stem formants (§5.2.3), some of them have less clear status on this dimension, such as the object suffixes (§5.2.2.1) and the perfect and distal past TAM markers, which also inflect for subject (§§5.2.4.1 and 5.2.4.2). The more important distinction for the purposes of this description is between root, stem, and inflection.

5.1.7 The alternating shapes of verbal morphemes

All morphemes that can occur non-word-finally in verbs have three basic allomorphs which I call *shapes*. I use this term to distinguish these allomorphs from other allomorphy caused by phonological processes such as monophthongization (§2.5.2) or the velar-alveolar alternation. The shapes are defined by their segmental phonology (ignoring stress), and their distribution is conditioned by which morpheme follows them in the grammatical word.

The three shapes are the *default shape*, the *pre-realis shape*, and the *pre-imperative shape*. The pre-realis shape is used when the next morpheme is a member of the realis paradigm in final verbs (§5.3.1), and the pre-imperative shape is used when the next morpheme is a member of the imperative paradigm in final verbs (§5.3.4). The default shape is used otherwise, and I use it as the citation form when talking about morphemes in prose, when labeling rows in tables, etc. The only

morpheme with a fourth shape is the root *we-* 'go', whose extra shape *wi* is mentioned in the relevant sections.

As an example, (100)-(102) illustrate the shapes of the verb root *isi-* 'sleep'. (100) shows its default shape *isi-* when followed by a different-subject switch reference suffix (§5.4.4), (101) shows its pre-realis shape *ise-* when followed by a realis conjugation suffix, and (102) shows its pre-imperative shape *isa-* when followed by an imperative suffix.

(100) ***isi****-pa*
sleep-1SG.DS
'I slept and ... (DS)'

(101) ***ise****-ra*
sleep-1SG.R
'I slept'

(102) ***Isa****-ni!*
sleep-2PL.IMP
'Y'all sleep!'

An important consequence of this distribution, in which the pre-realis and pre-imperative shapes are only used before morphemes in Slot 6 (inflection), is that it is only the last morpheme of a stem that undergoes alternation. In (103) and (104), the root *ukumi-* 'grab' is used in the realis status. However, it only has its pre-realis shape *ukume-* in (103), in which the realis conjugation marker *-ra* '1SG.R' immediately follows it. In contrast, in (104), the continuative marker *-ma* intervenes, so the root is kept in default shape.

(103) ***ukume****-ra*
grab-1SG.R
'I grabbed it'

(104) ***ukumi****-ma-ra*
grab-CONT-1SG.R
'I am grabbing it'

In (105) and (106), the root is used in the imperative mood. However, it only has its pre-imperative shape *okoma-* in (105), in which the imperative conjugation marker *-ni* '2PL.IMP' immediately follows it. In contrast, in (106), the continuative marker *-maka* (which is itself in a pre-imperative shape) intervenes, so the root is kept in default shape.

(105) ***Okoma****-ni!*
grab-2PL.IMP
'Y'all grab it!'

(106) ***Ukumi****-maka-ni!*
grab-CONT-2PL.IMP
'Y'all keep grabbing it!'

The phonologically null suffix *-Ø* '2SG.IMP' triggers the pre-imperative shape just as well as its overt non-singular counterparts do. Thus, one has *okoma-Ø* (grab-2SG.IMP), *okoma-khi* (grab-2DU.IMP), and *okoma-ni* (grab-2PL.IMP). However, while null affixes can trigger ablaut like this, they cannot block it from being triggered on an overt morpheme farther to the left. Thus, the verb

nge-na-ra (see-2SG.OBJ-1SG.R) 'I saw you' has the root *nge-* 'see' in its default shape because there is an overt suffix between it and the realis marker, but the verb *ngo-Ø-ra* (see-3SG.OBJ-1SG.R) 'I saw it / that person' has the root in its pre-realis shape *ngo-*, despite the definite null object marker intervening.

There are some patterns in how the three shapes are formed. Most of the alternations are realized by vowel changes, which I refer to as ablaut. I name the patterns of ablaut based on the final vowels found in the three shapes (in the order of default, pre-realis, pre-imperative). Thus, *isi-* 'sleep', shown in (100)-(102), has *i/e/a ablaut* because its shapes are *is**i***-, *is**e***-, *is**a***-. Some morphemes do not change shape (or, equivalently, all three of their shapes have the same form) and some have idiosyncratic changes. Such morphemes do not have a named ablaut pattern.

All morphemes ending in *i* undergo ablaut of some kind[7]. The pattern of *i/e ablaut* has the vowel pattern *-i*, *-e*, *-e* in the three shapes. The pattern of *i/e/a* ablaut seen in *isi-* 'sleep' does not involve changes to the morpheme-internal vowel. However, the pattern of *i/e/a ablaut with lowering*, which does change morpheme-internal vowels, is found in a small number of morphemes with a similar phonological form. It can only apply to morphemes which end in *i* and have additional high vowels in preceding syllables. In the pre-realis shape, only the final vowel changes (*i*>*e*), but in the pre-imperative shape, all of the morpheme's vowels are lowered (final *i*>*a*, internal *i*>*e*, internal *u*>*o*). All pre-realis shapes ending in *e* are involved in one of these types of ablaut from a default shape ending in *i*. Whether an *i*-final morpheme has *i/e* ablaut, *i/e/a* ablaut, or *i/e/a* ablaut with lowering is a lexical property of that morpheme.

All *e*-final morphemes also undergo ablaut. These all use the same type, *e/o ablaut*, which has the final vowel pattern *-e*, *-o*, *-e*. The only exception is the root *we-* 'go', which has *-o* in the pre-realis shape but irregularly has *-a* in the pre-imperative shape (in addition to its extra shape *wi* mentioned earlier). Not all pre-realis shapes that end in *o* are involved in *e/o* ablaut. For instance, the roots *ko-* 'say' and *yo-* 'come' do not change stem shape, so all of their shapes have *o* in the root.

Table 5.3 shows the named ablaut patterns and some morphemes illustrating each one.

7 For these purposes, morphemes ending in the diphthong *ai* (e.g., *khai-* 'be located', *yai-* 'climb, build') do not count as ending in *i-*. No verbal morphemes end in *ui*.

Table 5.3: Named ablaut types

ablaut type	**default shape**	**pre-realis shape**	**pre-imperative shape**
i/e ablaut	*-i*	*-e*	*-e*
• example: 'be, do'	*(h)**i**-*	*(h)**e**-*	*(h)**e**-*
• example: 'RECIP'	***-́ni***	***-́ne***	***-́ne***
• example: 'DES'	***-́Ki***	***-́Ke***	(none)
i/e/a ablaut	*-i*	*-e*	*-a*
• example: 'die'	*um**i**-*	*um**e**-*	*um**a**-*
• example: 'peel'	*raph**i**-*	*raph**e**-*	*raph**a**-*
• example: 'stab'	*p**i**-*	*p**e**-*	*p**a**-*
i/e/a ablaut with lowering	*-V_{hi}Ci*	*-V_{hi}Ce*	*-V_{mid}Ca*
• example: 'cook'	***u**k**i**-*	***u**k**e**-*	***o**k**a**-*
• example: 'pick'	*r**i**k**i**-*	*r**i**k**e**-*	*r**e**k**a**-*
• example: 'throw'	*kh**i**m**i**-*	*kh**i**m**e**-*	*kh**e**m**a**-*
e/o ablaut	*-e*	*-o*	*-e*
• example: 'take'	***e**-*	***o**-*	***e**-*
• example: 'see'	*ng**e**-*	*ng**o**-*	*ng**e**-*
• example: 'give.3PL'	*nok**e**-*	*nok**o**-*	*nok**e**-*

Table 5.4 shows the shapes for various morphemes, both roots and suffixes, to illustrate that the ablaut patterns occur throughout the verbal template.

Table 5.4: Alternating shapes for some verbal morphemes

meaning	slot	default	pre-realis	pre-imperative	syncretism pattern
'die'	0 (root)	*umi-*	*ume-*	*uma-*	all different (*i*/*e*/*a*)
'grab, hold'	0 (root)	*ukumi-*	*ukume-*	*okoma-*	all different (*i*/*e*/*a* with lowering)
'eat'	0 (root)	*i-*	*ire-*	*ha-*	all different (idiosyncratic, but resembling *i*/*e*/*a*)
'put into bag'	0 (root)	*ki-*	*ke-*	*oka-*	all different (idiosyncratic, but resembling *i*/*e*/*a*)
'go'	0 (root)	*we-*	*wo-*	*wa-*	all different (idiosyncratic: *e*/*o*/*a*)
'be located'	0 (root)	*khai-*	*khai-*	*khaka-*	default = pre-realis (idiosyncratic)
'see'	0 (root)	*nge-*	*ngo-*	*nge-*	default = pre-imperative (*e*/*o*)
'be, do'	0 (root)	*i-*	*e-*	*e-*	pre-realis = pre-imperative (*i*/*e*)
'come'	0 (root)	*yo-*	*yo-*	*yo-*	all the same
'23PL.OBJ'	1 (object)	*-iki*	*-ike*	*-eka*	all different (*i*/*e*/*a* with lowering)
'RECIP'	1 (object)	*-ńi*	*-ńe*	*-ńe*	pre-realis = pre-imperative (*i*/*e*)
'1SG.OBJ'	1 (object)	*-ya*	*-ya*	*-ya*	all the same
'all'	2 (adverbial)	*-Kape*	*-Kapo*	*-Kape*	default = pre-imperative (*e*/*o*)
'fast'	2 (adverbial)	*-KHókhi*	*-KHókhe*	*-KHókhe*	pre-realis = pre-imperative (*i*/*e*)
'CONT'	3 (aspect)	*-maki*	*-make*	*-maka*	all different (*i*/*e*/*a*)

5.2 Verb stem formation

This section treats the slots of affixes from Table 5.2 (and others that were omitted in it for simplicity) that can be added to a root to create an inflectable stem. These are organized in order from earlier in the word to later in the word. §5.2.1 describes the repetitive, marked by reduplication of the root. §5.2.2 describes object marking (Slot 1), which is mostly done via suffixes, but which has a number of complications such as fusional root+object combinations. §5.2.3 describes the adverbial stem formants (Slot 2). §5.2.4 describes aspect marking (Slot 3), which has sub-slots with a fixed ordering. §5.2.5 describes negation (Slot 4). §5.2.6 describes the conditional (Slot 5), which can be used for both subordination and imperative meaning.

5.2.1 The repetitive

Verb roots can be fully reduplicated to encode a meaning of pluractionality or repetition. While this could be considered a prefix whose form is reduplicative, I see no reason to privilege one of the two copies as being the "real" root while the other is called an affix. Instead, I consider these reduplicated roots to simply be modified roots, consisting of two copies of the phonological material found in the unmodified root. Despite this, for readability of glossing, I gloss the second copy as 'REP'.

(107) shows a simple example of this. The root *nge-* 'see, look' is in the repetitive form *nge~nge-* to signify that the onlookers are all looking at the rope multiple times.

(107) *Khaingu* ***nge~nge****-pina* *uso* *khaingu* *pérĭ_a-ne* *khaingu* *nge-pi ...*
this see~REP-3PL.DS rope this tie-3SG.DS this see-3PL.SS.R
'They kept looking, and he tied this rope, and they looked, and ...' (Graun 9:55-10:02)

(108) shows a similar usage, telling someone not to keep grabbing something again and again. While the two copies of the root have different stresses (primary vs secondary), it is not yet clear what factors determine the location of primary stress in such cases, so I do not consider this to bear on the question of which copy of the root is the affix and which is the original (if one is theoretically disposed to such a question).

(108) ***Ukúmĭ~ukùmĭ****-wápho-Ø.*
grab~REP-PROH-2SG.IMP
'Don't grab it too much.' (elicited, I produced first, consultants accepted)

(109) is a bit different. While it has the same reduplicated verb root *ukumi-* 'grab' as seen in (108), this is used in a serial verb construction (§8.3) together with a verb indicating a process. As a result, the repeated action is understood to be done throughout the course of the process as it progresses, rather than before the process starts. That is, the command is to keep grabbing multiple things as you go, not to grab a bunch of things and then go.

(109) ***Ukúmĭ~ukùmĭ*** *wa-Ø!*
grab~REP go-2SG.IMP
'Grab things as you go (e.g., grass/trees, to avoid falling down on a jungle trail).' (elicited, consultants produced first)

5.2.2 Object marking

Hurukui verbs use morphology to mark their objects, if they have one, as well as their subjects. The semantic role of the object is often patient, beneficiary, or recipient. Hurukui is a secundative language, so the recipient of a giving event is marked the same as the patient of a prototypically transitive event. Most object marking is agglutinative, using one paradigm of suffixes that immediately follow the root (§5.2.2.1). There is also a reciprocal suffix (§5.2.2.2). Both the object suffixes and the reciprocal can act as valence-increasing derivational morphology that adds a beneficiary.

There are also two special verbs that mark their objects in unique ways involving something like fusion of the object suffix with the root, or suppletion in a few cases. The first is the set of 'affect' verbs (§5.2.2.3), used to mean 'hit', 'kill', and various other senses involving the patient being negatively affected by something. The second verb with special object marking is the set of 'give' verbs (§5.2.2.4), which are mostly used for events of giving but can be extended to various ditransitive-like situations.

5.2.2.1 Object suffixes

Hurukui mostly uses suffixes to mark the object of a transitive verb. These suffixes immediately follow the root (or reduplicated root, if it is marked as repetitive). Table 5.5 shows the normal object marking paradigm. Notably, it distinguishes 2DU from 3DU and 2PL from 3PL, unlike most of the subject marking paradigms. It does, however, still syncretize the inclusive and exclusive categories in the first person non-singular. None of these suffixes changes shape.

Table 5.5: The normal object marking paradigm

person/number	**OBJ**
1SG	*-ya*
2SG	*-na*
3SG	*-Ø*
1DU	*-hia ~ -ya*
2DU	*-nia*
3DU	*-nua*
1PL	*-hita ~ -ta*
2PL	*-neka*
3PL	*-noka*

(110) shows a simple usage of object marking for a semantic role of patient or theme. The verb *nukhitamu* 'she left us (SS)' shows 1PL.OBJ marking.

(110) *Mokono-ito* *ahe* *kaya* *ahe*
sister-1PL.POSS INDEF long.ago INDEF

naita *nukhi-**ta**-mu* *wé-nekhai-Ø.*
1PL.INCL leave-1PL.OBJ-3SG.SS.R go-3SG.DPST-3SG.R
'Long ago one of our sisters left us and went.' (Nadzab 0:09-0:14)

In (111), the object marking expresses a semantic role of recipient, the person to whom something is being shown (rather than the person/thing that is being shown to someone).

(111) *Ahe* *hupúri-**ta**-pina* *hita* *noka* *keku* *wé-ke-pu.*
another show-1PL.OBJ-2PL.DS 1PL.EXCL 3PL to go-DES-1PL.R
'Y'all show us some others [coffee buyers] and we'll go to them.' (Kofi 6:21-6:25)

There is also an object marking suffix *-(i)ki* '23PL.OBJ', which I refer to as 'anomalous' because it seems to exist as an extra piece of the system despite the presence of the object suffixes *-neka* '2PL.OBJ' and *-noka* '3PL.OBJ'. Unlike the rest of the second- and third-person non-singular object markers, it does not resemble a personal pronoun. It also shows the syncretism pattern of 23PL like many of the other morphemes in the verb system. It has no dual counterpart.

The suffix *-(i)ki* '23PL.OBJ' often eclipses the final vowel of the root to which it attaches. For some roots, this is obligatory. For others, it is prohibited. And for some, it is optional. The vowels do not combine into a diphthong or sequence; either the last vowel of the root or the first vowel of *-(i)ki* is dropped. For roots ending in *i*, the results are the same. Table 5.6 shows some roots and their forms when suffixed with *-(i)ki* '23PL.OBJ'. Disallowed cells are grayed out.

Table 5.6: Some verb roots and their forms with *-(i)ki* '23PL.OBJ'

root	**with *-(i)ki*, non-eclipsing**	**with *-(i)ki*, eclipsing**
nge- 'see'		*ng-iki-* 'see them'
e- 'take'	*e-ki-* 'take them' (rare)	*Ø-iki-* 'take them' (common)
pa- 'stab'	*pa-ki-* 'stab them'	*p-iki-* 'stab them'
hina- 'watch over'	*hina-ki-* 'watch over them'	

The suffix *-(i)ki* '23PL.OBJ' undergoes *i/e/a* ablaut with lowering, so its pre-realis shape is *-(i)ke* and its pre-imperative shape is *-(e)ka*. Both of these also obey the same restrictions on vowel eclipsis that the default shape does. For instance, the verb *pa-* 'stab', when suffixed with *-(i)ki* in the imperative, can be either *pa-ka* 'stab them' (no eclipsis) or *p-eka* 'stab them' (with eclipsis).

Object markers can also function to add a beneficiary argument. In this usage they are more like derivational affixes. (112) shows the object markers *-neka* '2PL.OBJ' and *-(i)ki* '23PL.OBJ' optionally used for this purpose. However, the object-unmarked verb is also acceptable here because the overt pronoun *neka* '2PL' is present.

(112) *Ima-kha* *neka* *pahe*
1SG-FOC 2PL house

[yai-ra / *yai-**neka**-ra* / *yai-**ke**-ra].*
build-1SG.R build-2PL.OBJ-1SG.R build-2PL.OBJ-1SG.R
'I built y'all a house.' (elicited)

Only some verbs allow this. Many, such as *yo-* 'come', do not.

5.2.2.2 The reciprocal

The reciprocal suffix is *-ńi*, which undergoes *i/e* ablaut. It has the same basic functions as the object markers (affected, beneficiary). (113) and (114) show its usage.

(113) shows a reciprocal patient, about a collision between two bikers. The verb *pukhúmbi-* 'collide' is marked for reciprocal object, because it was not only one participant that notably crashed into the other, but both of them were moving and crashed into each other.

(113) *Harae* *para* *pukhúmbì-**ni**-khu* *kha ...*
together LOC collide-RECIP-3DU.SS.R and
'They two collided, and ...' (TP1 2:42-2:49)

(114) shows a reciprocal recipient or goal, in which people are shouting to each other.

(114) *"Ima* *wana-ra* *o!"* *kó-**ni**-sa ...*
1SG miss.shot-1SG.R faraway.call say-RECIP-SS.IRR
'We would tell each other, "I missed it!", and ...' (GiamanPik 0:48-0:51)

Like the object suffixes, the reciprocal can also function to add a beneficiary argument, but this is fairly rare. In this usage it is more like a derivational affix. (115) shows a usage of this with an idiomatic meaning.

(115) *Yoho* *kokà-**né**-ni!*
fire make.fire-RECIP-2PL.IMP
'Have a verbal altercation with one another!' (lit. 'Make a fire for one another!') (elicited)

Only some verbs allow this. Many, such as *yo-* 'come', do not.

5.2.2.3 The 'affect' verb

There is a verb in Hurukui which means various sorts of prototypically transitive actions, depending on the context. It is perhaps most generally and accurately glossed as 'affect', and it is generally associated with negative effects. Its form is unusual because it has a separate fusional root for each of the person/number categories of object. For most of them, this fusional root is homophonous with the normal object marker, as if the verb has zero root. For 3SG object, however, it has the form *a-*, while the normal 3SG.OBJ marker is zero (§5.2.2.1). Under the zero-root analysis, we would expect the 3SG object form to still have the form of zero: *Ø-Ø- (affect-

3SG.OBJ-), but this is not correct. Thus, it is a more unified analysis to treat these 'affect' roots as fusional rather than agglutinative, although the difference is only seen in the 3SG form. Table 5.7 lists them, alongside the corresponding normal object markers for comparison. None of them change shape. There is no form of 'affect' corresponding to the anomalous object marker *-(i)ki* '23PL.OBJ'.

Table 5.7: Object-marked fusional roots of 'affect'

object	**normal OBJ marker**	**'affect'**
1SG	*-ya*	*ya-*
2SG	*-na*	*na-*
3SG	*-Ø*	***a-***
1DU	*-hia ~ -ya*	*hia- ~ ya-*
2DU	*-nia*	*nia-*
3DU	*-nua*	*nua-*
1PL	*-hita ~ -ta*	*hita- ~ ta-*
2PL	*-neka*	*neka-*
3PL	*-noka*	*noka-*

The following examples show some of the major uses of this verb. In (116), it is used with the meaning 'hit'.

(116) *Akuime* *tata* *kha* *mokho-yo* *kha-pi*
mother father FOC anger-1DU.POSS be.angry-3PL.SS.R

***hia**-pina* *áipe-khu ...*
hit.1DU-3PL.DS cry-1DU.SS.R
'Our parents were mad at us, and they hit us, and we cried, and ...' (SkinDiwai 3:32-3:39)

In (117), the verb is used with the meaning 'kill'.

(117) *Numisa* *kha* *yo-mu* *kha* *ngu* *haikhu* *khaingu*
eel FOC come-3SG.SS.R and 3SG woman this

***a**-mu* *hoka-mu* *i-mu ...*
kill.3SG-3SG.SS.R cut-3SG.SS.R eat-3SG.SS.R
'The eel came and it killed and ate the woman, and ...' (Raunwara 1:26-1:32)

A common everyday usage is seen in (118). The verb in this example still has its generic meaning of 'negatively affect'. With the subject of 'smoke' this becomes a prototypical type of effect that smoke from a fireplace has on people, based on world knowledge.

(118) *Yoho* *mua* ***a**-re.*
fire smoke affect.3SG-3SG.R
'The smoke is getting in that person's face and making their eyes sting.' (overheard)

Another common usage is seen in (119). This is a so-called *experiencer predicate* (§7.3.8), in which a stimulus noun is the subject, and the people being affected by that stimulus are the object. In Hurukui, the 'affect' verb is used in such predicates.

(119) *Aráya-maki-puna* *ókV*[8] ***tà**-mukhài-Ø.*
wander-CONT-1PL.DS hunger affect.1PL-3SG.PRF.R-3SG.R
'We've been wandering around and we've gotten hungry.' (Kaikai 11:59-12:02)

There is another verb with fusional roots whose object marking resembles that of the 'affect' verb. This is *ámukha-* 'fall down', in which the portion of the fusional roots preceding *-mukha-* is identical to the 'affect' roots. For instance, *yámukha-re* (fall.down.1SG-3SG.R) means 'I fell down'. Table 5.8 lists its realis forms, which contain a non-referential *-re* '3SG.R' suffix.

8 The stimulus noun *okV* 'hunger' is so highly collocated with the following 'affect' verbs that its final vowel has been lost and it has become phonologically univerbated with whatever follows it. Sometimes a short and lax epenthetic *e* or *o* is heard at the end of it. The official orthography has not yet prescribed how this univerbation and others like it should be spelled/punctuated.

Table 5.8: Object-marked fusional roots of 'fall down'

experiencer	**'fall.down'**
1SG	*yámukha-re*
2SG	*námukha-re*
3SG	*ámukha-re*
1DU	*hiámukha-re*
2DU	*niámukha-re*
3DU	*nuámukha-re*
1PL	*hitámukha-re*
2PL	*nekámukha-re*
3PL	*nokámukha-re*

This is likely from univerbation of an earlier multi-verb construction involving the 'affect' verb and may be related to the root *ekha-* 'go down'. If this etymology is correct, then *ámukha-* 'fall down' used to be an experiencer predicate (§7.3.8) with no overt subject.

5.2.2.4 The 'give' verb

There is another verb in Hurukui which has fusional roots depending on the person/number category of the object. I refer to it as the 'give' verb. It usually means 'give' in the sense of transferring an item from one person to another, but there are also some common semantic extensions. The semantic role of its object is recipient, not theme.

The forms of the 'give' verb are more irregular than those of the 'affect' verb. The 1SG form is completely suppletive, with obscure etymology. This verb syncretizes 2SG and 3SG objects, which is unusual in the language. All of the stems marked for non-singular object are transparently derived from the personal pronoun, with their final *a* replaced by *e*. Table 5.9 lists these root forms. All the *e*-final roots undergo *e/o* ablaut. Forms of obscure structure are bolded.

Table 5.9: Object-marked fusional roots of 'give'

recipient	personal pronoun	'give'
1SG	*ima*	***opa-***
2SG	*na*	*ne-*
3SG	*ngu*	***ne-***
1DU	*hia* (EXCL) *naya* (INCL)	*hie-*
2DU	*nia*	*nie-*
3DU	*nua*	*nue-*
1PL	*hita* (EXCL) *naita* (INCL)	*hite-*
2PL	*neka*	*neke-*
3PL	*noka*	*noke-*

The first-person non-singular object forms are derived only from the full pronouns *hia* '1DU.EXCL' and *hita* '1PL.EXCL', not the reduced object suffix forms *-ya* '1DU.OBJ' and *-ta* '1PL.OBJ'. That is, there are no 'give' verbs of the forms **ye-* and **te-*. Additionally, the 'give' roots are only derived from the exclusive 1DU and 1PL pronouns, not the inclusive ones, so there are no 'give' verbs of the forms **naye-* and **naite-*.

(120) shows a typical usage of the verb with the basic meaning of 'give'. The morphologically marked object of the verb *ne-* 'give to 3SG', which is fused into the verb stem, is the recipient, *akaime* 'mother', not the theme (the thing being given to her, which in this story is betel nut).

(120) *Pihe* *hapoka* *e* *akaime* ***ne-Ø.***
piece cut.small:SER take.SER mother give.3SG-2SG.IMP
'Cut part of it and give it to mother.' (Kofi 2:54-2:57)

(121) shows a bit more metaphorical usage of the 'give' verb, since sweet potatoes cannot actually transfer possession of strength from themselves to someone else, but such extensions seem very natural and cross-linguistically common.

(121) *Mburumu [...]* *ngu-kha* *ona-ito* *mbangiri*
sweet.potato 3SG-FOC bone-1PL.POSS strong

***hite**-ne* *kha* *kha-mai-pu.*
give.1PL-3SG.DS and live-HAB-1PL.R
'Sweet potato gives us strength and we stay alive.' (PpM3 2:21-2:31)

In (122), the verb *ne-* 'give to 3SG' is used in a more metaphorically extended sense. The structure used, shown in brackets, can be more literally translated as 'you bend your knees and give them to it' but means 'you pray to it' or 'you worship it'.

(122) *Na* *kupisu* *kheko* *nge* *[kapira-na* *apisa-mu* *nge*
2SG cane that TOPIC [knee-2SG.POSS bend-2SG.SS.R TOPIC

***ne**-ma-ma]* *ngae* *ngu-kha* *hi-ne*
give.3SG-HAB-2SG.R] thus 3SG-FOC make-3SG.DS

akine-na *máisa* *ì-mukhai-Ø.*
thought-2SG.POSS bad become-3SG.PRF.R-3SG.R
'This cane of yours, you worship it, and because of that you can't think clearly.' (Kanda 1:53-1:59)

Extra object markers cannot be added after the fusional object-marked roots of the 'give' verb. Thus, conceivable forms like **opa-noka-ma* (give.1SG-3PL.OBJ-2SG.R) 'you gave them to me' do not exist. Only the recipient, not the theme, can be marked as object on the 'give' verb.

5.2.3 Adverbial stem formants

There is a small set of two-syllable derivational suffixes that turn a verb stem into another verb stem, but with added adverbial meaning. These are *-Kape* 'all', *-KHókhi* 'quickly', and *-khóphi* 'a little'. Only verbs with zero or one of these morphemes are attested; whether they can stack will require further research.

The suffix *-Kape* 'all' undergoes *e/o* ablaut. It has scope over the absolutive argument. In (123), it quantifies the subject of the intransitive verb *umi-* 'die', while in (124), it quantifies the object of the transitive verb *ngáe-* 'tell'.

(123) *Umi-**tape**-pínakhai-Ø.*
die-all-3PL.DPST-3SG.R
'They all died (long ago).' (Kikuape 1:57-2:00)

(124) *Moko* *ngu* *ngáe-**kapo**-ma* *khaingu.*
speech 3SG tell-all-2SG.R this
'You told the whole story.' (Graun 15:06-15:09)

The suffix *-KHókhi* 'fast, quickly' undergoes *i/e* ablaut. It has the allomorphs *-khókhi* and *-sókhi* as specified by its morphophonemic notation (§5.1.1). However, it also has an allomorph *-tókhi*, found in the Arapipi dialect after stems whose last syllable is *si*. In dialects without this allomorph, the morphophonologically predictable *-sókhi* is found in these environments instead. The *t*-initial allomorph is shown in (125).

(125) *Wi* *isi-**tòkhé**-ke!*
go.SER sleep-fast-2SG.IMP
'Go sleep immediately!' (Kaikai.Tr 193)

The allomorph *-tókhi* can also occur if the stem consists of a highly collocated combination of two words, whose last syllable is *si* after vowel deletion and univerbation. In (126), the predicate *maisa i-* 'go bad, rot' is marked with *-KHókhi* 'fast'. This predicate is very common and often shortened to *máisi-*. Because the last syllable of this resultant complex stem is *si*, and the speaker speaks Arapipi, the adverbial stem formant *-KHókhi* has allomorph *-tókhi*. The surface realization of the sentence is then further reduced by the deletion of more vowels.

(126) *Máisă* *ĭ-**tòkhĭ**-tòmu.*
bad become-fast-3SG.FUT
'It (peanuts) will go bad quickly.' (Pinat 0:30-0:32)

The *-tókhi* allomorph is also sometimes found after stems that end in *i* but have no *s* to motivate the dissimilation. In (127), the verb *áuni-* 'finish (transitive)' is not followed by the expected *s*-initial allomorph, but instead by the *t*-initial one. I have no explanation for this.

(127) *Àuni-**tókhe**-re.*
finish.TR-fast-3SG.R
'That person finished it quickly.' (elicited, consultants produced first)

The suffix *-khóphi* 'a little, somewhat' undergoes *i/e* ablaut. It is seen in (128), where the verb *khai* 'be located' is in its sense of 'stay (in a place)'. The adverbial stem formant adds the meaning that the subject only stayed for a short time.

(128) *Kha-**khóphi**-mai-pu* *araiku* *harawohe_harawohe* *yo-pi ...*
stay-a.little-CONT-1PL.R earlier a.few come-3PL.SS.R
'We stayed there for a bit, and a few (more people) came, and ...' (Bibuai 1:42-1:49)

(129) illustrates the position of the adverbial stem formants relative to object marking. Here, the adverbial stem formant *-Kape* 'all' follows the fusional 'affect' verb root *noka-* 'kill them'. Notably, the meaning of 'all' is also redundantly expressed in the sentence by both the quantifier *khuha* 'all' and the intensifying particle *pore*, which often has semantics of completeness (§6.5.4).

(129) *Hita* *hamunge* *khuha* *khekheko* *pore* *noka-**kapo**-pi.*
1PL.EXCL man all those INTENS kill.3PL-all-3PL.R
'They killed all of our men.' (Graun 13:31-13:34)

5.2.4 Aspect marking

The aspect slot (Slot 3) is not really only one slot, but I presented it as such for ease of understanding of the verb template in Table 5.2. I will thus refer to it as having *sub-slots*. There are two sub-slots within the aspect slot: PRF/DPST/DES and continuative (in that order). The PRF/DPST/DES sub-slot contains the realis and irrealis perfect paradigms (PRF, §5.2.4.1), the distal past paradigm (DPST, §5.2.4.2), and the desiderative (§5.2.4.3). The continuative sub-slot contains only the continuative (§5.2.4.4). The orderings PRF-CONT, DPST-CONT, and DES-CONT are attested.

5.2.4.1 The perfect

The perfect aspect (PRF) expresses a state that obtains as the result of a past event or change. It can also express the state of doing an action, but the circumstances in which this is allowed require further investigation. It distinguishes realis from irrealis status, agreeing in status with the TAM of the clause chain in which it occurs. The perfect agrees with the marked verb's subject in the realis, but not in the irrealis.

(130) and (131) show basic usages of the perfect, with realis markers following.

(130) *Yoho* *óme-**mukhai**-Ø.*
fire light.up-3SG.PRF.R-3SG.R
'The fire has lit (and is still lit).' (elicited)

(131) *Ima* *ihe* ***i-khukhay**-a.*
1SG pig eat-1SG.PRF.R-1SG.R
'I have eaten pork (recently).' (elicited)

In the realis, the perfect attaches to the default stem. Table 5.10 presents the realis perfect paradigm, including what it looks like when followed directly by the realis markers for both positive and negative polarity (§5.3.1). The 2SG forms *-máikhai* and *-́mukhai* are in free variation.

Table 5.10: The realis perfect conjugation paradigm, with its realis completions

subject	**PRF.R**	**PRF.R-R**	**PRF.R-NEG-R**
1SG	*-́khukhài*	*-́khukhày-a*	*-̀khukha-wá-pa*
2SG	*-máikhài* *-́mukhài*	*-máikhài-ma* *-́mukhài-ma*	*-màikha-wá-ma* *-̀mukha-wá-ma*
3SG	*-́mukhài*	*-́mukhài-Ø*	*-̀mukhái-wa-Ø*
1DU	*-́khukhài*	*-́khukha-ròu*	*-̀khukha-wá-u*
23DU	*-́khukhài*	*-́khukha-rèi*	*-̀khukha-wá-i*
1PL	*-́pukhài*	*-́pukhài-pu*	*-̀pukha-wá-pu*
23PL	*-́pikhài*	*-́pikhài-pi*	*-̀pikha-wá-pi*

The realis perfect suffixes are the same in the default and pre-realis shapes. Due to a mismatch of status, the realis PRF suffixes cannot be used in imperatives, so they do not have pre-imperative shapes.

In the irrealis, the perfect has only two forms, which are used in free variation: *-́khai* ~ *-khai* 'PRF.IRR'. These forms do not inflect for person/number of the subject, and they attach to the default form of the stem. The verb root *we-* 'go', however, uses its fourth shape *wi-* in this situation. The irrealis perfect has the pre-imperative shape *-̀khaka*. Due to a mismatch of status, the irrealis PRF cannot be used in the realis, so it does not have a pre-realis shape.

(132) and (133) show some examples of the perfect followed by future tense marking. In (132), the perfect is used in its irrealis form because Hurukui classifies the future tense suffixes as irrealis. The entire suffix complex has future perfect semantics.

(132) *Hindaho* *amúka-kanu* *nge* *puhi* *phomo* *ahe*
later count-2SG.FUT TOPIC money big INDEF

*háyi-**kha**-komu* *khaingu* *nge.*
amount.to-PRF.IRR-3SG.FUT this TOPIC
'If you count it later, it will have added up to a lot of money.' (Kimbe 3:01-3:04)

In (133), the irrealis perfect is seen followed by a future suffix with conditional semantics.

(133) *Hupí* *na* *ihe* *apono* *khaingu* *é-**kha**-kanu* *ni ...*
if 2SG pig[9] three this take-PRF.IRR-2SG.FUT TOPIC
'If you've bought three horses, then ...' (Hos 1:15-1:20)

The perfect comes from univerbation of certain medial verb forms (§§5.4.1 and 5.4.2) with the former auxiliary verb *khai-* 'be located' (Jones submitted). This is reflected in the allomorphy of the *-khai ~ -kha* part of the suffixes, the pre-imperative shape *-̀khaka* of the irrealis perfect (cf. pre-imperative shape *khaka-* of *khai-*), as well as in the placement of secondary stress. For the vowel allomorphy, the perfect has the same distribution of the shapes *khai* and *kha* as the lexical verb *khai-* 'be located' does. For the stress, the perfect has stress in the same place as the corresponding form of lexical *khai-* 'be located' does. The magnitude of the stress is less predictable. In some forms, primary stress remains on the root while the perfect receives secondary stress, in some forms it is the other way around, and some forms can accept either of these options. Table 5.11 shows some examples of forms of lexical *khai* versus the corresponding forms of the perfect. The parts of the PRF forms that correspond to the lexical form are bolded.

9 Any large hoofed animal is called 'pig' in Hurukui, because things like horses and cows were only recently introduced to New Guinea.

Table 5.11: Correspondence between forms of lexical *khai* 'be located' and the perfect

following suffixes	**lexical *khai* 'be located'**	**PRF with root *nge-* 'see'**	**stress magnitude on PRF and following suffixes**
-1PL.R	*khái-pu*	*ngé-pu**khài-pu***	secondary
-NEG-1PL.R	*kha-wá-pu*	*ngè-pu**kha-wá-pu*** *ngé-pu**kha-wà-pu***	primary or secondary
-1PL.FUT	*kha-kópu*	*ngé-**kha-kòpu***	secondary
-SS.IRR	*khái-sa*	*ngé-**khài-sa***	secondary
-2SG.IMP	*kháka-Ø*	*ngè-**kháka-Ø***	primary

5.2.4.2 The distal past

The distal past expresses past events that are either in the remote past or are noticeably physically separated from the speech event. It does not distinguish realis from irrealis status.

Table 5.12 shows the distal past paradigm when followed by realis markers in both positive and negative polarity (§5.3.1). Unlike the perfect, in which the following suffixes agree in person/number with the subject, for the distal past these subsequent suffixes must all be marked 3SG, regardless of the actual subject. This suffix is zero in both the positive and negative polarity, but I still posit it because in other TAM categories following the DPST paradigm, the 3SG forms are used (e.g., future *-komu*, different-subject *-ne*). Additionally, the zero suffix in both polarities of the realis is expected here because the 3SG realis suffix is zero after the root *khai-* 'be located', from which the PRF and DPST grammaticalized, and the 3SG realis suffix is zero in the negative for all verbs.

Table 5.12: The distal past conjugation paradigm, with its realis completions

subject	**DPST**	**DPST-R**	**DPST-NEG-R**
1SG	*-́pakhài*	*-́pakhài-Ø*	*-̀pakhái-wa-Ø*
2SG	*-mánakhài*	*-mánakhài-Ø*	*-mànakhái-wa-Ø*
3SG	*-́nekhài*	*-́nekhài-Ø*	*-̀nekhái-wa-Ø*
1DU	*-púakhài*	*-púakhài-Ø*	*-pùakhái-wa-Ø*
23DU	*-píakhài*	*-píakhài-Ø*	*-pìakhái-wa-Ø*
1PL	*-púnakhài*	*-púnakhài-Ø*	*-pùnakhái-wa-Ø*
23PL	*-pínakhài*	*-pínakhài-Ø*	*-pìnakhái-wa-Ø*

The distal past comes from the univerbation of medial verbs marked for different subject (§5.4.4) with the former auxiliary verb *khai-* 'be located'. This is why other final or medial verb endings must follow the distal past morpheme in a word, and why more than one syllable in the word has some quantity of stress.

(134)-(136) show some basic uses of the distal past. In (134), the story is set in the far past. This temporal distality can be clearly marked by the distal past paradigm.

(134) *Maha* *khaingu* *hapoka-**pínakhài**-Ø.*
ground this cut-3PL.DPST-3SG.R
'They cut this ground [in ancestral times].' (Raunwara 3:22-3:23)

In (135), the distal past is used because of separation of the expressed event from the reference point, in some combination of time and space. The person has just arrived at the garden and seen for the first time that it has been destroyed. The reference point is the person arriving at the garden just now, and the expressed event is a pig destroying it, which may or may not have occurred a longer time ago than the bare realis would be appropriate for. But it is clear that the pig that did this has left and is long gone, providing spatial separation of the subject from the reference point.

(135) *Mburumu* *ihe* *kha* *yo-mu* *ĭ-tapé-**nekhài**-ne ...*
sweet.potato pig FOC come-3SG.SS.R eat-all-3SG.DPST-3SG.DS
'A pig had come and eaten all the sweet potatoes and ...' (PikGaden 0:11-0:16)

(136) shows a primarily spatial distality meaning of the distal past. Someone$_i$ had been around earlier that morning, but then disappeared after some time. Another person$_j$ then asked where they$_i$ had gone, and someone else$_k$ replied with (136), saying that they$_i$ had gone to the garden. This example therefore has little temporal distality, but because the subject has left and is now significantly spatially separated from the speech event, the distal past can still be used.

(136) *Kanapa* *wé-**nekhài**-Ø.*
garden go-3SG.DPST-3SG.R
'He went to the garden.' (overheard)

When the distal past is followed by future suffixes, the verb expresses counterfactual meaning. (137) shows this. The distal past suffix agrees for the 2SG subject, and the following future suffix agrees for 3SG subject as must all suffixes following the distal past.

(137) *Na* *kepu* *yo-**mánakha**-kòmu* *ni* *ngu* *epu.*
2SG yesterday come-2SG.DPST-3SG.FUT TOPIC 3SG all.right
'If you had come yesterday, that would have been good.' (elicited)

5.2.4.3 The desiderative

The desiderative expresses desires or immediate future tense. It can be formed in two ways, both originally from multi-verb constructions. The more common form, which I call the *do-desiderative*, uses the suffix *-́Ki* 'DES'. The other, which I call the *say-desiderative*, uses the suffix *-́Kăko*, in which the *a* of the first syllable is nearly always elided in contemporary Hurukui, and which is thus pronounced with a geminated *k*, as[-k:o]. The names of the do-desiderative and say-desiderative come from their historical source as a combination of the 1SG potential suffix *-Ka* (§5.3.5) with the matrix verb roots *i-* 'do' and *ko-* 'say'. The do-desiderative undergoes *i/e* ablaut, and because it lacks an imperative form for pragmatic reasons, I cannot say whether it would be classified as *i/e/a*-ablauting or not (but the root *i-* 'do' has *i/e* ablaut, not *i/e/a* ablaut). The say-desiderative does not change shape.

The desiderative suffixes can be followed by medial or final TAM marking. (138)-(140) show realis marking on desideratives.

(138) *Moko* *khaingu* *é-**ke**-rèu.*
speech this take-DES-1DU.R
'We want to talk about this.' (BW 0:17-0:21)

(139) *Na* *ì-**té**-ma?*
2SG eat-DES-2SG.R
'Do you want to eat it?' (elicited)

(140) *Ngu* *wé-**kăkò**-re.*
3SG go-DES-3SG.R
'That person wants to go.' (elicited)

The do-desiderative can also mark immediate future, shown in (141). The say-desiderative cannot mark this meaning.

(141) *E* *umo* *khahu* *oke-ni;* *opopo* *horórà-**kè**-re.*
take.SER tree COMIT tie-2PL.IMP food spill.INTR-DES-3SG.R
'Tie it to the tree; the food is about to spill.' (MejikDiwai 4:03-4:08)

(142) shows future tense marking on a desiderative. This is a say-desiderative used in a clause with conditional semantics, which the future tense can mark (§5.3.2).

(142) *Sono* *í-**tăko**-kànu* *nge_ngu* *sono* *i* *a-kha ...*
water eat-DES-2SG.FUT TOPIC water eat.SER PFV-SS.IRR
'If you want to drink water, then you'll drink it and ...' (Lae 3:30-3:34)

The do-desiderative with present tense semantics is often accompanied by a continuative marker. (143) shows this. In some contexts the accompanying continuative is optional, as shown in (144). Whether there is a semantic difference between these will require future research.

(143) *E* *inòngi-**ti-má**-re* *kho?*
take do.what-DES-CONT-3SG.R I.wonder
'I wonder what he's going to do with it?' (VD2 1:07-1:09)

(144) *Na* *[wé-**ki-mà**-ma* ~ *wé-**kè**-ma]?*
2SG go-DES-CONT-2SG.R go-DES-2SG.R
'Do you want to go?' (elicited)

Some contexts allow only one of the two desiderative types. The reasons for this will require further investigation. (145) shows one such context, allowing a do-desiderative but not a say-desiderative.

(145) *Ima* *[wé-**ki**-mà-ra* / **wé-**kăkò**-ra]* *nemo* *we-wa-o.*
1SG go-DES-CONT-1SG.R go-DES-1SG.R but go-NEG-1SG.FUT
'I want to go, but I won't go.' (elicited)

5.2.4.4 The continuative

The continuative marks progressive and habitual meanings. It is marked by a suffix *-maki* which undergoes *i/e/a* ablaut. The Eastern dialect uses the form *-ma(i)* for the default and pre-realis shapes, rather than *-maki* and *-make*. Which allomorph of *-ma(i)* is used (i.e., *-ma* vs *-mai*) is selected by the following suffix, and it will be specified in the sections on those suffixes (§5.3.1 for realis, §5.3.2 for future, §5.4.4 for different-subject).

(146) shows a progressive usage.

(146) *Hauphe* *hangiki-**maka**-Ø!*
betel.nut cut-CONT-2SG.IMP
'Keep cutting the betel nut!' (Kofi 2:12-2:15)

(147) shows a habitual usage.

(147) *Puhi* *keme* *e-pi* *kuniĕ_apono* *khaingu* *ne-**mai**-pi?*
money how.much take-3PL.SS.R car this give.3SG-CONT-3PL.R
'How much is the fare?' (lit. 'How much money do they give to the car?') (Lae 0:17-0:22)

5.2.4.5 The unmarked aspect

There is conflicting evidence regarding whether there is a zero-marked aspect in Hurukui. Final verbs with no overt aspect marking are often interpreted as being punctual or perfective, rather than allowing interpretations of perfect, distal past, desiderative, or continuative. For instance, in (148) and (149), the predicate *epu a-* 'be enough, be all right' is seen unmarked for aspect, in which case it means that the state has just been entered, versus marked for perfect aspect, in which case it means that the state was entered some time in the past and is currently still inhabited.

(148) *Épŭ_à-**re**.*
be.enough-3SG.R
'That's enough.'
(it just reached an appropriate state)

(149) *Épŭ_à-**mukhai-Ø**.*
be.enough-3SG.PRF.R-3SG.R
'It's all right as-is.' (both overheard)
(it is already in an appropriate state)

However, in some situations, final verbs unmarked for aspect do have meanings that are marked by the aspect markers, such as continuative or immediate future (which is marked by the desiderative). This is seen in (150), where the verb *wora* 'I go' is in a form that more typically means 'I went', but it has immediate future meaning because the person is about to embark on the journey home.

(150) *Anao,* *hindaura* *maya,* *ima* *pahe* *weka* ***wo-ra**.*
vocative.name afternoon good 1SG house over.there go-1SG.R
'Anao[10], good afternoon, I'm going home.' (text message)

Additionally, evidence on this from medial verbs is conflicting. Some medial verbs unmarked for aspect have the same aspectual semantics as the final verb in the chain. This is seen in (151), in which all the verbs have semantics of continuous action, but this is only marked on the last verb.

(151) *Humingi* *renga-khu* *akai-khu* *we-**make**-rei.*
louse pick.out-3DU.SS.R continue-3DU.SS.R go-CONT-3DU.R
'They two were in the process of picking lice out of their hair.' (Yombo 2:03-2:06)

10 This is the vocative name assigned to me and others with the Christian given name "Wesley", so a more accurate English translation would address the listener as "Wesley" because English does not have vocative names.

Other medial verbs unmarked for aspect have punctual or perfective semantics despite the final verb marking a different aspect. This is seen in (152), where the medial verb *amu* 'he put it, and (SS)' does not have continuative semantics because the person only put the water bottle in their mouth once, and that action is finished.

(152) *More-Ø* *a-mu* *ekhono* *i-**ma**-re.*
mouth-3SG.POSS put-3SG.SS.R still eat-CONT-3SG.R
'He put it (water bottle) in his mouth, and he's drinking it.' (VD1 2:43-2:46)

It is quite possible that pragmatics and world knowledge about the typical structure of certain events is responsible for much of the availability of different interpretations of unmarked aspect. The question of unmarked aspect in Hurukui will therefore have to be left for future investigation. For the purposes of this description, I treat the aspect slot as optional and do not include a null aspect morpheme.

5.2.5 Negation

Verbal negation in Hurukui is marked by a suffix *-wa* 'NEG', which does not change shape. It conditions a good deal of allomorphy in the conjugation paradigms of Slot 6 which follow it. For instance, in the realis (§5.3.1), the 1SG suffix is *-ra* in the positive but *-pa* in the negative, and the 3SG suffix is *-re* in the positive but zero in the negative. These allomorphs will be described in their respective sections. In this section, I have chosen examples for which the person marking does not change with polarity, for simplicity of exemplifying negation itself.

Arguably, by the logic that the lack of the negation marker has a particular meaning (i.e., positive polarity), one could consider there to be a zero suffix in paradigmatic opposition to the negative *-wa* (as discussed in §5.1.4). For brevity of parsing, though, I will omit it and treat positive-polarity examples as simply lacking a polarity marker.

In (153), the verb *yoma* 'you came' is not marked for negation. It consists only of the root *yo-* 'come' and the agreement suffix *-ma* '2SG.R'. In (154), the same verb is found but negated: *yowama* 'you did not come'. It has the negative suffix *-wa* following the root.

(153) *Sawao* *nge-ka* ***yo-ma?***
given.name see-PUR come-2SG.R
'Did you come to see Sawao?' (Nadzab 8:10-8:13)

(154) *Inóngi-ne* *ngae* *yo-**wa**-ma?*
do.what-3SG.DS thus come-NEG-2SG.R
'Why didn't you come?' (BW 0:44-0:49)

In (155), we have a medial verb, *yone* 'it comes, and ...', in the positive polarity. It consists only of the root *yo-* 'come' and the switch reference suffix *-ne* '3SG.DS' (§5.4.4). In (156), we have the same medial verb but negated: *yowane* 'that person did not come, and ...'.

(155) *Kherekhere* *a-mana* ***yo-ne*** *uso* *ngu* *para*
insert put-2SG.DS come-3SG.DS rope 3SG LOC

ukumi-khai-sa *a-kanu.*
grab-PRF.IRR-SS.IRR hit.3SG-2SG.FUT
'You'll stick the rope through (lit. 'you insert it and it comes') and hold it here and hit it.' (Susap 1:19-1:23)

(156) *Ngu* *yo-**wa**-ne* *hita* *ngu* *maríhe-pu.*
3SG come-NEG-3SG.DS 1PL.EXCL 3SG search.for-1PL.R
'That person didn't come, and we looked for them.' (elicited)

The negative suffix *-wa* is usually easy to spot in the various TAM categories, because its form rarely changes. In some cases, though, it is more obscured. An example of this is when it follows the continuative. When this happens, the continuative's default shape *-maki* is used in all persons except 3SG (where *-mai* is used instead), and the complex *-maki-wa* (CONT-NEG) is reduced to /maku.a/. The orthography reflects this, writing *-makua*.

5.2.6 The conditional

The conditional suffix *-NGe* does not change shape. It attaches to the default shape of the preceding morpheme, except for the root *we-* 'go', where it instead attaches to the fourth shape *wi-* (creating the form *winde*). It has two main uses: subordination creating conditional clauses, and stem formation creating polite imperatives.

When the conditional acts as a subordinator, the marked verb does not inflect for subject or further TAM. (157) shows a usage that means a bit like 'if' and a bit like 'when' or 'whenever'. The subject of the conditional is 2SG, but this is not shown in the morphology.

(157) *Nuimi* *ahe* *yó-kara* *ko-**nge*** *khakono* *yo-kha ...*
tomorrow INDEF come-1SG.FUT say-COND right.here come-SS.IRR
'If/whenever you want to come (lit. 'if you say "I will come"'), then you'll come here and ...' (Nadzab 6:51-6:54)

(158) shows a subordinating usage of the conditional with a meaning more clearly like 'when'. It is notably followed by a topic marker *nge ngu*, which is common (Haiman 1978), and its form even resembles the *nge* part of that topic marker (§6.5.2).

(158) *Kohe* *yo* *anga* *khaingu* *yo-**nge*** *nge_ngu*
okay come.SER day this come-COND TOPIC

no *yo* *khaka-ni.*
here come.SER stay-2PL.IMP
'Okay so when that day comes, y'all come stay here.' (Graun 9:44-9:48)

The second use of the conditional suffix is to mark politeness in imperatives, in which case it forms a stem that takes the imperative inflectional paradigm (§5.3.4). This use is syntactically a matrix-clause final verb.

The imperative suffixes used are *-Ø* '2SG.IMP', *-khi* '2DU.IMP', and *-ni* '2PL.IMP', agreeing with the number of the subject. The imperative suffix *-ke* '2SG.IMP' is not used together with the conditional. Table 5.13 shows some verb stems and their polite imperative forms.

Table 5.13: Some verb stems and their polite imperative forms

stem	**2SG**	**2DU**	**2PL**
yo- 'come'	*yo-**nge**-Ø*	*yo-**nge**-khi*	*yo-**nge**-ni*
e-ya- 'take me'	*e-ya-**nge**-Ø*	*e-ya-**nge**-khi*	*e-ya-**nge**-ni*
khai- 'stay'	*khai-**nde**-Ø*	*khai-**nde**-khi*	*khai-**nde**-ni*
isi- 'sleep'	*isi-**nde**-Ø*	*isi-**nde**-khi*	*isi-**nde**-ni*
we- 'go'	*wi-**nde**-Ø*	*wi-**nde**-khi*	*wi-**nde**-ni*

(159) shows the conditional marker used in a polite imperative. In this text, a consultant shows how to make a traditional comb from bamboo. Once he has made it, he wants to demonstrate it on someone who has more hair than him, so he improvises and calls over one of the audience members.

(159) *Anai* *no* *yo-**nge-Ø*** *ngarangi* *ahe* *no*
vocative.name here come-POL.IMP-2SG.IMP comb INDEF here

raphe-ra *na* *pira-na* *akhurúkha-kara.*
carve-1SG.R 2SG head-2SG.POSS comb-1SG.FUT
'Hey Paul, come here, I've made a comb and I'll comb your hair.' (Kom 6:45-6:50)

There is no polite prohibitive. When the conditional *-NGe* follows the negative *-wa*, this can only have negative conditional meaning.

5.3 Inflection of final verbs

Final verbs (§5.1.3) mark multiple TAM categories with suffixes, most of which inflect for the subject of the marked verb. These are:

- the realis (§5.3.1)
- the future (§5.3.2), including an extra way of marking negative future (§5.3.3)
- the imperative (§5.3.4)
- the potential (§5.3.5)
- an extra way of marking negative habitual (§5.3.6)

5.3.1 Realis marking

The realis is used to express events in the present and recent past. It is also used as a default or unmarked tense in narratives, once the remote past setting has been established.

(160) is the first sentence of a conversation between two people about what they did yesterday.

(160) *Na* *kepu* *nga* *wo-**ma**?*
2SG yesterday where go-2SG.R
'Where did you go yesterday?' (Asde 0:06-0:11)

In (161), the story is about a school being built in 2010, 11 years before the story was recorded. The first sentence of the story uses the distal past, as well as mentioning the explicit year, to establish the setting. Subsequent sentences are in the realis with unmarked aspect, even though they occurred at that remote past time.

(161) *Sukuru* *pai* *yo-**re*** *nake* *kopohe* *amie* *yo-**pi**.*
school new come-3SG.R because child many come-3PL.R
'Because a new school came, many children came.' (Skul 2) (written only)

The paradigm of realis suffixes is shown in Table 5.14. Negative realis markers that differ from the positive are bolded.

Table 5.14: The realis conjugation paradigm

subject	**R**	**NEG-R**
1SG	*-ra*	*-wa-**pa***
2SG	*-ma*	*-wa-ma*
3SG	*-re*	*-wa-**Ø***
1DU	*-rou ~ -reu*	*-wa-**u***
23DU	*-rei ~ -rui*	*-wa-**i***
1PL	*-pu*	*-wa-pu*
23PL	*-pi*	*-wa-pi*

The stem directly preceding the realis suffix will be in its pre-realis shape. Importantly, this means that the verb root is only in the pre-realis shape if a realis suffix directly follows it, because it is only the last morpheme of a stem that undergoes the changes defining the stem form type. If any intervening material separates the verb root from the realis suffix, then the root will be in its default shape.

(162) shows the root *péwe-* 'stop existing' in pre-realis shape because the suffix *-re* '3SG.R' directly follows it. (163) shows the same root and suffix, but with an intervening continuative marker, so the root is instead in default shape.

(162) *Mie* ***péwo-re*** *hi,* *pahe* *wa-ni.*
earthquake stop.existing-3SG.R EMPH village go-2PL.IMP
'The earthquake is over; y'all go to the village.' (Bibuai 8:03-8:06)

(163) *E-pu* *humia* *a-puna* *kha* *mbukhu* ***péwe-ma-re.***
take-1PL.SS.R sun put-1PL.DS and wet stop.existing-CONT-3SG.R
'We put it in the sun and it dries.' (Bilum 0:31-0:36)

In (164) and (165), we observe form changes in the morpheme *-(i)ki* '23PL.OBJ', which undergoes *i/e/a* ablaut with lowering. The root *nge-* 'see' which precedes it is in its default shape in both examples because it is not stem-final. In (164), *-(i)ki* is stem-final and immediately followed by a realis conjugation marker, so it takes its pre-realis shape *-ike*.

(164) *Neka* *wo-pi* *khaingu* *ndapono* *khaingu* *ngě-**ike-pi*** *o*
2PL go-2PL.R this those.people this see-3PL.OBJ-2PL.R TAG.Q
'When y'all went, did y'all see these people too?' (Dokta 3:14-3:19)

In (165), *-(i)ki* is still stem-final, but the negator *-wa* intervenes between it and the realis conjugation marker, so it remains in its default shape *-iki*.

(165) *Haikhu* *khaingu* *ngu* *ngě-**iki-wa-i**?*
woman this TOPIC see-3PL.OBJ-NEG-2DU.R
'Did you two not see these women?' (Asde2 2:20-2:23)

The positive-polarity forms marking 1DU and 23DU in Table 5.14 have allomorphs in which the first vowel can be either *e* or *o*. This is an example of a vowel harmony process in Hurukui, albeit irregular and very restricted in its distribution. For '1DU.R', the allomorph *-rou* is by far the most common. The allomorph *-reu* is only used after stems ending in *i*, and even most of these can optionally take *-rou* instead. The place where *-reu* is most reliably found is following the desiderative *-ʹKi*. Table 5.15 shows some examples of stems taking *-rou* only, stems taking both, and stems taking *-reu* only. Grayed-out cells indicate unacceptable forms.

Table 5.15: Allomorphy of the 1DU.R suffix *-rou ~ -reu*

stem	**with *-rou***	**with *-reu***
khai- 'be located'	*kha-rou*	
ko- 'say'	*ko-rou*	
isi- 'sleep'	*ise-rou* (less common)	*ise-reu* (more common)
pi- 'stab'	*pe-rou* (less common)	*pe-reu* (more common)
we-maki- 'be going'	*we-make-rou* (less common)	*we-make-reu* (more common)
wé-ki- 'want to go'		*wé-ke-rèu*

As for the 23DU.R suffix *-rei ~ -rui*, this preference is reversed. The allomorph *-rei* is by far the most common, and in the Eastern dialect it is the only option. In some speakers of the Arapipi dialects, the allomorph *-rui* is found after certain stems ending in *o*. The most common of these are *yo-* 'come' and *we-* 'go', which has pre-realis shape *wo-*.

The verbs *i-* 'eat' and *khai-* 'be located' have irregular conjugations in the realis in the positive polarity (but regular in the negative). These are shown in Table 5.16, with irregular forms bolded.

Table 5.16: Verbs with irregular realis conjugations

subject	***i-* 'eat'**	***khai-* 'be located'**
1SG	*ire-ra*	***khay-a***
2SG	*ire-ma*	*khai-ma*
3SG	*ire-re*	***khai-Ø***
1DU	***ire-u***	*kha-rou*
23DU	***ire-i***	*kha-rei*
1PL	*ire-pu*	*khai-pu*
23PL	*ire-pi*	*khai-pi*

For *i-* 'eat', the irregular forms are 1DU *ireu* and 23DU *irei*. The simplest analysis posits that the syllable *re* is part of the root's pre-realis shape for all forms in this paradigm, and that the dual

suffixes are *-u* and *-i*. The existence of these suffixes is corroborated by the negative realis conjugation: *-wa-u* '-NEG-1DU.R' and *-wa-i* '-NEG-23DU.R' (Table 5.14). For *khai-* 'be located', an anomalous agreement suffix is found: *-a* '1SG.R', which is not found anywhere else in the language (other than with the realis perfect, §5.2.4.1, which grammaticalized from auxiliary *khai-*), and the 3SG.R suffix is zero.

Table 5.17 shows the allomorphy of the continuative *-maki ~ -ma(i)* (§5.2.4.4) when followed by the realis. For the dialectal variant *-maki*, the pre-realis shape *-make* combines regularly with the following realis suffix. For the dialectal variant *-ma(i)*, however, the *-mai* allomorph is only used in the plural (bolded). The reason for this will require further research.

Table 5.17: Combination of the continuative with the realis

subject	***-maki***	***-ma(i)***
1SG	*-make-ra*	*-ma-ra*
2SG	*-make-ma*	*-ma-ma*
3SG	*-make-re*	*-ma-re*
1DU	*-make-reu*	*-ma-rou*
23DU	*-make-rei*	*-ma-rei*
1PL	*-make-pu*	***-mai-pu***
23PL	*-make-pi*	***-mai-pi***

5.3.2 The future tense

The future tense conjugation in Hurukui is used to express events that will happen in the future. It is also one way to mark conditional clauses, particularly if they are followed by the topic particles *nge*, *ngu*, or *nge ngu* (§6.5.2) (Haiman 1978). The future conjugation is shown in Table 5.18.

Table 5.18: The future tense conjugation paradigm

subject	**FUT**	**NEG-FUT (regular)**	**NEG-FUT (alternate, * = Pre-Hurukui)**
1SG	*-ʹKara*	*-wá-kara*	*-wa-o*
2SG	*-Kanu*	*-wa-kanu*	
3SG	*-Komu*	*-wa-komu*	**-wa-mu > -waumu*
1DU	*-Kohu*	*-wa-kohu*	**-wa-hu > -wauhu*
23DU	*-Kekhu*	*-wa-kekhu*	
1PL	*-Kopu*	*-wa-kopu*	**-wa-pu > -waupu*
23PL	*-Kepi*	*-wa-kepi*	

The 1SG.FUT suffix *-ʹKara* is the only one of these that creates a word with antepenultimate stress. The reason for this probably involves analogy with the desiderative *-ʹKe-ra* (-DES-1SG.R) (§5.2.4.3), but this requires further research.

As seen in the "NEG-FUT (regular)" column in Table 5.18 shows, future-tense verbs can be negated regularly, with the negator *-wa* preceding the future suffix. As the next column shows, however, some person/number values allow the future to be negated in more than one way. These are simply alternative markers of the negative future. Any semantic or pragmatic differences between them and the regular negative future markers will require further investigation. For 3SG, 1DU, and 1PL, the "underlying" form is *-wa* 'NEG' plus the second syllable of the future suffix, but the "surface" form involves anticipatory semivowel insertion, changing the *a* to *au* (§2.6, Table 2.21).

For 1SG, the alternate negative future suffix instead is composed of *-wa* 'NEG' plus *-o*, which is form-identical to the 1SG potential marker *-o* (§5.3.5). These are likely etymologically the same, but their semantics have become somewhat differentiated in the two polarities.

(166) shows a typical usage with future meaning. The speaker is showing how to make a trap to catch wild animals, and he is telling what he will do before he does it.

(166) *Uso* *hapoká-**kara**.*
rope cut-1SG.FUT
'I will cut the rope.' (Trap 1:22-1:24)

With first-person non-singular subjects, the future suffixes can also express hortative meaning (i.e., making suggestions to the group, or asking if it is okay to do something). (167) shows such a use.

(167) *Ngáhitomu* *sono* *humekhe* *e* *wi* *wa-ka* *a-**kopu**?*
all.right water container take:SER go:SER upland-there put-1PL.FUT
'All right, so should we take the water jug and put it up there?' (ObjP1 0:28-0:31)

The verb *we-* 'go' has regular future forms as well as some irregular ones. The irregular forms all occur in cells whose future suffixes contain the vowel *o*, viz. 3SG, 1DU, and 1PL. They are all formed by using the root shape *wo-* instead of *we-* by a process of regressive vowel harmony. For each of these there is a longer form, formed by this root change alone, and a shortened one formed by subsequent contraction of the sequence *oko* to just *o*. These are compared in Table 5.19.

Table 5.19: Regular and irregular future forms of *we-* 'go'

subject	**FUT (regular)**	**FUT (irregular, longer)**	**FUT (irregular, shorter)**
1SG	*wé-kara*		
2SG	*we-kanu*		
3SG	*we-komu*	*wo-komu*	*womu*
1DU	*we-kohu*	*wo-kohu*	*wohu*
23DU	*we-kekhu*		
1PL	*we-kopu*	*wo-kopu*	*wopu*
23PL	*we-kepi*		

For *we-* 'go', the 1DU and 1PL hortative meanings must use the shortened forms *wohu* and *wopu*, never the longer irregular forms nor the regular forms. An example is shown in (168).

(168) *Hauphe* *i-sa* ***wopu*** *o!*
betel.nut eat-SS.IRR go:1PL.HORT EMPH
'Let's chew betel nut and then go!' (Kofi 2:25-2:28)

(169) shows an "if-then" sentence, in which both the protasis and the apodosis are marked with future-tense final verbs (§8.4.3.3). The protasis of such sentences (in which the future marking has conditional meaning) is also commonly followed by topic markers, and (169) is no exception (topic marker *nge ngu* underlined). The apodosis illustrates a negative future.

(169) *Hai* *sono* *angere* *eko* *we-ma-**kanu*** *nge ngu*
only water close that.near.you go-CONT-2SG.FUT TOPIC

mbukura *ukumi-na-ne* *sono* *ekha-wa-**kanu**.*
cold hold-2SG.OBJ-3SG.DS water bathe-NEG-2SG.FUT
'If you go close to the river, you'll feel very cold, and you won't bathe.' (Dokta 7:25-7:33)

Table 5.20 shows the allomorphy of the continuative *-ma(i)* (§5.2.4.4) when followed by the future. The dialectal variant *-maki* is not used in this context.

Table 5.20: Combination of the continuative with the future

subject	***-ma(i)*-FUT (regular)**	***-ma(i)*-FUT (alternate, * = Pre-Hurukui)**
1SG	*-mái-tara*	*-may-o* *-mai-ta*
2SG	*-ma-kanu*	
3SG	*-ma-komu*	**-ma-mu > -maumu*
1DU	*-ma-kohu*	**-ma-hu > -mauhu*
23DU	*-ma-kekhu*	
1PL	*-ma-kopu*	**-ma-pu > -maupu*
23PL	*-ma-kepi*	

5.3.3 The negative future in *-ome*

There is another set of negative future markers, this one not involving the negator *-wa*. Instead, it is composed (mostly) of the normal future suffixes, minus their final vowels, plus the

sequence *ome*. This form is homophonous with the negative predicate particle *ome* 'is not' (§6.5.1). Table 5.21 lists these suffixes. The difference in meaning between this and the normal negative future, if any, will require further research.

Table 5.21: The negative future paradigm in *-ome*

subject	**NEG.FUT**
1SG	*-ŏ-ome* > *-ome*
2SG	*-Kanŭ-ome* > *-Kanome*
3SG	*-Komŭ-ome* > *-Komome*
1DU	*-Kohŭ-ome* > *-Kohome*
23DU	*-Kekhŭ-ome* > *-Kekhome*
1PL	*-Kopŭ-ome* > *-Kopome*
23PL	*-Kepĭ-ome* > *-Kepome*

As can be seen in Table 5.21, the 1SG form does not use the 1SG.FUT suffix *-ʹKara* before *-ome*. It instead uses the 1SG potential marker *-o*. All of the forms except 1SG and 2SG have raised mid vowels (§2.2) in their future suffixes, and these are retained in the *-ome* forms, but the orthography does not reflect this because the high vowels conditioning the change have been deleted here.

(170) shows a usage of this paradigm.

(170) *Opŏ-ahe* *pi-**tanome*** *wi* *itana*
thing-INDEF stab-2SG.NEG.FUT go.SER 2SG.EMPH

hokere *oita* *khaka-Ø.*
side down.there be.located-2SG.IMP
'You won't be able to shoot anything, just go stand aside down there.' (Raunwara 4:14-4:21)

5.3.4 The imperative

The imperative is used for issuing commands. Their subjects can be either second or third person. These suffixes attach to the imperative stem form. Table 5.22 lists them. The singular imperative has two options for suffixes, *-ke* or zero, but the prohibitive cannot use *-ke*.

Table 5.22: The imperative paradigm

subject	**IMP**	**PROH**
23SG	*-Ø* *-ke*	*-wapho-Ø*
23DU	*-khi*	*-wapho-khi*
23PL	*-ni*	*-wapho-ni*

Most verbs lexically specify which 23SG.IMP suffix they take. The most notable examples are *i-* 'be, do' (23SG.IMP *e-ke*), *e-* 'take' (23SG.IMP *e-ke*), and *a-* 'put' and the homophonous *a-* 'affect.3SG' (23SG.IMP *a-ke*). Any verbs derived from these via light verb constructions (§7.3.11) also take *-ke*. Some other verbs admit a bit of free variation, but there is usually still a strong preference for one 23SG.IMP suffix over the other. For instance, *hie-* 'give.1DU' strongly prefers *hie-Ø*, but *hie-ke* is also marginally accepted.

Here are some basic examples of second-person imperatives and prohibitives. (171)-(173) show 2SG, 2DU, and 2PL subjects, respectively.

(171) *Na-kha* *no* *yo-**Ø!***
2SG-FOC here come-2SG.IMP
'Come here, you!' (Kom 6:43-6:46)

(172) *Kohe* *wopa-**khi**.*
okay go.past-2DU.IMP
'All right, you two go first.' (Kofi 0:51-0:53)

(173) *Khaingu* *khanìngi-waphó-**ni**.*
this do.thus-PROH-2PL.IMP
'Y'all don't do this!' (Man3 2:30-2:33)

In the imperative, second and third person are syncretized in each of the three numbers. Here are some examples of third-person imperatives. They all actually involve situations where the speaker is commanding the listener about something, but the grammatical subject of the prohibitive-marked verb is third person.

(174) shows a prohibitive with 3SG subject. This predicate, *eme- ra-* 'be sleepy', marks the experiencer as a possessor on *eme-* 'eye' and marks the verb with 3SG subject. For instance, in the positive realis, one has *eme-o ra-re* (eye-1SG.POSS be.sleepy-3SG.R) 'I am sleepy'. Therefore, the verb in (174) also has a 3SG subject.

(174) *Eme-nă* *ra-**wapho-Ø!***
eye-2SG.POSS be.sleepy-PROH-3SG.IMP
'Don't be sleepy!' (lit. 'May your eyes not be sleepy!') (elicited, consultants produced first)

(175) shows a 3SG imperative in the final verb of a clause chain. The entire chain has imperative semantics even though the medial morphology does not show this, so the medial verb *khéngepia* 'you two call that person and ... (DS)' is a command as well as the final verb *yo* 'let that person come!'.

(175) *Khéngĕ-pìă* *yó-**Ø!***
call.out-2DU.DS come-3SG.IMP
'You two call that person over!' (lit. 'You two call that person, and may they come!') (elicited, consultants produced first)

In (176), there are two imperative sentences. The first is 2SG.IMP, meaning 'don't give him food'. The second chain has one clause with 2SG subject and another with 3SG subject. The chain agrees in mood, with the imperative mood of the final suffix *-ke* '3SG.IMP' having scope over both *ok-a-* 'be hungry' (an experiencer predicate, §7.3.8) and *ere-* 'touch'.

(176) *Opopo* *e* *ne-**wapho-Ø;***
food take.SER give.3SG-PROH-2SG.IMP

ere-wa-mana *òk-á-**ke!***
touch-NEG-2SG.DS hunger-affect.3SG-3SG.IMP
'Don't give him any food, so he'll be hungry.' (lit. 'Don't give him food; don't touch him and may hunger affect him.') (elicited, consultants produced first)

(177) is similar to the second chain of (176) in that it has two verbs with different subjects and imperative semantics that has scope over the whole chain. The negative polarity of the prohibition also has scope over both verbs.

(177) *Ima* *hi-mana* *mbukura* *ya-**wapho-Ø!***
1SG make-2SG.DS cold affect.1SG-PROH-3SG.IMP
'Don't make me cold!' (lit. 'You (don't) make and may cold not affect me!') (elicited, consultants produced first)

(178) shows another chain in which the imperative semantics of the final verb has scope over both verbs. Here the final verb is marked for 3PL imperative.

(178) *Noka-pina* *nenge* *wa-**ni!***
hit.3PL-2PL.DS outside go-3PL.IMP
'Y'all hit them (dogs) so they go outside!' (lit. 'Y'all hit them and may they go outside!') (overheard)

When an object suffix is followed by *-Ø* '23SG.IMP', the word may optionally take final stress. For instance, one has *é-ya-Ø* ~ *e-yá-Ø* (take-1SG.OBJ-2SG.IMP) 'Take me!'.

5.3.5 The potential markers *-o* and *-Ka*

There is a suffix *-o*, with future-like semantics, which is used when making a suggestion or asking a question 'shall I do X'? It is only used with 1SG subjects. After a stem-final *o*, this suffix can be either *-o* (creating a sequence *-o-o* pronounced [o̱ː]) or it can take the allomorph *-wo*.

In (179), the speaker wants to hear something and is telling people of this intent.

(179) *Ima* *epha-**o***.
1SG hear-1SG.POT
'Let me hear (it).' / 'I will hear (it).' (elicited)

In (180), the speaker is asking if they should do something, so the potential is used. The semantics of this potential-marked final verb has scope over all three verbs in the chain: *e-* 'take', *yo-* 'come', and *a-* 'put'.

(180) *Mui* *khaino* *e* *yo* *khaino* *a-**o?***
cucumber this take.SER come.SER here put-1SG.POT
'Shall I bring this cucumber and put it here?' (ObjP1 1:20-1:23)

There is also a roughly synonymous suffix *-Ka*, also only used with 1SG subjects. (181) and (182) show it. These uses have semantics of future or hortative, expressing the speaker's intent for something to happen.

(181) *Ngae* *na-kha* *moko* *e-mana* *epha-**ka**.*
thus 2SG-FOC speech take-2SG.DS hear-1SG.POT
'So you tell about it and I'll listen.' (Lae 0:13-0:15)

(182) *Ima* *ná* *opitanŭ* *i-**ta**.*
1SG 2SG food:2SG.POSS eat-1SG.POT
'I'm going to eat your food.' (elicited, consultants produced first)

This suffix selects the continuative's allomorph of *-mai*. This is seen in (183).

(183) *Opópŏ* *i-**mái-ta**!*
food eat-CONT-1SG.POT
'Let me keep eating!' (elicited)

Based on the semantics of clauses and chains like (181) and (182), these potential markers could be thought of as belonging in the imperative paradigm, alongside *-ke* / *-Ø* '23SG.IMP', *-khi* '23DU.IMP', and *-ni* '23PL.IMP'. However, I treat them separately for analytical convenience for a few reasons. First, the potential markers do not condition the pre-imperative shape of preceding morphemes, instead using the default shape. Thus, in (182) we have ***i**-ta* 'eat-1SG.POT', not ****ha**-ka* 'eat:IMP-1SG.POT'. Second, the marker *-o* '1SG.POT' patterns together with future markers, not imperative markers, in several constructions. This can be seen especially well in irrealis complement clauses, such as (406) in §8.4.1. Third, the first-person non-singular counterparts of the 1SG potential markers (i.e., those expressing suggestions and asking 'shall we X') are the future markers *-Kohu* '1DU.FUT' and *-Kopu* '1PL.FUT'. I therefore find it more analytically satisfying to treat the potential markers as separate from the imperative markers.

5.3.6 The negative habitual with possessive marking

There is another way to express negative habitual meaning besides negating the continuative. It is formed with *-wa* 'NEG', then an element *-phu*, then the possessive suffixes used on inalienable nouns (§3.1.1, Table 3.1). These forms are listed in Table 5.23.

Table 5.23: The negative habitual paradigm with possessive marking

subject	**NEG.HAB**
1SG	*-wa-phu-yo*
2SG	*-wa-phu-na*
3SG	
1DU	*-wa-phu-yo* (?)
2DU	*-wa-phu-nia*
3DU	*-wa-phu-nua*
1PL	*-wa-phu-ito*
2PL	*-wa-phu-neka*
3PL	*-wa-phu-noka*

Because of the use of the possessive suffixes, this paradigm distinguishes 2DU, 3DU, 2PL, and 3PL, so it is an exception to the syncretism pattern described in §5.1.2. No form is found for 3SG subject. Both of the expected options *-wa-phu-Ø* and *-wa-phu-e* were rejected by consultants. The form for 1DU subject was only marginally accepted.

(184) shows a usage of this kind of negative habitual marking.

(184) *Puhi* *pite* *e-**wa-phu-yo**.*
money big take-NEG-HAB-1SG.POSS
'I don't make a lot of money.' (Komplen 5:16-5:18)

5.4 Inflection of medial verbs

On medial verbs (§5.1.3), the most important distinction is that of *switch reference*, which tells whether the next matrix-clause verb in the chain has the same subject or a different subject from the marked verb. In Hurukui, many of the switch-reference suffixes also agree with the subject of the marked verb (but never the subject of the following verb).

Same-subject marking differentiates between realis status (§5.4.1) and irrealis status (§5.4.2). A verb form called the *serial form* (§5.4.3) also plays a central role in switch reference in

both statuses, but especially in the irrealis. Different-subject marking, in contrast to same-subject marking, is the same regardless of reality status (§5.4.4).

Switch reference in Hurukui is unmarked for relative tense by default, but there do exist switch reference suffixes/paradigms marking simultaneity between clauses (§5.4.5) as well as delayed sequentiality between clauses (§5.4.6). When neither simultaneity nor sequentiality is explicitly marked, there is instead an implicature that the events are listed in chronological order, with each event starting after the last one started.

§5.4.7 describes how switch reference morphology combines with other TAM markers that precede it. §5.4.8 describes how switch reference morphology combines with negation.

5.4.1 Realis same-subject

In realis TAM categories, Hurukui same-subject marking agrees with the same seven person/number categories found in other verb conjugation paradigms. Table 5.24 shows this. The 2SG.SS.R forms *-mai* and *-mu* are in free variation.

Table 5.24: Realis same-subject conjugation paradigm

subject	**SS.R**
1SG	*-khu*
2SG	*-mai* *-mu*
3SG	*-mu*
1DU	*-khu*
23DU	*-khu*
1PL	*-pu*
23PL	*-pi*

(185) shows a long clause chain, all of whose verbs are marked for a 3SG subject which refers to the same person. Thus, all of them except the last one bear the marker *-mu* '3SG.SS.R', and the last one is marked for the TAM that has scope over the whole chain (here, realis with unmarked aspect).

(185) *Khaingu* *ihe* *ahe* *e-**mu*** *yo-**mu*** *aráya-**mu***
this pig INDEF take-3SG.SS.R come-3SG.SS.R wander-3SG.SS.R

kanapa *khaingu* *aráya-**mu*** *hambúkha-**mu*** *kha* *wo-re.*
garden this wander-3SG.SS.R finish-3SG.SS.R and go-3SG.R
'He brings a goat and walks around in this garden and then goes (back home).' (WP2 2:17-2:27)

5.4.2 Irrealis same-subject

In the irrealis TAM categories, Hurukui same-subject marking does not inflect for person/number categories. There are only two forms available to mark this function. The first, the irrealis same-subject form (glossed SS.IRR), is treated in this section. The other, the serial form, is treated in §5.4.3.

The SS.IRR suffix is *-KHa*, which attaches to the default shape of a stem. The verb root *we-* 'go', however, instead uses its fourth shape *wi-* in this situation, creating *wi-sa* instead of the expected **we-kha*[11]. Table 5.25 shows some verb stems and their SS.IRR forms.

Table 5.25: Some verb stems and their SS.IRR forms

stem	**SS.IRR**
yo- 'come'	*yo-**kha***
e- 'take'	*e-**kha***
hapoka- 'cut'	*hapoka-**kha***
nge-nua- 'see them two'	*nge-nua-**kha***
i- 'eat'	*i-**sa***
ukumi- 'grab, hold'	*ukumi-**sa***
ngĕ-iki- 'see them'	*ng-iki-**sa***
we- 'go'	***wi-sa***

(186) shows some uses of the SS.IRR suffix. This chain is in the irrealis mood because the final verb is in the future tense, which Hurukui classifies as irrealis. The switch-reference markers are same-subject because the same person is the subject of each of these verbs.

11 The more regular form *we-kha* is also attested, but much more rarely.

(186) *Uso* *nakamui-**sa*** *ngu_para* *ruka-**kha*** *hindaho* *pahe* *wé-kara.*
rope tie.loosely-SS.IRR and.then tighten-SS.IRR later house go-1SG.FUT
'I'll tie the rope and tighten it, and then after that I'll go home.' (Trap 5:19-5:27)

5.4.3 Serial marking

Verb stems can occur in a form that I call the *serial form*. The serial form of a Hurukui verb has a few different functions. It is most often used to express same subject in medial clauses in the irrealis.

The serial form is based on the default shape of a stem. There are two options for creating the serial form: no change, or a change of stress position to final. The verb root *we-* 'go' uses its fourth shape *wi-* to create the serial form *wi* 'go.SER'. I gloss all serial forms by appending '.SER' to the gloss (or ':SER' after glosses already containing periods). I do this even if they are zero marked, to clearly show that they should be considered to be in serial form based on their role in the sentence.

The two options (zero marking or final stress) are used in free variation. For monosyllabic stems, their outputs are the same. Table 5.26 shows some verb stems and their serial forms.

Table 5.26: Some verb stems and their serial forms

stem	**unmarked serialization**	**final-stress serialization**
a- 'affect.3SG'	*a*	*a*
uki- 'cook'	*uki*	*ukí*
hina- 'watch over'	*hina*	*hiná*
e-ya- 'take me'	*e-ya*	*e-yá*
ě-iki- 'take them'	*iki*	*ikí*
we- 'go'	***wi***	***wi***

(187) shows simple irrealis same-subject uses of serial forms. In this example, the whole clause chain is irrealis because the final verb is marked for future tense, which Hurukui classifies as irrealis, and the medial verbs *we-* 'go' and *humba-* 'fill' are in their serial forms because each of them has the same subject as the following verb.

(187) *Ima* *weka* ***wi*** *sono* ***humba*** *yo-mamo* *í-tara.*
1SG over.there go.SER water fill.SER come-SS.SEQ eat-1SG.FUT
'I'll go fill up (a container with) water and come back and then I'll drink it.' (Kaikai 2:26-2:30)

(188) shows a serial form bearing final stress. The verb *e-* 'take' is marked with the object marker *-(i)ki* '3PL.OBJ' (§5.2.2.1) and serialized by final stress. This choice of serialization marking over the zero-marking strategy does not change the meaning; it marks same subject in the irrealis status.

(188) ***Ĕ-ikí*** ***wì*** *wekă* *á-ke!*
take-3PL.OBJ:SER go.SER over.there put-2SG.IMP
'Take them away and put them over there!' (elicited, consultants produced first)

Stems ending in the continuative and desiderative cannot be serialized. Thus, forms such as **we-maki* or **we-makí* (go-CONT.SER) 'be going (SER)' do not exist, nor do forms such as **wé-ki* or **wè-kí* (go-DES.SER) 'want to go (SER)'.

The serial form can also be used in the realis status in certain situations. This usually relates to when the verbs that are linked by it are working together to describe a multi-part event that is nonetheless conceived of as a single type of action, rather than a sequence of actions. For more on the syntax of these (called *serial verb constructions*), see §8.3.

(189) illustrates the serial form used in the realis status. In the bolded phrase, the serial form is used instead of SS-marked forms in a realis chain, because "going back and forth" is treated as a single type of action here.

(189) *Itae* *masohome~masohome* ***wi*** ***yo***
3SG.EMPH many~PL go.SER come.SER

wi *yo-make-pi* *nge.*
go.SER come-HAB-3PL.R TOPIC
'There are a lot of them (cars) that go back and forth.' (Lae 1:00-1:04)

5.4.4 Different-subject

The Hurukui different-subject paradigm inflects for person/number categories, but not for reality status. Table 5.27 shows it.

Table 5.27: Different-subject conjugation paradigm

subject of the marked verb	**DS**
1SG	*-pa*
2SG	*-mana*
3SG	*-ne*
1DU	*-pua*
23DU	*-pia*
1PL	*-puna*
23PL	*-pina*

(190) and (191) show basic usages of the different-subject paradigm. In (190), the people creating the new villages are the subject of the verb *a-* 'put', and the next verb, *khai-* 'be located', has a subject with a different referent. The verb *a-* 'put' is thus marked with *-pina* '3PL.DS' to show that its own subject is 3PL but has a different referent from the subject of the following verb.

(190) *Hahe* *nenge* *hi-pi* *angimu* *oita*
some.PL village make-3PL.SS.R valley down.there

*a-**pina*** *kha* *khai-pu.*
put-3PL.DS and be.located-1PL.R
'Some of them made villages down in the valley, and we lived there.' (Bibuai 1:38-1:43)

In (191), the subject of *khimi-* 'throw' is 1SG, but the subject of the next verb, *ekha-* 'go down' is 3SG, so *khimi-* 'throw' is marked with *-pa* '1SG.DS'. Additionally, the sentence goes on, and the verb following *ekha-* 'go down' has yet another subject, so *ekha-* 'go down' is marked with *-ne* '3SG.DS' to signal this. Note that what is translated into English here as '(I) put a stick into the ground' is more literally 'I throw a stick and it goes down into the ground'.

(191) *Ngohu* *ruka* *umo* *khimi-**pa*** *maha* *khane* *ekha-**ne** ...*
trap make.trap:SER tree throw-1SG.DS ground inside go.down-3SG.DS
'I'll make a trap and put a stick into the ground, and ...' (Trap 3:35-3:38)

5.4.5 Simultaneous switch reference

As mentioned in the introduction to §5.4, Hurukui switch reference morphology can mark relative tense in addition to same vs different subject. Simultaneous switch reference morphology expresses that the events of the marked clause and those of the next clause are happening at the same time.

Simultaneous SR does not inflect for reality status in either SS or DS. The SS simultaneous morphology is a special morpheme *-hVhV*. The DS simultaneous morphology is more transparent: it consists of a continuative marker plus the normal DS paradigm.

The morpheme *-hVhV* 'SS.SIM' copies the last vowel of the stem to which it attaches. The stem is in its default shape. However, the root *we-* 'go' instead uses its fourth shape *wi-* here, giving the SS.SIM form *wi-hihi*, not **we-hehe*. Table 5.28 shows some verb stems and their SS.SIM forms.

Table 5.28: Some verb stems and their SS.SIM forms

stem	**SS.SIM**
i- 'eat'	*i-hihi*
e- 'take'	*e-hehe*
a- 'affect.3SG'	*a-haha*
yo- 'come'	*yo-hoho*
i-ya- 'do to me'	*i-ya-haha*
we- 'go'	***wi-hihi***

The SS.SIM form of a single-vowel root, such as *i-* 'eat', *i-* 'do', or *e-* 'take', often undergoes deletion of the unstressed initial vowel (§2.5.1). The SS.SIM morpheme cannot attach to a stem containing an overt aspect or polarity marker.

(192) shows a usage of the SS.SIM marker.

(192) *Moko* *ě-**héhe*** *khúhuya-ma-re*
speech take-SS.SIM jump-CONT-3SG.R
'That person is jumping while talking.' (elicited, consultants produced first)

Table 5.29 shows the simultaneous DS paradigm, consisting of a continuative plus the normal DS suffixes. There are some co-occurrence restrictions on which continuative morpheme, i.e., *-maki* vs *-ma(i)*, can be used with which DS suffixes. Disallowed cells are grayed out.

Table 5.29: Simultaneous different-subject conjugation paradigm

subject of the marked verb	**CONT-DS with *-maki***	**CONT-DS with *-ma(i)***
1SG	*-maki-pa*	*-mai-pa*
2SG	*-maki-mana*	
3SG	*-maki-ne*	*-mai-ne*
1DU	*-maki-pua*	
23DU	*-maki-pia*	
1PL	*-maki-puna*	
23PL	*-maki-pina*	

The form of *-ma(i)* in these combinations is always *-mai*, never *-ma*.

(193) shows a usage of simultaneous different-subject marking.

(193) *Na* *sono* *ekha-**maki-mana*** *ngu* *mua* *húke-ma-re*
2SG water go.down-CONT-2SG.DS 3SG smoke blow-CONT-3SG.R
'While you were bathing, that person was smoking.' (elicited)

5.4.6 Sequential switch reference

As mentioned in the introduction to §5.4, Hurukui switch reference morphology can mark relative tense in addition to same vs different subject. Sequential switch reference morphology expresses that the events of the marked clause happen first, then there is some delay, and then the events of the next clause start. It could therefore more accurately be called 'delayed sequential', but I just call it 'sequential' for short because there is no other relative tense category with that name in Hurukui.

Sequential switch reference marking in Hurukui makes use of an element *-amo* 'SEQ', which is homophonous with a particle *amo* that means 'at a future time' and follows time adverbs. For same subject sequential, there is a paradigm that is transparently composed of realis SS

suffixes plus *-amo*, as well as a morpheme *-mamo* that is invariant for person/number categories. This non-agreeing morpheme *-mamo* 'SS.SEQ' is much more common, but its meaning is the same as that of the agreeing paradigm in the realis. It appears to be a generalization of the 3SG form *-mŭ-amo* > *-mamo* to all person/number categories. In the irrealis, only *-mamo* can be used. Table 5.30 shows the agreeing paradigm.

Table 5.30: Same-subject sequential conjugation paradigm

subject	**SS.SEQ**
1SG	*-khŭ-amo*
2SG	(unattested)
3SG	*-mŭ-amo*
1DU	(unattested)
23DU	(unattested)
1PL	*-pu-amo* *-pŭ-amo*
23PL	*-pi-amo* *-pĭ-amo*

These forms are very often, but not obligatorily, followed by the connective word *kha* 'and' (§6.6). (194) shows a usage of the non-agreeing suffix *-mamo* 'SS.SEQ'.

(194) *Uita* *we-**mamo*** *kha* *hamame* *uita* *para*
down.there go-SS.SEQ and again down from

mamai *yai-pu* *nge, ...*
slope climb-1PL.R TOPIC
'We went down and we climbed up a mountain again, and ...' (Dokta 4:23-4:29)

(195) shows a usage of the agreeing paradigm for same-subject sequential.

(195) *Imisi* *khakono* *para* *khainóngì-pi* *mamai* *yaka* *we-**pǐ-amo***
road this.here LOC do.this-3PL.SS.R slope up.there go-3PL.SS.R-SEQ

kha *yo-pi* *uini* *ekha-mai-pi.*
and come-3PL.SS.R down.here go.down-HAB-3PL.R
'On this road, people go up the mountain first and then come down.' (Kofi 4:37-4:43)

Sequential DS is marked by the normal DS paradigm plus *-amo* 'SEQ'. Table 5.31 shows this.

Table 5.31: Different-subject sequential conjugation paradigm

subject of the marked verb	**DS.SEQ**
1SG	*-pă-amo*
2SG	*-mană-amo*
3SG	*-ně-amo*
1DU	*-puă-amo*
23DU	*-piă-amo*
1PL	*-pună-amo*
23PL	*-pină-amo*

(196) shows some usages of the DS.SEQ paradigm (bolded), as well as a usage of the particle *amo* 'at a future time' (underlined). This particle is likely part of the etymological source material for the sequential switch reference affixes. This example is from a text where three men are working together to demonstrate how to cook food in the traditional way. They are discussing what they will do once the food is finished cooking inside a tube of bamboo. These events must occur in a clear sequence, with some delay before the next event starts, so the sequential markers are used.

(196) *Hindaho* *amo* *re-**ně-amo*** *pukhu* *pi-pa*
later at.future.time burn-3SG.DS-SEQ hit stab-1SG.DS

uita *ekha-**ně-amo*** *kha ...*
down go.down-3SG.DS-SEQ and
'Once it's cooked, I'll hit it (bamboo) and it (food) will fall down, and ...' (Kaikai 16:41-16:46)

5.4.7 Switch reference combined with other TAM

Switch reference can also be used in medial clauses whose verbs have aspect markers (Slot 3). This section describes the marking of these combinations, in the same order that the aspects were presented in §5.2.4: perfect, distal past, desiderative, and continuative.

5.4.7.1 The perfect plus switch reference

The realis perfect suffixes are straightforwardly followed by regular realis SS suffixes. The irrealis perfect can only be marked for SS using the suffix *-KHa* 'SS.IRR', not by serialization. All perfect suffixes can be followed by regular DS suffixes. Table 5.32 shows these combinations that mark person.

Table 5.32: Inflecting combinations of perfect and switch reference suffixes

subject	**PRF.R-SS.R**	**PRF.R-DS**	**PRF.IRR-DS**
1SG	*-́khukhài-khu*	*-́khukhài-pa*	*-khài-pa* *-́khài-pa*
2SG	*-máikha-mài* *-́mukhài-mu*	*-máikha-màna* *-́mukha-màna*	*-kha-màna* *-́kha-màna*
3SG	*-́mukhài-mu*	*-́mukhài-ne*	*-khài-ne* *-́khài-ne*
1DU	*-́khukhài-khu*	*-́khukha-pùa*	*-kha-pùa* *-́kha-pùa*
23DU	*-́khukhài-khu*	*-́khukha-pìa*	*-kha-pìa* *-́kha-pìa*
1PL	*-́pukhài-pu*	*-́pukha-pùna*	*-kha-pùna* *-́kha-pùna*
23PL	*-́pikhài-pi*	*-́pikha-pìna*	*-kha-pìna* *-́kha-pìna*

The irrealis PRF plus *-KHa* '[SS.IRR]' yields the combination *-khài-sa ~ -́khài-sa* for all subjects.

5.4.7.2 The distal past plus switch reference

Because the morphological subject marked after the distal past suffix paradigm is always 3SG, and because it cannot be followed by same-subject marking due to its non-referential dummy subject, the only switch reference marker that can follow the DPST is *-ne* '[3SG.DS]'.

Table 5.33: Combinations of distal past and switch reference suffixes

subject	**DPST-DS**
1SG	*-ˊpakhài-ne*
2SG	*-mánakhài-ne*
3SG	*-ˊnekhài-ne*
1DU	*-púakhài-ne*
23DU	*-píakhài-ne*
1PL	*-púnakhài-ne*
23PL	*-pínakhài-ne*

5.4.7.3 The desiderative plus switch reference

With the desiderative, same-subject is marked by a special construction called an internal speech construction. These are covered instead in Chapter 8 on multi-verb constructions, in §§8.4.1.2-8.4.1.3. The desiderative with different-subject morphology is much more straightforward, consisting simply of the desiderative suffix *-ˊKi* plus the normal DS paradigm (§5.4.4). Table 5.34 lists the combined forms.

Table 5.34: Desiderative different-subject conjugation paradigm

subject of the marked verb	**DES-DS**
1SG	*-ˊKi-pa*
2SG	*-ˊKi-mana*
3SG	*-ˊKi-ne*
1DU	*-ˊKi-pua*
23DU	*-ˊKi-pia*
1PL	*-ˊKi-puna*
23PL	*-ˊKi-pina*

(197) shows a DS-marked desiderative, in the meaning of immediate future.

(197) *Ara haringúni yà-**ki-ne** kha hóphophi-pu kha yo-pu.*
place night night.fall-DES-3SG.DS and hurry-1PL.SS.R and come-1PL.R
'Night was about to fall, and we hurried back.' (Bibuai 4:09-4:17)

5.4.7.4 The continuative plus switch reference

The continuative cannot be used with SS suffixes. Instead, the SS.SIM form of the verb is used (§5.4.5). For continuative plus DS, see Table 5.29 in §5.4.5. These forms of the continuative are conditioned by the following DS suffix, so they have the same distribution when the continuative is preceded by a perfect or desiderative suffix and followed by DS.

5.4.8 Switch reference combined with negation

Medial verbs can be negated. For same subject, this is done with a morpheme *-wapi* 'NEG.SS'. It could be thought of as subordinating, with semantics of 'without doing X', or it could be thought of as the same kind of thing as a switch reference marker. Syntactic tests to distinguish between these possibilities must await further research.

The NEG.SS marker is used in both realis and irrealis status. It does not inflect for any person/number categories. The more agglutinative structure of *-wa* 'NEG' plus same subject suffixes does not occur. That is, forms such as **isi-wa-khu* (sleep-NEG-1SG.SS.R) do not exist.

(198) and (199) show the NEG.SS marker in the realis and irrealis status, respectively.

(198) *Ngu khainíngi-**wapi** phai e-khu yó-khukhay-a ko.*
3SG do.thus-NEG.SS knife take-1SG.SS.R come-1SG.PRF.R-1SG.R DISC.PRTC
'I didn't do that, I brought the knife!' (Kaikai 12:13-12:35)

(199) *Sono ekha-**wápĭ** wa-Ø!*
water go.down-NEG.SS go-2SG.IMP
'Go without bathing!' / 'Don't bathe, do go!' (elicited, consultants produced first)

For different-subject verbs, negation is straightforward. The negator *-wa* is followed by the normal DS suffixes. These forms are shown in Table 5.35.

Table 5.35: Different subject negative conjugation paradigm

subject of the marked verb	**NEG-DS**
1SG	*-wa-pa*
2SG	*-wa-mana*
3SG	*-wa-ne*
1DU	*-wa-pua*
23DU	*-wa-pia*
1PL	*-wa-puna*
23PL	*-wa-pina*

(200) illustrates a usage of negative DS marking. The first clause commands the listeners not to touch something. Despite its prohibitive semantics, the correct SR marking is negative DS. This is because it is a medial verb in a chain where the next clause has a different subject, and there is no such thing as a medial imperative or prohibitive marker in Hurukui.

(200) *Ere-**wa-pína*** *noká-kha* *e-képi!*
touch-NEG-2PL.DS 3PL-FOC take-3PL.FUT
'Y'all don't touch it, let them take it!' (elicited)

5.5 Subordinating morphology

There are a few suffixes that create subordinated verbs. These are the *purposive -Ka* (§5.5.1), which creates adverbial clauses, and the *adjectivizer -(i)khu* (§5.5.2), which creates clauses that function as adnominal modifiers. Also, as described in §5.2.6, the conditional can create subordinate clause chains.

5.5.1 The purposive *-Ka*

The purposive suffix *-Ka* is used to form a subordinate adverbial clause with purposive semantics. It is homophonous with the 1SG potential suffix *-Ka* (§5.3.5), but it creates subordinate verbs while the potential suffix creates final verbs.

In (201) there is a subordinate clause (bracketed) whose final verb is marked in the purposive.

(201) *Ima* *[ena* *e-**ka**]* *wé-kara.*
1SG [leaf take-PUR] go-1SG.FUT
'I'll go get a leaf.' (Kaikai REC2: 0:20-0:22)

Hurukui classifies the purposive as irrealis. A subordinate clause chain whose final verb is marked in the purposive will use irrealis SR marking on its medial verbs. This is true regardless of whether the matrix clause is realis or irrealis. In this way the purposive is similar to the desiderative in terms of status marking.

5.5.2 The adjectivizer *-(i)khu*

The suffix *-(i)khu* 'ADJZR' derives adjectives from verbs. It has the allomorph *-ikhu* after *a* and *-khu* elsewhere. The resulting adjectives describe the absolutive argument of a verb, such as *umi-khu* (die-ADJZR) 'dead', or *hòhi-wá-ikhu* (paint-NEG-ADJZR) 'unpainted'. The adjective follows the noun that it modifies (see §7.1.1 for more on noun phrase syntax).

In (202), the S=P ambitransitive verb *re-* 'burn' is adjectivized to mean 'cooked'. It modifies the noun *opopo* 'food' to make a noun phrase (bracketed) meaning 'food that has been cooked' (not 'food that has cooked/burned something'). That is, the absolutive argument (the S/P) is the one whose role in the burning event is being attributed to whatever this adjective modifies.

(202) *[Opopo* *re-**khu**]* *e-ne* *kha* *pahe* *yo-pu.*
[food burn-ADJZR] take-3SG.DS and house come-1PL.R
'He took the cooked food, and we came home.' (Nadzab 13:06-13:08)

In (203), the transitive verb *wana-* 'miss a shot, misbehave' is adjectivized to mean something like 'depraved, delinquent'. It modifies the noun *imisi* 'custom, method' to make a noun phrase (bracketed) meaning 'bad behavior'.

(203) *[Imisi* *wana-**ikhu**]* *hoka-pi.*
[custom misbehave-ADJZR] use.method-3PL.R
'They behaved badly.' (Man3 8:42-8:46)

In (204), the transitive verb *rukhi-* 'split (tr.)' is adjectivized to mean 'split (in the state of having been split)', not 'having split something'.

(204) *uho* *rukhi-**khu***
firewood split.TR-ADJZR
'split firewood' (elicited, consultants produced first)

6 Other word classes

In addition to the word classes described in the preceding chapters, Hurukui of course has a number of others that are smaller, less often used, or of dubious analytical status (in the case of adjectives). This chapter is about those word classes. It covers adjectives, quantifiers, adverbs, adpositions, particles, connectives, and ideophones. Some of these are quite small but are important for understanding the light verb construction system described in Part II (e.g., ideophones).

Description of minor word classes of lesser relevance for Part II will be left out until a later iteration of this grammar. Omitted word classes include the additive and restrictive particles (*umo* 'also', *hai* 'only', and *kanara* 'only'), interjections (e.g., *ih* 'oops!, oh no!'), and a few words with their own idiosyncratic behavior that do not resemble any word class (e.g., *minde*, used only in the construction *minde NP khahu* which means 'a group including NP').

6.1 Adjectives

Adjectives are words that can modify a nominal by following it. Adjectives take very little morphology, but they optionally undergo full reduplication to mark non-singular number.

The most common use of adjectives is adnominal modification. In (205), the adjective *ninira* 'black' modifies the noun *ihe* 'pig'.

(205) *Ihe* ***ninira*** *ni_nge_ngu,* *ihe* *ngamba* *ko-mai-pi.*
pig black TOPIC pig pig.sp say-HAB-3PL.R
'Black pigs are called *ngamba* pigs.' (KainPik 0:41-0:47)

Adjectives can also serve as nonverbal predicates. In (206), the adjective *mbangiri* 'strong' is the predicate.

(206) *Kuni-ŏ* ***mbangiri.***
leg-1SG.POSS strong
'My legs/feet are strong.' (overheard)

Adjectives can also act as NPs themselves (§6.1.3).

6.1.1 Reduplication of adjectives

Adjectives can be fully reduplicated to mark non-singular number of the referent being modified, but this reduplication is not obligatory in all contexts. An example of a non-reduplicated adjective modifying a non-singular referent is seen in (206) above. In that example, the speaker has two legs, each with one foot, but the adjective *mbangiri* 'strong' is not reduplicated.

Reduplication marking non-singular number occurs in adnominal adjectives as well as in predicate adjectives. In (207), the noun phrase *hamunge khomorae-khomorae* 'short men' includes the fully reduplicated adjective *khomorae* 'short'. It is clear from the linguistic co-text, including the free pronoun *hita* '1PL.EXCL' and the agreement marker *-pu* '1PL.R', that multiple men are being talked about. Thus, adjective reduplication is not necessary for ensuring the listener knows about the referent's plurality, but it is used anyway.

(207)	*Hita*	*hamunge*	***khomoráe~khomoràe***	*kanara*	*khai-pu.*
	1PL.EXCL	man	NSG~short	only	be.located-1PL.R
	'Only we short men are here.'				(Chameleon.Tr 17)

In (208), the noun phrase *monu pite-pite* 'big birds' includes the fully reduplicated adjective *pite* 'big'.

(208)	*Monu*	***píte~pìte***	*nge_ngu*	*monu*	*wahawe,*	*monu*	*kaipa, ...*
	bird	NSG~big	TOPIC	bird	hornbill	bird	cockatoo
	'The big birds are the hornbill, the cockatoo, ...'						(BusAbus 1:54-2:05)

Examples of reduplicated predicate adjectives are not very common. In (209), the noun phrase *opopo akhamonoka ngungia-ngungia* 'things with long fur' contains a subordinate clause *akhamonoka ngungia-ngungia* 'their fur is long'. Here the predicate adjective *ngungia* 'long' is reduplicated because it is describing the fur of multiple animals, construed as a non-singular number of furs.

(209) *Na* *moko* *ĕ-mai* *hai* *opopo* *akhamo-noka* ***ngungia~ngungia***
2SG speech take-2SG.SS.R only thing fur-3PL.POSS NSG~long

nge *opopo* *ipe~ipe* *khaingu* *ngáĕ_e-mana* *ngae* *ima* *menae ...*
TOPIC thing NSG~name this tell-2SG.DS thus 1SG I.think
'You talked about this animal as if it had long fur, so I thought that ...' (Fluffy 8:13-8:18)

In (210), however, the predicate adjective *ambirou* 'soft, weak' is not reduplicated even though the referent of the head noun is plural and animate. The reasons why it is not marked as such will require further investigation.

(210) *Mbumbuma* *numu-noka* ***ambirou***.
white.person body-3PL.POSS soft
'White people have soft/weak skin/bodies.' (Misin 5:49-5:53)

6.1.2 Intensive suffix *-ko ~ -ngu*

Only one adjective takes intensive morphology: *pite* 'big'. In the Arapipi dialect, the intensive suffix is *-ngu*, while in the Eastern dialect it is *-ko*. Thus the dialectal variants *pitengu* and *piteko* both mean 'very big'. It is shown adnominally in (211), modifying the noun *yoho* 'fire'.

(211) *Yoho* *koka-puna* *yoho* ***pite-ko*** *ahe*
fire make.fire-1PL.DS fire big-INTENS INDEF

mirihi *a-mu* *mokepe* *we-make-re.*
flame put-3SG.SS.R above go-cont-3SG.R
'We make a big fire and it lights and goes up into the air.' (Pot2 6:03-6:07)

These can also be reduplicated. In (212), the intensive adjective *piteko* 'very big' is fully reduplicated to express non-singular number.

(212) *E* *kehe* ***piteko~pite-ko*** *noka-komu* *ni* *ngu* *kehe,*
eh cough NSG~big-INTENS hit.3PL-3SG.FUT FOC TOPIC cough

umo *ariki* *i-tepi.*
tree juice eat-3PL.FUT
'If they have bad coughs, they'll drink tree sap/juice.' (Sik 2:28-2:33)

Adjectives can also be intensified using particles (§6.5.4).

6.1.3 Adjectives used nominally

The adjective-noun distinction in Hurukui is often unclear. Adjectives can act as noun phrases without a head noun, and nouns can act as adnominal modifiers (which is distinct from the noun-noun possession construction, §7.1.1.1). In (213), the adjective *humbukura* 'cold' acts as a nominal, because it is the object of the verb.

(213) ***Humbukura*** *e-mai-pu.*
cold take-CONT-1PL.R
'We started to feel cold.' (lit. 'We were taking cold.') (Bibuai 5:14-5:18)

In (214), the noun *arakua* 'jungle, wilderness', a univerbation of the phrase *ara kua* 'hidden place', is used as an adnominal modifier. It is fully reduplicated because the modified noun's referent is nonsingular. This is adjective-like behavior, but the word is otherwise clearly a noun.

(214) *Opopo* ***arakua~arakua*** *e-khu* *yo-khu ...*
thing NSG~jungle take-1SG.SS.R come-1SG.SS.R
'I brought things from the jungle, and ...' (Purpur 1:46-1:50)

In (215), the inalienable noun *angere* 'vicinity', whose word class is seen by its ability to take possessive suffixes (§3.1.1), is used as an adnominal modifier and fully reduplicated.

(215) *Ngu_para* *hoka-pa* *akepa-pu*
and.then cut-1SG.DS distribute-1PL.SS.R

pahe ***angereito~angere-ito*** *noke-pu* *kha ...*
house NSG~vicinity-1PL.POSS give.3PL-1PL.SS.R and
'And then I cut it (meat), and we shared it with the houses nearby, and ...' (Hunt5 3:57-4:07)

However, there are still distributional asymmetries in the modification construction. More noun-like words, such as *pahe* 'house', tend to be the modified and thus come first. More adjective-like words, such as *pite* 'big', tend to be the modifier and thus come second. But these overlaps in morphosyntax between nouns and adjectives require further investigation, and it may turn out that positing a separate word class of adjectives in Hurukui is not the best analysis.

6.2 Quantifiers

Quantifiers are words that occur adnominally to add meaning about how much/many of the referent there are. They occur after the head noun. The subtypes of quantifiers are numerals, general quantifiers, indefinite quantifiers, and an interrogative quantifier.

6.2.1 Numerals

The only monomorphemic numerals in the language are *harawohe* 'one', *apoki* 'two', and *apono* 'three'. The numeral *apono* 'three' can also be used with the broader meaning 'a few, several'.

In (216), the numeral *apono* 'three' quantifies the noun *kopohe* 'child, young man'.

(216)

Hita	*kopohe*	***apono***	*we-púnakhai-Ø.*
1PL.EXCL	child	three	go-1PL.DPST-3SG.R

'We three boys went.' (Mario 0:29-0:33)

There is a system of composing higher numbers using words for fingers for units of one, and hands/feet for units of five (e.g., *omaito anukera* 'our hand with many children' for 'four', *omaito harawohe* 'one of our hands' for 'five'). In (217), the noun *hamunge* 'man' is quantified by a phrase literally translating to 'one of our hands (and) two of our fingernails', meaning five fingers plus two more fingers, meaning the number seven.

(217)

hamunge	***oma-ito***	***harawohe***	***hunginga-ito***	***apoki***
man	hand-1PL.POSS	one	fingernail-1PL.POSS	two

'seven generations' (Histry 9:17-9:22)

There is also a system in which Arabic numeral digits can be referred to by the names of objects that they visibly resemble (e.g., *ngorane* 'egg' for *0*, *saporo* 'axe' for *7*), and a system in which amounts of money are referred to by the names of objects on the corresponding banknote (e.g., *hamunge pirone* 'head of man' for 50 kina, because the banknote has a picture of Michael Somare). A detailed description of these conventions is omitted for the purposes of this dissertation.

6.2.2 General quantifiers

I use the term *general quantifiers* to refer to words such as *amie* 'many, much' and *khuha* 'all' which quantify an amount which is definite in a given context but varies depending on context. This differs from numerals, which quantify the same amount in all contexts.

The quantifier *amie* means 'many' or 'much'. It can be used to modify nouns with count or mass semantics, because Hurukui does not distinguish between these. In (218), it modifies the count noun *kopohe* 'child' to make a noun phrase meaning 'many children'.

(218) *Sukuru pai yo-re nake kopohe* ***amie*** *yo-pi.*
school new come-3SG.R because child many come-3PL.R
'Because a new school came, many children came.' (Skul 2) (written only)

In (219), *amie* 'many, much' modifies the mass noun *opopo* 'food'.

(219) *Opopo* ***amie*** *i-make-pu.*
food much eat-HAB-1PL.R
'We eat a lot of food.' (Man3 6:02-6:05)

The quantifier *masawohome ~ masohome* has a similar meaning to *amie*. I translate it as 'a lot'. It is optionally fully reduplicated when the modified referent is nonsingular and countable. This is shown in (220).

(220) *Na kui* ***masohómĕ~masohòme****.*
2SG bag NSG~a.lot
'You have a lot of bags.' (overheard)

In (221), however, it is not reduplicated even though the referent of *ihe* 'pig' is animate and nonsingular.

(221) *Ihe khaingu* ***masohome*** *yo-pi.*
pig DEM a.lot come-3PL.R
'A lot of pigs came.' (PikGaden 0:35-0:39)

Animacy does not seem to be a factor in whether this quantifier will be reduplicated. The head noun *kui* 'bag' in 220 is inanimate, but the quantifier is reduplicated, while the head noun *ihe* 'pig' in 221 is animate, but the quantifier is not reduplicated. This is the opposite of the usual

pattern in Hurukui, in which adnominal modifiers are more likely to be reduplicated when the noun's referent is nonsingular (§6.1.1).

The quantifier *khuha* 'all' can mean all individuals of some population or all the mass of a substance. Hurukui does not mark a distinction between these meanings. In (222), *khuha* 'all' is used twice, once in the phrase *hamunge khuha* 'all men' and once in the phrase *opopo khuha* 'everything'.

(222)

Hamunge	***khuha***	*moko*	*harawohe*	*e-pi*	*opopo*	***khuha***
man	all	language	one	take-3PL.SS.R	thing	all

e-pi	*a-pi*	*pahe*	*ngungia*	*ahe*	*yai-pi.*
take-3PL.SS.R	put-3PL.SS.R	house	tall	INDEF	build-3PL.R

'All men spoke one language and took everything and built a tall house.' (Histry 0:53-1:11)

6.2.3 Indefinite quantifiers

The indefinite quantifiers are *ahe* 'INDEF.SG' and *hahe* 'INDEF.NSG'. They can encode both specificity and non-specificity of the referent. The quantifier *ahe* 'INDEF.SG, one, some (SG), another' has a few different uses. In (223), *ahe* modifies the noun *pahe* 'house' and contributes an indefinite non-specific meaning of 'one, a(n)'. This is the meaning that I gloss simply as 'INDEF.SG'.

(223)

Ima	*nge*	*pahe*	***ahe***	*yai-Ø*	*ngu*	*isí-tè-ra.*
1SG	TOPIC	house	INDEF.SG	build-SER	that.place	sleep-DES-1SG.R

'I want to build a house and sleep there.' (BikHevi 14:46-14:51)

The quantifier *ahe* also has a meaning like 'some (SG), an undetermined, a certain' that is still indefinite but has a specific referent, which may or may not be known. In (224), it modifies the noun *opopo* 'thing' to mean that the speaker is about to mention a particular thing that has not been introduced yet in the story.

(224)

Opopo	***ahe***	*ekhono*	*háye-re*	*ni_ngé_ngu,*
thing	a.particular	now	arise-3SG.R	TOPIC

hauyŭ	*mokepe*	*para*	*we-ma-re.*
wind	above	LOC	go-HAB-3SG.R

'One thing that has now come about, is the cell phone network.' (Laiftaim 2:47-2:56)

In (225), *ahe* modifies the noun *anga* 'day' to make a noun phrase meaning 'another day, one more day'.

(225) *Anga* ***ahe*** *khai-sa* *we-kepi.*
day another stay-SS.IRR go-2PL.FUT
'Y'all will stay another day and then go.' (Guria3 4:09-4:13)

The quantifier *hahe* 'some (NSG)' has plural indefinite meaning. In (226), it modifies the noun *hamunge* 'man' and contributes the meaning of 'some but not all'.

(226) *Hamunge* ***hahe*** *nukhi-ta-pi* *yo-pi.*
man some.NSG leave-1PL.OBJ-3PL.SS.R come-3PL.R
'Some men left us and came.' (Bibuai 1:32-1:36)

Its form is likely a shortening of *áhě-àhe* (surface form *áhàhe*), which is the fully-reduplicated form of *ahe*. This more conservative form is still commonly found, e.g., in the phrase *mokono-ito áhàhe* (sister-1PL.POSS some.NSG) 'some of our sisters'.

6.2.4 Interrogative quantifiers

The quantifier *keme* asks the question 'how many, how much'. In (227), it occurs in adnominal position, modifying the noun *anga* 'day'.

(227) *Anga* ***keme*** *kha-rei?*
day how.many stay-2DU.R
'How many days did you two stay?' (Nadzab 12:08-12:10)

6.3 Adverbs

Adverbs are words that modify verbs, clauses, and clause chains to add information such as time, place, reason, and manner. Adverbs are placed relatively freely in the clause, but they must precede the verb(s) they modify. They have no morphology.

Adverbs expressing time include *ekhono* 'now, today', *kepu* 'yesterday', *nuimi* 'tomorrow', *kaya* 'long ago', *araiku* 'already, earlier', and *kemeka* 'when'.

In (228), the adverb *araiku* 'earlier, already' modifies the entire clause chain to express that all the verbs already happened.

(228) ***Araiku*** *nd-apono* *ngáě-iki-pa* *wo-pi.*
already those.people-several tell-3PL.OBJ-1SG.DS go-3PL.R
'I already told those people and they left.' (Pinat 0:42-0:44)

Adverbs expressing place are to be distinguished from both inalienable nouns (some of which express spatial relations, §3.1.1) and deictics (Chapter 4), which have different morphosyntactic behavior. Adverbs of place include words like *mokepe* 'up, above', *apakhane* 'indoors', *nenge* 'outdoors', *khukhui* 'hidden', and *haya* 'in the open, visible'.

In (229), the adverb *haya* 'in the open' modifies only the verb *a-* 'put' which follows it.

(229) *Humekhe* *eko* *e-Ø* ***haya*** *a-wapho-Ø.*
container near.you take-SER in.the.open put-PROH-2SG.IMP
'Don't put that container out in the open.' (Kaikai 13:45-13:47)

Adverbs expressing reason include *ínake* 'why', a univerbation of the phrase *ino ake* 'for what'. In (230), it is used to modify the second clause chain, containing the verbs *e-* 'take', *hapoka-* 'cut', and *khaníngi-* 'do thus'.

(230) *Kheremu* *i-mukhai-wa-Ø* *ni* ***ínake*** *phai* *e-mu*
knowledge do-3SG.PRF.R-NEG-3SG.R EMPH why knife take-3SG.SS.R

ópŏ~opo *hapoka-mu* *khainíngi-ma-re?*
NSG~thing cut-3SG.SS.R do.thus-CONT-3SG.R
'He doesn't know what he's doing; why is he using a knife and cutting things like that?' (VD2 1:00-1:05)

Adverbs expressing manner include *arare ~ araremae* 'slowly', and *pena* 'quickly'. In (231), *araremae* 'slowly' modifies the verb *yákha-* 'stand up'.

(231) *Ima* ***ararĕmae*** *uita* *para* *yákha-khu ...*
1SG slowly down.there LOC stand.up-1SG.SS.R
'I stood up slowly from down there, and ...' (PutimBebi 11:12-11:15)

Other adverbs include *hamame ~ hameme* 'again, back'. In (232), it modifies the clause chain containing the verbs *ekha-* 'descend' and *yo-* 'come'.

(232) *Ngu_para* ***hameme*** *ekha-pu* *yo-pu.*
and.then back descend-1PL.SS.R come-1PL.R
'After that, we came back.' (Pik 2:18-2:22)

6.4 Adpositions

Adpositions are words that must take a nominal complement to create a phrase that has some kind of modification meaning. Adpositional phrases usually modify nouns or verbs. These phrases usually express meanings such as location, direction, instrument, comitative, or benefactive.

Almost all adpositions in Hurukui are postpositions. An example is *para* 'LOC' (usually meaning 'from' but also often 'at'). In (233), its argument is the noun *imisi* 'road', creating the postpositional phrase *imisi para* 'from the road'.

(233) *Ngu* *khongokho* *imisi* ***para*** *e-pi*
3SG hat road from take-3PL.SS.R

ngu *e-pi* *we-pi* *no-pi.*
3SG take-3PL.SS.R go-3PL.SS.R give.3SG-3PL.R
'They took his hat from the road and took it to him and gave it to him.' (WP2 6:38-6:49)

Other postpositions include *koko* 'COMIT', *khahu* 'COMIT', *ake* 'for', and *kha* 'INSTR'. These are variously used in the marking of oblique noun phrases that act as adjuncts to verbs (§7.3.5).

There is one preposition in the language: *himu* 'like, similar to'. In (234), the prepositional phrase *himu khaino* 'like this' modifies the noun *kunisu* 'corner'.

(234) *Kunisu* ***himu*** *khaino* *a-pina* *kha* *itae* *kha-komu.*
corner like this put-3PL.DS and 3SG.EMPH stay-3SG.FUT
'They put it in a corner like this one, and it will stay there.' (Pot 4:31-4:36)

6.5 Particles

I use *particle* as a convenient "wastebasket" term for a group of words that do not have a unified nature. Generally speaking, what I here refer to as particles are words with no morphology that do not fit well into the other categories and are members of small closed classes. They have various functions, such as negation (§6.5.1), topicalization (§6.5.2), focus (§6.5.3),

intensification (§6.5.4), adding emphasis to statements (§6.5.5), adding emphasis to questions (§6.5.6), and answering polar questions (§6.5.7).

6.5.1 Negative particles

There are two particles used for negating nonverbal predicates[12]. These particles do not inflect for person or number. The first negative particle, *ome* 'is not', negates nonverbal predicates headed by nominals, adjectives, and quantifiers.

In (235), the subject is null, and the negated nonverbal predicate is the noun phrase *ara humaya ahe* 'a good place'.

(235) *Ara humaya ahe* ***ome***.
place good INDEF is.not
'It's not a good place.' (Dokta 0:52-0:56)

In (236), the negated nonverbal predicate is the adjective *mbangiri* 'strong, tough, sturdy'.

(236) *Hita opopo khuha ekhono e-mai-pu ni,*
1PL.EXCL thing all now take-HAB-1PL.R RELZR

Saina para e-mai-pu ni, ngu hai opopo
China from take-HAB-1PL.R RELZR 3SG only thing

ambirou wohe, mbangiri ***ome***.
weak very strong is.not
'Everything we get from China now is just weak stuff, it's not strong.' (Komplen 2:51-3:02)

In (237), the subject is *nambehe* 'banana', and the negated nonverbal predicate is the quantifier phrase *awohe amie* 'too many'.

(237) *Nambehe awohe amie* ***ome***.
banana too.much many is.not
'There aren't too many bananas.' (Dokta 6:30-6:34)

The other negative particle, *pe* 'NEG.EXIST', expresses negative existential meaning. In (238), this particle negates the existence of any referents of the noun *hamunge* 'man'.

12 Here I discuss only their use for negating nonverbal predicates in final clauses. For negation of nonverbal predicates in medial clauses, see §8.1.7.

(238) *Ngu* *pahe* *khaingu* *yo-re* *ningu* *hamunge* ***pe***.
3SG house this come-3SG.R TOPIC man NEG.EXIST
'He came to his house, and there was no one there.' (HausPisin 1:56-2:00)

These negative existential predicates are also used to express the meaning of not having something. In (239), the pronoun *hita* '1PL.EXCL' is the possessor, the noun phrase *kar imisi* 'car road' is the possessed, and the possession relation is negated.

(239) *Hita* *kar* *imisi* *ahe* ***pe***.
1PL.EXCL car road INDEF NEG.EXIST
'We don't have a car road.' (Wok 1:39-1:43)

6.5.2 Topic particles

There is a common sequence of particles that is used for topicalizing: *ni ngé ngu* 'TOPIC'. In addition to this three-particle sequence, any subsequence of it that preserves the particles' linear order may also be used for topicalizing. There are thus seven possibilities:

- *ni* 'TOPIC'
- *nge* 'TOPIC'
- *ngu* 'TOPIC'
- *ni ngé* 'TOPIC'
- *ni ngu* 'TOPIC'
- *ngé ngu* 'TOPIC'
- *ni ngé ngu* 'TOPIC'

The differences in meaning among these must await future investigation.

In (240), the topicalizing particle sequence *ni ngé ngu* is used to introduce another kind of frog in a list of various kinds of frogs.

(240) *Humura* *khokora* ***ni_ngé_ngu*** *itae* *kaku.*
frog green.frog TOPIC 3SG.EMPH different
'The green frog is another kind.' (Abus 7:04-7:08)

In (241), only *ngé ngu* is used, to mark as topical a village that the speakers had already been discussing.

(241) *Mbimbiu* ***ngé_ngu*** *makara* *khai-Ø.*
Bimbiu.village TOPIC middle be.located-3SG.R
'Bimbiu village is in the middle.' (Rot 1:35-1:38)

In (242), only *ni* is used. It marks the deictic *khaingu* 'that one' as topical. The deictic here refers to something that the other speaker had just finished telling about, so it was old, given information.

(242) *Kohe,* *khaingu* ***ni,*** *nia* *khane-nia* *humaya*
okay that.one TOPIC 2DU inside-2DU.POSS good

e-khu *harae* *a-khu* *ngu_para* *na* *no* *yo-ma?*
take-2DU.SS.R together put-2DU.SS.R and.then 2SG here come-2SG.R
'Okay, about that, you two reached an agreement and then you came here?' (Nau 1:28-1:38)

6.5.3 The contrastive focus marker *kha*

The particle *kha* 'FOC' marks contrastive focus. It takes the form of a suffix after pronominals and deictics, and it is a separate word elsewhere. Table 6.1 shows some examples.

Table 6.1: Occurrence of *kha* 'FOC' as a suffix versus as a particle

form	with *kha* 'FOC'	suffix or particle	gloss	word class
ima	*imá-kha*	suffix	1SG	personal pronoun
na	*ná-kha*	suffix	2SG	personal pronoun
ngu	*ngú-kha*	suffix	3SG	personal pronoun
hia	*hiá-kha*	suffix	1DU.EXCL	personal pronoun
naya	*nayá-kha*	suffix	1DU.INCL	personal pronoun
nia	*niá-kha*	suffix	2DU	personal pronoun
nua	*nuá-kha*	suffix	3DU	personal pronoun

hita	*hitá-kha*	suffix	1PL.EXCL	personal pronoun
naita	*naitá-kha*[13]	suffix	1PL.INCL	personal pronoun
neka	*neká-kha*	suffix	2PL	personal pronoun
noka	*noká-kha*	suffix	3PL	personal pronoun
ino	*inó-kha*	suffix	what	interrogative pronoun
neku	*nékŭ-khă*[14] [nɛ́x]	suffix	who	interrogative pronoun
khaino	*kha(i)nó-kha*	suffix	this	deictic
khaingu	*kha(i)ngú-kha*	suffix	this one / that one	deictic
kheko	*khekó-kha*	suffix	that near you	deictic
ihe	*íhe kha*	particle	pig	alienable noun
mba-noka	*mba-nóka kha*	particle	grandfather-GENER	inalienable noun
itana	*itána kha*	particle	2SG.EMPH	emphatic pronoun
Girson	*Girsón kha*	particle	(given name)	proper noun

The following examples illustrate the basic uses of *kha* 'FOC'. Whether it is a suffix or particle has no effect on its meaning.

The particle is primarily used to mark contrastive focus. This is shown in (243), where the subject *ima* '1SG' is focused to encode that it was me, not someone else, who bathed.

(243) ***Ima-kha*** *sono* *ekha-ra.*
1SG-FOC water go.down-1SG.R
'**I'm** the one who bathed.' (elicited, consultants produced)

The particle also commonly marks agents. In (244), the subject *ipoe* 'that person's father' is marked as agentive using *kha* 'FOC', glossed here as 'AGT' to reflect this meaning.

(244) *Ipo-e* ***kha*** *sono* *ngunu* *khaingu* *we-mu* *hupúri-mu ...*
father-3SG.POSS AGT water pool this go-3SG.SS.R show-3SG.SS.R
'His father went and showed him this pool.' (Mira 1:58-2:02)

13 The form **natá-kha*, which would be expected due to monophthongization of unstressed *ai* (§2.5.2), is not used.

14 This form retains the original stress position on the penult of the root *neku* 'who', rather than shifting it to **nekú-kha*, as would be expected from suffixing. Nonetheless, I analyze this form as suffixed because of the obligatory lenition of the *kh* in *-kha* 'FOC'.

Another common function of the particle is to mark instruments. This is seen in (245), where the noun *hengemu* 'fork, nail' is marked as instrumental. The particle is glossed here as 'INSTR' to reflect this meaning.

(245) *Kaya* *hengemu* ***kha*** *pi-mai-pi.*
long.ago nail INSTR stab-CONT-3PL.R
'Long ago they poked it with nails/darts.' (Nadzab 9:55-9:58)

Finally, the particle is used in a construction where a question is asked of the form [NP *kha*], meaning 'What about NP?'. This is illustrated in (246).

(246) *Khaino* ***kha****?*
this FOC
'What about this?' (Mbaruha1 3:14-3:18)

Deictics are suffixed with *-kha* 'FOC' when they are the last word in a noun phrase that is marked as focused. In this way the focus marker could be considered an enclitic on whole noun phrases, but I retain the suffix-vs-particle analysis due to the fact that its phonological attachment is conditional on the word class of the host.

In (247), the phrase *kono yaka* 'that mountain up there' is focus-marked.

(247) *Nenge* *khai-no* *nukhi-Ø* *we-kopu,* *hindahu* *hingáhisa*
village DEM-here leave-SER go-1PL.FUT later lest

kono ***ya-ka-kha*** *khape-ta-komu.*
mountain up-there-FOC cover.up-1PL.OBJ-3SG.FUT
'Let's leave this village, lest that mountain cover us up (i.e., landslide).' (Guria3 19-21)

The non-oriented deictics (§4.1.1) cannot take the suffix *-kha* 'FOC'. While the focused form *ngu-kha* is common, it is formed from the personal pronoun *ngu* '3SG', not the deictic *ngu* 'that place'.

Focus marking is more common on deictics with the demonstrative prefix than on those without it. Monophthongization of unstressed *ai* to *a* (§2.5.2) is optional here (bolded in the examples).

- *khaino* 'this' + *-kha* 'FOC' > *kh**ai**nokha* ~ *kh**a**nokha* 'this (FOC)'

- *khaingu* 'that one' + *-kha* 'FOC' > *kh**ai**ngukha* ~ *kh**a**ngukha* 'that one (FOC)'
- *kheko* 'that near you' + *-kha* 'FOC' > *khekokha* 'that near you (FOC)'

In (248), *khanokha* 'this (FOC)' is used to express that there is only one thing in a scene: snow/ice.

(248) *Opopo* *humbukura* *kanara,*
thing cold only

umo *ameha* *umo* *ínŏ~ìno* *pe,*
tree leaf tree PL-what NEG.EXIST

hai ***kha-no-kha*** *kha-ma-re.*
only DEM-here-FOC exist-CONT-3SG.R
'Only cold stuff, there are no leaves or trees or anything, there is only this.' (Chameleon 2:27-2:35)

In (249), *khaingu* 'that one' is focus-marked for instrumental meaning. It also has contrastive focus meaning here: the speaker is telling the listener to stop using a modern knife because they are supposed to be showing how food was made in earlier times.

(249) *Ihe* *ona-Ø* *e-Ø* *yo-kanu* *ngu* *epu.*
pig bone-3SG.POSS take-SER come-2SG.FUT TOPIC all.right

Khai-ngu-kha *mburumu* *hangiki-tanu.*
DEM-that.place-FOC sweet.potato slice-2SG.FUT
'If you bring a pig bone, that'll be okay. You can slice the sweet potatoes with that.' (Kaikai 12:08-12:13)

6.5.4 Intensifying particles

Several particles add some sort of intensifying meaning to adjectives. Some also intensify the whole predicate or sentence.

The particle *wohe* 'very, completely' intensifies adjectives. In (250), it modifies the-adjective *khomorae* 'short'. This adjective may look adverbial, as if it modifies the verb *umitepi* 'y'all will die', but actually it is a predicate adjective whose subject is a headless relative clause (§8.4.2.2) *neka umitepi* '(the time when) y'all will die', which is an NP.

(250) *Ngáe-ta-pi* *neka* *umi-tepi* *nge*
tell-1PL.OBJ-3PL.SS.R 2PL die-2PL.FUT TOPIC

hai *khomorae* ***wohe*** *khaino* *ko-pi.*
only short very this say-3PL.R
'They told us that our deaths would come very soon.' (BikHevi 12:40-12:44)

This particle is homophonous with an adjective *wohe* 'native, true' which modifies nouns to mean that the referent is the real deal, more authentic, or native as opposed to imported. An example is the phrase *moko wohe* (speech native) 'local language', often used to refer to Hurukui as opposed to Tok Pisin. These two uses of the form *wohe* are probably historically related.

The particle *awohe* 'too much, a lot' is phonologically similar to but distinct from *wohe* 'very'. It intensifies adjectives and verbs but precedes them. In (251), *awohe* 'too much' modifies the adjective *maya* 'good'. This creates an adjective phrase *awohe maya* 'too good' which is used as an NP meaning 'something/anything very good'.

(251) ***Awohe*** *maya* *hite-wa-Ø.*
too.much good give.1PL-NEG-3SG.R
'He didn't give us (anything) very good.' (Kofi 8:54-8:59)

In (252), *awohe* 'too much' modifies the verb *noroka-* 'think'.

(252) *Mokono-ito* *khaingu* ***awohe*** *noroka-khu ...*
sister-1PL.POSS this too.much think-1SG.SS.R
'I was thinking about this sister of ours a ton, and ...' (Nadzab 0:55-0:58)

There is also a particle *akahe* 'somewhat, a little' which behaves syntactically the same as *awohe* 'too much, a lot' but has the opposite meaning: lower intensity. I include it in this section for that reason, although the label *intensifier* does not seem very good for this function.

In (253), *akahe* 'somewhat' modifies the adjective *anute* 'small'.

(253) *Nge-kanu* *mbo* *ane,* *khaporo* *khahu*
see-2SG.FUT animal whatchamacallit wallaby COMIT

haráě_i-mukhai-Ø *nemo* *ngu* ***akahe*** *anute.*
be.similar-3SG.PRF.R-3SG.R but 3SG somewhat small
'You'll see that it resembles a (what's the name of that animal), a wallaby, but it's a bit small.' (Abus 14:48-14:54)

In (254), *akahe* 'somewhat' modifies the verb *khai-* 'stay' to mean that the subject stayed there for a short amount of time.

(254) *Hita* ***akahe*** *kha-puna* *sono* *húkhe-ma-re* *yo-pu* *ni.*
1PL.EXCL a.little stay-1PL.DS rain stop-CONT-3SG.R come-1PL.R EMPH
'We stayed a bit, and once the rain started abating, we came.' (Asde 3:00-3:06)

A common way of intensifying adjectives is by adding *maisa* 'bad' after them. This often, but not always, means that the intensity is to a bad extent.

In (255), *maisa* 'bad' modifies the adjective *mbukura* 'cold' to mean unpleasantly cold.

(255) *Sono* *mbukura* ***maisa***.
water cold bad
'The water is very cold (too cold to bathe in).' (Dokta 7:03-7:10)

Less commonly, the adjective *(hu)maya* 'good' can be used in this way as well, to mean intensity to a good extent. In (256), *maya* 'good' modifies the adjective *mbukura* 'cold' to mean pleasantly cold.

(256) *Sono* *mbukura* ***maya***.
water cold good
'The water is nice and cold (refreshing to drink).' (overheard)

An exception to this pattern is found in the adjectives *anu maisa* 'small (SG)' and *musu maisa* 'small (PL)'. Here, the element *maisa* does not mean 'bad' or 'very'. These phrases have simply become frozen, and the *maisa* is required but no longer carries a meaning of its own. These adjectives cannot receive a further *maisa* to add intensive meaning.

The particle *pore ~ poro* 'INTENS' has a general meaning of intensification or certainty, whose more specific interpretation depends on context. When modifying motion predicates, it means 'all the way'. In (257), the particle is placed before the whole verb phrase and contributes the meaning of 'all the way'.

(257) ***Pore*** *khaingu* *para* *we-mu ...*
INTENS that.place LOC go-3SG.SS.R
'He went all the way there, and ...' (Huakono 4:37-4:40)

When modifying negative existential predicates, *pore ~ poro* 'INTENS' means 'at all'. This is shown in (258).

(258) *Imisi* *humaya* *nama* *humaya* *ahe* *ningu* ***pore*** *pe.*
road good lifestyle good INDEF TOPIC INTENS NEG.EXIST
'There is no good road and no good amenities at all.' (Aser 1:13-1:19)

When modifying adjectives, *pore ~ poro* 'INTENS' precedes them, and it has a generic intensive meaning like 'very'. In (259), the predicate *maisa pa-* 'feel bad' is intensified in this way.

(259) *Yo-khu* ***pore*** *maisa* *pa-khu ...*
come-1DU.SS.R INTENS bad feel-1DU.SS.R
'We came and we felt terrible, and ...' (Pik 1:46-1:49)

When modifying future-tense verbs, *pore ~ poro* 'INTENS' adds the meaning of obligation ('must'). This is shown in (260).

(260) *Ngáhitomu* ***pore*** *wé-kara.*
all.right INTENS go-1SG.FUT
'All right, I've got to go.' (Guria3 8:09-8:12)

However, *pore ~ poro* 'INTENS' is also used in many contexts in which its meaning is less clear to me. In (261), it does not contribute any obvious meaning.

(261) *Kopohe* *umo* ***pore*** *nare-pi* *e-pi* *yo-pina ...*
young.man tree INTENS carry-3PL.SS.R take-3PL.SS.R come-3PL.DS
'The young men brought the tree by carrying it, and ...' (BusHaus 8:36-8:40)

Examples like this abound. The various uses of this particle will require further investigation.

6.5.5 Emphatic particles

Hurukui has various particles which are placed at the end of an utterance to express some kind of emphasis. These include *hi*, *ni*, and *o*, all of which I gloss as 'EMPH' for lack of understanding of their more specific meanings.

The emphatic *hi* can add general emphasis to a sentence. In (262), the speaker seems to be expressing frustration or exasperation that he cannot remember what he had planned to say.

(262) *Ahe* *ekékĭ* *ya-re* ***hi.***
INDEF forgetting affect.1SG-3SG.R EMPH
'I forgot some of them.' (TokGusap 6:19-6:22)

The emphatic *hi* is also often used at the end of direct quotes, as shown in (263), but this is not done for all direct quotes, and *hi* 'EMPH' has other uses, so it is not solely a quote-ending particle.

(263) *Ima* *rame-o* *ngu* *khai-Ø* ***hi*** *ko-mu ...*
1SG husband-1SG.POSS 3SG be.located-3SG.R EMPH say-3SG.SS.R
'She said, "My husband is here!" and ...' (Huakono 8:14-8:19)

The emphatic *ni* seems to have shades of meaning in which the speaker thinks the listener does not realize something. This is shown in (264), in which the speaker is telling people about a premonition.

(264) *Mie* *pite-ko* *e-komu* ***ni***.
earthquake big-INTENS take-3SG.FUT EMPH
'There will be a huge earthquake.' (MejikDiwai 2:39-2:42)

The particle *ni* also functions as a relativizer. It can optionally be used at the end of relative clauses (§8.4.2). A simple example is shown in (265), where the bracketed part is a relative clause internally headed by *imisi* 'road, path, custom, method'.

(265) *[Noka* *imisi* *ekhono* *khai-pi* ***ni]***,
3PL custom now be.located-3PL.R RELZR

ngu *nukhi-pi* *kha ...*
3SG abandon-3PL.SS.R and
'They left behind the customs they had been following, and ...' (Man3 6:12-6:19)

The emphatic *o* sometimes seems to have shades of meaning in which the speaker wants the listener to relent about something. This is shown in (266), where the listener keeps trying to take over the job of handling the lighter, and the speaker resists.

(266) *Ima* *kaisa* *húngè-kè-ră* ***o***.
1SG bamboo.lighter pull-DES-1SG.R EMPH
'I'll pull the bamboo lighter!' (Kaikai 3:08-3:10)

The emphatic *o* is also often used in greetings and is phonetically long. For instance, *imuhu maya* 'good morning' is often said as *imuhŭmáyoo*, and *arare wa* 'bon voyage' (lit. 'go slowly') is often said as *ararĕwáoo*.

6.5.6 Question particles

Some particles are placed at the end of a question. These are *kho* 'I wonder', *khuní* 'hell if I know', and *o* 'TAG.Q'.

The particle *kho* is used when the speaker is wondering about the answer to the question. In (267), the speaker is being told what is happening in a video that he cannot see, and he becomes confused about what the video's subject is doing.

(267) *E-Ø* *inòngi-tĭ-má-re* ***khò?***
take-SER do.what-DES-CONT-3SG.R I.wonder
'I wonder what he's (taking it and) trying to do with it?' (VD2 1:07-1:09)

The particle *khuní* is used when the speaker is expressing that they haven't the slightest clue about the answer. It is probably a combination of *kho* 'I wonder' and *ni* 'EMPH'. In (268), the interlocutors are naming and describing animals, but one speaker cannot think of a native Hurukui name for a recently introduced species.

(268) *Ino* *ko-kopu* ***khuní?*** *Opopo* *pai* *yó-mukhai-Ø.*
what say-1PL.FUT hell.if.I.know thing new come-3SG.PRF.R-3SG.R
'What on earth should we call it? It's a new thing.' (Abus 4:48-4:51)

The particle *o*, possibly a loan from Tok Pisin *o* 'or', is used to create a tag question, as if to add 'or is it not so?'. This is seen in (269).

(269) *Neka* *wo-pi* *khaingu*
2PL go-2PL.R TOPIC

nd-apono *khaingu* *ng-ike-pi* ***o?***
those.people-several this see-3PL.OBJ-2PL.R TAG.Q
'When y'all went, y'all saw those people too, didn't you?' (Dokta 3:14-3:19)

6.5.7 Polar answer particles

The polar answer particles are the words *eo* 'yes' and *pe* 'no'. The latter is homophonous with the negative existential particle *pe* (§6.5.1).

In Hurukui, polar questions are answered with *eo* 'yes' if the statement that has been interrogativized is true, and *pe* 'no' if it is false. That is, answering 'yes' to a negative question means that the negative proposition is correct.

In (270), the question and answer are both in the positive polarity.

(270) A: *Na* *hauhu* *e-ma* *khaingu* *nukhe-ma* *ngu?*
2SG game do-2SG.R that.one lose-2SG.R DISC.PRTC

B: ***Eo,*** *na* *puhi* *asakha* *a-ma-ma* *ngu.*
yes 2SG money waste put-CONT-2SG.R DISC.PRTC
A: '(If the horse you bet on loses), you lose the game you play?' (Hos 1:33-1:39)
B: 'Yes, you waste your money.'

In (271), the question is in the negative polarity, notably using the particle *pe* 'NEG.EXIST' which is homophonous with *pe* 'no'. However, because it is correct that speaker B does not have a small knife, he answers with *eo* 'yes'.

(271) A: *Na* *phai* *mbene* *pe?* B: ***Eo.***
2SG knife small NEG.EXIST yes
A: 'You don't have a small knife?' B: 'No, I don't.' (Kaikai 2:03-2:06)

In addition to answering polar questions, *eo* 'yes' and *pe* 'no' can be used to assent or object to something that someone says or does. *Eo* 'yes' is particularly commonly used for backchanneling.

6.6 Connectives

Clauses can be connected by *kha* 'and' and *nemo ~ namo* 'but'. In addition, the connective *o* 'or', a loan from Tok Pisin, is widely used, but this is viewed as less correct Hurukui due to a widespread language ideology disfavoring loanwords.

The connective *kha* 'and' is used only after medial verbs, not final verbs. It can be used after both same-subject and different-subject verbs. In (272), it is used twice.

(272) *Opo* *e-khu* *ngu* *hambe-Ø* *a-pia* ***kha,***
taro take-3DU.SS.R 3SG top-3SG.POSS put-3DU.DS and

re-ne, *re-mu* *sáhi-mu* *i-ne* ***kha,*** ...
burn-3SG.DS burn-3SG.SS.R do.well-3SG.SS.R do-3SG.DS and
'They put the taro on top of it, and it cooked well, and ...' (Huakono 2:31-2:39)

Why *kha* 'and' is used after some medial verbs and not others is a topic for future investigation.

The connective *nemo ~ namo* 'but' is used between clause chains, i.e., after final verbs or final nonverbal predicates. It is shown between the two clauses in (273).

(273) *Hongoro* *mukhí-mukhai-Ø* ***nemo*** *hongoro* *ome.*
lizard.sp resemble-3SG.PRF.R-3SG.R but lizard.sp is.not
'It looks like a *hongoro* lizard, but it's not.' (Chameleon 0:13-0:16)

6.7 Ideophones

Ideophones are words that do not occur on their own but can be used with the light verb *ko-* 'say' to create a predicate expressing some meaning involving sound or motion. Examples are *phih ~ phihi* 'knock', *phum* 'explode', and *khum* 'sprout'.

In (274), the predicate *phihi ko-* 'knock' uses the ideophone *phihi* 'knock'. This construction does not mean that anyone is saying "*phihi*" or making a knocking-like sound with their mouth, it means that they are actually knocking on an object even if they are not making any sound with their body.

(274) *Nangasa* *khaingu-kha* *hambe-Ø* *yani* ***phihi*** *ko-pu*
coconut.shell this-FOC top-3SG.POSS up.here knock say-1PL.SS.R

eko *i-puna* *we-mu* *épŭ_a-make-re.*
whatchamacallit do-1PL.DS go-3SG.SS.R suffice-CONT-3SG.R
'This coconut shell, we knock on the top of it until it's all right.' (Pot2 4:35-4:42)

The combination of an ideophone with *ko-* 'say' can also express predicates that have no sound. In (275), the predicate *khum ko-* 'sprout' uses the ideophone *khum* 'sprout'. To my knowledge, plants do not make audible sounds when sprouting[15].

(275) *Amburine* *kha* *pháipha-pina* *kha* ***khum*** *ko-ma-re.*
seed FOC spread-3PL.DS and sprout say-HAB-3SG.R
'They spread its seeds and it sprouts.' (Plan2 1:05-1:09)

Ideophones can be reduplicated to create pluractional meaning. Some of them undergo full reduplication, such as *phihi* 'knock', seen in (276).

15 But then again, I grew up in cities and have never farmed or gardened.

(276) *Arárě* *ake* ***phihi~phihi*** *ko-khu* *yo-ra.*
slow for PL~knock say-1SG.SS.R come-1SG.R
'I knocked on it lightly (multiple times) and came.' (elicited, consultants produced)

Some ideophones instead undergo reduplication with ablaut. The most common such pattern is that an ideophone containing the vowel *u* will use *i* in the reduplicated form. The ablauted reduplicant precedes the unaltered root. For instance, *khum* 'sprout' becomes *khim~khum*, shown in (277).

(277) ***Khim~khum*** *kó-mukhai-Ø.*
PL~sprout say-3SG.PRF.R-3SG.R
'They have sprouted.' (Pinat 0:07-0:10)

7 Syntax of the clause

This chapter describes the basics of Hurukui syntax at and below the level of a single clause. §7.1 describes some phrase types, but many more minor phrase types are omitted for the purposes of this sketch grammar. §7.2 describes nonverbal clauses, defined as those clauses that are not headed by a verb. §7.3 describes verbal clauses. §7.4 describes how interrogatives are formed, which is almost exclusively via intonation in the case of polar questions.

7.1 Phrases

In this section I briefly describe the structure of noun phrases (§7.1.1) and adjective phrases (§7.1.2). Instead of dedicating a section to verb phrases, I treat the structures of verbal clauses in §7.3.

7.1.1 Noun phrase

The following basic phrase structures most noun phrases in Hurukui. Each of them that starts with "NP > " provides one option, and within that, further optionality is represented by parentheses around optional items. Lower-level phrase structures, such as N', are also provided as needed. The head of each, if it is clear which constituent that is, is bolded. The symbol ^ means "and/or", requiring at least one of the options and preserving their order if more than one is present. The asterisk here means zero or more of the preceding constituent.

- N' > **N** (Adj | Q | PP | $\text{Clause}_{\text{nmzd}}$)*
- NP > **N'** (Deic)
- NP > **NProp** (PP) (Deic)
- NP > **Pro** (Q) (Deic)
- NP > [**Adj** | **Q** | **PP**] (Deic)
- NP > **Deic**
- NP > NameVoc ^ NameGiven
    - headedness unclear

- NP > NP_{possr} N'_{possd} (Deic)
    - This is the noun-noun possession construction (§7.1.1.1).
- NP > **N** NP_{mod} (Deic)
- NP > NP NP_{appos}
    - This is an appositive construction, which is double-headed.

There are surely more fine-grained co-occurrence restrictions that these rules do not account for. For instance, the various subtypes of pronouns (§3.2) likely do not have the same syntactic distribution. The ordering of different kinds of modifiers, such as adjectives, quantifiers, and PPs, is also not accounted for here. This section presents only the basic, coarse-grained syntactic facts, and further details of noun phrase syntax will have to await further investigation.

Noun compounds (§7.1.1.2) are composed of two noun roots, but they behave as single nouns in the phrase structure. As such, I do not separate the two parts of a noun compound (head and adnominal modifier) in the PSRs.

A noun phrase may consist solely of a nominal, such as a noun or pronoun, or a deictic. In (278), the noun *yoho* 'fire' is a noun phrase by itself, acting as the object of the first verb *a-* 'affect.3SG, hit.3SG, kill.3SG'.

(278) ***Yoho*** *a-pina* *umi-ne* *we-pínakhai-Ø.*
fire kill.3SG-3PL.DS die-3SG.DS go-3PL.DPST-3SG.R
'They put out the fire and went.' (San 0:46-0:51)

Nouns may be modified by adjective phrases, quantifiers, and/or deictics that follow them. In (279), the noun *pahe* 'house, village' is modified by the adjective *khiri* 'old (of objects)'.

(279) *We-khu* ***[pahe*** ***khiri]*** *we-khu* *isi-khu ...*
go-1DU.SS.R village old go-1DU.SS.R sleep-1DU.SS.R
'We two went and slept in the old village, and ...' (Hunt8 0:27-0:31)

In (280), the noun *anga* 'day, time' is quantified by the numeral *apono* 'three'.

(280) ***[Anga apono]*** *wo-re* *khaingu.*
day three go-3SG.R that
'Three days passed.' (BusHaus 4:21-4:23)

In (281), the noun phrase *puhi akharara khaino* 'this white stone' has the order [N Adj Deictic].

(281) *Ngu_para* ***[puhi akharara khaino]*** *nge*
and.then stone white this TOPIC

a-pina *phi~phéke-ne* *ngu_para ...*
hit.3SG-3PL.DS REP~break-3SG.DS and.then
'And this white stone, they break it into pieces, and ...' (Planim 5:01-5:07)

Pronouns can head NPs and can form entire NPs by themselves. In (282), the pronoun *neka* '2PL' is a noun phrase by itself, acting as the subject.

(282) ***Neka*** *ínake* *khaníngе-pi?*
2PL why do.thus-2PL.R
'Why did y'all do that?' (KilimPkn 2:18-2:22)

Pronouns can be modified by quantifiers (§6.2). In (283), the pronoun *neka* '2PL' is quantified by the numeral *apono* 'three'. This NP in turn is the possessor in a noun-noun possession construction (§7.1.1.1), the possessed being *ipeneka* 'y'all's names'.

(283) ***[Neka apono]*** *ipe-neka* *e-pu* *akakóya-pu.*
2PL three name-2PL.POSS take-1PL.SS.R lift.up-1PL.R
'We lift up the names of you three (the Holy Trinity).' (Kaikai 21:12-21:16)

Pronouns can also be modified by deictics (Chapter 4). In (284), the vocative pronoun *ndapono* 'you three', used in its 3PL referential sense (§3.2.3), is modified by the deictic *khaino* 'this'.

(284) *O* *mbénako* ***[ndapono*** ***khaino]*** *umo* *amburine*
oh sorry those.people this tree seed

kono *wa-ka* *para* *e-pi* *yo-mai-pi.*
mountain upland-there LOC take-3PL.SS.R come-CONT-3PL.R
'Oh what a shame, these people carry their coffee through the mountains.' (Kofi 5:13-5:19)

Pronouns cannot, however, be modified by adjective phrases.

Deictics can head NPs and can form entire NPs by themsleves. In (285), the deictic *khaino* 'this' is a noun phrase by itself, acting as the object of the verb *e-* 'take' (and, via zero anaphora, it is also the object of *a-* 'put' semantically but not syntactically).

(285) ***Khaino*** *e-khu* *ara* *woyo* *a-khu* *khanóngi-khu ...*
this take-1SG.SS.R place aside put-1SG.SS.R do.thus-1SG.SS.R
'I put this aside like so, and ...' (Purpur 0:43-0:47)

Adjective phrases may also act as NPs by themselves. In (286), the adjective *pai* 'new' is used as an NP meaning 'a new one'. It is describing a battery, which had been introduced earlier in the story context, so the noun is recoverable and thus omissible.

(286) ***Pai*** *é-kara* *ko-mu* *pahe* *wo-re.*
new take-1SG.FUT say-3SG.SS.R house go-3SG.R
'He wanted to get a new one and went home.' (Windua 0:39-0:44)

7.1.1.1 Noun-noun possession construction

An NP in which one nominal constituent possesses another takes the form [possessor possessed]. I call this the *noun-noun possession construction*. It consists simply of juxtaposition, with no morphosyntactic marking of the possession relation (unless the second noun is inalienable, §3.1.1). The possessum cannot itself be a noun-noun possession construction; equivalently, it cannot have a possessor NP in it. I account for this in the phrase structure by calling an unpossessed nominal constituent an N', and omitting possessors from the rule for N', so that the possession construction has structure [NP_{possr} N'_{possd}].

(287) exemplifies the noun-noun possessive construction, with the noun *wongu* '*kundu* drum' as the possessor and the noun *more-* 'mouth' as the possessed.

(287) *wongu* *more-Ø*
kundu.drum mouth-3SG.POSS
'the opening of the *kundu* drum' (Kundu 1:00-1:03)

The noun-noun possession construction is equally valid when the possessed is an inalienable noun (§3.1.1), except the possessor is optional if it is a personal pronoun. For instance, the phrase *ngu hambe-Ø* (3SG top-3SG.POSS) 'its top, above it' can also be expressed as just *hambe-Ø* (top-3SG.POSS), but the phrase *Martin ipo-e* (Martin father-3SG.POSS) 'Martin's father' cannot omit mention of Martin unless he is topical and recoverable from context.

7.1.1.2 Noun compounds

Noun compounds consist of two nouns juxtaposed, used as a single nominal. Usually, one of the nouns is the head, and the other acts more like an adnominal modifier. An example is *opopo ona* (food bone) 'starchy food', in which the head noun is *opopo* 'food', and the other noun *ona* 'bone' (an inalienable noun seen in its zero-marked generic form, §3.1.1) is used more like a modifier. The phrase refers to a kind of food, not a kind of bone. There are a few types of noun compound, which I call *possessive-like*, *associative*, *hypernym-hyponym*, and *symmetrical*.

In *possessive-like compounds*, the modifier noun is first and the head noun is second. Their semantic relationship resembles possession of the head by the modifier. Examples include *umo amburine* (tree seed) 'fruit, coffee bean, medicine', *ihe hoisi* (pig grease) 'lard', and *mbuma hunui* (white.person kunai.grass) 'tin roofing'[16]. However, these compounds are not the same as the noun-noun possession construction (§7.1.1.1). The main reason for this is non-referentiality of the modifier noun, whereas a possessor is referential. That is, *ihe* 'pig' in *ihe hoisi* (pig grease) 'lard' does not refer to any particular pig(s), but instead is specifying the kind of grease/fat. The semantic relationship is often more like association with the modifier, rather than outright possession.

In *associative compounds*, like in some possessive-like compounds, the semantic relationship is one of association with the modifier, but the modifier instead occurs in

16 Kunai grass is a very commonly used roofing material, so the name of the grass has been generalized to mean 'roof'.

postnominal position like adjectives do (§7.1.1). An example is *hamùnge páhe* (man house) 'villager'.

In *hypernym-hyponym compounds*, the first noun is a hypernym of the second, and the whole compound denotes the smaller subset of denotata (i.e., that of the hyponym). Examples include *mbo itepa* (small.animal bandicoot) 'bandicoot' and *umo mbemba* (tree tree.sp) 'tree.sp'.

In *symmetrical compounds*, neither noun is more head-like than the other. These are often used to refer to a collective by juxtaposing the names for two of its subsets, such as *hamunge haikhu* (man woman) 'people', *ihe mbo* (pig small.animal) 'various animals', or *ipe nayau* 'sorcery', in which each of the nouns refers to a different subtype of sorcery.

Intonation often helps to identify the head of a noun-noun compound, which do not have a consistent ordering of head and modifier. The head is intonationally unmarked (lower pitch) while the modifier is marked (higher pitch). Table 7.1 shows some examples. The high-pitch modifiers are bolded.

Table 7.1: Headedness and intonation in noun compounds

Hurukui	**gloss**	**meaning**	**compound type**
hamùngě ***páhe***	man **house**	'villager'	associative
hamúngě *pàhe*	**man** house	'men's house'	possessive-like
mbò ***itépa***	animal **bandicoot**	'bandicoot'	hypernym-hyponym
nòhě ***wayómu***	bamboo **bamboo.sp**	'strong bamboo'	hypernym-hyponym
ùmŏ ***mbémba***	tree **tree.sp**	'tree.sp'	hypernym-hyponym
úmŏ *mbò*	**tree** animal	'tree kangaroo'	possessive-like
mbúmă *hunùi*	**white.person** kunai.grass	'tin roofing'	possessive-like
pàhě ***mbangúnu***	house **shade**	'church'	associative

The use of higher pitch / primary stress to mark the modifier noun is especially reliable when the head is a hypernym of the modifier, as in *mbò* ***itépa*** 'bandicoot' which is expressed as 'small.animal **bandicoot**'. This makes sense from the perspective of focus and informativity. The more specific information is more desirable as a clarification. If you know I'm talking about a

bandicoot, you also know that it's a small animal, but if you only know I'm talking about a small animal, you may wonder what kind.

However, some compounds instead have higher pitch on this head noun, the reverse of the normal pattern. These are possessive-like compounds which consist of high-frequency collocations, so they are likely more lexicalized. Examples include *ùmŏ* ***amburíne*** (tree **seed**) 'fruit, coffee bean, medicine', *ùmŏ* ***arakúi*** (tree **root**) 'root', and *òră* ***amburíne*** (grass **seed**) 'peanut'.

Noun compounds are used as nominals. They can be quantified, e.g., with the indefinite quantifier *ahe* 'INDEF.SG' (§6.2.3). This is shown in (288).

(288) ***Opopo*** ***ona-Ø*** *ahe* *nge* *é-kara* *ko-mu ...*
food bone-GENER INDEF TOPIC take-1SG.FUT say-3SG.SS.R
'That person wanted to get some root vegetables, and ...' (Fluffy 4:36-4:40)

7.1.2 Adjective phrase

Adjective phrases are headed by an adjective (§6.1) and can have the following structures:

- AdjP > (PP) Adj (Intensifier)
- AdjP > *awohe* Adj

In (289), the object is *ngu* '3SG', referring to manufactured goods. The verb *sáhi-* 'create' is modified by an adjective phrase containing an intensifier. The adjective phrase has resultative meaning.

(289) *Ngu* ***[ambirou*** ***wohe]*** *sáhi-pina ...*
3SG weak very create-3PL.DS
'They make it very weak, and ...' (Komplen 3:38-3:42)

The intensifier *awohe* (§6.5.4) is the only ad-adjectival modifier that precedes the head adjective. In (290), the noun *umo* 'tree' is modified by an adjective phrase containing this preposed intensifier.

(290) *Umo* ***[awohe*** ***pite~pite]*** *a-khukhai-pa* *ngae* *ngáhe-re.*
tree too.extent big~PL put-1SG.PRF.R-1SG.DS thus be.thus-3SG.R
'I've put trees that are too big, so it's like this.' (Trap 6:08-6:12)

In (291), the adjective *harae* 'same, similar' is modified by a postpositional phrase. The PP and the AdjP containing it are both bracketed.

(291) *Na* *numu-na* ***[[hia*** ***numu-yo*** ***khahu]***$_{\textbf{PP}}$
2SG body-2SG.POSS 1DU.EXCL body-1DU.POSS COMIT

harae]$_{\textbf{AdjP}}$ *ì-mukhái-wa-Ø.*
same become-3SG.PRF.R-NEG-3SG.R
'Your body is not the same as ours.' (Chameleon 4:04-4:07)

So far, no adjective phrases are attested with the structure [PP Adj Intensifier] or [*awohe* Adj Intensifier]. Further details of this phrase type are a topic for future investigation.

7.2 Nonverbal clauses

In Hurukui, there are two types of nonverbal clause. The first, which I call *copular clauses* (§7.2.1), involve a nonverbal predicate which is a noun phrase, adjective phrase, deictic, or other contentful constituent. The other type, *negative existential clauses* (§7.2.2), have a predicate headed by the negative existential particle *pe* (§6.5.1).

All other clauses are verbal, meaning that they are headed by verbs. This includes positive-polarity existential predicates, which use the verb *khai-* 'be located, exist, stay, live'. Verbal clauses are treated in §7.3.

7.2.1 Copular clauses

Nominal and adjectival predicates are often expressed by a zero-copula construction. In (292), the noun *khiwinu* (a kind of plant) is the predicate.

(292) *Kohe,* *khaino* *ngu* ***khiwinu****.*
okay this TOPIC plant.sp
'Okay, this is *khiwinu*.' (Plan2 2:49-2:53)

In (293), the adjective *khomorae* 'short' is the predicate.

(293) *Akine-Ø* ***khomorae***.
thought-3SG.POSS short
'She has a bad memory.' (lit. 'Her thought/intellect is short.') (Shop1 4:21-4:24)

In addition to the zero-copula construction, nominal and adjectival predicates can be expressed as arguments to the verb root *(h)i-* 'be'. Even though clauses using this copular verb are thus not nonverbal, I discuss them here due to their relationship to the nonverbal clauses, rather than in §7.3 with the rest of the verbal clauses.

In (294), the predicate *mbangiri maisa* 'too strong' is marked for perfect aspect. The semantic contribution of this verb, as opposed to simply using the zero-copula predicate, will require further research.

(294) *Khaingu umo mbangiri maisa* ***í-mukhai-Ø***.
that.one also strong bad be-3SG.PRF.R-3SG.R
'That one's also too strong.' (Trap 2:05-2:08)

In some situations, the semantics of the TAM-marked copula is more like 'become' rather than the stative 'be'. In (295), the adjectival predicate *maisa* 'bad' is marked for continuative aspect with semantics of habitual becoming.

(295) *Maisa* ***i-ma-re***.
bad become-HAB-3SG.R
'It tends to deteriorate.' (Komplen 3:59-4:03)

The copula *(h)i-* 'be' is also used even in some situations where the TAM is realis with unmarked aspect. The reason for this will require further research. In (296), the adjectival predicate *khomou* 'short' is accompanied by a realis copula.

(296) *Na khomou* ***e-mă*** *eko.*
2SG short be-2SG.R near.you
'You're short.' (ObjP2 0:52-0:55)

In counterfactuals, the copular *(h)i-* 'be' does not inflect for person, instead taking an invariant morphological subject of 3SG. In (297), the subject of the clause meaning 'if I were a woman' is clearly shown to be 1SG from the presence of the pronoun *ima* '1SG' and from the translation, but the copula marks 3SG subject.

(297) *Ima* *hioho* ***ĭ-tomu*** *ngu,* *aukhu* *oke-o.*
1SG woman be-3SG.FUT TOPIC skirt tie-1SG.POT
'If I were a woman, I would wear a skirt.' (elicited)

Negative copular clauses use the particle *ome* 'is not' (§6.5.1) at the end of the clause in the realis mood with unmarked aspect. This kind of clause is nonverbal. In (298), the noun phrase *na pahe* 'your house' is the predicate.

(298) *Khaino* *no* *na* *pahe* ***ome.***
this here 2SG house is.not
'This isn't your house.' (Sik 1:39-1:41)

In TAM categories besides realis with unmarked aspect, negative copular predicates replace *ome* 'is not' with the auxiliary verb *(h)i-* 'be' marked for negative polarity. In (299), the nonverbal predicate *humaya* 'good' is in the negative and marked for continuative aspect.

(299) *Opitohu* *wo* *a-pua* ***humaya*** ***ĭ-mai-wa-Ø.***
food:1DU.POSS garden put-1DU.DS good be-HAB-NEG-3SG.R
'We two plant our food and it generally is not good.' (Mun 3:21-3:25)

7.2.2 Negative existential clauses

Negative existential clauses use the particle *pe* 'NEG.EXIST' (§6.5.1) at the end of the clause in the realis mood with unmarked aspect. In (300), the subject is *uraho* 'fire' (Arapipi dialect).

(300) *Uraho* ***pe.***
fire NEG.EXIST
'There was no fire'. (Tamiok2 1:58-2:02)

In TAM categories besides realis with unmarked aspect, negative existential predicates add the auxiliary verb *(h)i-* 'be' after *pe* 'NEG.EXIST'. In this situation the initial *h* is always retained, so one never gets **pé i-*. (301) shows such a predicate, marked for future tense, with conditional meaning.

(301) *Kohe,* *puhi* ***pé*** ***hi-tomu*** *nge,*
okay money NEG.EXIST be-3SG.FUT TOPIC

ai!, *maha* *para* *aráya-kanu* *ni.*
oh! ground LOC wander-2SG.FUT EMPH
'However, if you don't have money, you'll have to go on foot!' (Lae 2:20-2:27)

7.3 Verbal clauses

Verbal clauses have the basic constituent order of SOV. In (302), the subject is *ima* '1SG', the object is *akaime* 'mother', and the verb *ngáa-* 'tell, talk to' is at the end.

(302) *Ima* *nge* *akaime* *ngáa-Ø-ra.*
1SG TOPIC mother tell-3SG.OBJ-1SG.R
'I told Mom.' (Gaden 0:54-0:57)

Existential clauses in the positive polarity are a type of verbal clause, in contrast to negative existential clauses (§7.2.2). Positive existential clauses are headed by the verb *khai-* 'be located, exist'. A simple example is shown in (303).

(303) *Moko* *ahe* ***khai****-Ø?*
speech another exist-3SG.R
'Is there anything else to say?' (Graun 15:04-15:07)

When the subject of an existential clause (the theme which is said to exist) has a possessive structure (either alienable or inalienable), these clauses can also be read as predicating possession. In (304), the inalienable noun *orape-* 'skin, body' is marked for 1SG possessor, so the clause can be read either as existential or possessive.

(304) *Orape-o* *khuha* ***khai-Ø****.*
body-1SG.POSS all exist-3SG.R
'My whole body exists.' / 'I have a whole body.' (Bodi 0:16-0:19)

This section starts with a discussion of the definitions and concepts related to argument structure that I employ in my description of clause structure and verbal valence. §7.3.1 defines the distinction between *core*, *oblique*, and *bare* participants/NPs in a verb's frame. §7.3.2 describes the *definite null test*, a way of determining whether certain NPs are core or bare and consequently whether a verb is transitive. §7.3.3 defines the distinction between *argument* and *adjunct*. §§7.3.4,

7.3.5, and 7.3.6 illustrate the three types of participants/NPs that occur in Hurukui based on the preceding definitions: arguments, oblique adjuncts, and bare adjuncts, respectively. §7.3.8 describes *experiencer predicates*, a clause type rather commonly found in the languages of New Guinea. §7.3.9 describes where adverbial modifiers occur in clauses. §7.3.11 introduces *light verb constructions* (LVCs), from which a large portion of the Hurukui verbal lexicon is built.

7.3.1 Core, oblique, and bare

I define *core* and *oblique* within the specific morphosyntactic system of Hurukui, without regard to how these terms have been used to describe other languages (albeit inspired by them). As discussed in Chapter 5, Hurukui verbs have polypersonal agreement, indexing their subjects and objects with affixation or root fusion/suppletion. I call a referent and its overt NP (if there is one) "core" if it is indexed on the verb in one of these ways.

In (305), both the subject and object of the verb *ngáe-* 'tell, talk to' are marked by affixes on the verb. They are also present as overt pronouns. I thus refer to the 3SG and 3PL referents (the semantic participants), as well as the NPs *ngu* '3SG' and *noka* '3PL' (the syntactic constituents), as "core".

(305) ***Ngu*** ***noka*** *ngàě-**iki**-wáu-**mu**.*
3SG 3PL tell-3PL.OBJ-NEG-3SG.FUT
'That person will not tell them / talk to them.' (elicited)

In (306), both clauses consist only of a verb, with no overt NPs. The first clause, with the verb *ngáe-* 'tell', has a 1SG object suffix in addition to its subject-marking switch-reference suffix *-pina* '3SG.DS'. It thus has two core participants, but there are no core NPs.

(306) *Ngáe-**ya**-**pina*** *epha-ma-ra.*
tell-1SG.OBJ-3PL.DS hear-HAB-1SG.R
'They tell me / talk to me, and I listen.' (Nadzab 0:43-0:47)

Hurukui does not allow more than one object to be indexed on the verb, as discussed in §5.2.2. As such, Hurukui verbs cannot have more than two arguments. It is a question for future research whether verb forms that do not inflect for subject, such as the subordinating suffix marking purposive meaning (§5.5.1), should be considered to have a morphological subject.

As discussed in §6.4, Hurukui uses adpositions (mostly postpositions) for various purposes. I call a referent and its overt NP *oblique* if it is the argument of an adposition[17] and the adpositional phrase is adverbial. It is thus not possible to have an oblique participant without an overt NP, because then there would be nothing to mark with an adposition. In (307), the pronoun *nia* '2DU' is marked as oblique by the comitative postposition *khahu*, and this postpositional phrase modifies the verb *háuhi-* 'play'.

(307) *Ngae* ***[nia*** ***khahu]*** *hàuhi-wá-o.*
thus 2DU COMIT play-NEG-1SG.FUT
'So, I can't play with y'all two.' (SickGirl 1:23-1:28)

These two criteria (core = indexed on verb, oblique = marked by adposition) are mutually exclusive. Hurukui does not mark the same participant/NP in both ways at the same time. For instance, in (307), the oblique-marked pronoun *nia* '2DU' is not indexed on the verb. However, NPs are not required to be either core or oblique. Many NPs in a verb's frame are not marked in either of these ways. I refer to such NPs as *bare*.

In (308), the bare noun *eya* 'stinging nettle' is used in a semantic role of instrument, not patient. This is clear from context because they are using the stinging nettle to 'cook' their legs (i.e., cause a burning sensation) to reduce muscle pain.

(308) ***Eya*** *uki-topu* *ni.*
stinging.nettle cook-1PL.FUT EMPH
'Let's rub stinging nettles on our legs.' (Kofi 1:55-1:57)

As described in §§5.2.2.1-5.2.2.2, the object suffixes and reciprocal suffix can function to add a beneficiary participant. (112), reproduced here as (309), shows the object markers *-neka* '2PL.OBJ' and *-(i)ki* '23PL.OBJ' optionally used for this purpose. However, the object-unmarked verb is also acceptable here because the overt pronoun *neka* '2PL' is present.

17 I do not consider the contrastive focus particle *kha* ~ *-kha* (§6.5.3) to be an adposition syntactically, even though it also often has instrumental meaning.

(309) *Ima-kha* *neka* *pahe*
1SG-FOC 2PL house

[yai-ra / *yai-**neka**-ra* / *yai-**ke**-ra].*
build-1SG.R build-2PL.OBJ-1SG.R build-2PL.OBJ-1SG.R
'I built y'all a house.' (elicited)

When the 2PL participant is indexed by verb morphology (i.e., the verb form is *yai-neka-ra / yai-ke-ra* 'I built (it) for y'all'), it is core. When the 2PL participant is not indexed by verb morphology (i.e., the verb form is *yai-ra* 'I built (it)'), the structure is ambiguous. One possibility is that the pronoun *neka* '2PL' is part of the verb's frame, thus being a bare NP with benefactive meaning. The other possibility is that the phrase *neka pahe* (2PL house) is a noun-noun possession construction (§7.1.1.1) meaning 'y'all's house'. Syntactic tests to distinguish between these will require further research.

7.3.2 The definite null test for covert 3SG objects

An analytical issue arises with the distinction between core and bare NPs when one considers that 3SG object marking is almost always zero (§5.2.2), the only exceptions being the fusional object-marked verb roots for 'affect' (§5.2.2.3) and 'give' (§5.2.2.4). As a result of this, it is often uncertain whether non-oblique NPs with 3SG animate or third-person inanimate reference should be considered objects.

(310) shows a very common structure found in Hurukui. It consists of two clauses, each with a bare NP and a verb that marks only 1SG subject and has no overt object marker. In the first clause, headed by the verb *e-* 'take', the NP *mburumu ikhunu* 'sweet potato stem' has the semantic role of theme, the thing taken. In the second clause, headed by the verb *a-* 'put', the NP *wo* 'garden' has the semantic role of location or goal, the place where something was put. The thing being put (the theme) is the sweet potato stem mentioned in the first clause, but no overt markers refer to it in the second clause.

(310) ***[Mburumu*** ***ikhunu]*** *e-khu* ***[wo]*** *a-khu ...*
sweet.potato stem take-1SG.SS.R garden put-1SG.SS.R
'I took a sweet potato stem and planted it and ...' (ManDai 0:32-0:36)

The overt NP *mburumu ikhunu* 'sweet potato stem' may be an object that is zero-marked on the verb, or it may be something else. Should the verbs *e-* 'take' and *a-* 'put' be considered to have objects or not? One test that can be used for this is something I call the *definite null test*. A verb passes the definite null test if, when it has no overt object NP and no overt object marking, it must nonetheless be interpreted as having a semantic object that refers (i.e., the object is 'that person' or 'it') rather than being unspecified for such a participant.

An example of a verb that passes the definite null test is *i-* 'eat', whose object-unmarked form *ire-re* (eat-3SG.R) must be interpreted as 'that person ate **it**', with the thing eaten recoverable from context. If one wants to express the equivalent of English 'that person ate', with no specification as to what was eaten, one needs a dummy object NP *opopo* 'food, thing': *opopo ire-re* (food eat-3SG.R) 'that person ate (food / something)'. Thus, even though the verb *i-* 'eat' often appears with no marked object and no overt object NP (and often with no other words in the clause at all), I consider it to pass the definite null test.

Applying this test to *e-* 'take' and *a-* 'put' from (310), it can be seen that clauses whose sole word is one of these verbs still have a definite referent for the participant in the semantic role of theme. For instance, *o-ra* (take-1SG.R) means 'I took **it**', not 'I took (something / stuff)', and *a-pi* (put-3PL.R) means 'they put **it** (somewhere)', which is unspecified for the goal of motion but not for the theme. Because both of these verbs always have definite reference for the theme when they lack an overt object NP or suffix, they both pass the definite null test.

Further evidence for this conclusion comes from uses of these verbs in which the theme is animate, triggering object agreement. In (311), a group of people is the object of both the verb *e-* 'take' and the verb *a-* 'put', shown by the object suffix *-(i)ki* '23PL.OBJ'.

(311) *Ĕ-**iki**-mu* *hitae~hitae* *nenge* *a-**ki**-mu* *wo-re.*
take-3PL.OBJ-3SG.SS.R 3SG.EMPH~REP village put-3PL.OBJ-3SG.SS.R go-3SG.R
'It (helicopter) took them and put them in each of their villages.' (Bibuai 8:22-8:25)

From this kind of evidence, assuming the verb uses the same mapping from participant roles to morphosyntactic marking in different contexts, we can draw conclusions about the roles of NPs in examples with a more ambiguous surface structure. Applying this reasoning to (310), the NP *mburumu ikhunu* 'sweet potato stem' is the object of both *e-* 'take' and *a-* 'put', and so *wo* 'garden' is an adjunct, not an object.

7.3.3 Argument and adjunct

I use the definition of an *argument* of a verb as a participant that must be expressed in the clause headed by a given verb. In contrast to this, an *adjunct* of a verb is a participant that is not required to be expressed in the clause headed by that verb. This says nothing about how involved the participant is in the event, or even if it is a participant at all. There are impersonal "dummy" subjects in various constructions in Hurukui, realized as 3SG subject marking on the verb but not referring to anything at all, and I still consider these to be arguments of their verbs. Conversely, something that is heavily involved in the event, or without which the event would not even be possible, are not guaranteed to be required by a given verb's syntax. When such an NP is not required, I consider it an adjunct. Because Hurukui verb agreement marks the person and number of the subject, there is no such thing as a Hurukui verb that has zero arguments. There is always at least one: the subject.

I consider a Hurukui verb to be transitive if it passes the definite null test when it has no overt object marker (§7.3.2). For instance, *i-* 'eat' has definite null object whenever it lacks an object NP or object suffix[18], so I consider it strictly transitive, not ambitransitive.

I consider a Hurukui verb to be intransitive if it cannot take overt object NPs nor object marking (including definite null 3SG) for the roles of patient, theme, or recipient. For instance, *yo-* 'come' cannot ever take the object or reciprocal suffixes, and there is no definite null object

18 In fact, *i-* 'eat' cannot take any object markers in any context, even when the things being eaten are animate. It is unusual in this respect.

referred to (i.e., it does not mean 'come to that person / it' or similar). Thus, *yo-* 'come' is strictly intransitive.

If a verb has the ability to mark an object as core using object suffixes (including definite null 3SG) or the reciprocal, as well as the ability to lack object marking altogether (meaning that the lack of object marker is **not** interpreted as definite null in that context), then I consider the verb to be *ambitransitive*.

The following pair of examples illustrates well the line of reasoning that I use to determine the valence of a verb. In (312), none of the the nominals *ima* '1SG', *hengemu* 'fork', or *opopo* 'food' are marked by an adposition. Their roles are agent, instrument, and patient/theme, respectively, of the verb *pi-* 'stab' that heads the first clause. Only the agent *ima* '1SG' is overtly indexed on the verb. This is done by the subject marker *-ra* '1SG.R', so *ima* '1SG' is the subject and is an argument. The instrument, *hengemu* 'fork', can optionally be followed by the focus marker *kha ~ -kha* (§6.5.3), but this is often omitted and is not syntactically a postposition anyway. The patient/theme, *opopo* 'food', cannot be marked by a postposition and immediately precedes the verb.

(312) ***Ima*** ***hengemu*** ***opopo*** *pi-khu* *no-ra.*
1SG fork food stab-1SG.SS.R give.3SG-1SG.R
'I gave that person food using the fork.' (lit. 'I stabbed the food$_i$ with the fork$_j$ and gave it$_i$ to them$_k$.') (elicited, consultants produced first)

(313) is the same sentence but without the NP *opopo* 'food'. The verb *pi-* 'stab' still has no overt marker to index this participant, but there is a definite null object. There was a particular object that was stabbed with the fork and then given to the recipient. This example cannot have an antipassive-like meaning in which the unmentioned patient/theme is generic or unknown.

(313) ***Ima*** ***hengemu*** *pi-khu* *no-ra.*
1SG fork stab-1SG.SS.R give.3SG-1SG.R
'I gave it to that person with the fork.' (lit. 'I stabbed (it$_i$) with the fork$_j$ and gave it$_i$ to them$_k$.') (elicited, I produced first, consultants accepted)

Thus, *pi-* 'stab' passes the definite null test and has a core object: the patient (the thing stabbed), not the instrument. The verb *ne-* 'give.3SG' presents a more complicated case. Its root shape fusionally marks 3SG recipient (§5.2.2.4), so the recipient is core. However, this verb also

passes the definite null test for a zero-marked theme (the thing given, coreferent with the patient of *pi-* 'stab'), because it means 'I gave **it** to them'. One could thus consider the recipient as the primary object and the theme as the secondary object. But the 'give' verbs never overtly mark such a secondary object via affixation, nor do any other semantically ditransitive verbs such as *hupúri-* 'show'. It is analytically strange to posit that the language has morphologically ditransitive verbs but that these always index one of their arguments with a zero marker. As a result, I instead refer to such verbs as *quasi-ditransitive*, because they pass the definite null test for a second object (the T) in addition to marking their primary object (the R) via affixation, but never allow overt marking of both.

For stems created by the valence-increasing uses of the object suffixes (§5.2.2.1) and the reciprocal suffix (§5.2.2.2) which add a beneficiary argument, I treat the argument structure separately from that of the root to which they attach. I consider these to derive new verb stems with increased valence, and so these uses do not bear on the transitivity of the root/lexeme.

The definitions of core/oblique/bare and argument/adjunct adopted here lead to a simple categorization of participants and the NPs that refer to them. While many options are conceivable, such as oblique arguments (in which a verb requires a participant to be present as well as marked by an adposition, as found in various languages including English), only three of the possibilities are used in Hurukui. Table 7.2 lists these and their defining features. Verb indexation includes the zero marker *-Ø* '3SG.OBJ' for definite null objects.

Table 7.2: Arguments, oblique adjuncts, and bare adjuncts

	indexed on verb	**marked by adposition**	**required**
(core) argument	yes	no	yes
oblique adjunct	no	yes	no
bare adjunct	no	no	no

These types are illustrated in §§7.3.4, 7.3.5, and 7.3.6, respectively.

7.3.4 Arguments

Based on the discussion in §§7.3.1-7.3.3, all arguments are core and all core NPs are arguments. In (314), the two pronouns *ngu* '3SG' and *nia* '2DU' are core because they are indexed by suffixes on the verb. The suffix *-re* '3SG.R' agrees with the subject *ngu* '3SG', and the suffix *-nia* '2DU.OBJ' agrees with the object *nia* '2DU'.

(314) ***Ngu*** ***nia*** *nge-**nia**-**re**.*
3SG 2DU see-2DU.OBJ-3SG.R
'That person sees y'all two.' (elicited)

Semantically nonsingular inanimate referents are generally treated as 3SG by verb indexation. In subject marking, there is a bit of gradience along the scale from human/animal referents that require 3DU/3PL subject agreement according to their number, to very inanimate referents that require 3SG agreement regardless of their number. In the middle are somewhat-animate referents that optionally trigger number agreement in subject markers, but data on this is inconsistent and largely reliant on elicitation. As such, it will require further research for more solid conclusions to be drawn. Nonetheless, I show some examples of this gradience here.

In (315) and (316), the animate nouns *hamunge* 'man' and *ipiha* 'fish' are indexed by 3PL subject marking on the verb.

(315) *Ara* *hamunge* *khuha* *khai-**pi*** *ro-ka ...*
place man all be.located-3PL.R downland-there
'The place where all the men are, ...' (Nadzab 6:43-6:46)

(316) *Ipiha* *khai-**pi**.*
fish exist-3PL.R
'There are fish.' (elicited)

In (317) and (318), the noun *umo* 'tree' only marginally triggers 3PL subject marking in elicitation, and instead triggers 3SG subject marking in natural speech.

(317) *? Umo* *khuha* *khai-**pi**.*
tree all be.located-3PL.R
'All the trees are there.' (elicited)

(318) *Amie* *ya-ka* *khai-**Ø**.*
many up-there be.located-3SG.R
'There are many (trees) up there.' (BusHaus 0:37-0:41)

In object marking, this gradience in number agreement is not observed. Instead, the use of 3SG object marking for inanimate referents is strictly required. In (319), the NP *yoho khekhaingu* 'these fires' is explicitly marked nonsingular by the prefix *khe-* on the deictic (§4.4), but the verbs of which this referent is the object have no overt object marker. I explicitly parse and gloss -*Ø* '3SG.OBJ' in this example for clarity.

(319) *Nohe* *naure-mu* *yoho* ***khe**-khaingu* *naure-**Ø**-mu*
bamboo carry-3SG.SS.R fire PL-this carry-3SG.OBJ-3SG.SS.R

*e-**Ø**-mu* *kha* *yó-nekhai-Ø.*
take-3SG.OBJ-3SG.SS.R and come-3SG.DPST-3SG.R
'He carried bamboo and he carried these fires and brought them.' (Huakono 8:18-8:23)

We can tell that the verb *nge-* 'see' used in (314) is strictly transitive by using the definite null test. In (320), Speaker A is showing a particular picture to Speaker B and asking if she sees it. Both verbs have a covert object referring to the picture, and cannot be interpreted as meaning seeing in general.

(320) A: ***Ngo**-ma?* B: *Mm,* ***ngé**-khŭkhày-a.*
see-2SG.R yes see-1SG.PRF.R-1SG.R
A: 'Do you see it?' B: 'Mhm, I see it.' (Chameleon 0:05-0:07)

(320) also shows that we can have arguments which are not realized by an overt NP but are only shown by verb indexation (here, the 2SG and 1SG subjects of the verbs). We can also have arguments which are not realized by any linguistic form at all (here, the zero-marked 3SG object of both verbs), but we know it is there because of the definite null test.

Clauses headed by 'give' verbs (§5.2.2.4) can only have up to two overt NP arguments, even though the 'give' verbs are quasi-ditransitive (§7.3.3). When all of the agent, recipient, and theme are expressed by overt NPs, the theme must be the object of a verb (usually *e-* 'take') in an earlier clause. Then the recipient is the primary object (fused into the choice of root) of the 'give' verb in a later clause, and the theme is a definite null. This is shown in (321), where the first clause

is headed by *e-* 'take' and has agent *haikhu* 'woman' and theme *opopo* 'food'. Then a later clause expresses the recipient *kokopenoka* 'ghost' as the object of *ne-* 'give to 3SG'.

(321) ***Haikhu***$_{\mathbf{A}}$ *araiku* ***opopo***$_{\mathbf{T}}$ *e-****Ø***$_{\mathbf{T}}$*-****mu***$_{\mathbf{A}}$ *yo-****mu***$_{\mathbf{A}}$
woman earlier food take-3SG.OBJ-3SG.SS.R come-3SG.SS.R

kokope-noka$_{\mathbf{R}}$ ***no***$_{\mathbf{R}}$*-****re***$_{\mathbf{A}}$.
ghost-GENER give.3SG-3SG.R
'Earlier a woman brought some food and gave it to the ghost.' (Lakka 5:51-5:57)

7.3.5 Oblique adjuncts

In many situations, adjuncts must be marked as oblique using postpositions. NPs in the role of source of motion cannot be bare. In (322), the noun *kui* 'bag' is the source from which the stone was taken. This noun must occur in a PP headed by *para* 'from'.

(322) *Puhi* ***kui*** *[****para*** */ *Ø]* *e-khu* *khimi-pa* *wo-re.*
stone bag from take-1SG.SS.R throw-1SG.DS go-3SG.R
'I took the stone from the bag and threw it (stone) away.' (elicited)

Adjuncts have freer placement within the clause. In (323), the same sentence from (322) is seen with the source oblique fronted.

(323) ***Kui*** *[****para*** */ *Ø]* *puhi* *e-khu* *khimi-pa* *wo-re.*
bag from stone take-1SG.SS.R throw-1SG.DS go-3SG.R
'I took the stone from the bag and threw it (stone) away.' (elicited)

In (324), the pronoun *na* '2SG' in the role of comitative must be marked with the postposition *khahu* 'COMIT'.

(324) *Yakham* ***ná*** *[****khahu*** */ *Ø]* *khai-Ø?*
Yakam.given.name 2SG COMIT be.located-3SG.R
'Is Yakam with you?' (elicited)

In (325), the pronoun *na* '2SG' in the role of beneficiary must be marked with the postposition *ake* 'for'. Without this, the sequence *na kanapa* is instead interpreted as 'your garden' because it forms a noun-noun possession construction (§7.1.1.1).

(325) ***Ná*** ***ake*** *kanapa* *wo-ra.*
2SG for garden go-1SG.R
'I went to the garden for you.' (elicited, consultants produced first)

7.3.6 Bare adjuncts

Bare adjuncts are frequent in Hurukui, especially in the roles of location and goal. In (326), the noun *khapia* 'bed' is a bare adjunct in the role of location.

(326) ***Khapia*** *náma-mu ...* *humekhe* *sono* *koko* *e-mu ...*
bed sit-3SG.SS.R container water COMIT take-3SG.SS.R
'He sits on a bed, takes a water bottle, and ...' (VD1 1:49-1:53)

In (327), the noun *kui* 'bag' is used as a goal adjunct to the verb *ekha-* 'go down'.

(327) *E* *a-mana* ***kui*** *ekha-komu.*
take.SER put-2SG.DS bag go.down-3SG.FUT
'You will put it into the bag.' (Kimbe 2:59-3:02)

With motion verbs, only the goal can be bare. Any source or path adjuncts must be marked with the postposition *para* 'LOC, from'. In (328), the noun *pahe* 'house' can only be interpreted as a goal of motion, not a source or path.

(328) *E-pi* ***pahe*** *yo-mai-pi.*
take-3PL.SS.R house come-HAB-3PL.R
'They bring it home.' (BGPW 2:04-2:08)

In (329), the noun *khinanga* 'stairs, ladder' is the path of motion, marked as oblique by *para* 'LOC'.

(329) ***Khinanga*** ***para*** *yo-mu* *nenge* *yó-kara* *ko-mu ...*
stairs LOC come-3SG.SS.R outside come-1SG.FUT say-3SG.SS.R
'She came via the stairs in order to come outside, and ...' (SickGirl 0:48-0:52)

In (330), the place name *Mbawa* is the source, marked as oblique by *para* 'from'.

(330) ***Mbawa*** ***para*** *we-pi* *Rumpaya* *we-pi* *kha ...*
Bawa.village from go-3PL.SS.R Rumpaya.village go-3PL.SS.R and
'They went from Bawa to Rumpaya, and ...' (Kofi2 1:19-1:25)

The definite null test is useful for distinguishing between objects and bare adjuncts. For instance, the clause *pahe wo-re* (house go-3SG.R) means 'that person went home', and the noun *pahe* 'house' is non-oblique and not overtly indexed on the verb. So is it an argument, with a corresponding *-Ø* '3SG.OBJ' on the verb, or an adjunct? Looking at other uses of the verb can help resolve this. The verb *wo-re* (go-3SG.R) can also be a clause by itself, meaning 'that person went'. Notably, this does not mean 'that person went [there / to it]'. The goal of motion could be anything/anywhere and does not have to be something recoverable from context. Therefore, there is no definite null object marker on *wore* 'that person went'. Due to the absence of uses of *we-* 'go' with overtly indexed objects or definite null objects, we can conclude that it is strictly intransitive. As a result, the noun *pahe* 'house' in *pahe wore* 'that person went home' should be analyzed as a bare adjunct.

7.3.7 Adjuncts of uncertain obliquity

Some adjuncts involve a relational noun, such as *hambe-* 'top', which could be analyzed either as a possessor-marked noun or as a postposition that inflects for its argument.

In (331), the noun *sia* 'chair' could be marked as oblique, or the phrase *sia hambe-Ø* 'the top of the chair' could be a bare adjunct in the role of goal.

(331) *Kui* *kheko* *e* ***sia*** ***hambe-Ø*** *a-ke.*
bag that.near.you take.SER chair top-3SG.POSS put-2SG.IMP
'Take that bag and put it on the chair.' (ObjP4 0:04-0:06)

Some evidence that these are noun phrases (and thus bare adjuncts, not obliques) comes from the fact that they can themselves be the arguments of the more certain postpositions (which never behave as nouns either morphologically or syntactically).

In (332), the phrase *eme-o hambe-Ø* (lit. 'on top of my eyes') is the argument of the postposition *para* 'LOC'.

(332) *Anga* *ahe* *khaníngì-pi*
time INDEF do.thus-3PL.SS.R

eme-o ***hambe-Ø*** ***para*** *yo-pi ...*
eye-1SG.POSS top-3SG.POSS LOC come-3PL.SS.R
'Sometimes they do that and I see them, and ...' (Hapapia 8:05-8:09)

Another piece of evidence in favor of the noun-phrase analysis is that these relational nouns can be used without an overt NP argument, unlike the true postpositions. In such cases, the relational noun still bears a suffix showing its possessor.

In (333), the relational noun *khane-* 'inside' is marked for 3PL possessor. This is interpreted as the goal of motion, meaning 'into them'.

(333) *Hamunge* *Pite* *moko* *ni_ngé_ngu* *e-pi*
man big speech TOPIC take-3PL.SS.R

khane-noka *a-pi ...*
inside-3PL.POSS put-3PL.SS.R
'They took the word of God to heart, and ...' (PapaLaif 1:12-1:21)

I thus conclude that the best analysis of this sort of adjunct is as a bare NP, not an oblique.

7.3.8 Experiencer predicates

Hurukui has a phenomenon called *experiencer predicates* (EPs). In these, the situation described is that someone is experiencing some state, usually things like hunger, sickness, pain, etc., rather than actively doing anything. There are two types of EPs based on the construction they use: *stimulus-affect experiencer predicates*, in which the experiencer is the object of the verb (§7.3.8.1), and *possessive experiencer predicates*, in which the experiencer is the possessor of an inalienable noun (§7.3.8.2). In both types, the verb always agrees for 3SG subject. In some EPs, the stimulus subject is impersonal (non-referential), lacking an overt NP subject. I translate the stimulus as 'it' in such cases.

7.3.8.1 Stimulus-affect experiencer predicates

In stimulus-affect EPs, a stimulus noun phrase is the subject of the 'affect' verb, and the experiencer is its object. In (334), the adjective/noun *mbukura* 'cold' is the stimulus subject, and the 1SG experiencer object is shown by the object-marked 'affect' verb.

(334) *Mbukura* *ya-re.*
cold affect.1SG-3SG.R
'I feel cold.' (overheard)

Table 7.3 shows some nouns that can be the subject of such an EP.

Table 7.3: Examples of stimulus-affect EPs

stimulus noun	**EP**	**EP meaning**
okV 'hunger'	*okV ya-re*	I am hungry
mbukura 'cold'	*mbukura ya-re*	I feel cold
ukhu 'heat'	*ukhu ya-re*	I feel hot
iki 'pain'	*iki ya-re*	I am in pain
amomora 'sweat'	*amomora ya-re*	I am sweating
mbumi 'choking'	*mbumi ya-re*	I am choking
kehe 'cough'	*kehe ya-re*	I am coughing
yoho mua 'smoke'	*yoho mua ya-re*	the smoke is getting me (in the face/eyes)

The EP may be accompanied by a modifier, which is usually the adjective/intensifier *maisa* 'bad, very' (§6.5.4), but others do occur. In (335), *maisa* 'bad' modifies the stimulus *kehe* 'cough'.

(335) *Kehě* ***máisă*** *a-re.*
cough bad affect.3SG-3SG.R
'That person has a bad cough / is coughing a lot.' (overheard)

In (336), the adjective *maya* 'good' modifies the stimulus *nahe* 'taste'.

(336) *Nahě* ***máyă*** *na-re?*
taste good affect.2SG-3SG.R
'Does it taste good (to you)?' (overheard)

The structure in (335) and (336) is ambiguous between the adjective modifying the stimulus noun in a noun phrase (§7.1.1), vs being an adverbial modifier in pre-verbal modifier position (§7.3.9). This will require further investigation.

7.3.8.2 Possessive experiencer predicates

In possessive EPs, a noun is marked as inalienably possessed, and the possessor is the experiencer. It is ambiguous whether this noun is the subject or the object of the verb because there is no second overt NP to show the SOV order, and the inanimate referent would trigger 3SG agreement whether it is the subject or the object.

In (337), the noun *ona-* 'bone' is marked for 1SG possessor.

(337) *Ona-o* *a-re.*
bone-1SG.POSS affect.3SG-3SG.R
'I am full.' (lit. 'It affects/hits my bone(s). / My bone(s) affect/hit it.')[19] (overheard)

Table 7.4 shows some more examples of possessive EPs.

Table 7.4: Examples of possessive EPs

noun	**verb**	**EP**	**EP meaning**
ona- 'bone'	*a-* 'affect'	*ona-o a-re*	I am full
orape- 'skin'	*i-* 'do'	*orape-o e-re*	I am ashamed
eme- 'eye'	*ra-* 'be sleepy'[20]	*eme-o ra-re*	I am sleepy

Modifiers may also occur between the noun and the verb in possessive EPs. In (338), the adjective *maya* 'good' is used to indicate that the experience is a positive one.

19 The noun *ona-* 'bone' could even be an adjunct here, making the literal translation 'It_i affects/hits it_j on my $bone(s)_k$', with two impersonal arguments. Morphosyntactic tests about this are a topic for future research.

20 This verb is only found in this experiencer predicate, so it is not possible to give a gloss separate from the EP's meaning.

(338) *Ona-ŏ* ***máyă*** *a-re.*
bone-1SG.POSS good affect.3SG-3SG.R
'I'm nice and full. / I'm full to a good extent.' (overheard)

In (338), it seems more semantically plausible that the modifier is adverbial (affecting/hitting in a good way) rather than adnominal (good bones), but because this EP is already fairly idiomatic, this remains uncertain.

7.3.9 Placement of adverbial modifiers

Constituents that modify the verb are placed fairly freely in the clause, as long as they are after the subject and before the verb. A common location for them is immediately preceding the verb.

In (339), the temporal adverb *araiku* 'earlier, already' modifies the verb *ngáe-* 'tell'. It immediately precedes the verb and follows the object *opopo* 'thing'.

(339) *Ima* *opopo* ***araiku*** *ngáe-pa* *wo-pi* *nge*
1SG thing earlier tell-1SG.DS go-2PL.R TOPIC

a-pi *yo-pi.*
kill.3SG-2PL.SS.R come-2PL.R
'The thing (eel) I told y'all about, you killed it and brought it.' (Mario 3:18-3:23)

Other word classes can also be used in this slot. In (340), the intensifier *awohe* 'too much' (§6.5.4) is modifying the verb *epha-* 'hear, know'. It is not modifiying the noun *moko* 'language' because it cannot modify nouns, and it is not modiying the adjective *wohe* 'native' because it precedes any adjectives that it modifies.

(340) *Ima* *moko* *wohe* ***awohe*** *ephà-khukha-wá-pa.*
1SG language native too.much know-1SG.PRF.R-NEG-1SG.R
'I don't know Hurukui well.' (Tamiok2 0:07-0:11)

(341) has two adverbials: a modifier expressing time, *arakaya ahe* 'some time ago', and a modifier expressing place, *naita nenge no* 'here in our village'. They are used before the second verb of the chain, *khai-* 'live', rather than the first verb, *yo-* 'come'. The reason for this will require further research.

(341) *Hamunge* *wo* *yo-mu* ***[arakaya*** ***ahe]***
man work come-3SG.SS.R long.ago INDEF

[naita ***nenge*** ***no]*** *khai-mu* *kha ...*
1PL.INCL village here live-3SG.SS.R and
'A doctor came and lived here in our village some time ago, and ...' (Dokta 5:04-5:08)

The preverbal modifier position is not strict. Adverbials can be placed elsewhere in the clause. In (342), the temporal modifier *kaya* 'long ago' precedes the object *hepa* 'trap' of the first verb *a-* 'put' (not to be confused with the homophonous verb stem *a-* 'kill.3SG' in the second clause).

(342) *Hamunge* *pite* ***kaya*** *hepa* *a-pi* *mbo* *a-mai-pi.*
man big long.ago trap put-3PL.SS.R animal kill.3SG-HAB-3PL.R
'Long ago, the elders set traps and killed animals (bandicoots).' (Trap7 0:05-0:08)

7.3.10 Periphrastic passive

For the purposes of this description I consider *passive* to be a functional notion, not a structural one, in which the agent is not mentioned explicitly, either because the agent's role is being downplayed in favor of talking primarily about what the patient experienced, or because the agent's identity is unknown.

Hurukui encodes a passive function by using an indefinite noun phrase as the subject. A common choice of dummy subject for human agents is *hamunge ahe* (man INDEF) 'someone'. This is shown in (343).

(343) ***Hamunge*** ***ahe*** *opopo* *khúnŭ* *e-pi* *we-pínakhai-Ø.*
man INDEF food theft take-3PL.SS.R go-3PL.DPST-3SG.R
'Our food got stolen.' (lit. 'Some men stole our food and went.') (elicited)

For non-human referents, one available choice is *opahe* 'something'. This is seen in (344), where the agent of the biting was likely a bug or small animal.

(344) ***Opahe*** *haringuni* *aka-yá-mukhai-Ø.*
something night bite-1SG.OBJ-3SG.PRF.R-3SG.R
'I got bitten in the night.' (lit. 'Something bit me in the night.') (elicited)

7.3.11 Light verb constructions

The verbal lexicon in Hurukui is highly enriched by *light verb constructions* (LVCs), in which a verb root with little or no semantic content is combined with another element which I call the initial, following Harris (2008). The initial cannot be an inflecting verb. LVCs are the primary topic of Part II of this dissertation, but I give a simple introduction here for completeness of the description of clausal syntax.

The initial is often a noun, but plenty of other word classes and phrase types can be used as initials in Hurukui LVCs. The verb of an LVC is called a *light verb* (LV), because it has less semantic content than a more typical lexical verb (which could conversely be referred to as semantically "heavy"). Most but not all LVs in Hurukui have lexical uses as well. Initials are more divided as to whether they are also used lexically outside of the LVC or not. LVCs are the primary strategy for creating new verbal lexical items in Hurukui, as there is no derivational morphology that can turn a noun or any other word class into a verb.

An example of an LVC is *moko e-* 'talk, speak, tell a story'. Its initial is the noun *moko* 'speech, story, language', and its LV is *e-* 'take'. The LV does not actually have any meaning of 'take' here and is just used as a verbalizer to turn the noun 'speech' into a verb meaning 'do speech'. A simple example is shown in (345), in which *waipu* 'chicken' is the subject of the LVC.

(345)	*Isí-mukhài-ne*	*we-mu*	*waipu*	***moko***	***ĕ-make-re***.
	sleep-3SG.PRF.R-3SG.DS	go-3SG.SS.R	chicken	speech	take-CONT-3SG.R
	'He slept until the rooster crowed.'				(MejikDiwai 1:09-1:14)

The two parts have combined to make a new lexeme; it is not the case that the noun 'speech' is the object of 'take', although this is a plausible diachronic source. This can be seen by the syntactic test of object NP ellipsis, in which a second clause omits an already-mentioned shared object. To see how this works, consider (346), where the noun *mburumu* 'sweet potato' is the object of the verb *i-* 'eat'. In the response to the question, this object can be omitted because it is recoverable from context, and the verb retains its lexical meaning of 'eat'.

(346) A: *Mburumu* *ire-ma?* B: *Eo,* ***(mburumu)*** *ire-ra.*
sweet.potato eat-2SG.R yes sweet.potato eat-1SG.R
A: 'Did you eat a/the sweet potato?' B: 'Yes, I did.' (elicited)

LVCs, in contrast, do not allow this. They need the initial to be present for the LVC meaning to be possible. In (347), a similar surface structure is used to that in (346), but the noun *moko* 'speech' cannot be omitted in the response.

(347) A: *Tata* *moko* *e-ma-re?* B: *Eo,* ***[moko / *Ø]*** *e-ma-re.*
father speech take-CONT-3SG.R yes speech take-CONT-3SG.R
A: 'Is Dad talking?' B: 'Yes, he is.' (elicited)

(348) shows another example of this, with the LVC *khikho- re-* 'urinate' (lit. 'urine burn').

(348) A: *Khikhŏ-Ø* *ro-re?* B: *Eo,* ***[khikhŏ-Ø / *Ø]*** *ro-re.*
urine-3SG.POSS burn-3SG.R yes urine-3SG.POSS burn-3SG.R
A: 'Did that person urinate?' B: 'Yes, they did.' (elicited)

This omissibility test is also evidence against analysis of the LVC initials as adjuncts. Adjuncts are omissible, expressing extra information such as time, location, or instrument, that acts as extra specification to the predicate rather than an inextricable part of it. For instance, in (349), speaker A uses an oblique adjunct *maha para* 'on the ground', which speaker B omits when repeating the clause. The fact that the initials in (347) and (348) cannot be omitted like this is evidence that they are not adjuncts.

(349) A: ***Maha*** ***para*** *kuni-e* *ya-mu* *yo-re.*
ground LOC leg-3SG.POSS step-3SG.SS.R come-3SG.R

B: *Eo,* *kuni-e* *ya-mu* *yo-ne ...*
yes leg-3SG.POSS step-3SG.SS.R come-3SG.DS
A: 'She came on foot along the ground.' (Shop1 5:09-5:14)
B: 'Yes, she came on foot, and ...'

A similar test is omission in tail-head linkage (THL), a discourse strategy in which the last clause of one chain is repeated at the beginning of the next chain. Object and adjunct NPs can be omitted in the recapitulatory clause of THL, but LVC initials cannot. In (350), the final clause of the first chain is headed by the verb *yo-ra* 'I came', which has a bare adjunct NP *Nanambu no* 'Nanambu here, this place Nanambu' (bracketed) in the semantic role of goal. The next clause

chain begins with a clause consisting only of the verb *yo-khu* 'I came, and (SS)', omitting the adjunct.

(350) *Ima* *hamame* *ekha-khu* ***[Nanambu*** ***no]*** ***yo-ra.***
1SG again go.down-1SG.SS.R place.name here come-1SG.R

Yo-khu *kha* *hindaura* *ni_ngé_ngu ...*
come-1SG.SS.R and afternoon TOPIC
'I came down again to Nanambu here. I came, and in the afternoon ...' (Asde2 2:44-2:52)

In contrast, in (351), the final clause of the first chain is headed by the LVC *ngáe-* 'tell', which is completely fused from an earlier word *ngaV* (final vowel no longer recoverable) plus the LV *e-* 'take'. In the first clause of the next chain, both the initial and LV must be repeated.

(351) *Uso* *oke-pi* *a-pi* *kha* *yo-pi* ***ngáV*** ***e-ta-pi.***
rope tie-3PL.SS.R PFV-3PL.SS.R and come-3PL.SS.R tell take-1PL.OBJ-3PL.R

NgáV ***e-ta-pina*** *kha* *wo* *numu-e* *nge* *hoka-pu.*
tell take-1PL.OBJ-3PL.DS and work body-3SG.POSS TOPIC do.work-1PL.R
'They made a decision and came and told us. They told us, and we did the work.' (Bibuai 9:52-9:59)

More evidence that the initial is not an object comes from the fact that other core NPs are marked as objects on the verb, while the initial is retained in order to create the LVC meaning. In (352), the subject is a 3SG referent that is not shown by an overt NP in this clause chain, and the object is *ima* '1SG'. Both of these arguments are indexed on the final verb. In addition, there is an initial *khenge* 'call out', without which the verb *i-* 'do' cannot express the predicate of calling to someone.

(352) *Yo-mu* *ima* ***khénge*** ***ĭ-ya-re.***
come-3SG.SS.R 1SG call.out do-1SG.OBJ-3SG.R
'He came and called (to) me.' (Toilet 0:51-0:53)

For these morphosyntactic and semantic reasons, the LVC initials are neither objects nor adjuncts. Elsewhere in this description, I often write the more fused LVCs as one root with irregular stress marked, e.g., *hupúri-* 'show' (rather than *hupurV i-* or *hupur-i-*) and *ngáe-* 'tell, talk to' (rather than *ngáV e-* or *nga-e-*). If the initial is still usable as a constituent separate from the LV

in other contexts, I write them separately, e.g., *moko e-* 'talk' (rather than *móke-*) and *khikho- re-* 'urinate' (rather than *khíkhŏre-*).

There is much more to say about these constructions, and a much wider array of morphosyntactic behaviors than I have shown here. The reader is referred to Part II for a detailed description and analysis of Hurukui LVCs.

7.4 Interrogative clauses

Interrogative clauses are morphosyntactically the same as indicative (declarative) clauses. Polar questions (§7.4.1) differ from statements only in intonation. Content questions (§7.4.2) have interrogative words in place of some constituent, and their intonation is the same as a statement. When sentence-final question particles are present, though, there are some intonational differences (§7.4.3).

7.4.1 Polar questions

Polar questions are the same as the corresponding statement, but with rising intonation toward the end. They are not marked by any changes to morphosyntax. In this section, I show the intonation using diacritics. I use the acute and grave accents similarly to how they are used for primary and secondary stress elsewhere in this description. However, I use them here with more focus on pitch alone. Thus, the acute accent means a syllable with length and high pitch, while the grave accent means a syllable with length and low pitch. This is phonetically the same as how primary and secondary stress are defined, but I ignore the phonological stress in favor of the phonetic realization for the purposes of this section. I also use the hook diacritic (*ả ẻ ỉ ỏ ủ*) to indicate high pitch on an unstressed (and therefore short) vowel, because it looks like a question mark.

In (353), the realis verb *ngoma* 'you see (it)' is unaltered except for pitch. The highest pitch is found on the final syllable, and the preceding syllable, which has primary stress, is demoted to a lower pitch. The prosody of this verb in a statement would be *ngóma*.

(353) *Ngò-**mả?***
see-2SG.R
'Do you see (it)?' (Chameleon 4:25-4:27)

7.4.2 Content questions

Hurukui is a *wh-in-situ* language. Interrogative words, such as the pronouns *ino* 'what' and *neku* 'who' (§3.2.4), the deictics *nga* 'where' and *ngapa* 'what place' (§4.1), the quantifier *keme* 'how much, how many' (§6.2.4), and the adverb *kemeka* (§6.3), are not fronted but are kept in the same place in the sentence where their answer would be located. No morphosyntactic or intonational changes are made to the clause.

In (354), the pronoun *ino* 'what' is the object of the verb, and it is found in the usual position between the subject and the verb.

(354) *Eha* *oro-ka* ***ino*** *a-mai-pi?*
dog downland-there what kill.3SG-CONT-3PL.R
'What are the dogs killing down there?' (Hunt5 1:44-1:48)

In (355), *nga* 'where' is used as an adnominal modifier (more like 'which' in this context) modifying *amenge* 'side'.

(355) *Hamiha* *amenge* ***nga*** *akhu* *para* *yó-nekhai-Ø?*
clan side where LOC from come-3SG.DPST-3SG.R
'What clan did she come from?' (Klen 0:34-0:37)

7.4.3 Intonation with question-final particles

Some differences in intonation are found when questions are ended with the particles *o* 'TAG.Q', *kho* 'I wonder', or *khuní* 'hell if I know'. Before *o* 'TAG.Q', and *kho* 'I wonder', the rest of the sentence has the same rising intonation found in polar questions that don't have final particles, even if the question is a content question. The particles *o* and *kho* themselves have low pitch and are often long.

This is shown in (356) and (357), where the syllable before the final particle has high pitch. In (356) the pitch before *o* 'TAG.Q' is the same as if the particle were absent, because it is a polar question.

(356) *Ngu* *okhàrí-**wả**-Ø* ***ò?***
3SG be.confused-NEG-3SG.R TAG.Q
'She wasn't confused, or was she?' (Nadzab 9:46-9:48)

However, in (357), the pitch is different from the particle-less content question, which has the prosodic pattern *inóngì-tǐ-mà-re?*

(357) *È* *inòngì-tǐ-má-**rẻ*** ***khò?***
take.SER do.what-DES-CONT-3SG.R I.wonder
'I wonder what he's trying to do?' (VD2 1:07-1:09)

Before the particle *khuní* 'hell if I know', questions have the same intonation found when the particle is absent, but the particle itself has an unusual pitch pattern. The particle *khuní* itself has low pitch on the first syllable and falling pitch (not simply high) on the second. I notate this with circumflex accent (*î*), for this section only. This is shown in (358).

(358) *Íno* *ko-kòpu* ***khunî?***
what say-1PL.FUT hell.if.I.know
'What on earth should we call it?' (Abus 4:48-4:51)

8 Multi-verb constructions

Hurukui has various constructions involving multiple verbs. Some of these combine multiple verbal clauses into one sentence, while others combine multiple verbal words within one clause. The most common multi-verb construction, a very Papuan feature, is clause chaining (§8.1). In addition, the language uses simple juxtaposition of equipollent clauses (§8.2), serial verb constructions (§8.3), and subordination (§8.4).

8.1 Clause chains

Hurukui makes extensive use of a morphosyntactic device called *clause chaining*. A *clause chain* (which I often call a *chain* for short) is a series of clauses in which the matrix verbs of the non-final clauses are marked with switch reference morphology (§5.4). The last clause in a chain is called the *final clause*, its main verb is called a *final verb*, the other clauses are called *medial clauses*, and their main verbs are called *medial verbs* (Roberts 1997). The verb of a free-standing clause (i.e., in a chain of only one clause) is final.

Medial verbs inflect for switch reference (SR), a type of verbal morphology described in §5.4. SR morphology is only found on matrix verbs, not subordinate ones. SR encodes whether the referent of the marked verb's subject is the same as that of the next matrix verb (same-subject, SS) or not (different-subject, DS). In Hurukui, most of these morphemes also agree with their own subject.

A simple chain consisting of only two clauses is shown in Example (359). The first clause is *Martin ipoe wene* 'Martin's father went', and the second clause is *khéngepera* 'I asked him'. The verb *wene* is marked with a morpheme *-ne* '3SG.DS', indicating that its subject is 3SG and the next verb's subject has a different referent. The final verb *khéngepera*, in contrast, does not have SR marking, only a TAM marker *-ra* '1SG.R'.

(359) *[Martin ipo-e we-**ne**] [khéngepe-ra].*
Martin father-3SG go-3SG.DS ask-1SG.R
'Martin's father went, and I asked him.' (BW 1:03-1:07)

8.1.1 Basic use of switch reference

The switch reference markers (§5.4) are grouped into same-subject (§§5.4.1-5.4.2) vs different-subject (§5.4.4). (360) shows a simple same-subject clause chain. Each clause (bracketed) consists of only a single verb. Both of their subjects are 1SG. The first clause has a medial verb form, conjugated with *-khu* '1SG.SS.R', so it is a medial clause. The second clause has a final verb form, conjugated with *-ra* '1SG.R', so it is a final clause.

(360) *[E-**khu**]* *[a-ra].*
take-1SG.SS.R put-1SG.R
'I took it and put it.' (elicited)

(361) shows a simple different-subject clause chain. Each clause (bracketed) consists of only a single verb. The medial verb's subject is 1SG, while the final verb's subject is 3SG. To show that the subject changes, the medial verb is conjugated with *-pa* '1SG.DS'.

(361) *[A-**pa**]* *[ume-re].*
kill.3SG-1SG.DS die-3SG.R
'I killed it.' (overheard)

Hurukui switch reference markers are also divided into two groups along a dimension called *reality status*, or *status* for short. Its values are *realis* (glossed R), generally describing events that have actually happened or are actually happening, and *irrealis* (glossed IRR), describing other events. The way that TAM categories are categorized into one of the two statuses is language-specific, but this general semantic characterization works well for Hurukui. All matrix verbs in a clause chain must agree in reality status, and the TAM of the final verb determines which status is used in the medial verbs. In Hurukui, same-subject markers distinguish reality status (§§5.4.1-5.4.2), while different-subject markers do not (§5.4.4). Only the same-subject realis and different-subject markers inflect for the person of the marked verb's subject.

Example (362) shows a same-subject realis chain. The subject of all three verbs is 1SG. The two medial verbs thus bear the 1SG.SS.R suffix *-khu*, while the final verb bears the final TAM suffix *-ra* '1SG.R'.

(362) *[E-**khu**]* *[we-**khu**]* *[amenge* *ahe* *wo-ra].*
take-1SG.SS.R go-1SG.SS.R place another go-1SG.R
'I took it and went to another place.' (Samting 3:06-3:09)

Example (363) shows a different-subject realis chain. The first verb's subject is 3PL, while the second verb's subject is 1PL. The medial verb thus bears the 3PL.DS suffix *-pina*, while the final verb bears the final TAM suffix *-pu* '1PL.R'.

(363) *[Imisi* *khaingu* *pore* *nukhi-**pina**]*
custom that completely abandon-3PL.DS

[ekhono *khai-pu* *khanoku].*
now be.located-1PL.R this
'They stopped doing that, and now we live here.' (Man3 2:32-2:36)

In the irrealis status, same-subject marking does not inflect for person. Instead, there are two ways of marking same-subject irrealis: the suffix *-KHa* (§5.4.2), and the serial form (§5.4.3). The suffix is used more often when two verbs are construed as separate events, while the serial form is used more often when they are seen as parts of one complex event. In many contexts they are both acceptable.

Examples (364) and (365) show clause chains in the irrealis. In (364), both verbs have 2SG subject, so the medial verb uses one of the strategies for marking same subject in the irrealis (here, serialization).

(364) *Hamiha* *khaingu* ***yai*** *ya-ka* *we-**kanu**.*
stairs that climb:SER up-there go-2SG.FUT
'You will climb up those stairs.' (Lae 1:53-1:56)

In (365), the medial verb has 1SG subject while the final verb has 3SG subject, so the suffix *-pa* '1SG.DS' is used on the medial verb. Only one different-subject paradigm exists and is used in both realis and irrealis.

(365) *Umo* *amburine* *umo* *hi-**pa*** *kui* *khane-Ø* *ekha-**komu**?*
tree seed also cause-1SG.DS bag inside-3SG.POSS go.down-3SG.FUT
'Shall I also put the fruit into the bag?' (ObjP1 3:32-3:35)

While I follow the convention from Papuanist literature in using the terms *same-subject* and *different-subject* to refer to the switch reference categories, these terms are somewhat

misleading for Hurukui, because it is not always the subject (for any given definition of *subject*) whose sameness is tracked by SR morphology. More accurately, the term *pivot* means whichever referent is tracked by SR. Often, criteria for morphosyntactic subject as well as pivothood all identify the same participant. For instance, it is often the case that the the first argument also controls verb agreement and also is tracked by SR. This is referred to as *canonical* SR. Often most or all clauses in a chain use canonical SR. However, there are enough cases where SR behaves otherwise, called *non-canonical* SR, that it is worth making the distinction between subject and pivot.

The pivot often differs from the morphological subject in the case of experiencer predicates (§7.3.8), where the more topical/animate referent (the experiencer) takes precedence over the inanimate subject (the stimulus) for the purposes of SR marking. (366) describes a situation in which one group of people is visiting another village, whose inhabitants are indexed *i*. The visiting group sends out two representatives, indexed *j*, and the rest of the visiting group is indexed *k*. The first clause has subject *i*. The second clause is an experiencer predicate whose experiencer is *i* and whose morphological subject is 3SG to agree with the syntactic subject noun *khane-* 'inside', marked for possessor *i*. This noun and its agreement on the verb are indexed as *x*. The following clauses have subject *i* again. The SR marking from the first clause into the second is same subject, despite the morphological and syntactic subjects being different. That is, the same subject marking tracks the topical referent here. The SR marking from the second into the third clause, however, is different subject, even though the same relationship holds as that between the first and second clauses (experiencer = topical referent). This pattern, in which SS is used into experiencer predicates but DS is used out of them, is common in Hurukui.

(366) *[Nua nge-nua-**pi**] [khane-noka máyă i-**ne**]*
3DU$_j$ see-3DU$_j$.OBJ-3PL$_i$.SS.R inside$_x$-3PL$_i$.POSS good become-3SG$_x$.DS

[khaingu para opopo e-pi] [hito-pi ni].
this LOC food take-3PL$_i$.SS.R give.1PL$_{j+k}$-3PL$_i$.R EMPH

'They$_i$ saw the two of them$_j$, and they$_i$ were happy, and so they$_i$ gave us$_{j+k}$ food. (Dokta 5:43-5:53)

8.1.2 Referential overlap

Sometimes the referents of the subjects of adjacent chained clauses overlap only partially, a situation called *referential overlap*. As a general rule, Hurukui uses different-subject marking if any new referents have been added in the second subject, and same-subject marking otherwise. Thus, same-subject marking is used when the second subject is a subset of the first (called *contracting subject*), and different-subject marking is used when the first subject is a subset of the second (called *expanding subject*).

In (367), only Ruben is the subject of the first two verbs, but both Ruben and the speaker are the subject of the next one, so DS marking is used on the second verb to indicate expanding subject.

(367)	*Ruben*	*we-mu*	*nge-ya-**ne**,*	*ngu_para*	*hia*	*we-khu ...*
	Ruben	go-3SG.SS.R	see-1SG.OBJ-3SG.DS	and.then	1DU.EXCL	go-1DU.SS.R
	'Ruben went and saw me, and then we two went and ...'					(Pik 0:41-0:47)

Contracting subject is seen in (368). The first subject is 'you and I' while the second is just 'I'. Since no new referents were added, SS marking is used, and DS marking is not allowed.

(368)	*Naya*	*[wi-**sa***	/	**we-pua]*	*ima*	*opopo*	*í-tara.*
	1DU.INCL	go-SS.IRR		go-1DU.DS	1SG	food	eat-1SG.FUT
	'Let's go and I'll eat.'						(elicited)

8.1.3 Relative tense

Another important piece of information about Hurukui clause chains concerns the temporal relationship between the verbs in successive clauses. While most medial verbs in Hurukui do not mark relative time, there are two kinds that can be marked.

The first is *strong sequentiality*, in which a delay separates the end of the first event from the start of the second event. Strong sequentiality is marked by the addition of the suffix *-amo* after the SR suffix. For same-subject verbs with this marker, person distinctions are often collapsed, leading to the use of the invariant form *-mamo*, derived from *-mu-amo* (3SG.SS.R-SEQ), regardless of the subject.

The second is *simultaneity*, in which the two events are occurring at the same time for part of their durations. Simultaneity is marked for same-subject verbs by the suffix *-hVhV* 'SS.SIM' (§5.4.5), where *V* are copies of the previous vowel (e.g., *i-hihi* 'while eating'). Simultaneity is marked for different-subject verbs by the continuative marker *-ma(i) ~ -maki* (§5.2.4.4) preceding the DS suffix.

Between the extremes of simultaneity and strong sequentiality, there are other kinds of relative time meanings. An important one is *weak sequentiality*, in which the second event starts after the first event starts, but not necessarily after it ends. In some situations, this can be understood as the more narrow meaning of *immediate sequentiality*, in which the second event starts right at the instant when the first event ends. Immediate sequentiality is thus a subset of the meaning of weak sequentiality.

These types of sequentiality are illustrated in Figure 8.1. In the figure, time runs from left to right, and each category has two events, represented by parallel horizontal lines. A solid line indicates an event in progress (or stasis). A circle on the left or right end of a line indicates the start or end, respectively, of an event. A question mark on the left or right end of a line indicates underspecification for when the event starts or ends.

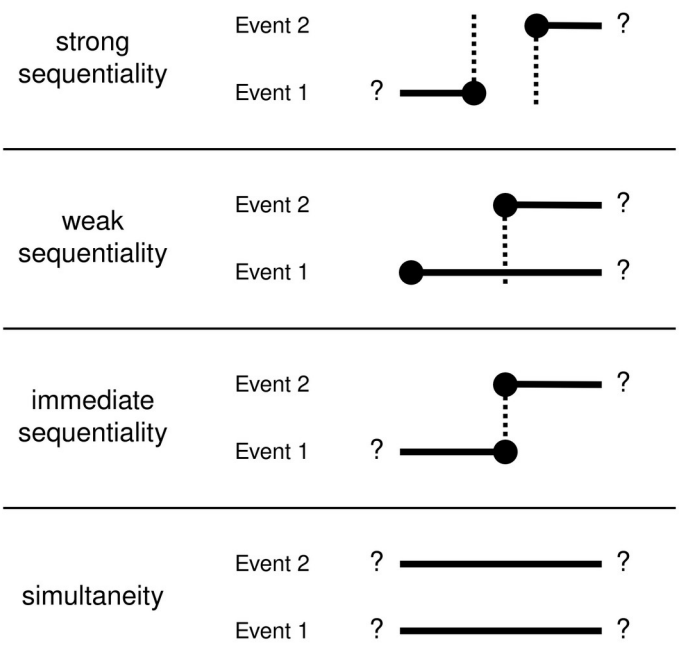


Figure 8.1. Relative time meanings

These are etic meanings, not specific to Hurukui, but I use them to describe what the Hurukui forms mean. The terms for the three subtypes of sequentiality are of my own creation, not derived from previous literature.

Because there are dedicated markers for simultaneity and strong sequentiality, the choice not to use them carries an implicature that the events are neither simultaneous nor strongly sequential. This is an example of an implicature arising from speakers' knowledge of the linguistic system as a whole. In addition to this, speakers know that events are usually listed in chronological order, called *tense-iconicity* (Haiman 1983: 813), so it is assumed that clause chains without simultaneous or strong sequential marking follow this order by default.

The events in most chains are thus understood to have the temporal relationship of weak sequentiality. Depending on the context and world knowledge about how different kinds of events work, the narrower meaning of immediate sequentiality may make the most sense. For instance, the events in the clause chain in (369) are understood to be immediately sequential, because hitting a mosquito is a very punctual event that causes it to die essentially instantly.

(369) *Sisu* *a-pa* *ume-re.*
mosquito hit.3SG-1SG.DS die-3SG.R
'I hit the mosquito and it died.' (overheard)

8.1.4 Polarity within chains

Clauses within a chain may share polarity or not. When all the clauses in a chain have positive (unmarked) polarity, no special marking is added beyond normal switch reference. (369) is a simple example of this. Similarly, when all the clauses have negative polarity, typically only the last verb must be marked as negative, and the negative marker has scope over the whole chain. In (370), the chain *we- hi-* (go- become-) is used to describe change of state, but the change did not happen as expected in the story context. In this story, three chameleons are traveling together, and two of them change color, but the third one fails to do so. The medial verb *we-* 'go' here has a metaphorical meaning that is common in Hurukui with predicates of becoming or changing. The *we-* 'go' part of the chain indicates the ongoing change process, and the *hi-* 'become' part indicates the entrance of the final state, the moment of becoming the new thing. Because the chameleon did not even start to change color at all, both the process (*we-* 'go') and the entrance of state (*hi-* 'become') failed to occur. Thus, *we-* 'go' is not marked for polarity but is still interpreted as negative.

(370) ***We-mu*** *nua* *khahu* *harae* *hi-**wa**-Ø.*
go-3SG.SS.R 3DU COMIT same become-NEG-3SG.R
'He didn't become like them.' (Chameleon 2:54-2:57)

An exception to this pattern is seen in same-subject irrealis chains. Here, polarity has scope over preceding verbs in their serial form (§5.4.3), but not over those with the SS.IRR suffix *-KHa* (§5.4.2). This distinction is shown in (371) and (372), respectively.

In (371), both verbs are negated by the prohibitive suffix *-wapho*, because the first verb is in its serial form and thus falls under the scope of the negation.

(371) ***Wi*** *yo-**wapho**-Ø!*
go:SER come-PROH-2SG.IMP
'Don't go and come! (back and forth)' (elicited)

In (372), only the verb bearing the prohibitive suffix is negated. The preceding verb, which bears the SS.IRR suffix, is interpreted as having positive polarity (the unmarked default).

(372) ***Wi-sa*** *yo-**wapho**-Ø!*
go-SS.IRR come-PROH-2SG.IMP
'Go and don't come back!' (elicited)

When a negative medial verb is followed by a positive verb with the same subject, the marker *-wapi* 'NEG.SS' is used. It is invariant for person. In (373), the first clause is negative, the second is positive, and they have the same subject, so this marker is used.

(373) *Ngu* *khongokho* *nge-**wapi*** *wé-mukhai-ne* *kha ...*
3SG hat see-NEG.SS go-3SG.PRF.R-3SG.DS and
'He didn't see his hat, and he went, and ...' (TP1 3:14-3:19)

However, when the subject is 1SG, one often finds *-wa-pa* (the negative realis for 1SG) rather than *-wapi* 'NEG.SS'. This may be due to analogy aided by final vowel loss making the two forms sound the same much of the time.

When a negative medial verb is followed by a positive verb with a different subject, the negator *-wa* is followed by the usual DS suffix agreeing with the subject of the former clause. This is seen in (374), where the first clause (excluding the topicalized/subordinated verb *khaipu*) has the verb *yo-* 'come' marked as negative, and the second clause is headed by the positive-polarity verb *ko-* 'say'.

(374) *Khai-pu* *nge* ***yo-wa-ne*** *kha* *haringuni* *kopohe* *na*
stay-1PL.R TOPIC come-NEG-3SG.DS and night youth and

akhure *ngu* *marihí-tara* ***ko-pi**.*
wife:3SG.POSS 3SG search-1SG.FUT say-3PL.R
'We were there, he didn't come back, and at night the young men and women wanted to look for him.' (Robert 0:47-0:54)

Some combinations of verbs are strongly collocated, leading to them being treated more like a single verb by negation. An example of this is *e- yo-* (take- come-) 'bring'. In (375), the whole chain is negated even though, strictly speaking, only the first verb is interpreted as negative in context. This is because the speaker **did** come, but failed to bring the object (a stone he meant

to show to someone). This can also be expressed as *e-wapi yo-ra* (take-NEG.SS come-1SG.R) 'I didn't take it and I came'.

(375) *Opopo harawohe* ***e-khu*** ***yo-wa-pa.***
thing one take-1SG.SS.R come-NEG-1SG.R
'There's something I didn't bring.' (Planim 0:58-1:02)

The case of negating *e- yo-* shows an example of gradience in wordhood and the incipient univerbation of a strongly collocated/chunked phrase. The option using *-wapi* 'NEG.SS' shows an analysis of the two verbs as separate: I didn't take it, but I did come. The option where the whole chain is negated shows an analysis of *e-*[SS] *yo-* as more like a single verb, 'bring', which has a bipartite stem and two agreement sites (one for the SS marking and one for the final subject marking). This complex verb can then be negated in the usual way by adding *-wa* 'NEG' before the final marking, to mean 'not bring' without necessarily meaning 'not come'.

8.1.5 Mood within chains

All clauses in a chain must have the same mood (e.g., realis, imperative, desiderative). Realis clause chains are very common and shown by various examples throughout this description, such as (384) and (375) in §8.1.1. Another is seen in (376).

(376) *Isi-puna* *ánga-re.*
sleep-1PL.DS dawn.break-3SG.R
'We slept and dawn came.' (Mario 3:27-3:30)

Chains that are entirely in the imperative mood may have the same subject in all their clauses. Because Hurukui classifies imperative as irrealis, all the medial verbs are either in their serial form or marked with *-KHa* 'SS.IRR'. Examples (377) and (378) show this.

(377) *Sono* *humba* *e* *yo-Ø!*
water fill:SER take:SER come-2SG.IMP
'Fetch water!' (Kaikai 2:32-2:34)

(378) *Nembake* *epha-kha* *noroka-kha* *khaka-ni.*
well hear-SS.IRR think-SS.IRR stay-2PL.IMP
'Y'all listen well and think about it and stay (here).' (Bibuai 1:04-1:08)

Some imperative chains, however, involve different-subject marking, but all the verbs still have an imperative illocutionary force. An example is seen in (379). At first glance, this sentence may appear to involve a direct quote *yo yoho hukeke* 'come blow the fire!', but because the verb *ngáe-* 'tell' is marked for 2SG.DS rather than final imperative, there is no subordinate quotative clause here.

(379) *[Na kheko ngáe-mana] [yo yoho huke-ke].*
2SG that.near.you tell-2SG.DS come:SER fire blow-3SG.IMP
'Tell him to come blow the fire!' (elicited)

A similar example with three clauses, with both medial verbs marked for DS, is seen in (380).

(380) *[Yoho húke-mana] [óme-ne] [yoho mua pe-wa-Ø]!*
fire blow-2SG.DS light.up-3SG.DS fire smoke NEG.EXIST-go-3SG.IMP
'Blow the fire so it lights and the smoke goes away!' (overheard)

Another mood that has scope over entire clause chains is desiderative, as in (381), where both *nukhi-* 'lose' and *pé-we-* 'become absent' are interpreted as immediate future, one of the meanings of the desiderative (§5.2.4.3).

(381) *Moko wohe araiku nukhi-puna pé-wè-ki-ma-re.*
language native soon lose-1PL.DS NEG.EXIST-go-DES-CONT-3SG.R
'Soon we will lose the local language and it will go away.' (LusimTp 1:47-1:51)

8.1.6 Shared object marking

When multiple chained transitive verbs share their object, only one of them has to mark it using an object suffix. In (382), the verbs *e-* 'take' and *a-* 'put' share a 1SG object, but it is only marked on the first verb.

(382) ***E-ya-mu*** *maha* ***a-ne*** *pípa-pa* *kha ...*
take-1SG.OBJ-3SG.SS.R ground put-3SG.DS sit.down-1SG.DS and
'It (bus) dropped me off and I sat down, and ...' (Nadzab 8:30-8:33)

In (383), the same two verbs share a 1PL object, but it is only marked on the second verb.

(383) *Kaya* *ipo-ito* ***e-mu*** ***a-ta-ne***
long.ago father-1PL.POSS take-3SG.SS.R put-1PL.OBJ-3SG.DS

maha *khaino* *yo-pu* *ni.*
ground this come-1PL.R EMPH
'Long ago, our father [God] put us on this ground.' (Hevi 0:06-0:14)

The factors determining which verb carries the object marker are a topic for future research.

8.1.7 Medial nonverbal predicates

Nonverbal predicates in medial clauses must use the verb *(h)i-* 'be'. In copular clauses (§7.2.1), this verb is already used when the TAM is anything other than the unmarked realis. In the unmarked TAM category, final zero copula corresponds to medial *(h)i-* in the positive polarity, and final *ome* 'is.not' corresponds to medial *(h)i-* in the negative polarity. This is shown in (384) and (385), respectively.

(384) *Arahaiwe* *maisa* ***i-ne*** *hamame* *we-pu*
far bad be-3SG.DS again go-1PL.SS.R

imisi *makara-Ø* *ise-pu.*
road middle-3SG.POSS sleep-1PL.R
'It was very far, and we slept again along the way.' (Dokta 3:36-3:42)

(385) *Awohe* *maya* ***ĭ-wa-ne*** *kha ...*
too.extent good be-NEG-3SG.DS and
'It wasn't too good, and ...' (Graun 12:29-12:33)

In negative existential clauses (§7.2.2) the particle *pe* 'NEG.EXIST' is retained, and *hi-* 'be' is added. This creates a combination *pé-hi-* 'be absent' that acts somewhat like a light verb construction (§7.3.11), a new complex predicate of its own. The *h* in *(h)i-* 'be' is retained because *pe* ends in a vowel that cannot be elided. (386) shows it.

(386) *Hauphe* ***pé-hi-ne*** *hauphe* *ake* *eme-ito* *ra-ne ...*
betel.nut NEG.EXIST-be-3SG.DS betel.nut for eye-1PL.POSS be.sleepy-3SG.DS
'There was no betel nut, and our eyes felt sleepy for (the lack of) betel nut, and ...' (Buai 0:07-0:11)

8.1.8 Interaction of experiencer predicates with switch-reference

Experiencer predicates (EPs) of both the stimulus-affect type (§7.3.8.1) and the possessive type (§7.3.8.2) allow switch reference morphology to treat either their experiencer or their morphological subject as the pivot. This variation can be seen in switch-reference marking both into and out of the EP clause.

(387) shows some stimulus-affect EPs preceded by same-subject marking. The example is a bit complex but illustrative. It has three clauses (bracketed): a normal verbal clause and two stimulus-affect EPs. The subject of the verbal clause (headed by *khai-* 'live') has the same referent as the experiencer of both EPs. The verbal clause is marked for same-subject into the first EP, so the pivot is the experiencer.

In addition, same-subject marking is used from the first EP into the second one, even though the stimulus nouns (and thus the 'affect' verbs' morphological subjects) are different. The pivot for this same-subject marker is therefore also the experiencer. It is notable that the same morpheme *-mu* '3SG.SS.R' reflects the person/number of one referent (3SG, agreeing with the stimulus noun *ua* 'sickness') but the sameness of a pivot other than that referent.

(387) *[Mba* *phephe* *nua* *pahe* *maisa* *khai-**khu**]*
grandfather grandmother 3DU house bad live-3DU.SS.R

[ua *awohe* *nua-**mu**]* *[kehe* *nua-ne] ...*
sickness too.much hit.3DU-3SG.SS.R cough hit.3DU-3SG.DS
'An old man and old woman lived in a bad house, and they were sick a lot, and they coughed, and ...' (Mangi 1:00-1:06)

Stimulus-affect EPs prefer to be followed by different-subject marking even if the next clause's subject is the experiencer of the EP. In (388), the first bracketed clause is headed by an EP with stimulus noun *nungurV* 'lack of desire' and 1SG experiencer. The next bracketed clause is headed by the verb *noroka-* 'think', whose subject is the same 1SG referent. Different-subject marking is used from the former into the latter, so the experiencer is not the pivot of the EP clause.

(388) *We-ka* *nge* *puhi* *apokha* *nake*
alongland-there TOPIC money place because

[no *yo-wo* *nungúrV* *ya-**ne**]* *ngae*
here come-1SG.POT lack.of.desire hit.1SG-3SG.DS thus

[khái-tara *ko-khu* *noroka-ra].*
stay-1SG.FUT say-1SG.SS.R think-1SG.R
'It's a place with lots of money over there, so I didn't want to come back here and was thinking of staying.' (Kimbe 3:45-3:55)

In (389), both clauses are stimulus-affect EPs. The morphological subject of the first EP is the stimulus noun *mbukura* 'cold', and that of the second EP is *nungurV* 'lack of desire'. The experiencer of both clauses is 1SG, but different-subject marking is preferred between them. Same-subject marking was only marginally accepted.

(389) *[Ima* *mbukura* *ya-**ne*** *(?ya-**mu**)]* *[nungúrV* *ya-re].*
1SG cold hit.1SG-3SG.DS hit.1SG-3SG.SS.R lack.of.desire hit.1SG-3SG.R
'I feel cold and I'm sick of it.' (elicited)

Whether (389) would have been acceptable instead with *ya-* 'affect.1SG' marked with *-khu* '1SG.SS.R' (such that the same-subject affix reflected the 1SG experiencer rather than its own verb's morphological 3SG subject) is a question for future elicitation.

(390) shows a stimulus-affect EP that must be followed by different-subject marking, disallowing same-subject altogether. The 3SG morphological subject of the EP, the stimulus noun *nungurV* 'lack of desire', is not the same as the subject of *we-* 'go' in the next clause. Thus, the pivot is the morphological subject, not the experiencer.

(390) *[NungúrV* *a-**ne*** *(*a-**mu**)]* *[wo-re].*
lack.of.desire hit.3SG-3SG.DS hit.3SG-3SG.SS.R go-3SG.R
'That person got sick of it and left.' (elicited)

(391) shows a possessive EP preceded by same-subject marking. The subject of the verb *nge-* 'see' is 3PL. The next clause uses the possessive EP *khane- maya i-* 'be happy' (lit. 'one's inside becomes good'). Its experiencer is the same 3PL referent as the subject of *nge-* 'see', but its morphological subject is the noun *khane-* 'inside'. The verb *nge-* 'see' is marked for same subject, so the pivot is the experiencer.

(391) *[Nua nge-nua-**pi**] [khane-noka máyă i-ne] ...*
3DU see-3DU.OBJ-3PL.SS.R inside-3PL.POSS good become-3SG.DS
'They$_i$ saw the two of them$_j$ and were happy, and ...' (Dokta 5:43-5:49)

(392) shows a possessive EP preceded by different-subject marking. The subject of the verb *i-* 'eat' is 3PL. The next clause uses the possessive EP *ona- a-* 'be full from eating' (lit. 'it hits one's bones'). Its experiencer is the same 3PL referent as the subject of *i-* 'eat' in the preceding clause, but its morphological subject is impersonal 3SG. The verb *i-* 'eat' is marked for different subject, so the experiencer is not the pivot here.

(392) *Opopo uki-pa [i-**pina**] [ona-noka a-ne] kha ...*
food cook-1SG.DS eat-3PL.DS bone-3PL.POSS hit.3SG-3SG.DS and
'I cooked food, and they ate until they were full, and ...' (HG 2:25-2:28)

A more systematic study of the factors determining when the experiencer is treated as the pivot is a topic for further research.

8.1.9 Clause chains expressing manner

Some clause chains combine a predicate with a following verb that expresses the manner in which the predicate is done. In (393), the verb *khonóngi-* 'do thus' (derived from a deictic, §4.6) adds manner meaning to the preceding verb in the chain: *moko e-* 'tell a story' (a light verb construction, §7.3.11).

(393) *Ngu moko e-pi **khonónge-pi**.*
3SG story take-3PL.SS.R do.thus-3PL.R
'This how the story goes.' (lit. 'They tell its story and do like this.') (Papa 0:18-0:22)

In (394), the verb *sáhi-* 'make, create, fix, prepare, do well' is used in its sense of 'do well' to add manner meaning to the preceding verb in the chain: *khupi-* 'sharpen, carve'.

(394) *Nohe* *khupi-pi* ***sáhe-pi***.
bamboo sharpen-3PL.SS.R do.well-3PL.R
'They sharpen the bamboo well.' (Kaikai 2:30-2:33)
(lit. 'They sharpen the bamboo and do well.')

8.1.10 Clause chains expressing periphrastic aspect

Some subtypes of clause chains express aspectual meanings using a structure similar to a main verb followed by an auxiliary verb. However, the two verbs are in separate clauses rather than the same clause. To reflect this difference, I refer to the aspect-marking verb as a *chained auxiliary*.

8.1.10.1 Clause chains expressing durative

Durative meaning emphasizes that an event or state keeps going on for a long time. This meaning is expressed in Hurukui using the verb *we-* 'go' as a chained auxiliary.

Sometimes the durative chained auxiliary is marked for the same subject as the main verb. In (395), the verb *i-* 'eat, drink' and the chained auxiliary *we-* 'go' have the same subject, shown by SS marking on *i-* 'eat'.

(395) ***I-mu*** ***we-mu*** *ambúkha-mu* *kha ...*
eat-3SG.SS.R go-3SG.SS.R finish-3SG.SS.R and
'He drank it until he finished it, and ...' (VD1 2:48-2:51)

Sometimes, however, the durative chained auxiliary is marked for impersonal 3SG subject. In (396), the durative-marked verb is *khai-* 'stay' (seen here in its allomorph *kha-*), and it bears DS marking leading into the next clause. The chained auxiliary *we-* 'go' itself bears SS marking even though the next clause resumes the 1PL subject. This is an example of non-canonical SS marking.

(396) ***Kha-puna*** ***we-mu*** *haringuni* *apoki* *ya-ka*
stay-1PL.DS go-3SG.SS.R night two up-there

isí-pukhai-pu *kha ...*
sleep-1PL.PRF.R-1PL.SS.R and
'We stayed for a while, we slept up there for two nights, and ...' (Dokta 2:52-3:01)

The difference in meaning between SS and DS marking before the durative chained auxiliary, if any, is a question for future investigation.

8.1.10.2 Clause chains expressing perfective

Perfective meaning emphasizes that an event is finished. This contrasts with *perfect* meaning (§5.2.4.1), in which there is an emphasis that the end state of a finished event still endures, and with simple realis marking (§5.3.1), in which the main function expressed is that the event happened.

Perfective meaning is expressed in Hurukui by clause chains in which a chained auxiliary follows the content verb. The most common verb root used for this purpose is *a-* 'PFV'[21]. In (397), the chained auxiliary is preceded by a chain of SS-marked verbs.

(397) *Opitopu* ***uki-pu*** ***i-pu*** ***a-pu*** *kha* *ise-pu.*
food:1PL.POSS cook-1PL.SS.R eat-1PL.SS.R PFV-1PL.SS.R and sleep-1PL.R
'We finished cooking and eating our food and then we went to sleep.' (Bibuai 4:50-4:53)

The perfective chained auxiliary *a-* can also mark aspect in other kinds of clause chains. In (398), the chained auxiliary *a-* 'PFV' has scope over the preceding chain that expresses a periphrastic causative (§8.1.11). The first clause's subject is the causing agent with 1PL reference, and the second clause's subject is the theme whose motion is caused, agreeing for 3SG subject. The perfective chained auxiliary marks 1PL subject, agreeing with the causing agent, not the theme, because the point being made is that the action of sending was complete, not necessarily that the theme had arrived.

21 Whether this derives from *a-* 'put', *a-* 'affect.3SG', or something else is a question for future research.

(398) *Moko* *hi-puna* *hamiha* *apoki* *Ragidzaria* *we-ne*
speech cause-1PL.DS clan two Ragidzaria go-3SG.DS

a-pu *kha ...*
PFV-1PL.SS.R and
'Once we had sent word to the two clans in Ragidzaria, (Bibuai 10:28-10:36)
[we did something else].'

Perfective meaning can also be expressed using the verb *áuni-* 'finish' as a chained auxiliary. An example is seen in (399).

(399) *Re-mu* ***áune-re*** *ni_ngé_ngu ...*
burn-3SG.SS.R finish-3SG.R TOPIC
'Once it's burned up, ...' (Sol 0:56-0:59)

8.1.11 Clause chains expressing periphrastic causative

Clause chains using different-subject marking often have causative meaning. In some cases, the DS-marked verb specifies the causing action. In (400), the verb *a-* 'hit.3SG' describes the causing action, and *umi-* 'die' describes the caused event.

(400) *A-mana* *uma-Ø!*
hit.3SG-2SG.DS die-3SG.IMP
'Kill it!' (lit. 'You hit it and may it die!') (overheard)

In many situations, however, the causing action is not expressed, but causative meaning is still expressed. This is done by using *hi-* 'do, make' as the DS-marked verb. I gloss this use as 'cause'. An example is shown in (401), where the caused event is denoted by the verb *we-* 'go', creating the meaning 'send'.

(401) *Martin* ***hi-puna*** *pahe* *we-mu* *hauphe* *riki-mu* *yo-re.*
Martin cause-1PL.DS village go-3SG.SS.R betel.nut pick-3SG.SS.R come-3SG.R
'We sent Martin to the village, and he picked betel nut and (Buai 0:14-0:19)
came back.'

8.2 Clause juxtaposition

Clauses in Hurukui can be juxtaposed directly, with no morphology or other marking joining the two. I consider it juxtaposition rather than subordination when neither clause

functions as a modifier or argument of the other. It is a question for future research whether prosody can elucidate the difference between juxtaposed final clauses in one sentence versus two independent sentences spoken close together in time.

8.2.1 Comparative clause juxtaposition

Hurukui expresses comparative meaning by juxtaposing clauses expressing opposite property concepts. This is shown in (402), with the clauses bracketed.

(402) *[Ima pahe pite], [na pahe anute].*
1SG house big 2SG house small
'My house is bigger than yours.' (elicited)
(lit. 'My house is big; your house is small.')

8.2.2 Imperative clause juxtaposition

Imperative and hortative clauses can be juxtaposed. This is used to command others to take part in an action that the speaker wants the whole group to do. An example is seen in (403), with the clauses bracketed.

(403) *[Opopo i-topu] [ha-khi]!*
food eat-1PL.FUT eat-2DU.IMP
'Let's eat!' (said to two others) (Kaikai>REC2 4:00-4:03)

8.3 Serial verb constructions

The serial form of a verb (§5.4.3) is mostly used as a medial verb marker, encoding same-subject meaning in the irrealis. However, it is also found in the realis, in which case I speak of *serial verb constructions* (SVCs). A serial verb construction is a string of verbs, of which all but the last one are found in their serial form. The last one carries medial, final, or subordinating morphology. The whole string of verbs denotes a complex predicate. The subject cannot intervene between the serialized verbs; it must precede them all.

Serial verb constructions are quite rare in Hurukui, as the language strongly prefers to use clause chains. When SVCs do occur, its verbs are usually working together to describe a multi-part event that is nonetheless conceived of as a single type of action, rather than a sequence of

actions. (189), reproduced below as (404), illustrates an SVC. In the bolded phrase, the serial form is used instead of SS-marked forms in a realis chain, because "going back and forth" is treated as a single type of action here.

(404) *Itae* *masohome~masohome* ***wi*** ***yo***
3SG.EMPH many~PL go.SER come.SER

wi ***yo-make-pi*** *nge.*
go.SER come-HAB-3PL.R TOPIC
'There are a lot of them (cars) that go back and forth.'[22] (Lae 1:00-1:04)

Many strings of verbs do not allow serialization in the realis, even if they are very highly collocated and semantically related. (405) shows that the verbs *e-* 'take' and *opa-* 'give.1SG' must be treated as a clause chain, not a SVC, even though 'take' and the 'give' verbs are so often found together to express events of giving.

(405) ***[E-mu*** ***/ *E]*** *opa-re.*
take-3SG.SS.R take:SER give.1SG-3SG.R
'That person gave it to me.' (elicited)

8.4 Subordinate clauses

Subordinate clauses in Hurukui can be divided into complement clauses (§8.4.1), relative clauses (§8.4.2), and adverbial clauses (§8.4.3).

8.4.1 Complement clauses

Several verbs in Hurukui can take clausal complements. One is *ko-* 'say', seen already in the sections on direct speech (§8.4.1.1) and internal speech (§8.4.1.2).

In (406), the clause *na pahe yayo* 'I shall build you a house' is the subject of the LVC *epu a-* 'be enough, be able, be all right'.

22 The final verb *yomakepi* is structurally ambiguous between *yo-make-pi* (come-HAB-3PL.R) 'they come' and *yo-ma-kepi* (come-CONT-3PL.FUT) 'they will keep coming'. Habitual meanings are also often expressed using future morphology, because repeated past behavior is the best predictor of future behavior. If this verb form is in fact future rather than realis, then the serial marking on the preceding verbs could just be irrealis same-subject marking.

(406) ***[Na*** ***pahe*** ***yay-o]*** *épŭ_a-mukhai-Ø.*
2SG house build-1SG.POT be.enough-3SG.PRF.R-3SG.R
'I can build you a house.' / 'It's okay for me to build you a house.' (elicited)

The LVC *hambukha- i-* 'like' can take complement clauses, but unusually, these follow the LVC. This is seen in (407), where the LVC is marked with a suffix *-khu*. This suffix may be 1SG.SS.R *-khu*, in which case this is non-canonical SS marking, or may be the adjectivizer *-khu* (§5.5.2), in which case this is a zero-copula construction rendered as 'the liked-thing is [you will take it away]'. The translation suggests the former analysis, but Hurukui-internal evidence corroborating this would be preferable.

(407) *Hambukhă-Ø* *i-khu* ***[na-kha*** ***e*** ***we-kanu]***.
liking-3SG.POSS do-1SG.SS.R 2SG-FOC take:SER go-2SG.FUT
'I want you to take it away.' (Nadzab 7:30-7:33)

Complement clauses are often offset with an intervening *nge* 'TOPIC', *ngu* '3SG', or their combination *ngé ngu* 'TOPIC'. In (408), the complement clause (bracketed) is marked for future tense and has conditional meaning. The structure is ambiguous between (1) the first clause being a complement subject of the nonverbal predicate *epu* 'all right' and (2) the first clause being an adverbial conditional (§8.4.3.3).

(408) *[Ihe* *ona-Ø* *e* *yo-kanu]* ***ngu*** *epu.*
pig bone-3SG.POSS take.SER come-2SG.FUT TOPIC all.right
'If you bring a pig bone, that would be good.' / (Kaikai 12:08-12:11)
'It would be a good idea for you to bring a pig bone.'

Several structures that one might expect to involve complement clauses (from a European perspective) are instead encoded using clause chains. In (409), the clause *khapekhe imamu* 'it will be drying up' is not an object of *nge-* 'see', but a separate clause in the chain.

(409) *Kanapa* *nge* *nge-**puna*** *[khapekhe* *i-ma-mu]*.
garden TOPIC see-1PL.DS dry become-CONT-3SG.FUT
'We'll see that the garden is drying up.' (Gaden3 1:25-1:30)

8.4.1.1 Direct speech construction

The *direct speech construction* is used to report direct quotes. It uses a quotative clause as a complement to the verb *ko-* 'say'. An example is shown in (410).

(410) ***"Tata ima umo wé-kè-ra"*** ***ko-ne*** *kha harae para wo-rou.*
father 1SG also go-DES-1SG.R say-3SG.DS and together LOC go-1DU.R
'He said, "Dad, I want to go too!", and we went together.' (Nadzab 1:41-1:47)

Often the particle *hi* 'EMPH' is used between the quote clause and *ko-* 'say'. Its contribution to the meaning will require further investigation, as it often does not mark emphasis in this situation. (411) shows such a usage.

(411) *"Ngu nukhi-khu yo-ra* ***hi"*** *ko-ne ...*
that.place leave-1SG.SS.R come-1SG.R EMPH say-3SG.DS
'She said, "I left it there and came.", and ...' (Shop1 2:46-2:51)

Note that the assignment of referents to persons within the direct speech clause is from the perspective of the one who spoke the quote. That is, in (410) the verb *wékèra* 'I want to go' is marked for 1SG subject, which refers to the speaker of the quote. This speaker is coreferent with the subject of the following verb *ko-ne* (say-3SG.DS) 'he said and (DS)', not the one telling the story. This kind of deictic shift is natural for direct quotes cross-linguistically.

8.4.1.2 Internal speech construction

The *internal speech construction* is structurally similar to but semantically distinct from the direct speech construction (§8.4.1.1). It also uses a clause that looks like a quote as a complement to the verb *ko-* 'say'. However, it does not actually refer to speech that was said aloud. Instead, the internal quote clause expresses a thought or desire. This is a common phenomenon in Papuan languages, discussed by Reesink (1993). For instance, in (412), the internal speech clause is followed by *noroka-* 'think', showing that it expresses a thought, not speech.

(412) ***[Khái-tara]*** ***ko-khu*** *noroka-ra nemo ...*
stay-1SG.FUT say-1SG.SS.R think-1SG.R but
'I thought of staying, but ...' (Kimbe 3:52-3:55)

The internal speech construction can also indicate thought even when the next verb is not a verb of thought. In (413), the verb *ko-* 'say' is the last verb in the internal speech chain, and the next verb *yo-* 'come' simply describes the next/associated event.

(413) ***[Ngu*** ***menae*** ***Ruben]*** ***ko-khu*** *yo-ra* *nemo* *ngu* *ome.*
3SG surely given.name say-1SG.SS.R come-1SG.R but 3SG is.not
'I came thinking that person was Ruben, but it wasn't him.' (elicited, consultants produced first)

Like direct speech clauses, indirect speech clauses have a deictic shift in which first person refers to the one thinking/feeling what is said in the clause. Example 414 shows a 1SG possessive marker in the idiomatic phrase *pokhona-o pi-* (back-1SG.POSS stab-) 'help me'. This 1SG refers to the thinker of the internal speech clause, which is coreferent with the 3SG subject of the following two verbs, not the speaker of the story.

(414) *[Nékŭ-khă* *pokhona-**o*** *pi-tomu]* *ko-mu* *noroka-ne ...*
who-FOC back-1SG.POSS stab-3SG.FUT say-3SG.SS.R think-3SG.DS
'He wondered who would help him, and ...' (Yaingo 0:39-0:44)

8.4.1.3 Desiderative same-subject using internal speech construction

The internal speech construction is the primary strategy for encoding same-subject marking on medial desiderative clauses. It usually takes the form of a clause "I will {verb}" as the complement to *ko-* 'say', which is then marked for same subject. It does not actually express a direct quotation, but something like an internal monologue that attributes intent.

Often an internal speech clause with desiderative meaning marks the verb for 1SG subject, even when the number of desirers is dual or plural. In (415), the internal speech clause contains the verb *yókara* 'I will come' even though it is a group of people wanting to come, as if each of them is separately thinking the same thought of 'I will come'.

(415) *Hita* *umo* ***[yó-kara*** *hi]* *ko-pu*
1PL.EXCL also come-1SG.FUT EMPH say-1PL.SS.R

hamiha *khaingu* *yo-mai-pu* *ngé_ngu ...*
door that come-CONT-1PL.R TOPIC
'We also wanted to come, and we came toward the door, but ...' (Asde 2:03-2:08)

Sometimes, however, the internal speech clause retains the subject number. In (416), a 1PL desiderative form is used in the complement clause because the number of desirers is plural.

(416) *Yo-pi* *[**noká-kè-pu*** *hi]* *ko-pi ...*
come-3PL.SS.R kill.3PL-DES-1PL.R EMPH say-3PL.SS.R
'They$_i$ came and they$_i$ wanted to kill them$_j$, and ...' (Kundu 3:13-3:16)

Pronominal reference within desiderative internal speech clauses does not work like it does in a direct or indirect speech clause. In those clause types, all pronominals (including affixes) in the clause are deictically shifted to be from the internal-speech perspective, as in (410) and (414). In desiderative internal speech clauses, however, pronominals are referential from the perspective of the matrix clause, and only the final first-person verb ending is deictically shifted to be from the internal-speech perspective.

This is shown in (417), where the internal speech construction has desiderative and/or purposive semantics. The object noun phrase, *hita moko wohe* 'our local language', uses a 1PL pronoun from the perspective of the storyteller, while the suffix *-́Kara* '1SG.FUT' is used from the internal-speech perspective.

(417) ***[Hita*** *moko* *wohe* ***é-kara]*** ***ko-mu*** *yó-mukhai-Ø.*
1PL.EXCL language native take-1SG.FUT say-3SG.SS.R come-3SG.PRF.R-3SG.R
'He had come to learn our local language.' (Waitman 0:34-0:41)

This mismatch in pronominal reference is most starkly shown when multiple 1SG markers occur on the same verb but have different frames of reference. In (418), the verb *yákara* literally means 'I will kill me'. The 1SG marking in the stem *ya-* 'affect.1SG', as well as the preceding overt pronoun *ima* '1SG', refer to the speaker of this whole sentence, the person who ran away for fear of being killed. The 1SG marking in the suffix *-́Kara* '1SG.FUT', however, is coreferent with the 3SG subject of the verb *ko-mu*, the person who wanted to kill the speaker.

(418) *Hamunge* *khaingu* *kepu* *khaino* *khai-mu*
man that yesterday this be.located-3SG.SS.R

[ima ***yá-kara]*** *ko-mu* *hi-ne*
1SG kill.1SG-1SG.FUT say-3SG.SS.R cause-3SG.DS

ima *ura* *wé-khŭkhay-a.*
1SG fear go-1SG.PRF.R-1SG.R
'That man was here yesterday, and he wanted to kill me, and so I fled.' (Mira 2:03-2:08)

The same phenomenon is reported by Bromley (1981: 244) for Lower Grand Valley Dani, discussed by Reesink (1993: 220).

8.4.2 Relative clauses

Relative clauses in Hurukui are internally headed (§8.4.2.1) or headless (§8.4.2.2), in contrast to the externally headed relative clauses found in European languages. Relative clauses are often, but optionally, ended with the particle *ni* 'RELZR'.

8.4.2.1 Internally headed relative clauses

Internally headed relative clauses function as noun phrases, their clausal contents describe what kind of thing is being referred to, and the head nominal is located inside the clause.

In (419), the head noun is *kunio* 'my leg/foot'. It is an instrumental adjunct of the verb *yaphe-* 'step' in the relative clause and the argument of the existential verb *khai-* in the matrix clause.

(419) ***[Kuni-o*** *maha* *yaphe-ma-ra]* *khai-Ø.*
leg-1SG.POSS ground step-HAB-1SG.R exist-3SG.R
'I have legs, with which I walk on the ground.' (Bodi 1:30-1:35)
(lit. '[I step on the ground (with) my leg] exists.')

A more complicated but illustrative example is seen in (420). It has an internally headed relative clause, headed by *imisi* 'road, custom, method', which means 'I will record a video of the method'. The concept of 'record' is rendered in Hurukui as 'make-DS go.down'. The relative clause is ended with the relativizer *ni*. Following the relative clause, there is a complement clause *ngu*

hokao 'I shall do it' which is the object of the verb *upúri-* 'show'. Inside that complement clause, the pronoun *ngu* '3SG' resumptively refers to the method (*imisi*) from the headless relative clause.

(420) *Ima* *[**imisi*** *kokope* *hi-pa* *ekha-komu* *ni]*
1SG method picture cause-1SG.DS go.down-3SG.FUT RELZR

[ngu ***hoka-o]*** *upúri-tara.*
3SG use.method-1SG.POT show-1SG.FUT
'I will show how to use the method that I will record a video of.' (Trap 0:08-0:15)

The anomalous possessive paradigm of *opopo* 'food' (Table 3.2) very likely derives from an internally headed relative clause structure. This paradigm involves future-tense verb endings, despite being nominal inflection. For instance, *opitopu* 'our food' and *opitepi* 'their food' end with allomorphs of *-Kopu* '1PL.FUT' and *-Kepi* '3PL.FUT'. I reconstruct these as having been clauses: *opo i-topu* (food eat-1PL.FUT) 'the food that we will eat' and *opo i-tepi* (food eat-3PL.FUT) 'the food that they will eat', which then univerbated. The 1SG form used the potential *-o* rather than the future *-Kara*, yielding *opio* 'my food'. The noun was *opo*, which now primarily means 'taro', rather than the reduplicated *opopo*, which now primarily means 'food' or 'thing'.

8.4.2.2 Headless relative clauses

Headless relative clauses function as noun phrases, and their clausal contents describe what kind of thing is being referred to, but the head nominal is absent. In (421), a headless relative clause *épukhaipu* '(the things that) we have' is the subject of a negative existential predicate.

(421) ***[É-pukhai-pu*** ***ni]*** *noka* *kaya* *pe.*
take-1PL.PRF.R-1PL.R RELZR 3PL long.ago NEG.EXIST
'The things that we have (now), they didn't have before.' (Lifestyle 0:34-0:38)
(lit. '[We have taken (it)] didn't exist long ago (for) them.')

8.4.3 Adverbial clauses

Adverbial clauses often lack any special marking; they are just standalone clause chains inserted into other clauses as adverbs. They often occupy the pre-verbal adverbial position (§7.3.9). An example is seen in (422), where the clause *yoho pe* 'there is/was no fire' modifies the

verb *khai-* 'live' to mean the situation or condition in which people lived. This is not a clause chain because there is no switch-reference-marked verb after *pe* 'NEG.EXIST' and before *khai-* 'live', as one would expect in a chaining structure (§8.1.7).

(422) *Namo* *noka* ***[yoho*** ***pe]*** *kha-mai-pi.*
but 3PL fire NEG.EXIST live-HAB-3PL.R
'But they lived without fire.' (Paia 2) (written only)

In (423), the clause *ima khariha khaya* 'I was a baby / I lived as a baby' functions as a time adverb.

(423) ***[Ima*** ***khariha*** ***khay-a]*** *moko* *ĕ-maki-wa-pa.*
1SG baby live-1SG.R speech take-CONT-NEG-1SG.R
'When I was a baby, I couldn't talk.' (elicited)

Some subtypes of adverbial clauses are treated in the following subsections: 'while' clauses headed by a continuative final verb (§8.4.3.1), conditional clauses headed by verbs marked with the conditional suffix (§8.4.3.2) or future suffixes (§8.4.3.3), and purposive clauses (§8.4.3.4).

8.4.3.1 'While' clauses with continuative final verb

Clause chains headed by a final verb marked for continuative aspect can serve as adverbial clauses with the meaning of 'while'. I consider these embedded as adverbial clauses, rather than matrix clauses in a chain with the next clause, because they have the same structure as the adverbial clauses in (422) and (423), which have final verbs rather than using simultaneous SR marking (§5.4.5). An example is seen in (424), where the continuative verb expresses a state.

(424) ***[Mokepe*** ***ya-ka*** ***kha-ma-re]*** *hamunge* *khunu* *yo-mu ...*
up up-there be.located-CONT-3SG.R man theft come-3SG.SS.R
'While he was up there, a thief came, and ...' (TP2 1:41-1:46)

In (425), the continuative-marked chain expresses a process.

(425) ***[Mamae*** ***yai-khu*** ***yo-ma-ra]*** *haringuni* *ya-ne ...*
mountain climb-1SG.SS.R come-CONT-1SG.R night night.fall-3SG.DS
'As I was climbing the mountain, night fell, and ...' (AsdeRaun 1:10-1:14)

8.4.3.2 Conditional clauses with conditional marker

If-then sentences are often formed with the protasis marked by the conditional suffix *-NGe* (§5.2.6). A simple example is seen in (426).

(426) *[**Ngáě-hi-nde**]* *we-pia* *hita* *nge-niá-kè-pu.*
thus-be-COND go-2DU.DS 1PL.EXCL see-2DU.OBJ-DES-1PL.R
'Okay then (lit. if it will be thus), you two go and we'll watch you.' (SkinDiwai 1:54-1:59)

Despite the formal resemblance and likely etymological connection between the topicalizer *nge* and the conditional *-NGe* (Haiman 1978), the conditional suffix is often explicitly re-topicalized by *nge* or *ngé ngu* 'TOPIC'. This is seen in (427).

(427) *Kohe* *[yo* *anga* *khaingu* *yo-**nge**]* ***ngé_ngu***
okay come:SER day that come-COND TOPIC

no *yo* *khaka-ni.*
here come:SER stay-2PL.IMP
'So if/when that day comes, y'all come stay here.' (Graun 9:44-9:48)

8.4.3.3 Conditional clauses with future final verb

If-then sentences are also often formed by the juxtaposition of two future-marked clause chains. Future-marked conditional clauses (protases) are often followed by the topic marker *nge*, a resumptive pronoun *ngu* '3SG', or their combination *ngé ngu* (Haiman 1978). Because I consider these to be subordinating structures, I have placed this section within the description of subordinate clauses, rather than equipollent clause juxtaposition. An example is seen in (428), where it is not certain if the *ngu* '3SG' should be analyzed as resumptively referring to the first clause (as if to say '(given) it/that') or as referring to the fire and thus the subject of the verb *óme-* 'light up'.

(428) *[Khane-ito* *akine-ito* *e* *harawohe* *a-**kopu**]*
inside-1PL.POSS thought-1PL.POSS take:SER one put-1PL.FUT

[ngu *óme-**komu**].*
3SG light.up-3SG.FUT
'If we combine our thoughts and feelings, it (fire) will light.' (Kaikai 3:15-3:19)

In (429), the protasis is topicalized by *nge* 'TOPIC'.

(429) *[Eme ahe í-kha-**kòmu**]* ***nge**,*
kind another become-PRF.IRR-3SG.FUT TOPIC

*[neka ekhono khekóngi-**tèpi**]?*
2PL now do.that-2PL.FUT
'If it had turned out otherwise, would y'all behave like that?' (Robert 3:01-3:05)

The protasis is sometimes preceded by the subordinator *hupí ~ hupahí* 'if', which is derived from the light verb construction *hupV-(ahe)-i-* 'thing-(INDEF)-do' in its final-stress serial form (§5.4.3). Example (430) shows this.

(430) ***Hupí** [na ihe khuha khaingu-kha húkì harae para*
if 2SG horse all that-FOC run.SER together LOC

*apokha-**kepi**]* ***nge_ngu** [na puhi phomo ahe e-**kanu**].*
go.past-3PL.FUT TOPIC 2SG money huge INDEF take-2SG.FUT
'If all your horses run together and win the race, you'll get a lot of money.' (Hos 1:44-1:53)

Counterfactual conditionals also use future marking. In (431), the nonverbal predicate *ima hioho* 'I am a woman' is marked as counterfactual by a following future-tense copula. Notably, the inflection is for invariant 3SG subject in this situation.

(431) ***[Ima hioho ĭ-tomu]** ngu, aukhu oke-o.*
1SG woman be-3SG.FUT TOPIC skirt tie-1SG.POT
'If I was a woman, I would wear a skirt.' (elicited)

8.4.3.4 Purposive clauses

Purposive clauses are headed by a verb marked with the purposive suffix *-Ka* (§5.5.1). These are adverbial clauses expressing reason. Example (432) shows one.

(432) ***[Puhi e-ka]** wopu.*
money take-PUR go:1PL.HORT
'Let's go (in order to) get money.' (Kofi 0:12-0:16)

Part II: Syntactic gradience in light verb constructions in Hurukui

9 Introduction: Gradience, categorization, and complex predicates

This second part of the dissertation focuses on synchronic gradience in light verb constructions (LVCs) in Hurukui. The language's LVCs were introduced more briefly in §7.3.11, and this part of the dissertation will refer back to the grammatical description in Part I when describing the properties of LVCs. The LVCs analyzed here provide a detailed case study of how multiple source constructions can give rise to complex predicates with different properties in synchrony. This multifaceted nature of complex predicates due to their varied diachrony has been noted by multiple researchers already.

This chapter begins by motivating the current study in §9.1. I then situating the study in the literature on gradience in grammar (§9.2) and in the literature on complex predicates and LVCs (§9.3).

9.1 Motivation and point of departure of this study

My primary, initial interest in choosing to research Hurukui was documentation and description. Over the years of working with Hurukui, however, I also became interested in gradience and the ways in which simpler models of categorization fail to capture the nuance of the real world as well as of human mental representations of the real world. As I discussed phenomena from Hurukui with other linguists, I often became frustrated at others' use of definitions that involve necessary and sufficient conditions to determine whether a certain linguistic feature is "really" an X or not. One often finds linguistic work that attempts to argue for and codify a set of such conditions for the best or most correct use of a term. An example is Cleary-Kemp's work on defining what is and is not a serial verb construction (2015). Another is Haspelmath's work on creating clear definitions of the terms *inflection* and *derivation* in morphology (2024), although he does this only retroactively (in a rather usage-based way) based on how the terms are already entrenched in the literature, rather than claiming that these definitions reflect anything in the reality of how languages are structured. From reading and discussing works like these, I became motivated to stop myself (and ideally other linguists) from saying things like "if the phenomenon meets X and Y condition, then it's a Z, and otherwise it's

not". I want us instead to be better about embracing the gradient and clustering nature of things, especially in the highly decentralized, distributed network of brain activations that we call language.

LVCs in Hurukui are a case study, a natural experiment, for looking at how categories are gradient in a linguistic system. I could have chosen some other phenomenon in Hurukui for illustrating my theoretical points, and I did consider looking at gradience in other parts of the language, but LVCs seemed to be the best domain for these purposes. In the system of Hurukui LVCs and similar constructions, there is a lot of gradience and morphosyntactic variety. There are also a lot of tokens, because these predicates are very frequently used and make up a large proportion of the verbal lexicon. The LVCs in this language are just one place in which I have seen that kind of gradience and clustering, and where I thought I could make a good argument for my conceptualization of categorization.

I speak a lot here about *univerbation*, which is a phenomenon in which multiple words (by some language-specific criteria for wordhood) become one word. Univerbation is thus characterizable on a language-specific basis because it depends on language-specific notions of how to count words. While my points about the Hurukui LVC system concern multiple aspects of syntactic gradience more generally, I see univerbation as a good way to talk in the aggregate about how multiple different features correlate and evolve together while a multi-word construction (initial + LV) becomes more entrenched, lexicalized, and fused. The gradience I am most interested in here is therefore the scale from two-word LVC to monomorphemic verb root. However, I am also interested in gradience from various source constructions to the two-word LVCs, which doesn't necessarily involve univerbation.

I now briefly describe what I mean by the term *gradience* and some of its subtypes, as well as some terms of my own creation that characterize philosophical and methodological approaches to categorization in a system exhibiting gradience.

Gradience and its subtypes

- *Gradience*: fluidity within and/or between categories, as opposed to Aristotelian all-or-none categories (Hjørland 2017)
- *Subsective gradience* (Aarts 2004a): gradience within a category, including degrees of membership and prototype effects (Rosch 1978)
- *Intersective gradience* (Aarts 2004a): gradience between categories, blending of features of both categories

Approaches to gradient categorization

- *Gradientism*: an approach accepting and embracing the existence of some kind of gradience in a phenomenon
- *Supergradientism*: an approach insisting that everything is gradient always, with no boundaries even on categories with subsective gradience
- *Clusterism*: an approach positing that what appear to us to be categories are really clusters in some feature space, whose members behave similarly enough that they can sometimes be approximated as behaving the same

All that is exists is whatever is the most fundamental physics. Human minds essentially made up everything else in order to make predictive models that help us survive and reproduce because we are beings that evolved by natural selection (Darwin 1859). We perceive entities based on how much things correlate with each other and what helps us understand and predict the world. We perceive objects in that way, and we also perceive categories in that way, by grouping entities whose properties and behavior correlate well enough that this helps us predict, understand, and communicate about things. It is very useful to approximate such a group as if it was a collection of the same kind of thing. So categories do not "really" exist, in the same way that nothing emergent exists. We created these entities in our minds. Only in that sense is it useful to ask questions about whether categories exist (i.e., is it something we perceive and use as if it were a thing). So I will sometimes say "exist", but what I mean is "is a useful tool that the mind can

reasonably have to understand the world". That is, categories "exist" subjectively but do not exist objectively.

Regarding subjectively existent categories, I do think there are boundaries. It would be highly unexpected to find a group of people who think a dog and the ocean are the same thing, and whose language reflects this mental model. Having that equivalence would be a very crude approximation to the world and would not be very useful, and in fact would be very misleading and potentially dangerous for survival.

In terms of gradience, we have metalanguage as well as language about the world. The terms we use in metalanguage are also language, and so words like "verb" are subject to the same types of gradience and prototype effects that words like "dog" are. We see things that seem to have the same properties and we decide to treat them as the same thing so that we can understand what's going on. When talking about gradience, I do presuppose that one can create categories as an approximation. So we may perceive something like a verb in a language because it has some properties, and that helps us understand the system. This creates a crude approximation to the reality, which is all this clustery, highly gradient distributional information. So one can create sharp- or fuzzy-bounded categories as a first-order approximation, and they may start either more lumpy or more splitty, with or without prototype effects, you just start somewhere, and then you refine it based on more and more nuance.

So gradience in this way is just a matter of our perception of the language. If we had a huge perfect brain that could correctly memorize all these distributional quantities/vectors, then we could just perceive thousands of word classes rather than having to think about things like nouns and verbs. But because of the way that our minds work, it's useful to approximate things in that way when trying to understand something.

I believe supergradientism is closer to the truth, but it is too extreme for humans to use to understand things.

In clusterism, everything is different from everything else. What matters is not whether two things are the same or not, but how similar they are and in what ways. Members of a category highly correlate but are not exactly the same. We should gain understanding by delineating things

first based on some amount of similarity, then subdividing them as we gather more data and desire more nuance. This demands a quantitative approach of some kind, rather than a strictly qualitative taxonomy.

I wish here to make a quick point about what definitions are, and what they are for. When trying to understand some entity, we need to know where to start looking for it in a feature space that will tell us information about how to think about its nature and behavior. The entity could be a single data point or a subset of the space. Definitions are characterizations that give us this starting point from which to branch out into more detailed and nuanced understanding. Our mental model of the entity should not start and end with some definition, but we should use the definition as a first-order approximation. It can be useful to have definitions that allow for both gradience within the category and the complete exclusion of things that are so different that including them would cause more confusion. This is similar to, but distinct from, subsective gradience. We just treat the category as having boundaries, pragmatically for our understanding, regardless of whether it "actually" does.

9.2 Gradience in morphosyntax

There is a lot of interest in gradience within and between categories, in linguistics as well as other fields. Aarts (2004b: 343) loosely defines *gradience* as "a cover term to designate a spectrum of continuous phenomena in language, from categories at the level of grammar to sounds at the level of phonetics." In the realm of syntax, gradience has been reported in the syntax of English (De Smet 2010a, 2010b) and other languages. Gradience has also been discussed at a more general and typological level, such as studies calling into question the validity of binary notions of polysynthesis (Sadock 2017). For the current study, I consider gradience to mean the existence of continua along which the phonological and morphosyntactic behaviors of linguistic items can vary without necessarily belonging to exactly one sharply defined category.

Gradience of acceptability has been investigated in the generativist tradition. For some authors, the acceptability of a given utterance is still binary, but gradience is seen in that the same syntactic structure can be made acceptable or unacceptable depending on factors such as processing difficulty given which words are used (Francis 2021: 11). For other authors,

acceptability judgments can themselves be gradient, which can be modeled as "soft constraints" within Optimality Theory (Sorace & Keller 2005). However, generativist theories often insist that gradient phenomena should be relegated to the periphery of grammar while "core" phenomena are still thought to be governed by symbolic rule-like mechanisms, a view which hinders progress in understanding the richness of language (McClelland & Bybee 2007).

In the usage-based, emergentist tradition, gradience is much more widely embraced as a natural by-product of diachronic change in patterns of usage, and an integral part of a language rather than an aberration from a core governed by strict categorization. Linguists following this school of thought have investigated gradience in word order (Levshina et al. 2023), subjecthood (Keenan 1976), and transitivity (Hopper & Thompson 1980), among various other topics.

The ontological structure of synchronic gradience is debated. Aarts (2004b: 361) introduced the distinction between subsective and intersective gradience. Subsective gradience means gradience within a category, i.e., variation of members' behaviors as long as they still show properties sufficient for belonging to the category. Intersective gradience means gradience between categories, i.e., the existence of members that show properties sufficient for membership in more than one category. Some authors challenge this distinction, contending that there is no principled way to privilege some criteria as defining word classes while other criteria are relegated to signaling gradience within those classes (Croft 2007).

Another way in which discussion of these ontological structures is framed is in terms of the lumping-splitting divide. Some analysts prefer to keep elements together in large categories to capture useful generalizations. Other analysts prefer to break down categories into subclasses, or even into more microscopic disjoint categories (Croft 2001, Bisang 2010).

A middle-of-the-road approach to the lumping-splitting debate is also possible and has become more popular in linguistics, especially from the cognitivist and construction-based traditions. Construction-based approaches to grammar are only recently beginning to integrate gradience into the question of whether something is a new/separate construction or not (Ungerer 2023). Cognitive approaches to grammar have highlighted the ability of humans' domain-general cognition to perceive the same phenomena as either homogeneous or heterogeneous depending

on context and the scope of attention (Langacker 2006). This duality between discreteness and continuity appears in the ways that grammatical items are or are not mutually activated in the minds of speakers, as well as in debates among linguists about whether certain groups of things are the same kind of linguistic phenomenon or not. Langacker considers various examples in which the same phenomenon can be viewed as either continuous or discrete, such as the positing of grammatical constructions whose full characterization involves both a prototype and extensions (2006: 139).

In the usage-based, functional theoretical approach taken here, the grammatical system has no "core" and "periphery" into which phenomena can be divided, as is found in some theories (McClelland & Bybee 2007). Rather, all successful communication is equally valid for determining which form-meaning mappings are found in the language. Thus, gradience is not only invoked to deal with messiness in less-central parts of the language, as it would be in theories positing a distinction between core and peripheral grammar. Instead, in the approach I favor, gradience is necessarily an integral part of the entire system.

Approaches to categorization that allow for gradience have become more popular in recent decades. Wittgenstein introduced family resemblance categories (1953). Rosch introduced prototypes and radial categories (1978). Prototype-based categorization has been argued to be more useful than clear-cut categories in syntactic constructions (Taylor 1998). A general approach to ontology that remains mostly confined to linguistics, to my knowledge, is Canonical Typology (Corbett 2007). Some of these approaches see fuzzy boundaries as important, while others allow for gradience of how central a member is while maintaining sharp boundaries. Of the sharp-boundary type is the approach of clustering in which points are assigned to one and only one cluster. Langacker's cognitivist position insists that whether a group of things that are different from one another is perceived as "the same" is only a matter of how far in or out we "zoom" when considering them. This is the core of my gradientist-clusterist approach, a superposition between lumping and splitting, where only the degrees and dimensions of difference matter rather than whether something is the same as something else.

Assuming that one accepts that some sort of gradience like this exists in linguistic categories, there are multiple approaches to explaining how it came about. One is connectionist models of the mind and language structure. Some say that grammatical systems are an interconnected network of items that can be associated in various ways (Diessel 2019). De Smet (2010a) takes a connectionist perspective, saying that multiple grammaticalizing items can influence each other because of spreading of activation in speakers' minds. This kind of interconnectedness between different items in a grammatical system, which has a different structure depending on what kinds of similarities/connections you are considering, is important for the present study. A gradient notion of constructionhood allows for different strengths of the generalizations involved in schema creation (Ungerer 2023). Some groups of similar instances, or similar constructions, can form tighter or looser groups in terms of behavioral properties.

One can also take a diachronic approach to explaining how synchronic gradience arises in language. Properties of diverse source constructions can persevere as items grammaticalize or otherwise undergo morphosyntactic change. Multiple source constructions can evolve into the same target construction, or at least a cluster of similar-enough outcomes. They can do this along multiple paths as well. This sort of thing will be central to my study.

One should take caution, though, in positing source constructions. Bowern (2008) warns about making syntactic folk etymologies as a linguist. Hilpert (2010) says that there are quantitative artifacts left over from source constructions even once something is reanalyzed, so using evidence from these properties could help corroborate hypotheses for source constructions.

In considering how multiple source constructions evolve into similar outcomes, the question arises of whether diachronic sinks exist. How much convergence in behavior is there? Once a sink is reached, what happens after that? Do dead ends exist? Additionally, along the way to an outcome, is a point on the trajectory destined to reach its endpoint or can it stop midway and then go somewhere else? These questions are important for the current study, where we will see that subgroups are starting to form because groups of constructions can be subject to generalization even while part of the way along a diachronic trajectory. These generalizations can

separate the subgroups from each other and lead to them becoming distinct subclasses in their behavior.

9.3 Complex predicates and light verb constructions

Complex predicates have been defined loosely as 'multiword' sequences acting like a single predicate, in which each part of the sequence contributes to the argument structure and the semantic roles assigned by the predicate (Nash & Samvelian 2016). These have attracted a lot of attention in theoretical and descriptive linguistics and are a great testing ground for questions about gradience in synchrony due to diachrony. There has been much debate about the distinctions, if any, between different subtypes of complex predicates, such as AVCs vs SVCs vs LVCs (Anderson 2006, Seiss 2009), both language-internally and cross-linguistically. This brings up many of the same questions about categorization discussed in §9.2.

This study focuses specifically on light verb constructions (LVCs), in which a verb root with little or no semantic content is combined with another element to create a predicate that needs both elements to be expressed. There is a lot of debate about whether light verbs are a legitimate category that is different from normal verbs, with authors such as Butt (2010) in favor of LVs and others such as Bruening (2016) opposed. Jespersen (1954) analyzed light verbs as being different from normal verbs in English. LVCs in many languages form complex predicates, rather than simple predicates, because they involve resolution of argument structure and other elements between the initial and the LV (Pawley 2012; Nolan 2014). LVs often encode information relating to event structure (Harris 2008; Nolan 2014; Pompei & Piunno 2023). The preverbal element is usually nonverbal in Persian (Family 2014). Many of the preverbal elements in Persian LVCs descend from bare nominal objects, but their behavior has changed such that they are no longer analyzable as objects (Megerdoomian 2012).

A large part of the debate about whether LVCs exist is due to varying degrees of compositionality and lexicalization among them. *Compositionality* means that "the meaning of a complex expression is determined by the meanings of its parts and the way they are put together" (Pagin & Westerståhl 2010: 250). My definition of *lexicalization* corresponds largely to Himmelmann's *idiomatization* (Himmelmann 2004: 27), i.e., "creating a new lexeme out of two or

more existing ones, which may continue to exist independently." More compositionality makes LVCs look more like normal predicates. Less compositionality makes them look more like idioms, i.e., lexicalized rather than grammaticalized.

Another reason for the debate about the status of LVCs is different uses of terminology by different authors. Butt uses *LVC* for some things that Anderson would call *auxiliary verb constructions (AVCs)*. In Papuan linguistics, the term *verb-adjunct construction (VAC)* is often used for what I call an LVC (Pawley 2012).

I speak of different broad levels of organization of the points into sets. I call a single point a(n) *(L)VC lexical item*, which I often abbreviate to simply *(L)VC*. For instance *moko e-* (speech take-) 'speak' is an LVC lexical item. I call a set of points an *(L)VC group*. Usually I refer to a set of points as an LVC group if they have some properties in common that warrants treating them as a group for the sake of increasing our understanding (which I consider to be the entire nature of categorization itself). For instance, *moko e-* (speech take-) 'speak', *akhawaV e-* (loud.sound take-) 'make a loud sound', and *kawV e-* (cry.out take-) 'cry out (animal)' all have the structure [{initial} take-] and have something to do with making sound, so I could refer to them (and others like them) as an LVC group. However, just because I use the term *LVC group* does not mean that I am positing that all the members of the group inherit from the same source construction, nor that the group is exhaustive. It is just a way of referring to a subset of the points in the LVC feature space.

Finally, I use the term *LVC schema* to refer to the largest generalization across form-meaning mappings that contains all the LVCs in Hurukui. That is, the LVC schema is the smallest common ancestor construction (generalization across form-meaning mappings) from which all the LVCs descend and inherit properties. My definition of *light verb construction* in §7.3.11 intends to characterize this schema.

Assuming the reader accepts that such a thing as an LVC exists and is not the same as a normal predicate, how do LVCs get created? Various source constructions (Anderson 2006) can follow various paths to become LVCs. This means they will show different relics of their sources

and diachronic paths (Hilpert 2010), which the analyst can use to corroborate hypotheses (Bowern 2008).

It has been argued that a framework that focuses on gradience and construction-based analysis will be more fruitful for analysis of phenomena involving grammaticalization (Denison 2010). The gradience in synchrony of LVCs in this study shares properties of such phenomena.

Once LVCs exist in a language, what happens to them after that? Butt has argued that LVs are a historical dead end (Butt & Lahiri 2002; Butt 2003; Butt 2010). This position has been opposed by authors such as Harris (2008), who shows how LVs have become more grammaticalized as event classifiers in the Udi language, and Anderson (2006), who details how iterated univerbation and auxiliation often causes complex predicate structures to evolve into complex verb inflection. Complex predicates have been argued to be a diachronic sink with various sources (Bowern 2008). This study will look at some ways that leftover properties of source constructions can have a large impact on future trajectories.

Previous work has found ways in which univerbation can be identified and has recognized gradience in the continua from syntax to morphology and from free words to clitics to affixes (Haspelmath 2011). However, there is a notable lack of research that takes a quantitative approach to specifying the extents to which a given construction is univerbated according to a given set of criteria. This study attempts to define a multi-dimensional space in which various criteria of phonological and morphosyntactic behavior, largely informed by questions of what should be counted as a word (Aikhenvald 2007; Zingler 2020), are combined to gain a sense of how different LVCs behave during the process of univerbation. Insights from investigating gradient univerbation in turn will bear on larger questions about categorization in language in general.

10 Properties for diagnosing univerbation of light verb constructions

In this chapter I define some terms in §10.1 before moving on to the description of the properties that I use in my analysis of univerbation of LVCs. §10.2 describes phonological properties. §10.3 describes morphological properties. §10.4 describes syntactic properties. §10.5 describes lexical and semantic properties. These properties define a multidimensional space in which a given Hurukui LVC can be located and compared to other LVCs.

10.1 Terminology

I use the term *initial* to mean the first part of an LVC in Hurukui, the element that appears before the light verb (which I often refer to as the *LV*). This choice follows the terminology of Harris (2008).

Some of the lexical items and constructions dealt with in this study are, however, not as obviously suited for the label *LVC*. For instance, a verb together with its object which is clearly a noun, both of which have heavy lexical semantics, should probably be called something else. I call those *verb-object constructions* (VOCs). More generally, I often remain conservative when deciding whether to call a given complex predicate an LVC, a VOC, or something else, and so I often use the general term *verb-construction* (VC), which is an abstraction over the terms LVC, SVC, and AVC. VCs can be LVCs, VOCs, or other sorts of things, as long as they have some initial element and some verb root. I refer to the verb in a VC very generally as simply the *verb*.

I use the term *hapax-type* to mean a morpheme that only occurs in one complex lexical item and never independently from the other part(s) of that lexical item. For instance, the verb root *ra-* 'be sleepy' only exists in the LVC *eme- ra-* (eye- be.sleepy-) 'be sleepy', and nowhere else in the language. Initials can also be hapax-type, such as that of the fused LVC *aunV i-* 'finish'; there is no word *auna*[23] */ aune / auni / auno / aunu* anywhere else in the language.

Now I move on to the list of properties that I have used to define the nature of univerbation in Hurukui LVCs (or, more generally, VCs).

23 Not counting the clearly unrelated *au-na* (maternal.uncle-2SG.POSS) 'your maternal uncle'.

10.2 Phonological and morphophonological properties

This section deals with phonological and morphophonological properties of VCs. §10.2.1 describes the property of stress position and quantity. §10.2.2 describes the domain of the mid-vowel raising allophony process (§2.2.1). §10.2.3 describes the elision of the verb root, which is common for roots consisting only of unstressed vowels (§2.5.1).

10.2.1 Stress positions and quantities

One hallmark of recently univerbated verb complexes in Hurukui is irregular stress placement. In Examples (433) and (434), primary stress is not found on the penult, its regular position in Hurukui, but instead on the LVC stem complex.

(433) ***Háukhà****-re.*
vomit-3SG.R
'That person vomited.' (elicited)

(434) ***Aráya****-mà-re.*
wander-CONT-3SG.R
'That person is wandering.' (elicited)

In *háukhà-* 'vomit', the pre-univerbation initial was *haukhV* (with a now-lost final vowel), which would have stress on the *a* of the diphthong *au* by regular stress assignment. The LV *a-* would have been stressed on the root in the immediate-tense form *a-re*. Once they were univerbated, primary stress remained on the initial and was demoted to secondary stress on the LV, giving *háukh-à-re*.

In *aráyà-* 'wander', the presence of stress on the second *a* means that that vowel was penultimate in the pre-univerbation form. It is unclear whether the complex was *arayV-a-* or *araV-ya-*, since both *a-* and *ya-* are potential LVs[24]. Regardless, we see that its continuative form had stress on the syllable *ma*, which has now been demoted to secondary stress while the initial retains primary stress.

24 A speculative etymology is *ara ahe ya-* where *ara ahe* means 'somewhere' or 'somewhere else' and *ya-* means 'step', so the phrase would have literally meant 'tread somewhere (else)' and perhaps became extended to 'walk around various places', giving the modern meaning of the LVC.

The property of stress positions and quantities states where the stresses are in certain forms of a VC, and whether those stresses are primary or secondary. I focus on the 3SG realis form and the 2SG imperative, because these often differ in stress placement. For instance, the LVC haukhV a- (vomit put-) 'vomit' has 3SG realis form *háukh-à-re* (primary-secondary) and 2SG imperative form *hàukh-á-ke* (secondary-primary).

10.2.2 Vowel raising domain

The raising of mid vowels before high vowels or other raised mid vowels (§2.2.1) often does not extend leftward beyond an old word boundary onto the initial. For instance, the do-desiderative form of verbs (§5.2.4.3) consists of the 1SG potential suffix *-Ka* (§5.3.5) plus the light verb *i-* 'do'. These are always univerbated in terms of inseparability and elision of the final vowel on *-Ka* '1SG.POT', but the *i* of the LV root does not cause raising of mid vowels on the initial verb form. Thus, one has *wé-ki-mà-ra* (go-DES-CONT-1SG.R) 'I want to go' pronounced with the first *e* in its lower allophone [e̞], not its higher allophone [e̝]. Another good example is the near-minimal pair of words *hengemu* [he̞ŋgé̞mu] 'fork' versus *héngV-e-mu* [hé̝ŋge̞mu] 'that person jumped over it, and (SS)'.

10.2.3 Verb elision

Can the verb root be elided completely? This often happens with single-vowel LV roots when they are unstressed. In (435), the LV root *i-* 'do' of the LVC *kha(i)ningV i-* (thus do-) 'do thus' is optionally elided.

(435) *Khaning-**i**-wápho-Ø!* ~ *Khaning-**Ø**-wápho-Ø!*
thus-do-PROH-2SG.IMP thus-do-PROH-2SG.IMP
'Don't do that!' (elicited, I produced)

In (436), the verb of the LVC *moko e-* (speech take-) 'talk' is elided.

(436) *OkV* *kóko* *móko-**ĕ**-ma-re.*
hunger COMIT speech-take-CONT-3SG.R
'That person is talking while hungry.' (elicited, consultants produced)

Some LVCs which have single-vowel LV roots disallow this elision even when the LV is unstressed and not necessary for disambiguating LVCs. An example of such an LVC is *sáhi-* 'fix,

prepare, straighten', which requires the *i* to be realized. In fact, this LVC is so fused that it is not clear whether its LV is the *i-* or *hi-* variant of *(h)i-* 'do', and its initial is hapax-type. I tested elision of both of these, and both were rejected, shown in (437). For simplicity of parsing, I assume the shape *sahV i-* except in the case where I was testing elision of *hi-*.

(437) *Sah-**i**-wápho-Ø!* / **Sah-**Ø**-wápho-Ø!* / **Sa-**Ø**-wápho-Ø!*
fix-do-PROH-2SG.IMP fix-do-PROH-2SG.IMP fix-do-PROH-2SG.IMP
'Don't fix it!' (elicited)

LV roots consisting of anything other than a single vowel cannot be fully elided, such as that of *haringuni ya-* (night night.fall-) 'night fall', seen in (438).

(438) *Ara haringúnĭ [**ya**-ma-re* / **Ø-ma-re].*
place night night.fall-CONT-3SG.R LV-CONT-3SG.R
'It's getting dark out.' (elicited)

10.3 Morphological properties

This section describes the morphological properties used to assess univerbation of Hurukui LVCs. §10.3.1 describes the ability of the initial to take inalienable possessor suffixes. §10.3.2 describes the ability of the verb to take object suffixes. §10.3.3 describes LVCs that can alternate between possessor-marked initials and object-marked verbs. §10.3.4 describes the ability of the initial to be reduplicated in a few different ways. §10.3.5 describes the ability of the verb to be reduplicated or to take the somewhat reduplicative same-subject simultaneous suffix.

10.3.1 Possessor on initial

Inalienable possessor suffixes (Table 3.1 in §3.1.1) are found on some initials, meaning that they are inalienable nouns.

In (439) we see 1SG possessor marking on the initial of the LVC *hambukha i-* (liking do-) 'like'. The possessor of the initial indicates the one who is liked (likee), and the subject marking on the verb indicates the one who likes someone/something (liker).

(439) *Imă **hambukhá-ŏ** e-re*
1SG liking-1SG.POSS do-3SG.R
'That person likes me.' (elicited, I produced)

In (440) we see that it is possible, but not required, to double-mark the likee as both a possessor on the initial and an object on the LV.

(440) *Hambukha-**néka*** *[i-**ke**-ra* / *e-ra].*
liking-2PL.POSS do-2PL.OBJ-1SG.R do-1SG.R
'I like y'all.' (elicited, consultants produced)

This LVC also allows the likee to be encoded only as the object of the LV, with no possessor marking on the initial. This is discussed in §10.3.2.

The LVC *hongo i-* (dream do-) 'dream', in contrast, allows optional object marking on the verb, seen in (441), but disallows possessor marking on the initial, seen in (442).

(441) *Ima* *ná* *hongŏ* *[i-**na**-ra* / *Ø-**na**-ra* / *e-ra].*
1SG 2SG dream do-2SG.OBJ-1SG.R do-2SG.OBJ-1SG.R do-1SG.R
'I dream of you.' (elicited, consultants produced)

(442) **Ima* *ná* ***hongó-nă*** *e-ra.*
1SG 2SG dream-2SG.POSS do-1SG.R
*'I dream of you.' (elicited, I produced)

10.3.2 Object on verb

Can object marking be used on the verb? For LVCs with more transitive meaning, the object-marked participant has a more undergoer-like meaning, while for LVCs with more intransitive meaning, the object-marked participant has a meaning more like beneficiary/maleficiary or external possessor.

In (443), the LVC *hungV e-* (pull take-) 'pull' marks the one being pulled as the object of the LV.

(443) *Hung-ĕ-**ya**-wápho-Ø!*
pull-take-1SG.OBJ-PROH-2SG.IMP
'Don't pull me!' / 'Don't walk so far ahead of me!' (elicited, consultants produced)

In (444), the patient of the LVC *kha(i)ningV i-* (thus do-) 'do thus' is marked by an object suffix.

(444) *Ngu* *ima* *khaníng-i-**ya**-re.*
3SG 1SG thus-do-1SG.OBJ-3SG.R
'That person did that to me.' (elicited, consultants produced)

With some LVCs about weather, the maleficiary can be marked as object. In (445), the LVC is *hauyu e-* (wind take-) 'wind blow', and the object marks someone who got caught in the wind and blown a lot.

(445) A: *Hauyŭ* *e-**ná**-re?* B: *Eo,* *hauyŭ* *e-**ya**-má-re.*
wind take-2SG.OBJ-3SG.R yes wind take-1SG.OBJ-3SG.R
A: 'Are you caught in the wind?' (elicited)
B: 'Yes, the wind is blowing me.'

In (446), the LVC is *haringuni ya-* (night night.fall-) 'night fall', and the object marker indicates someone on whom night fell in an unpleasant way (e.g., they were still walking on a trail without a flashlight).

(446) *Haringúni* *ya-**na**-komu.*
night night.fall-2SG.OBJ-3SG.FUT
'Night will fall upon you.' (elicited, consultants produced)

The LVC *hambukha i-* (liking do-) 'like', seen in §10.3.1 as taking possessor marking on its initial, can also instead mark the likee only as an object on the LV. This is seen in (447).

(447) *Ná* *hambukha* *Ø-**ná**-kha* *opopŏ* *é* *né-o.*
2SG liking do-2SG.OBJ-SS.IRR food take:SER give.2SG-1SG.POT
'I will like you and give you food.' (elicited, consultants produced)

Many LVCs do not allow object marking. Example (448) shows that the LVC *yakhV a-* (get.up put-) 'get up' cannot take an object, neither as a beneficiary, nor as a causee.

(448) **Yákh-a-**na**-re.*
get.up-put-2SG.OBJ-3SG.R
*'That person gets up for you.' / *'That person gets you up.' (elicited, I produced)

10.3.3 The possessor/object alternation

Some LVCs mean the same thing when either the initial is inalienably possessed (§10.3.1) or the verb is marked for object (§10.3.2).

The LVC *hambukha i-* (liking do-) 'like' is a good example of this. The meaning 'I like you' can be expressed literally as either 'I do your liking' or 'I liking do-you', shown in (449).

(449) *Ima [hambukhá-**nă** e-ra / hambúkhă Ø-**na**-ra].*
1SG liking-2SG.POSS do-1SG.R liking do-2SG.OBJ-1SG.R
'I like you.' (elicited, I produced)

The LVC *henga e-* 'jump over', seen in a typical use in (450), has a more unusual possessive alternant, shown in (451), in which the initial is possessed but the LV is also object marked (and its root fully elided).

(450) *Khúhu ya-khu **héngă** **e-na-ra**.*
jump go.up-1SG.SS.R jump.over take-2SG.OBJ-1SG.R
'I jumped over you.' (elicited, consultants produced)

(451) *Khúhu ya-khu **hengá-nă** **ĕ-na-ra**.*
jump go.up-1SG.SS.R jump.over-2SG.POSS take-2SG.OBJ-1SG.R
'I jumped over you.' (elicited, consultants produced)

The noun *henga* means a type of spear, so if it is the source of this initial, the semantic relationship has been significantly obscured. Perhaps there could be a metaphor between the arced flight path of a launched spear and the arced path of a person hopping over an obstacle. This etymology of course remains highly speculative. At any rate, a spear is not an inalienable noun, so the initial's ability to take possessor marking here may be due to analogy with other LVC initials that could be possessed at an earlier stage in the language, or the initial *henga* 'jump over' may be etymologically distinct from 'spear'.

The LVC *khunukha pa-* (smell stab-) 'smell' shows this alternation quite well in (452) and (453).

(452) *Eha ná khunukhà-**nă** pa-re.*
dog 2SG smell-2SG.POSS stab-3SG.R
'The dog smelled you.' (elicited, consultants produced)

(453) *Eha na khunúkhă pa-**na**-re.*
dog 2SG smell stab-2SG.OBJ-3SG.R
'The dog smelled you.' (elicited, I produced)

The LVC *ura i-* (fear do-) 'fear' shows this alternation as well, shown in (454).

(454) *Úră* *Ø-**na**-re.* ~ *Urá-**nă*** *e-re.*
fear do-2SG.OBJ-3SG.R fear-2SG.POSS do-3SG.R
'That person is afraid of you.' (elicited, consultants produced)

10.3.4 Initial reduplication

Can the initial be reduplicated? If so, what form does this take, and what meaning is expressed? Some initials can be reduplicated in full. For longer initials, this is often the only option, if reduplication is allowed at all. An example of this is the initial of the LVC *pipikiri we-* (tip.over go-) 'tip over'. In (455), the initial is fully reduplicated to encode repetitive meaning (a form-meaning mapping also found with verb roots, §5.2.1).

(455) *Hamunge* *ahe* ***pipikíri~pipikìri*** *we-má-re.*
man INDEF tip.over~REP go-CONT-3SG.R
'A man is walking and tipping from side to side.' (elicited, consultants produced)

Notice that the LV *we-* 'go' in (455) is doing double duty as the LV in the LVC 'tip over' as well as a lexical verb meaning 'go, walk'.

Some initials can be reduplicated with an ablaut process in which the first initial's vowel is *i* and the second's is *u*. When these are not reduplicated, the vowel is *u*. The initial of the LVC *khum ko-* (sprout say-) 'sprout' is an ideophone and shows this *i/u* ablaut. The meaning encoded is multiple participant number. In (456), the speaker commands a single plant to sprout, while in (457), they command multiple plants to sprout (not one plant to sprout multiple times). Note that the imperative suffix *-he* (an allomorph of *-ke* '2SG.IMP' particular to the root *ko-* 'say') is used in both of these, thus not marking participant number, because the plants are inanimate.

(456) *Khum* *ko-he!*
sprout say-2SG.IMP
'You sprout!' (elicited, I produced)

(457) ***Khim~khum*** *ko-he!*
PL~sprout say-2SG.IMP
'Y'all sprout!' (elicited, I produced)

Similar use of only reduplication, not verb indexation, to mark number is seen in (458).

(458) ***Khim~khum*** *ko-wa-Ø.*
PL~sprout say-NEG-3SG.R
'They didn't sprout.' (elicited, consultants produced)

The ideophone initial of the LVC *phum ko-* (explode say-) 'explode' behaves similarly, seen in (459) and (460).

(459) *Phúm* *kò-re.*
explode say-3SG.R
'It (one thing) exploded (once).' (elicited, consultants produced)

(460) ***Phim~phúm*** *ko-ma-re.*
PL~explode say-CONT-3SG.R
'There are many explosions / popping sounds.' (elicited, consultants produced)

The initial of the LVC *hungV e-* (pull take-) 'pull' is not as obviously an ideophone as *khum* 'sprout' and *phum* 'explode', particularly because it does not appear with the LV *ko-* 'say', but it can optionally undergo *i/u* ablaut, shown in (461). It also has a simple reduplicated variant, shown in (462). The meaning encoded is pluractional.

(461) ***Híng~hung*** *o-ra.*
PL~pull take-1SG.R
'I pulled and pulled them (e.g., plants, weeds).' (elicited, consultants produced)

(462) ***Húng-hung*** *e-khu* *yo-ra.*
PL~pull take-1SG.SS.R come-1SG.R
'I pulled and pulled them (e.g., plants, weeds) out (towards me).' (elicited, consultants, produced)

The similarly-shaped initial of the LVC *hukV i-* (run do-) 'run', however, can only be simply reduplicated and cannot undergo *i/u* ablaut, shown in (463). The meaning encoded is less clear but does not seem to mean starting and stopping running multiple times (while some reduplicated LVC initials **do** encode that kind of pluractionality).

(463) ***[Huk~huk*** / ****Hik~huk]*** *é-ke!*
run~REP PL~run do-2SG.IMP
'Run run!' (said to one person) (elicited)

Many LVC initials can be reduplicated but either lack a final vowel entirely or obligatorily lose it when reduplicated. In my dataset of LVC properties, I encode this kind of reduplication as "full except for final vowel", abbreviated as "except-V".

10.3.5 Verb reduplication

Can the verb root be reduplicated and/or take the SS.SIM morpheme *-hVhV* (§5.4.5)? Verb roots often allow full reduplication, sometimes with vowel quality changes. For example, (464) shows the root *ukumi-* 'grab, hold' reduplicated to mark repetition.

(464) ***Okómă-okòmă**-wapho-Ø.* ~ ***Ukúmĭ-ukùmĭ**-wapho-Ø.*
grab~REP-PROH-2SG.IMP grab~REP-PROH-2SG.IMP
'Don't grab it too much!' (elicited)

For the hapax-type verb root *ra-*, in the VC *eme- ra-* 'be sleepy', the SS.SIM form is the irregular *raraha* (expected **ra-haha*), shown in (465).

(465) *Eme-Ø* ***[ra-ráha*** / ****ra-háha]*** *móko* *ĕ-ma-re.*
eye-3SG.POSS be.sleepy-SS.SIM be.sleepy-SS.SIM speech take-CONT-3SG.R
'That person is talking while feeling sleepy.' (elicited, consultants produced)

The LVC *omV e-* (light.up take-) 'light up' uses the expected SS.SIM allomorph *-hehe*, but also allows *-hihi* even though there is no corresponding LVC **omV i-* (light.up do-). This is illustrated in (466).

(466) *Yoho* *ómV* ***[Ø-héhe*** / ***Ø-híhi]*** *re-ya-re.*
fire light.up take-SS.SIM do-SS.SIM burn-1SG.OBJ-3SG.R
'The fire was in the process of lighting and burned me.' (elicited, consultants produced based on my attempt)

10.4 Syntactic properties

The first group of syntactic properties involves whether certain kinds of material can be inserted between the initial and the verb. The types of insertable items considered here are modifiers of the initial (§10.4.1), time adverbs (§10.4.2), anaphora (pronominals or deictics, §10.4.3), focus-marked arguments (§10.4.4), and postpositional phrases (§10.4.5).

The second group of syntactic properties involves the behavior of the initial when used in utterances without the verb (uses where I say that the initial is *independent*). §10.4.6 describes the ability of initials to occur independently in the first place. §10.4.7 describes when initials are used clearly as nominals. §10.4.8 describes when initials can be modified by words that are known from elsewhere in the language to be adjectives (a property which is somewhat redundant with noun-like behavior, but worth investigating separately due to the restricted syntactic distribution of many LVC initials). §10.4.9 describes uses of initials clearly as adjectives.

The last subsections are not as easily grouped. §10.4.10 describes whether initials can be marked by topic or focus particles (not necessarily when independent from the verb). §10.4.11 describes when the verb can be used alone in response to a polar question, i.e., when the initial can be omitted in such responses.

10.4.1 Modifier insertion

The property of *modifier insertion* states whether the VC can be split by a modifier of the initial. This would be expected in cases where the VC is an object or adjunct noun followed by a verb, because adjectives follow nouns in Hurukui.

In (467), the VOC *opopo i-* (food eat-) 'eat (food)' is split by the quantifier *nembe* 'much', which modifies the noun *opopo* 'food'.

(467) *Opopo* ***nembe*** *ire-re.*
food much eat-3SG.R
'That person ate a lot.' (elicited)

In (468), the LVC *wo e-* (work take-) 'do work' is split by the adjective piteko 'very big', modifying the initial noun *wo* 'work, garden'.

(468) *Wo* ***pite-ko*** *e-ma-re.*
work big-INTENS take-CONT-3SG.R
'That person is doing a lot of work.' (elicited)

The indefinite article *ahe* 'a, some' can be inserted after some initials. In (469), we see this article with the LVC *pukhumb-i-* (collision do-) 'collide with, bump into'. The presence of the

article blocks the more literal/physical meaning (prototypical collisions such as running into a stone while walking) and instead allows only a more metaphorical related meaning.

(469) *Pukhúmb* ***ahě*** *ǐ-wá-u.*
collision INDEF do-NEG-1DU.R
'We did not meet with any resistance (elicited, consultants produced first)
(e.g., submitting a document).'
*'We did not collide with any physical object.'

Sometimes insertion of a modifier or other material can show what the verb is in a VC with a fully dependent initial and multiple candidate verb roots discernible from the surface form. For example, the LVC stem *aráya-* 'wander' could be from *aráV-ya-*, *aráyV-a-*, or possibly *arái-a-*. Modifier insertion, seen in (470), shows it to be *arai-ya-*, where the verb root is more semantically close to the verb *ya-* 'step' rather than the homophonous *ya-* 'go up'.

(470) *Arai* ***[yékhě*** / ***némbe]*** *ya-ma-re.*
wander a.lot a.lot step-HAB-3SG.R
'That person wanders around a lot.' (elicited)

Unsurprisingly, modifier insertion inside a monomorphemic verb root is not possible, as seen in (471) with *ekha-* 'go down'.

(471) **Sono* *ekh-* ***nembe*** *-a-ma-re.*
water (go.down-) a.lot (go.down)-HAB-3SG.R
*'That person bathes a lot.' (elicited, I produced)

Failing this test does not necessarily mean the VC is a monomorphemic root, as many LVCs also disallow modifier insertion.

10.4.2 Time adverb insertion

Time adverbs can sometimes intervene between the initial and the verb. Many LVCs prohibit time adverbs from intervening between the initial and the verb. It is a good test, however, for finding more loosely bound VCs.

The mostly univerbated LVC *omV e-* (light.up take-) 'light up' does not allow time adverbs to intervene. In (472), the adverb *araiku* 'already' is placed before the LVC stem, which is acceptable, but in (473), it is placed inside the stem, which was rejected.

(472) *Yoho* ***araiku*** *óm-e-mukhai-Ø.*
fire already light.up-take-3SG.PRF.R-3SG.R
'The fire has lit.' (elicited)

(473) **Yoho* *omV* ***araiku*** *e-mukhai-\O.*
fire light.up already take-3SG.PRF.R-3SG.R
*'The fire has lit.' (elicited)

The more loosely bound LVC *hongo i-* (dream do-) 'dream', however, does allow time adverb insertion. In (474), *kepu* 'yesterday' intervenes.

(474) *Ima* *hongŏ* ***képŭ*** *e-ra* *ni,* *ngu* *hurŭphá-kè-ra.*
1SG dream yesterday do-1SG.R RELZR 3SG describe-DES-1SG.R
'I want to tell a story about the dream I had yesterday.' (elicited)

10.4.3 Anaphora insertion

Some LVCs allow anaphora (pronominals and deictics) to intervene between the initial and the verb. In (475), the LVC *mbuye a-* (stop put-) 'stop, finish' is interrupted by the deictic *ngu* 'that place'. The context here is that the speaker wants to disembark from a bus, so the meaning of *mbuye a-* here is 'stop a journey'.

(475) *Hai* *mbuye* ***ngu*** *á-kara.*
only stop that.place put-1SG.FUT
'I'll just stop there.' (Nadzab 3:37-3:42)

10.4.4 Focused argument insertion

Can an argument marked with the focus marker *kha* intervene between the initial and the verb? In clear VOCs, this is common. In (476), the focus-marked agent *ima* '1SG' interrupts the VOC *opopo i-* (food eat-) 'eat'.

(476) *Opopŏ* ***imá-khă*** *ire-ra.*
food 1SG-FOC eat-1SG.R
'**I'm** the one who ate.' (elicited, I produced)

The VC *kuru hi-* (bird.blind make-) 'build a bird blind' is another good example, seen in (477). While it uses a verb root *hi-* 'make, do' that typically has light meaning, the verb here has more of a lexical meaning like 'build, construct'.

(477) *Kuru* ***imá-khă*** *[he-ra / e-ra].*
bird.blind 1SG-FOC make-1SG.R do-1SG.R
'**I'm** the one who built the bird blind.' (elicited, I produced)

For some LVCs too, this is allowed. This makes these LVCs seem more like VOCs. In (478), the LVC *akhawaV e-* (loud.sound take-) 'make a loud sound' is interrupted by the focused agent *ima* '1SG'. This is possible despite the phonological dependence of the initial *akhawaV* on the next word, one consequence of which is that its original final vowel has been completely lost.

(478) ***Ima-khă*** *akhawáV-o-ra.* ~ *AkhawàV* ***imá-khă*** *o-ra.*
1SG-FOC loud.sound-take-1SG.R loud.sound 1SG-FOC take-1SG.R
'I called out.' (elicited, consultants produced based on my attempt)

Most LVCs, however, disallow focused argument insertion. (479) and (480) show that the LVC *ongV i-* (rest do-) 'rest' cannot be interrupted by the focused argument *ima* '1SG'.

(479) ***Imà-khă*** *óng-i-khŭkhay-a.*
1SG-FOC rest-do-1SG.PRF.R-1SG.R
'I am resting.' (elicited, I produced)

(480) **OngV* ***imá-khă*** *i-khŭkhay-a.*
rest 1SG-FOC do-1SG.PRF.R-1SG.R
*'I am resting.' (elicited, I produced)

(481) and (482) similarly show that the LVC *yakhV a-* (get.up put-) 'get up' cannot have an inserted focused argument.

(481) ***Imá-khă*** *yákh-a-ra.*
1SG-FOC get.up-put-1SG.R
'I'm the one who got up.' (elicited, I produced)

(482) **YakhV* ***imá-khă*** *a-ra.*
get.up 1SG-FOC put-1SG.R
*'I'm the one who got up.' (elicited, I produced)

Even the idiomatic *pokhona- pi-* (back- stab-) 'help' (literally 'stab someone's back'), which is clearly syntactically a VOC and even has a more literal VOC sense 'poke/stab someone in the back' in addition to the idiomatic meaning 'help', disallows focused argument insertion. This is seen in (483) versus (484).

(483) ***Ngú-kha*** *pokhoná-o* *pe-re.*
3SG-FOC back-1SG.POSS stab-1SG.R
'That person helped me.' (elicited, I produced)

(484) **Pokhoná-o* ***ngú-kha*** *pe-re.*
back-1SG.POSS 3SG-FOC stab-1SG.R
*'That person helped me.' (elicited, I produced)

10.4.5 PP insertion

The property of PP insertion (postpositional phrase insertion) states whether a postpositional phrase can intervene between the initial and the verb.

The VC *pokhona- pi-* (back- stab-) has both a literal VOC sense 'stab in the back' and an LVC sense 'help'. Note that the translation 'stab in the back' here is meant entirely literally, not in the sense of the unrelated English idiom meaning 'betray'. This VC allows PP insertion, but the result can only be interpreted in the literal sense, as seen in (485) versus (486). These examples use the circumpositional collocation *minde ... khahu* 'together with', but I consider this to have similar enough syntax to the postpositions in Hurukui such that these examples illustrate the property of PP insertion.

(485) ***Minde*** ***akhure-Ø*** ***khahu*** *pokhoná-ŏ* *pe-rei.*
COMIT wife-3SG.POSS COMIT back-1SG.POSS stab-3DU.R
'He and his wife helped me.' (elicited, I produced)

(486) *Pokhoná-ŏ* ***minde*** ***akhure-Ø*** ***khahu*** *pe-rei.*
back-1SG.POSS COMIT wife-3SG.POSS COMIT stab-3DU.R
'He and his wife stabbed me in the back.' (elicited, I produced)
*'He and his wife helped me.'

Most LVCs disallow PP insertion. In (487) and (488), we see that the LVC *wo e-* (work take-) 'do work' cannot be split by a postpositional phrase even though *wo* 'work, garden' is a well known noun and this LVC probably came from a VOC.

(487) ***Kanápa*** ***keku*** *wó* *e-ma-re.*
garden LOC work take-CONT-3SG.R
'That person is working in the garden.' (elicited, I produced)

(488) **Wo* ***kanápa*** ***keku*** *e-ma-re.*
work garden LOC take-CONT-3SG.R
*'That person is working in the garden.' (elicited, I produced)

In (489) and (490), we see that the LVC *nama a-* (sit put-) 'sit' allows PP insertion, and that this gives a focus/emphatic meaning. The unmarked word order is in (489), where the PP precedes the complex predicate.

(489) ***Kháingu*** ***keku*** *námă-a-khŭkhay-a.*
this.place LOC sit-put-1SG.PRF.R-1SG.R
'I am sitting in this place.' (elicited, consultants produced)

(490) *Namă* ***kháingŭ*** ***keku*** *a-khŭkhay-a.*
sit this.place LOC put-1SG.PRF.R-1SG.R
'I am sitting in **this exact** place.' (elicited, consultants produced)

The LVC *mua*[25] *a-* (dig put-) 'dig up the ground (like a pig does)' allows PP insertion, as seen in (491).

25 This initial *mua* 'dig' is homophonous with the noun *mua* 'smoke, tobacco', but there is no clear semantic relationship between this and the LVC meaning, so I consider them different lexical items.

(491) *Ihe* *mua* ***kháino*** ***kekŭ*** *a-nekhai-Ø.*
pig dig here LOC put-3SG.DPST-3SG.R
'A pig dug up the ground here.' (elicited, consultants produced based on my attempt)

10.4.6 Initial independence

The property of initial independence states whether the initial is found in utterances that lack the verb. For instance, the VOC *kui hi-* (bag make-) 'make a bag', illustrated in (492), has an initial *kui* 'bag' which is a common noun that can occur in various contexts without *hi-* 'make, do'. In (493) the noun is the subject of an existential predicate.

(492) ***Kúi*** *hi-ma-re.*
bag make-CONT-3SG.R
'That person is making a bag.' (elicited, consultants produced)

(493) ***Kui*** *khai-Ø.*
bag exist-3SG.R
'There is a bag.' (elicited, I produced)

The LVC *hanguhV i/ya-* (slice.off.branch do/LV[26]-) 'slice off a branch using a knife' is an example where the two elements are fused and the initial cannot occur independently. The LVC is shown in (494), and (495) shows failed attempts to elicit the initial as an isolated word during a transcription session. The initial is now just an integral part of the verb stem and no longer has a life of its own.

(494) *Ima* *[hangúh-e-ra* / *hangúhV-ya-ra].*
1SG slice.off.branch-do-1SG.R slice.off.branch-LV-1SG.R
'I cut/remove a branch with a knife.' (Trap.Tr 29a-b)

(495) **hanguha / *hanguhe / *hanguhi / *hanguho / *hanguhu*
(not words) (Trap.Tr 32)

26 The LVs with the form *ya-* mean 'step', 'go up', and 'night fall', but these meanings are different enough from 'slice' that I cannot say which should be used for the gloss. As a result, I instead use LV meaning 'light verb'.

10.4.7 Initial used clearly as noun

The property of "initial used clearly as noun" indicates whether the initial can occur in the syntactic contexts where prototypical nouns such as *pahe* 'house' and *eha* 'dog' occur, such as the subject of a verb or nonverbal predicate, the object of a non-light verb, the argument of a postposition, a possessor, and a possessum.

In (496), the phonologically dependent noun *okV* 'hunger' is the subject, and the verb *a-* 'affect' that it goes with in the VC *okV a-* (hunger affect-) 'be hungry' is not present at all.

(496) ***OkV*** *pité-kŏ* *háy-i-mukhai-Ø.*
hunger big-INTENS apparent-become-3SG.PRF.R-3SG.R
'A big hunger arose/happened.' (elicited, consultants produced based on my attempt)

The initial *mua* of the LVC *mua a-* (dig put-) 'dig up the ground like a pig' can be used clearly as the subject of an existential predicate. In (497), it is the subject of an existential predicate, and it is also the possessum of the possessor noun *ihe* 'pig' as well as being modified by the adjective *pite* 'big' and the article *ahe* 'INDEF'. In this example, [4.181.24], it also heads an NP with three modifiers: the noun ihe 'pig', the adjective pite 'big', and the quantifier/article ahe 'a, some'.

(497) *Ihe* ***mua*** *pite* *ahe* *ara* *khaingu* *khai-Ø.*
pig dig big INDEF place this be.located-3SG.R
'There's a big area of pig-dug ground here.' (elicited, consultants produced based on my attempt)

The initial of the LVC *hongo i-* (dream do-) 'dream' is the noun *hongo*[27] 'dream'. In (498) it is seen as the subject of a nonverbal predicate, the negative existential.

(498) ***Hongo*** *pe!*
dream NEG.EXIST
'There is/was no dream.' (elicited, I produced)

In (499), the initial of the VC *kuru hi-* (bird.blind make-) 'build a bird blind house' is used independently of the LV, instead as the object of the non-light verb *nge-* 'see'.

27 Seemingly unrelated to the homophonous *hongo* 'fence'.

(499) ***Kuru*** *ngo-ra.*
bird.blind see-1SG.R
'I see a bird blind.' (elicited, I produced)

In (500), the initial of the VC *okV a-* (hunger affect-) 'be hungry' is the argument of a postposition *koko* 'COMIT'.

(500) ***OkV*** ***kóko*** *móko-ě-ma-re.*
hunger COMIT speech-take-CONT-3SG.R
'That person is talking while hungry.' (elicited, consultants produced)

10.4.8 Initial modification by adjective when independent from verb

If the initial can occur independently from the verb, can it be modified by an adjective when independent? This is a good sign that the initial is a noun, at least etymologically. The modification of *mua* 'dig' by *pite* 'big' in (497) is a good example of this.

Some initials that are not as obviously nouns can also occur both independently from the LV and modified by an adjective. In (501), the initial of *nama a-* (sit put-) 'sit' is part of a noun phrase and is modified by the adjective *humaya* 'good'. Note its meaning here of 'lifestyle, amenities', which is colexified with 'sit down' in both Hurukui and Tok Pisin (where the equivalent *sindaun* is both a noun and a verb).

(501) *Imisi* *humaya* ***nama*** ***humaya*** *ahe* *ni_ngu* *pore* *pe.*
road good sit good INDEF TOPIC INTENS NEG.EXIST
'There is no good road and no good lifestyle at all.' (Aser 1:13-1:19)

10.4.9 Initial used clearly as adjective

If the initial can occur independently from the verb, can it act as an adnominal modifier when independent? This is a good sign that the initial is an adjective, at least etymologically.

In (502), the LVC *khunu e-* (theft take-) 'steal' is the predicate of the first clause in the chain. The initial can also be used adnominally in phrases such as *hamunge khunu* (man theft) 'thief'.

(502) *Hamunge* *ahe* *opopo* ***khúnŭ*** *e-pi* *we-pínakhai-Ø.*
man INDEF food theft take-3PL.SS.R go-3PL.DPST-3SG.R
'Our food got stolen.' (lit. 'Some men stole our food and went.') (elicited)

Another good example of this is the initial *khunukha* 'smell, odor', which is found in the LVC *khunukha pa-* 'smell'. This LVC shows the possessor/object alternation, shown in (452) and (453) in §10.3.3. The fact that it can take possessor morphology diagnoses it as an inalienable noun by the criteria set forth in §3.1.1. However, it can also appear as an adnominal modifier, as in (503).

(503) *Opopo* ***khunúkhă*** *i-wapho-Ø!*
food smell eat-PROH-2SG.IMP
'Don't eat smelly food!' (elicited)

The initial *khunukha* 'smell' can also be used as a nonverbal predicate when combined with the copula *i-* 'become, be'. In this way it resembles both adjectives and nouns. This behavior is shown in (504).

(504) *Opopo* ***khunúkhă*** *i-mukhai-Ø.*
food smell become-3SG.PRF.R-3SG.R
'The food has gotten smelly.' (elicited)

10.4.10 Topic/focus marking of initial

The property of "topic/focus marking of initial" states whether the initial can be directly marked as topicalized by the topic particles (*nge*, *nge ngu*, *ni nge ngu*; §6.5.2) or focused by the focus particles (*kha*, *ni*; §§6.5.3, 6.5.5). Either of these would be a good indicator of the initial being a noun, at least originally. Verbs and adjectives do not receive this kind of topic or focus marking[28].

Many VCs disallow topic/focus marking of the initial. VCs in which the initial is a well-known separable noun, however, can sometimes do it. In (505), the noun *hauyu* 'wind', which is the initial of the LVC *hauyu e-* (wind take-) 'wind blow', is topicalized by *nge ngu*. Consultants said this would be a valid response to someone asking 'Is it also windy here?', so the topicalization of 'wind' can be taken to mean something like 'as for wind'.

28 Not counting the conditional suffix *-NGe* (§5.2.6), which is likely related to the topic marker *nge* (Haiman 1978).

(505) ***Hauyŭ*** ***ngé_ngu***, *e-ma-re.*
wind TOPIC take-HAB-3SG.R
'As for wind, (yes) it does blow.' (elicited, consultants produced based on my attempt)

In (506), the noun *hauyu* 'wind' is focused by *kha*. Consultants said this would be a valid thing to say if you saw a leaf moving, so the focus marking may have somewhat agentive semantics, a common element of the meaning of *kha* 'FOC'.

(506) ***Hauyŭ*** ***kha*** *e-ma-re.*
wind FOC take-CONT-3SG.R
'The wind is blowing.' (elicited, I produced)

The VC *eme- ra-* (eye- be.sleepy-) 'be sleepy' has the initial *eme-* 'eye', an inalienable noun. To place the experiencer in contrastive focus, the focus marker must be placed after an NP denoting the experiencer itself, as seen in (507), where the pronoun *ima* '1SG' is focused.

(507) *Na* *eme-nă* *rá-wa-Ø,*
2SG eye-2SG.POSS be.sleepy-NEG-3SG.R

nemo ***imá-khă*** ***emé-ŏ*** ***ra-re***.
but 1SG-FOC eye-1SG.POSS be.sleepy-3SG.R
'You're not tired, but **I** am.' (elicited, consultants produced based on my attempt)

In (508), we see that it is unacceptable to place the initial itself, the noun referring to one's eyes, in focus instead. Thus, the VC *eme- ra-* (eye- be.sleepy-) 'be sleepy' does not allow focus marking of the initial.

(508) **Na* *eme-na* *ra-wa-Ø,*
2SG eye-2SG.POSS be.sleepy-NEG-3SG.R

nemo *ima* ***eme-o*** ***kha*** ***ra-re.***
but 1SG eye-1SG.POSS FOC be.sleepy-3SG.R
*'You're not tired, but **I** am.' (elicited, I produced)

The structurally similar LVC *ipe- e-* (name- take-) 'call someone's name' allows its initial to be marked by the topic markers *nge* and *ngu*, but not the focus marker *kha*. This is seen in (509).

(509) *Ná* *ipè-nă* *é-wa-Ø,*
2SG name-2SG.POSS take-NEG-3SG.R

ímă ***ipè-ŏ*** ***[ngé / ngu / *kha],*** ***ó-re.***
1SG name-1SG.POSS TOPIC TOPIC FOC take-3SG.R
'That person didn't call your name, they called **my** name.' (elicited)

Notably, despite the overall contrastive focus-like meaning of the utterance in (509), the contrastive focus marker *kha* is disallowed here. The fact that *ipe-* 'name', in all of its inalienably possessed forms, is a commonplace noun that often occurs outside of LVCs likely makes it easier for speakers to mark it directly with a topic marker.

10.4.11 Verb alone in response (ellipsis)

Can the verb be uttered without the initial in a response to a polar question? In VOCs, this is often allowable (see §7.3.2 on the definite null test for covert objects).

(510) shows that the VOC *mburumu i-* (sweet.potato eat-) 'eat sweet potatoes' can optionally have its object elided in a response to a polar question.

(510) A: *Mburumu* *ire-ma?* B: *Eo,* ***(mburumu)*** ***ire-ra.***
sweet.potato eat-2SG.R yes sweet.potato eat-1SG.R
A: 'Did you eat sweet potatoes?' (elicited, consultants produced
B: 'Yes, I did.' based on my attempt)

The VC *wo hoka-* (garden cut-) has a more literal meaning 'clear an area of land for a garden' and a more idiomatic/generalized one 'do work'. In (511), we see that it allows the verb to optionally appear alone in the response to a polar question, with both meanings still available either way.

(511) A: *Na* *wo* *hoka-ma?* B: *Eo,* ***(wo)*** ***hoka-ra.***
2SG garden cut-2SG.R yes garden cut-1SG.R
A: 'Did you [cut a new garden / do work]?' (elicited, I produced)
B: 'Yes, I did.'

Most LVCs do not allow the verb to be alone in a polar response. An example is *khenge pi-* (ask stab) 'ask'. (512) shows that the initial is required in the response.

(512) A: *Khéngĕ* *pe-ma?* B: *Eo,* ***[khénge / *Ø]*** ***pe-ra***.
ask stab-2SG.R yes ask stab-1SG.R
A: 'Did you ask that person?' (elicited, I produced)
B: 'Yes, I did.'

10.5 Lexical and semantic properties

This section describes a few properties related to lexical selection and semantics. §10.5.1 describes variation in the selection of the verb for a given initial. §10.5.2 describes the availability of some more literal meaning of the verb in the VC, rather than the high semantic bleaching and opacity seen in many LVCs.

10.5.1 Variation in verb selection

Some LVCs allow multiple LVs to be used with the same meaning. The property of "variation in verb selection" states whether a VC allows this, and if so, among which set of verb roots.

The most common of these patterns is free variation between *(h)i-* 'become, make, do' and *e-* 'take'. For most speakers, both verbs have pre-imperative shape *e-*, though some speakers have a more regular pre-imperative shape *i-* for *(h)i-*, and both verbs take the 2SG.IMP suffix *-ke*. Thus, imperatives like that of the LVC *apu i/e-* (twist do/take-) 'twist rope', seen in (513), could belong to either of these LV roots.

(513) *Apu-**é**-ke!*
twist-do/take-2SG.IMP
'Twist the rope!' (elicited, consultants produced)

In (514), the same LVC shows free variation between the pre-realis shapes for each of those LV roots (from §5.1.7, recall that final *i>e* and final *e>o* in the pre-realis shape).

(514) *Uso* *[apu-**é**-re* ~ *apu-**ó**-re]?*
rope twist-do-3SG.R twist-take-3SG.R
'Did that person twist some rope?' (elicited, consultants produced based on my attempt)

For a given LVC exhibiting this variation between *i-* 'do' and *e-* 'take', however, there is often a preference for one or the other in certain parts of the paradigm. In (515) we see that the

prohibitive suffix selects the LV root *e-* 'take' and does not allow *i-* 'do'. Additionally, the LV root can be elided completely.

(515) *Usŏ* *apu-**Ø**-wápho-Ø.*
rope twist-LV-PROH-2SG.IMP

~ *apu-**e**-wápho-Ø.*
twist-take-PROH-2SG.IMP

apu-i**-wápho-Ø.*
twist-do-PROH-2SG.IMP

'Don't twist the rope.' (elicited)

This free variation is likely influenced by the facts that both of these LVs have 2SG imperative *eke* (for most speakers), and are both often subject to heavy reduction or full elision. Additionally, because of the *i*>*e* change in the pre-realis shape of *i-* 'do', the root vowel shapes of *i-* and *e-* are strongly connected in the grammatical representation network in speakers' minds (De Smet 2010a). Because of this connection, activation of part of the network is likely to spread to other parts, leading to the encroachment of one LV into the paradigm of another.

The unacceptability of forms with *i-* in parts of the paradigm where forms with *e-* are required is evidence that *e-* 'take' is the more conservative LV in *apu e-* 'twist rope'. The fact that this root is realized as the form *e-* in most of the paradigm (all except when followed immediately by a realis marker), together with the fact that many other LVCs have root vowel *e-* in the realis paradigm and also have imperative in *e-ke*, allowed the root vowel *e-* to encroach on the realis paradigm of *apu e-*. This in turn allowed analysis of these realis forms as having LV *i-* 'do' instead of *e-* 'take'. This has not, however, been extended to allowing the root vowel *i-* in other parts of the paradigm, as we would expect if there were wholesale variation between LVs *e-* and *i-* in this LVC.

10.5.2 Verb retaining some of its lexical semantics

Some LVCs have uses in which the verb is semantically more similar to its lexical counterpart, rather than just a semantically empty light verb that only exists to host verb morphology.

For example, the VC *wo hoka-* (garden cut-) has a VOC sense 'cut a new garden' along with an LVC sense 'do work'. The verb *hoka-* lexically means 'cut at a large object' or 'cut into but

not all the way through'. The noun *wo* colexifies 'garden' and 'work'. The availability of both meanings is seen in (516).

(516) A: *Na* *wo* *hoka-ma?* B: *Eo,* *(wo)* *hoka-ra.*
2SG garden cut-2SG.R yes garden cut-1SG.R
A: 'Did you [cut a new garden / do work]?' (elicited, I produced)
B: 'Yes, I did.'

11 Individual LVCs spanning the space

The purpose of this chapter is to illustrate that there are various ways in which LVCs in Hurukui can have constellations of the properties detailed in Chapter 9. There is not only one path through the space from two-word LVCs to fully univerbated LVCs. To show this, this chapter goes into detail about how four different LVCs instantiate the properties. Each of them is treated in a separate section that first illustrates the LVC's value(s) for each property, then tabulates them for easier reference.

§11.1 focuses on a very unfused LVC, *moko e-* (speech take-) 'speak', and shows how it behaves like two separate words with respect to most of the properties. §11.2 focuses on a very fused LVC, *sáhi-* 'fix, prepare, create, do well', showing how it behaves more like a single word with respect to most of the properties. §§11.3 and 11.4 each detail a different partially fused LVC so that they can be contrasted both with the first two LVCs and with each other. §11.5 combines the tables from the individual LVCs' sections into one so that they can be more easily compared.

11.1 Unfused LVC: *moko e-* 'speak'

The LVC *moko e-* (speech take-) 'speak' consists very clearly of two separate words. The initial *moko* 'speech' is a common noun (§3.1.1). The LV *e-* 'take' is semantically light, mostly contributing no meaning other than 'do', thus essentially acting as a verbalizer while the initial carries all the lexical meaning of this complex predicate. However, in some constructions a more literal meaning like 'take speech' is available.

The two words are phonologically and syntactically separate. Various kinds of material can intervene, and the initial often occurs independently (in utterances where the LV is absent), still referring to speech. When the verb *e-* 'take' is used without this initial, however, the predicate cannot mean 'speak', indicating that the initial contributes the majority or entirety of the lexical semantics. This is a quintessential light verb construction: a semantically full noun and a semantically empty verb are used together to make a predicate that can be inflected like a verb, but the two parts remain separate words and retain the usual properties of their word classes.

Stress positions and quantities

In the 3SG realis, the most usual stress pattern is primary-none, shown in (517).

(517) *M**ó**kŏ* ***o**-re.*
speech take-3SG.R
'That person spoke.' (elicited, I produced)

The none-primary pattern was mostly rejected, shown in (518).

(518) **?M**o**kŏ* ***ó**-re.*
speech take-3SG.R
'That person spoke.' (elicited, I produced)

In the 2SG imperative, the most usual stress pattern is secondary-primary, shown in (519).

(519) *M**ò**kŏ* ***é**-ke!*
speech take-2SG.IMP
'Speak!' (elicited, consultants produced)

The pattern of primary-none is also possible, shown in (520). The semantic meaning is the same, but consultants reported that this sounds angrier than (519).

(520) *M**ó**kŏ* ***e**-ke!*
speech take-2SG.IMP
'Speak!' (said angrily) (elicited, I produced)

In the 1SG potential with *-Ka* (§5.3.5), the stress pattern must be primary-none, seen in (521). The stress pattern none-primary is instead interpreted as having the pre-imperative ablauted form of the object marker *-(i)ki* '23PL.OBJ', shown in (522). Note that the structure of the verb in (522) is ambiguous between two analyses of which vowel is absent (the root *e-* 'take' or the beginning of *-(i)ki* '23PL.OBJ'), and these are shown as alternants in brackets although their surface form is the same.

(521) *M**ó**kŏ* ***e**-ka.*
speech take-1SG.POT
'I want to speak.' (elicited, consultants produced)

(522) *Mo**k**ŏ* *[**é**-ka-Ø* / *Ø-**é**ka-Ø].*
speech take-3PL.OBJ-2SG.IMP take-3PL.OBJ-2SG.IMP
'Speak to them!' (elicited, I produced)
*'I want to speak.'

Vowel raising domain

Mid-vowel raising does not spread leftward onto the initial *moko* 'speech'. Thus, forms like *moko e-ni* (speech take-2PL.IMP) have raising within the LV word only: [mo̱ko̱ é̱ni], not *[mo̱ko̱ é̱ni].

Verb elision

The verb root can be fully elided. Additionally, in accordance with Hurukui's final vowel deletion process (§2.5.1), the final vowel of the initial can also be elided, and both can be absent at the same time. These options exist in free variation, shown in (523).

(523) *Ngu* *ímă* *khahu* *[mók**ŏ*** ***ĕ**-ma-re* /
3SG 1SG COMIT speech take-CONT-3SG.R
*mók**o*** ***ĕ**-ma-re* /
speech take-CONT-3SG.R
*mók**ŏ*** ***e**-ma-re].*
speech take-CONT-3SG.R
'That person is talking with me.' (elicited, consultants produced)

In the prohibitive as well, all of these options are available, in addition to the fully enunciated realization in which neither vowel is elided. These are shown in (524). The first option produced by consultants, the most naturally occuring one, is that in which both vowels are elided: *mok-wápho*.

(524) *[Mok**o*** / *Mok**ŏ**]* *[**e**-wapho-Ø* / ***ĕ**-wapho-Ø].*
speech speech take-PROH-2SG.IMP take-PROH-2SG.IMP
'Don't speak.' (elicited)

Possessor on initial

The initial cannot be inalienably possessed. This is seen in the unacceptability of (525).

(525) **Ima na mokó-**nă** o-ra.*
1SG 2SG speech-2SG.POSS take-1SG.R
*'I spoke to/about you.' (elicited, I produced)

Object on verb

The LV can take object marking to indicate the person spoken to (semantic role of recipient). This is seen in (526), and the context given in the second clause makes the recipient role clear.

(526) *Ima na moko ě-**ná**-ra nemo epha-wa-ma.*
1SG 2SG speech take-2SG.OBJ-1SG.R but hear-NEG-2SG.R
'I spoke to you, but you didn't hear it.' (elicited, consultants produced based on my attempt)

The possessor/object alternation

This LVC does not show the possessor/object alternation, because its initial cannot take possessor marking.

Initial reduplication

The initial can be fully reduplicated to encode repetitive meaning, shown in (527).

(527) ***Moko~moko** awóhe e-ma-re.*
speech~REP too.much take-CONT-3SG.R
'That person is talking too much.' (elicited, consultants produced)

Verb reduplication

The LV *e-* 'take' cannot undergo true reduplication because of its phonological shape as just a single vowel. However, it can take the SS.SIM suffix *-hVhV*, with its root often elided in the process. This is seen in (528).

(528) *Moko* ***ĕ-héhe*** *khúhu-ya-ma-re.*
speech take-SS.SIM jump-go.up-CONT-3SG.R
'That person is jumping while talking.' (elicited, consultants produced)

This SS.SIM-marked form can also itself be used in a construction that appears to either be an LVC with the phrase *moko hehe* 'while talking' as its initial and an elided LV root (probably *i-* 'do'), or (less likely given the rest of the language's grammar) a verb form consisting of the SS.SIM form used as a stem and affixed by further TAM markers. This is seen in (529).

(529) *Sáhi-mu* *mokŏ* ***ĕ-héhe-Ø-ma-re.***
fix-3SG.SS.R speech take-SS.SIM-do-CONT-3SG.R
'That person is talking while fixing it.' (elicited, consultants produced)

Modifier insertion

Modifiers expressing extent or manner of the predicate can intervene. In (530), the adjective *pite-ko* (big-INTENS) 'very big' intervenes. This structure could be from modification of an initial noun by a modifier following it (the usual order in Hurukui, §7.1.1), or from modification of the verb by a modifier preceding it (the preverbal adverbial position, §7.3.9).

(530) *Mokŏ* ***pité-kŏ*** *o-re.*
speech big-INTENS take-3SG.R
'That person made a very big speech.' (elicited, consultants produced)

Time adverb insertion

This LVC allows time adverb insertion. In (531), the adverb *kepu* 'yesterday' intervenes. I produced this utterance intending it to be a full sentence, but it was interpreted as an internally-headed relative clause (§8.4.2.1) instead. Nonetheless, the example shows time adverb insertion.

(531) *moko* ***képŭ*** *o-ra*
speech yesterday take-1SG.R
'what I said yesterday' (elicited, I produced)

Anaphora insertion

Anaphora insertion is not observed with this LVC. Intervening tokens of *ngu* '3SG, TOPIC, that one, that place' are interpreted not as anaphora, but as topic markers. An example is seen in (532).

(532) *Sowo* *pi-tanu* *o* *moko* ***ngu*** *e*
demonstration stab-2SG.FUT or speech TOPIC take:SER

akai *we-mana* *kha* *ephá-kara.*
follow:SER go-2SG.DS and hear-1SG.FUT
'You demonstrate, or you talk through it, and I'll listen.' (Muruk 0:26-0:30)

Intervening deictics are interpreted instead as modifiers on the noun *moko* 'speech' itself, such as that seen in (533).

(533) *Moko* ***khaingu*** *é-kè-ra.*
speech this take-DES-1SG.R
'I want to tell this story.' (OlNu 0:16-0:21)
*'I want to talk to/about/with this one.'

Focused argument insertion

This LVC allows focused argument insertion. In (534), the agent *ima* '1SG', marked with *-kha* 'FOC', intervenes.

(534) *Moko* ***imá-khă*** *e-ka.*
speech 1SG-FOC take-1SG.POT
'I want to speak.' (elicited, I produced)

PP insertion

This LVC allows PP insertion. In (535), a PP occurs in preverbal position to encode contrastive focus, separating the initial from the LV.

(535) *Ngu* *moko* ***[ímă*** ***khahu]*** *e-ma-re.*
3SG speech 1SG COMIT take-CONT-3SG.R
'That person is talking with **me** (not you)!' (elicited, I produced)
(said to an interloper)

Initial independence

The initial *moko* 'speech' clearly behaves as an independent noun in Hurukui. The following examples illustrate its use in multiple roles associated with nominals (Chapter 3).

Initial used clearly as noun

The initial can be the subject of a verbal predicate. In (536), the initial *moko* 'speech' is the subject of multiple verbs in a clause chain, and the LV *e-* 'take' is absent.

(536) ***Mokŏ*** *isí-mŭ* *sahí-wa-Ø.*
speech sleep-3SG.SS.R do.well-NEG-3SG.R
'That doesn't make sense.' (elicited, consultants produced)
(lit. 'The speech doesn't sleep well.')

The initial can be the subject of a nonverbal predicate. In (537), *moko* 'speech' is the subject of a negative existential clause (§7.2.2).

(537) ***Moko*** *pe.*
speech NEG.EXIST
'There is nothing to say.' (elicited, I produced; also often overheard)
/ 'There is no speech.'

The initial can be a possessum. In (538), the initial is a noun possessed by the pronoun *ima* '1SG' in a noun-noun possessive construction (§7.1.1.1).

(538) ***[Ima*** ***moko]*** *epha-ma?*
1SG speech hear-2SG.R
'Did you hear what I said?' (elicited, I produced)

Initial modification by adjective when independent from verb

The initial *moko* 'speech' can be modified by the adjective *émeme* 'various', which is from reduplication of *eme* 'kind, sort'. This is shown in (539).

(539) *Moko* ***émeme*** *yakhé-re.*
speech various bring.out-3SG.R
'That person is bringing out all kinds of talk.' (elicited, consultants produced)

The initial can also be modified by a quantifier (§6.2). In (540), the quantifiers *yekhu* 'too much' and *awohe* 'too much' intervene.

(540) *Moko* ***yekhu*** ***awóhe*** *e-ma-re.*
speech too.much too.much take-CONT-3SG.R
'That person is talking too much.' (elicited, consultants produced)

Initial used clearly as adjective

The initial *moko* 'speech' is a noun and is not found as an adnominal modifier.

Topic/focus marking of initial

The initial can be marked with topic or focus particles, both when part of the LVC and when independent. In (541), the topic marker *ngé ngu* follows an independent noun *moko* 'speech'.

(541) *Moko* ***ngé_ngu*** *hai* *khakongu.*
speech TOPIC only that's.it
'That's it for this story.' (Hunt3 1:01-1:04)

In (542), the focus particle *kha* follows *moko* 'speech', the whole utterance expressing that even though the Hurukui tribe is spread across five different villages, they speak the same language.

(542) *Moko* ***kha*** *e-mai-pi.*
speech FOC take-HAB-3PL.R
'They speak a (one) language.' (Histry 2:27-2:31)

Verb alone in response

Despite largely acting unfused, this LVC does not allow the verb to occur alone in the response to a polar question. This is seen in (543), where only the reply containing *moko* 'speech' is accepted.

(543) A: *Tata* *mókŏ* *e-ma-re?*
father speech take-CONT-3SG.R

B: *Eo,* ***[mókŏ*** ***e-ma-re*** / ****e-má-re]***.
yes speech take-CONT-3SG.R take-CONT-3SG.R
A: 'Is father speaking?' B: 'Yes, he is.' (elicited, I produced)

Variation in verb selection

This LVC does not admit any variation in its selection of LV root (at least not without changing the meaning and/or argument structure). In (544), we see that the LV roots *i-* 'do' and *a-* 'put' are rejected.

(544) **Mókŏ* ***[e-re*** / ***a-re]***.
speech do-3SG.R put-3SG.R
*'That person spoke.' (elicited, I produced)

The verb *h(i)-* 'become, do' is used instead of *e-* 'take' in some contexts. In (545) and (546) we see two examples of this selection of verb when the initial is modified by the adjective *maisa* 'bad'. When the form with initial *h* is used, this is interpreted as a transitive predicate 'speak badly about someone' (and thus I gloss the LV as 'do'), and when the form without initial *h* is used, this is interpreted as an intransitive predicate 'be/become bad' whose subject is *moko* 'speech' (and thus I gloss the LV as 'become').

(545) *Mokŏ* *máisa* ***he****-re.*
speech bad do-3SG.R
'That person$_i$ spoke badly about that person$_j$.' (elicited, consultants produced)

(546) *Mokŏ* *máisa* ***e****-re.*
speech bad become-3SG.R
'The speech was bad.' (elicited, I produced)

Neither of these should be considered verb variation in the LVC *moko e-* 'speak', however, because it is intransitive and its subject is the agent who is speaking. In contrast, *moko maisa hi-* (speech bad do-) 'speak badly about someone', seen in (545) is transitive, with an agent subject and a theme object, and the noun phrase *moko maisa* 'bad speech' is a bare adjunct (§7.3.6). The other VC seen in (546) is just *maisa i-* (bad become-) 'go bad, be messed up', which is intransitive,

and its subject is the noun *moko* 'speech', which is thus independent from an LV. Neither of these VCs with *(h)i-* 'become, do' means the same thing as the LVC *moko e-* 'speak'.

Verb retaining some of its lexical semantics

The LV *e-* 'take' can be interpreted more literally, such that *moko e-* means something like 'get/take speech (that was spoken by someone else)' rather than 'speak'. In (547) and (548), we see a contrast in meaning in the same irrealis clause chain when serialization is used versus same-subject irrealis marking in *-KHa* (§5.4.2). In (547), the collocation *moko e-* (speech take) is interpreted as the LVC meaning 'speak', while in (548) the collocation *e- yo-* (take- come-) is interpreted as one predicate 'bring', whose object is *moko* 'speech'. This is one of the few instances in which these two ways of marking same-subject in the irrealis result in a difference in meaning, another being the difference in scope of negation seen in (371) and (372) in §8.1.4.

(547) *Moko* ***é-kha*** *yó-Ø!*
speech take-SS.IRR come-2SG.IMP
'Speak and come!' (elicited, consultants produced)

(548) *Moko* ***é*** *yó-Ø!*
speech take:SER come-2SG.IMP
'Bring word!' (elicited, I produced)

Summary

Table 11.1 shows a summary of the properties that have been shown for this LVC.

Table 11.1: Properties of the LVC *moko e-* 'speak'

Property	**Subproperty**	**Value(s)**
Stress positions and quantities	3SG.R	primary-none (p-n)
	2SG.IMP	primary-none (p-n), secondary-primary (s-p)
Vowel raising domain		separate
Verb elision		yes
Possessor on initial		no
Object on verb		yes
The possessor/object alternation		no
Initial reduplication		full
Verb reduplication		SS.SIM (2 syllables)
Modifier insertion		yes
Time adverb insertion		yes
Anaphora insertion		no
Focused argument insertion		yes
PP insertion		yes
Initial independence		yes
Initial used clearly as noun		yes
Initial modification by adjective when independent from verb		yes
Initial used clearly as adjective		no
Topic/focus marking of initial		yes
Verb alone in response		no
Variation in verb selection		no
Verb retaining some of its lexical semantics		sometimes

11.2 Fused LVC: *sáhi-* 'fix'

The LVC *sáhi-* 'fix, prepare, create, do well' is one of the most highly fused LVCs in Hurukui. In all ways except its stress position, it behaves as if it were a monomorphemic verb root. No material may intervene between the initial and the LV, and they are so fused that it is not even clear whether the parse should be *sahV i-* (fix do-) or *sa(V) hi-* (fix do-). The two must occur together at all times when the LVC's meaning is intended. The initial is hapax-type, occurring nowhere else in the language outside of this LVC. As a result, it is not clear what word class the initial originally was. If, at some point in the future of Hurukui, the stress in all forms of this complex predicate is moved to the regular penultimate position, it will be completely indistinguishable from multisyllabic verb roots such as *hina-* 'watch' and *ukumi-* 'hold'. Thus, *sáhi-* 'fix' is an example of a stem that is about as fused as one can get without being completely absorbed into the word class of verbs. Further, the behavior of this stem suggests the question of whether any modern multisyllabic verb roots came from such a source, their LVC origin now obscured. Such a question will have to await much future reconstructive work on the Madang language family.

Stress positions and quantities

In the realis with monosyllabic suffixes, the stress pattern is primary-none, shown in (549).

(549) *Kepu* ***sáhe**-ra.*
yesterday fix-1SG.R
'I fixed it yesterday.' (elicited, consultants produced)

In the 2SG imperative, the LV root must have primary stress, and the initial can be either primarily stressed as well, or unstressed. So the patterns available are primary-primary, shown in (550), and none-primary, shown in (551).

(550) *[Ná / ná-kha]* ***sáhé**-ke!*
2SG 2SG-FOC fix-2SG.IMP
'You fix it!' (elicited, consultants produced)

(551) *Yo* *s**a**h**é**-ke!*
come:SER fix-2SG.IMP
'Come fix it!' (elicited, consultants produced)

Vowel raising domain

This test does not apply to this LVC because the vowel *a* in the initial is not eligible for raising. Gaps like these can be dealt with by procedures such as *imputation* and *dimensional subsetting*, detailed in Chapter 12.

Verb elision

This LVC does not allow elision of the LV root (be it either *hi-* 'do' or *i-* 'do'). The full stem *sáhi-* must be pronounced, or sometimes the *h* alone can be elided, but the LV root's vowel *i* must be there. These options are shown in (552).

(552) *Sah-**i**-wápho-Ø.* / ***Sa**-**i**-wápho-Ø.*
fix-do-PROH-2SG.IMP fix-do-PROH-2SG.IMP
'Don't fix it.' (elicited, consultants produced)

In (553), we see that the stem cannot be realized only as either *sah-* or *sa-*.

(553) **Sah-**ǐ**-wápho-Ø.* / **Sa-**Ø**-wápho-Ø.*
fix-do-PROH-2SG.IMP fix-do-PROH-2SG.IMP
*'Don't fix it.' (elicited, I produced)

In case these restrictions on LV vowel elision are due to the stress position in the prohibitive, I also tested some continuative realis forms. The stem shape *sá-i-*, with the *h* elided as in (552), was not tested nor produced spontaneously by consultants. In (554), the most common and correct form is shown.

(554) *Sáh-**i**-ma-re.*
fix-do-CONT-3SG.R
'That person is fixing/making it.' (elicited, I produced)

In (555), we see that the form with *h* but without *i* is mostly rejected, with only marginal acceptance, and the form lacking both (and thus lacking the LV root entirely) is completely rejected.

(555) *?Sáh-***ǐ***-ma-re.* / **Sá-**Ø**-ma-re.*
fix-do-CONT-3SG.R fix-do-CONT-3SG.R
*'That person is fixing/making it.' (elicited, I produced)

Possessor on initial

The initial cannot be marked with inalienable possession morphology. This is seen in (556). Forms with an initial shape *sahV* (e.g., *saha-na*, *sahi-na*, etc.) were not tested.

(556) **Imă na sá-**nă** ì-tara.*
1SG 2SG fix-2SG.POSS do-1SG.FUT
*'I'll fix you.' (elicited, I produced)

Object on verb

The verb can be marked for object. In (557), *-na* '2SG.OBJ' marks a patient, the one being fixed.

(557) *Ima-kha sáh-i-**na**-ka.*
1SG-FOC fix-do-2SG.OBJ-1SG.POT
'I'll fix you.' (elicited, consultants produced)

The possessor/object alternation

Because the LVC does not allow possessor marking on the initial, it does not show the possessor/object alternation.

Initial reduplication

The initial can be reduplicated, with two variants. The first, shown in (558), reduplicates the *sah* portion of the stem. Its meaning is not clearly different from the non-reduplicated stem. I encode this kind of reduplication as "full except for final vowel", abbreviated as "except-V".

(558) *Némbakĕ **sah~sah**-é-ke!*
well ?~fix-do-2SG.IMP
'Fix it well!' (elicited, consultants produced)

The second reduplicated form, shown in (559), is idiosyncratic to this lexical item and gives pluractional meaning.

(559) *Némbakĕ* ***su~sah****-é-ke!*
well PL~fix-do-2SG.IMP
'Fix them well!' (elicited, consultants produced)

Verb reduplication

The LV root cannot be directly reduplicated because of its phonological form as a single vowel. However, it can received same-subject simultaneous marking with the monosyllabic allomorph *-hV* of *-hVhV* 'SS.SIM'. In fact, since the LV root may originally have had the shape *hi-* rather than just *i-*, it could be that the resulting form *sahihi* **does** reflect full reduplication of the LV root, but this is no longer synchronically clear. A usage of this form is shown in (560).

(560) ***Sàh-í-hi*** *mókŏ* *e-ma-re.*
fix-do-SS.SIM speech take-CONT-3SG.R
'That person is talking while fixing it.' (elicited, consultants produced)

Modifier insertion

This LVC does not allow insertion of any constituents between the initial and the LV at all. In (561), the modifier *maisa* 'bad, INTENS' (§6.5.4) cannot intervene.

(561) **SahV* ***máisă*** *e-re.*
fix bad do-3SG.R
*'That person fixed it badly.' / *'That person really fixed it.' (elicited, I produced)

In case the unacceptability of (561) is because the modifier *maisa* 'bad, INTENS' cannot be applied to whatever word class the initial is/was, I also tried placing known adverbs in the preverbal adverbial position (§7.3.9). However, this was also rejected. Examples (562) and (563) show that the adverb *némbake*, from the PP *nembe ake* (a.lot for) 'well, thoroughly' cannot intervene, but it can be placed before the whole stem.

(562) ***Némbake*** *[sáhe-re / sáhi-nekhài-Ø].*
well fix-3SG.R fix-3SG.DPST-3SG.R
'That person fixed it well.' (elicited, I produced)

(563) **SahV* ***némbake*** *[e-re / i-nekhài-Ø].*
fix well do-3SG.R do-3SG.DPST-3SG.R
*'That person fixed it well.' (elicited, I produced)

Time adverb insertion

Time adverbs are not allowed to intervene, shown in (564).

(564) **Sáh(i)* ***kepu*** *[e-ra / i-pakhài-Ø].*
fix yesterday do-1SG.R do-1SG.DPST-3SG.R
*'I fixed it yesterday.' (elicited, I produced)

Anaphora insertion

This property was not directly tested in elicitation, but it is completely unattested in the corpus of natural Hurukui speech. Given the complete rejection of any kind of material intervening inside this LVC, it would be unsurprising if any attempts to insert anaphora were also rejected by consultants. However, in the absence of conclusive negative data from elicitation, I have encoded this property's value as "unattested" rather than "no".

Focused argument insertion

Focused arguments must precede the stem *sáhi-*, as in (565). They cannot be inserted within it, as in (566).

(565) ***Imá-khă*** *sáhe-ra.*
1SG-FOC fix-1SG.R
'I fixed it.' (elicited, I produced)

(566) **Sah* ***ima-khă*** *e-ra. / *Sa* ***ima-khă*** *he-ra.*
fix 1SG-FOC do-1SG.R fix 1SG-FOC do-1SG.R
*'I fixed it.' (elicited, I produced)

PP insertion

This LVC does not allow PPs to intervene. This is seen in the acceptability of (567), where the PP precedes the stem *sáhi-*, and the rejection of (568), where it intervenes.

(567) ***Akhúre-Ø*** ***khahu*** *sáh-i-ma-rei.*
wife-3SG.POSS COMIT fix-do-CONT-3DU.R
'He and his wife are fixing it.' (elicited, consultants produced)

(568) **SahV* ***akhúrĕ-Ø*** ***khahu*** *i-ma-rei.*
fix wife-3SG.POSS COMIT do-CONT-3DU.R
*'He and his wife are fixing it.' (elicited, I produced)

Initial independence

The initial, be it either *sahV* or *sa(V)*, cannot occur independently of the LV. In (569), we see that attempts to use the initial as the subject of a nonverbal predicate were rejected, and consultants did not produce any corrected form from this.

(569) **[**SahV*** / ***Sa**]* *pé!*
fix fix NEG.EXIST
*'There is/was no fixing/creation.' (elicited, I produced)

The better way to express the intended meaning of (569) is to negate the whole complex predicate using verbal negation morphology (§5.2.5) and non-referential subject marking: *sah-i-**wá**-pi* (fix-do-NEG-3PL.R) 'they didn't fix it'.

Because the initial cannot occur independently of the LV, this LVC is undefined for all of the properties "initial used clearly as noun", "initial modification by adjective when independent from verb", and "initial used clearly as adjective", which are only able to be evaluated for independent initials.

Topic/focus marking of initial

This property was not directly tested in elicitation, but it is completely unattested in the corpus of natural Hurukui speech. As with the property of anaphora insertion, I cannot say for

sure whether such a thing would be rejected by consultants, although I highly suspect it, so I have encoded this property as "unattested" rather than "no".

Verb alone in response

The LV cannot appear without the initial in a response to a question. In (570), we see that the whole stem *sáhi-* must be used in the reply.

(570) A: *Nékŭ-khă* *sah-e-re?* B: *Imá-khă* ***[sah-e-ra*** / ****e-ra]***.
who-FOC fix-do-3SG.R 1SG-FOC fix-do-1SG.R do-1SG.R
A: 'Who fixed it?' B: 'I did.' (elicited)

Variation in verb selection

The LV cannot be changed; the whole stem must be *sáhi-*. In (571), we see that the LV roots *e-* 'take' (ablauted to *o-*) and *a-* 'put' are unacceptable with the initial shape *sahV*.

(571) *Sáh-**e**-re.* / **Sáh-**o**-re.* / **Sáh-**a**-re.*
fix-do-3SG.R fix-take-3SG.R fix-put-3SG.R
'That person fixed it.' (elicited, I produced)

Verb retaining some of its lexical semantics

The meanings of the LVC *sáhi-* 'fix, prepare, create, do well' all have to do with doing or making something. Because the LV *(h)i-* 'do, make' is already very semantically empty, it is difficult to say how much the lexical semantics is retained in the LVC's meaning as a whole.

Summary

Table 11.2 shows a summary of the properties that have been shown for this LVC.

Table 11.2: Properties of the LVC *sáhi-* 'fix'

Property	**Subproperty**	**Value(s)**
Stress positions and quantities	3SG.R	primary-none (p-n)
	2SG.IMP	primary-primary (p-p), none-primary (n-p)
Vowel raising domain		N/A
Verb elision		no
Possessor on initial		no
Object on verb		yes
The possessor/object alternation		no
Initial reduplication		full except for final vowel (except-V), idiosyncratic
Verb reduplication		SS.SIM (1 syllable)
Modifier insertion		no
Time adverb insertion		no
Anaphora insertion		unattested
Focused argument insertion		no
PP insertion		no
Initial independence		no
Initial used clearly as noun		N/A
Initial modification by adjective when independent from verb		N/A
Initial used clearly as adjective		N/A
Topic/focus marking of initial		unattested
Verb alone in response		no
Variation in verb selection		no
Verb retaining some of its lexical semantics		inconclusive

11.3 Partially fused LVC: *aipV e-* 'cry'

The LVC *aipV e-* (cry take-) 'cry' is intermediate between fused and unfused in various ways. It is similar in meaning and structure to *moko e-* 'speak', part of the sound take group of LVCs, but more fused. One of the ways in which it acts more like a complex verb stem than a VOC is that speakers strongly prefer to use the whole predicate, rather than just the initial, when referring to an event/bout of crying. This contrasts with the noun *moko* 'speech', which is often used to refer to speaking events/structures, such as stories or languages, with no LV in sight.

The initial *aipV* 'cry' can marginally be found independent from the LV, but it is always phonologically dependent on next word by the criterion of final vowel loss. This loss is complete; the erstwhile final vowel of the initial (if it ever had one) is unrecoverable. The initial can only very marginally be used as a noun, but largely lacks presence in any of the constructions that most readily define the word class of nominals (Chapter 3). Despite this reluctance of the initial to behave as a completely independent word, the LVC admits a variety of modifiers and arguments inserted between the initial and the LV.

Stress positions and quantities

By far the most readily produced stress pattern in the realis is primary-none, shown in (572).

(572) ***Á**ip-**o**-re.*
cry-take-3SG.R
'That person cries.' (various)

In the 2SG imperative, the most readily produced patterns are none-primary and secondary-primary, but primary-none is also used. These are both shown in (573).

(573) ***A**ip-**é**-ke!* ~ ***À**ip-**é**-ke!* ~ *(Na)* ***Á**ip-**e**-ke!*
cry-take-2SG.IMP cry-take-2SG.IMP 2SG cry-take-2SG.IMP
'Cry!' (elicited, consultants produced)

Primary-secondary was also accepted at my prompting, shown in (574).

(574) ***Áip-è-ke!***
cry-take-2SG.IMP
'Cry!' (elicited, I produced)

Vowel raising domain

As with *sáhi-* 'fix', this test does not apply to *aipV e-* 'cry' because the nucleus *ai* in the initial is not eligible for raising.

Verb elision

The LV root *e-* 'take' is often elided when unstressed. (575) shows this elision because of a following prohibitive suffix.

(575) *Aip-**e**-wápho-Ø!* ~ *Aip-**ĕ**-wápho-Ø!*
cry-take-PROH-2SG.IMP cry-take-PROH-2SG.IMP
'Don't cry!' (elicited)

In (576), elision is seen in the negative realis.

(576) *Aip-**ĕ**-wá-pi.*
cry-take-NEG-3PL.R
'They didn't cry.' / (elicited, consultants produced based on my attempt)
'There was no crying.'

Possessor on initial

The initial cannot receive inalienable possession marking. (577) shows that an attempt to do this was rejected. Part of the problem here is that the final vowel of the initial (if it ever had one) has been completely lost, and the initial is always phonologically attached to the next word by the criterion of final vowel deletion (§2.5.1). Consultants often write the word either without a final vowel or as *aipe*, the choice of silent written vowel likely influenced by the initial's strong collocation with the following LV root *e-* 'take'. I thus tested this form with the possessive suffix *-na* '2SG.POSS'.

(577) *Na aipé-**nă** o-re.
2SG cry-2SG.POSS take-3SG.R
*'That person is crying for/about you.' (elicited, I produced)

Object on verb

The beneficiary or theme, a person about which someone is crying, must be oblique, not marked using an object suffix (which can add beneficiary/maleficiary arguments for some verbs, §5.2.2.1). In (578), an attempt to add the object suffix *-na* '2SG.OBJ' to the LV was rejected, instead corrected to the PP in (579).

(578) *Na áip-ĕ-**na**-re.
2SG cry-take-2SG.OBJ-3SG.R
*'That person is crying for/about you.' (elicited, I produced)

(579) **Ná** **ake** áip-o-re.
2SG for cry-take-3SG.R
'That person is crying for/about you.' (elicited, consultants produced based on my attempt)

The possessor/object alternation

The LVC is strictly intransitive, allowing neither possessor marking on the initial nor object marking on the verb. Thus, it certainly does not exhibit the possessor/object alternation.

Initial reduplication

The initial has an idiosyncratic reduplicative form *kaipV kaipV* 'tears welling up', which is used with the verb *yo-* 'come' rather than *e-* 'take'. It is shown in (580).

(580) *(Emé-ŏ)* ***kaipV~káipV*** *yo-re.*
eye-1SG.POSS tears.welling.up come-3SG.R
'I'm about to cry.' (elicited, consultants produced)

These LVCs *aipV e-* 'cry' and *kaipV kaipV yo-* 'tears well up' are thus closely related, but the latter is not just an initial-reduplicated version of the former. Encoding this property as simply

"idiosyncratic" (meaning that the reduplicated form is irregular) would not account for the extra difference in LV, so I have instead encoded it as "idiosyncratic with verb change".

Regular full reduplication is rejected, shown in (581).

(581) **Khariha* ***aipV~aipV*** *o-re.*
baby PL~cry take-3SG.R
*'The baby cried multiple times.' / *'The baby kept crying.' (elicited, I produced)

Verb reduplication

The LV root *e-* 'take' cannot be reduplicated normally because of its shape consisting only of a single vowel. However, it does allow simultaneous same-subject marking with *-hVhV* 'SS.SIM'. An example with the shorter allomorph *-hV* is shown in (582), and an example with the longer allomorph *-hVhV* is shown in (583).

(582) *Áip-é-**he*** *wó-re.*
cry-take-SS.SIM go-3SG.R
'That person went while crying.' (elicited, consultants produced)

(583) *Àip-ĕ-**héhe*** *wo-re.*
cry-take-SS.SIM
'That person went while crying.' (elicited, consultants produced based on my attempt)

Modifier insertion

This LVC readily admits modifier insertion. In (584), the quantifier *nembe* 'a lot' intervenes.

(584) *AipV* ***némbĕ*** *e-ma-re.*
cry a.lot take-CONT-3SG.R
'That person is crying a lot.' (elicited, consultants produced)

In (585), the adjective *maisa* 'bad, INTENS' intervenes.

(585) *AipV* ***maisă*** *e-ma-re.*
cry INTENS take-CONT-3SG.R
'That person is really crying.' (elicited, consultants produced)

This LVC also admits insertion of some more idiosyncratic modifiers. In (586), *yekhu* 'too much' is used along with the more usual synonym *awohe* 'too much'. This modifier is also found with *moko e-* 'speak' (where it also appears with a following *awohe*) but nowhere else in the corpus.

(586) *AipV* ***yekhu*** *awohe* *e-mai-pi.*
cry too.much too.much take-CONT-3PL.R
'They are crying too much.' (elicited, consultants produced)

In (587), a hapax-legomenon modifier *kheikheikhV* 'softly' is used. Here I follow an unofficial transcriber convention and write *ei* instead of just *e* for its vowels because they are raised [ẹ], but the final vowel conditioning this raising is unknown.

(587) *AipV* ***khéikhéikhV*** *e-mà-re.*
cry softly take-CONT-3SG.R
'That person is crying softly/quietly.' (elicited, consultants produced)

Time adverb insertion

The LVC allows adverb insertion. In (588), we see that the adverb *kepu* 'yesterday' can optionally intervene.

(588) ***Kepŭ*** *aipV* *e-nekhai-Ø.* ~ *AipV* ***képŭ*** *e-nekhai-Ø.*
yesterday cry take-3SG.DPST-3SG.R cry yesterday take-3SG.DPST-3SG.R
'That person cried yesterday.' (elicited, consultants produced)

Anaphora insertion

This LVC is never attested with anaphora intervening between the initial and the LV.

Focused argument insertion

Focused arguments can either precede the complex predicate, as in (589), or intervene, as in (590).

(589) ***Imà-khă*** *àip-é-kara.*
1SG-FOC cry-take-1SG.FUT
'I'm going to cry.' (elicited, consultants produced)

(590) *AipV* ***imá-khă*** *e-ka.*
cry 1SG-FOC take-1SG.POT
'It's **my** turn to cry.' (elicited, I produced)

PP insertion

This LVC does not allow PP insertion. (591) and (592) show that the PP *ngu ake* (3SG for) 'for that person/reason' must precede the initial.

(591) ***Ngú*** ***ake*** *áipV-ĕ-ma-re.*
3SG for cry-take-CONT-3SG.R
'That person is crying about this.' (elicited, I produced)

(592) **AipV* ***ngú*** ***ake*** *e-ma-re.*
cry 3SG for take-CONT-3SG.R
*'That person is crying about this.' (elicited, I produced)

Initial independence

The initial *aipV* 'cry' is only very rarely used in sentences that lack the LV *e-* 'take'. This is discussed below in the subsection on the property "topic/focus marking of initial", where (601) has an independent initial. Notably, though, this was accepted and repeated only upon my production of it. Other constructions where a more independent initial would be expected to be used are rejected, shown in the subsections here on whether the initial can be used clearly as a noun, be modified by an adjective/adnominal, or be used clearly as an adjective (all of which it fails to do).

Initial used clearly as noun

The initial cannot be clearly used as a noun outside of a few constructions (such as focus marking, which is not readily produced without prompting, accounted for in the property

"topic/focus marking of initial"). In (593), we see that the initial cannot be the subject of a negative existential predicate.

(593) ***AipV*** *pe.*
cry NEG.EXIST
*'There is/was no crying.' (elicited, I produced)

In (594), we see that it also cannot head a possessive noun phrase in a negative existential predicate, a construction that normally is used to encode lack of possession, as in (239) in §6.5.1.

(594) **Ima* ***aipV*** *pe.*
1SG cry NEG.EXIST
*'I have no crying.' / *'I don't/didn't cry.' (elicited, I produced)

The correct way to express these meanings is to negate the whole complex predicate, as seen in (595) and (596) (in the latter, the 3PL subject can be interpreted as either referential or impersonal).

(595) *Aip-é-wa-Ø.*
cry-take-NEG-3SG.R
'That person didn't cry.' (elicited, consultants produced based on my attempt)

(596) *Aip-ĕ-wá-pi.*
cry-take-NEG-3PL.R
'They didn't cry.' / (elicited, consultants produced based on my attempt)
'There was no crying.'

Initial modification by adjective when independent from verb

The initial cannot occur with adnominal modification when independent (and it is rarely independent anyway). In (597), consultants rejected a noun phrase headed by *aipV* 'cry' being the subject.

(597) ****AipV*** ***kháino*** *máisă* *i-ma-re.*
cry this bad become-CONT-3SG.R
'This cry is getting bad.' (elicited, I produced)

Initial used clearly as adjective

This initial is never attested as an adnominal modifier.

Topic/focus marking of initial

The initial cannot be topic-marked. In (598), we see that trying to use the initial as the subject of a predicate is rejected. This attempt was corrected to (599), in which a whole clause (either a complement clause, §8.4.1, or a relative clause with the initial as its internal head, §8.4.2.1) is the subject.

(598) *Ímă **àipV** **ngé**, máisă e-re.
1SG cry TOPIC bad become-3SG.R
*'My cry was bad.' (elicited, I produced)

(599) Ímă **àipV** **ĕ-ma-ra** **ngé**, maisa.
1SG cry take-CONT-1SG.R TOPIC bad
'My cry was bad.' (elicited, consultants produced based on my attempt)

The LVC only reluctantly admits focus marking of the initial. In elicitation, I prompted consultants with a scenario in which a baby is crying too much and asked how they would express that the baby's crying is annoying. They first produced a clause chain, shown in (600). It shows modifier insertion, but not this property.

(600) Kopohe **aipV** awohe **e-pina** nungúrV-ya-re.
child cry too.much take-3PL.DS lack.of.desire-affect.1SG-3SG.R
'The children are crying too much and I'm sick of it.' (elicited, consultants produced)

I tried producing a structure in which the initial is independent (i.e., the LV is completely absent) and marked with *kha* 'FOC'. This was accepted and reproduced, shown in (601).

(601) Ngu **aipV** **kha** hi-ne ima nungúrV-ya-re.
3SG cry FOC cause-3SG.DS 1SG lack.of.desire-affect.1SG-3SG.R
'That person's crying is making me annoyed.' (elicited, I produced)

Verb alone in response

The LV cannot occur alone in the response to a question; the initial must be present. This is shown in (602) for positive polarity and (603) for negative polarity.

(602) A: *Áip-o-re?* B: *Eo,* ***[áip-o-re*** / ****ó-re]***.
cry-take-3SG.R yes cry-take-3SG.R take-3SG.R
A: 'Did that person cry?' B: 'Yes, they did.' (elicited, I produced)

(603) A: *Áip-o-re?* B: *Pe,* ***[aip-é-wa-Ø*** / ****é-wa-Ø]***.
cry-take-3SG.R no cry-take-NEG-3SG.R take-NEG-3SG.R
A: 'Did that person cry?' B: 'No, they didn't.' (elicited, I produced)

Variation in verb selection

The LVC does not allow variation in verb selection; the LV must be *e-* 'take'. No other LVs are attested in the corpus, and some common options are rejected in elicitation. In (604), none of the LV roots *i-* 'do', *a-* 'put', or *ko-* 'say' is accepted. The last of these indicates that the LVC does not belong to the ideophone-say group.

(604) **Áip-**e**-re.* / **Áip-**a**-re.* / **Áip-**ko**-ma-re.*
cry-do-3SG.R cry-put-3SG.R cry-say-CONT-3SG.R
*'That person cried / is crying.' (elicited, I produced)

Verb retaining some of its lexical semantics

The LVC *aipV e-* 'cry' has the LV *e-* 'take', which usually does not have any meaning of taking in this LVC. However, there is a blend between this LVC and the common use of *e-* 'take' as the first verb in clause chains expressing giving events (§8.3). Like other handling verbs, the 'give' verbs require a preceding serialized or chained *e-* 'take', unless the object being given to someone is already being held by the giver. This combines with the 'cry' LVC to create phrases calqueable as 'take a cry(ing) and give it to someone', shown in (605) and (606).

(605) *ÁipV* ***è*** *ne-kòmu.*
cry take:SER give.2SG-3SG.FUT
'That person will cry to you (so that you'll give them something).' (elicited, consultants produced)

(606) *ÁipV* ***è-mŭ*** *opá-re.*
cry take-3SG.SS.R give.1SG-3SG.R
'That person cried to me (so I would give them something).' (elicited, I produced)

Under the analysis of these examples in which a bout of crying is fictively being taken and given to someone, the verb *e-* 'take' has a combination of lexical and LVC meaning.

Summary

Table 11.3 shows a summary of the properties that have been shown for this LVC.

Table 11.3: Properties of the LVC *aipV e-* 'cry'

Property	**Subproperty**	**Value(s)**
Stress positions and quantities	3SG.R	primary-none (p-n)
	2SG.IMP	none-primary (n-p), secondary-primary (s-p), primary-none (p-n), primary-secondary (p-s)
Vowel raising domain		N/A
Verb elision		yes
Possessor on initial		no
Object on verb		no
The possessor/object alternation		no
Initial reduplication		idiosyncratic with verb change
Verb reduplication		SS.SIM (1σ)
Modifier insertion		yes
Time adverb insertion		yes
Anaphora insertion		no
Focused argument insertion		yes
PP insertion		no
Initial independence		marginal
Initial used clearly as noun		marginal
Initial modification by adjective when independent from verb		no
Initial used clearly as adjective		no
Topic/focus marking of initial	TOPIC	no
	FOC	marginal
Verb alone in response		no
Variation in verb selection		no
Verb retaining some of its lexical semantics		sometimes

11.4 Partially fused LVC: *(h)ambukha i-* 'like'

The LVC *(h)ambukha i-* (liking do-) 'like' is somewhat like a fused LVC and somewhat like an unfused LVC, but it resembles each of these in different ways than does *aipV e-* 'cry', which was showcased in §11.3. This LVC exhibits the possessor/object alternation, allowing either an inalienable possession suffix on its initial, an object suffix on its LV, or sometimes even both (double marking of its P argument). The initial can be independent, but only rarely, and in these cases it resembles a nominal. Despite this, though, it only receives inalienable possessor marking when it is part of the LVC, not in independent uses as a nominal. The LVC allows various modifiers to be inserted, but disallows insertion of heavier constituents such as PPs and argument NPs.

Stress positions and quantities

In the realis, the most common stress pattern is primary-none. In (607), this is seen with 1SG liker and 3SG likee. Note that the sequence *hambukha era* 'I like him' is structurally ambiguous between *hambukha-Ø e-ra* (liking-3SG.POSS do-1SG.R) and *hambukha e-Ø-ra* (liking do-3SG.OBJ-1SG.R) because of the zero markers for 3SG possessor and object. I thus do not gloss either of these zero markers here, nor in (609) and other examples where we have a similar situation.

(607) *Ngu hamungĕ máya nake ngu hambúkhă **e**-ra.*
3SG man good because 3SG liking do-1SG.R
'I like him because he is a good man.' (elicited, consultants produced)

Less commonly, a stress pattern of none-primary is seen, as in (608).

(608) *Ima humia ukhu ambukhă **é**-ra.*
1SG sun heat liking do-1SG.R
'I like the sun's heat.' (elicited, consultants produced)

In the 2SG imperative, the stress pattern is none-primary, shown in (609).

(609) *Ngu* *hamb**u**khă* ***é**-ke!*
3SG liking do-2SG.IMP
'Like him!' (elicited, consultants produced)

In contrast to the very avolitional semantics of 'like' in English, it seems that the liker has more agency in the Hurukui LVC *(h)ambukha i-*, so imperatives are felicitous. Perhaps better translations could include 'treat someone well' and 'be hospitable toward someone', based on the ways this LVC is used in natural-speech texts.

Vowel raising domain

As with *sáhi-* 'fix' and *aipV e-* 'cry', this test does not apply to *(h)ambukha i-* because the nuclei *a* and *u* are not eligible for raising.

Verb elision

The LV root *i-* 'do' is often elided when unstressed. In (610), for instance, the presence of the object suffix *-na* '2SG.OBJ' and realis suffix *-ra* '1SG.R' move the penultimate stress off of the LV root, and it is elided.

(610) *Ima* *hambúkhă* ***ĭ**-na-ra].*
1SG liking do-2SG.OBJ-1SG.R
'I like you.' (elicited, I produced)

Similarly, in (611), the presence of the negative suffix *-wa* 'NEG' and realis suffix *-pa* '1SG.R' move the penultimate stress off of the LV root, but here elision is optional.

(611) *Ima* *hambukhă* ***(i)**-wá-pa].*
1SG liking do-NEG-1SG.R
'I don't like that person / it.' (elicited, I produced)

Possessor on initial

This LVC often marks the liked participant as the possessor of the initial. This is shown with 1SG marking in (612) and 2SG marking in (613).

(612) *Imă* *hambukhá-**ŏ*** *e-re.*
1SG liking-1SG.POSS do-3SG.R
'That person likes me.' (elicited, I produced)

(613) *Ná* *hambukhá-**nă*** *e-ra.*
2SG liking-2SG.POSS do-1SG.R
'I like you.' (elicited, consultants produced)

Object on verb

The liked participant can also be encoded as an object on the LV *i-* 'do'. (614) shows an example of this.

(614) *Na* *ima* *hambukhă* *ĭ-**yá**-ke!*
2SG 1SG liking do-1SG.OBJ-2SG.IMP
'Like me!' (elicited, consultants produced)

The possessor/object alternation

In addition to allowing possessive marking on the initial and object marking on the LV, this LVC allows both options in many situations, with identical meaning. For instance, the meaning 'I like you' can be expressed literally as either 'I do your liking' (with possessive marking) or 'I liking do-you' (with object marking), shown in (615).

(615) *Ima* *[hambukhá-**nă*** *e-ra* / *hambúkhă* *Ø-**na**-ra].*
1SG liking-2SG.POSS do-1SG.R liking do-2SG.OBJ-1SG.R
'I like you.' (elicited, I produced)

In some situations, it is also possible to double-mark the liked participant as both a possessor on the initial and an object on the LV.

(616) *Hambukha-**néka*** *[i-**ke**-ra* / *e-ra].*
liking-2PL.POSS do-2PL.OBJ-1SG.R do-1SG.R
'I like y'all.' (elicited, consultants produced)

Initial reduplication

The initial can be fully reduplicated, seen in (617). This gives some sort of pluractional and/or intensive meaning, but the plural likee is also indexed by the object marker *-(i)ki* '23PL.OBJ' here.

(617) ***Hambúkhă~hambùkhă*** *i-kè-ra.*
PL~liking do-3PL.OBJ-1SG.R
'I like them a lot.' (elicited, consultants produced)

Verb reduplication

The LV root *i-* 'do' cannot be directly reduplicated because of its phonological form as a single vowel. However, it can received same-subject simultaneous marking with *-hVhV* 'SS.SIM', whereupon the initial vowel is elided (i.e., *i-* + *-hVhV* > *ihihi* > *hihi*). In fact, since the LV root may originally have had the shape *hi-* rather than just *i-*, it could be that the resulting form *hihi* **does** reflect full reduplication of the LV root, but this is no longer synchronically clear.

A usage of this form is shown in (618). The translation does not make too much sense in English, but the context is that the speaker heard there was something cool and went to see it, feeling excited about it along the way there.

(618) *Hambukhă* *ĭ-**híhi*** *wo-ra.*
liking do-SS.SIM go-1SG.R
'I went while liking/happy.' (elicited, consultants produced)

Modifier insertion

The LVC easily admits intervening modifiers. In (619), we see several options that all convey intensive meaning (including both *(hu)maya* 'good' and *maisa* 'bad, INTENS', which are lexically antonyms).

(619) *Hambukhă* ***[máisă / máyă / némbĕ]*** *e-ra.*
liking INTENS good a.lot do-1SG.R
'I really like that person / it.' (elicited)

Time adverb insertion

Intervention of the time adverb *kaya* 'long ago' in (620) was accepted and the sentence reproduced (as with any "I produced" examples that are not marked unacceptable), but consultants said that native Hurukui speakers do not talk this way. I thus treat this property as only marginally acceptable for this LVC.

(620) *?Ngu* *hambukhă* ***kaya*** *i-pakhai-Ø.*
3SG liking long.ago do-1SG.DPST-3SG.R
?'I used to like that person / it.' (elicited, I produced)

It is much preferred to have the time adverb before the whole complex predicate, as in (621).

(621) ***Kaya*** *hambúkhă* *i-pakhai-Ø,*
long.ago liking do-1SG.DPST-3SG.R

ekhono *nge* *hambukhă* *ĭ-wá-pa.*
now TOPIC liking do-NEG-1SG.R
'I used to like that person / it, but I don't anymore.' (elicited, consultants produced)

Anaphora insertion

This property was not directly tested for this LVC, but it is completely unattested in the corpus of natural Hurukui speech.

Focused argument insertion

This LVC does not permit focused argument insertion. We see in (622) that the focused liker *ngu* '3SG' must precede the whole complex predicate, and in (623) that it cannot intervene.

(622) ***Ngu-khă*** *hambukhá-ŏ* *e-re.*
3SG-FOC liking-1SG.POSS do-3SG.R
'That person likes me.' (elicited, I produced)

(623) **Hambukha-o* ***ngu-kha*** *e-re.*
liking-1SG.POSS 3SG-FOC do-3SG.R
*'That person likes me.' (elicited, I produced)

PP insertion

The LVC does not permit PP insertion. In (624) we see that the PP *ngu ake* (3SG for) 'for that reason' occurs before the whole complex predicate, and in (625) we see that it cannot intervene. Note that the other tokens of *ngu* '3SG' in the sentence refer to the man, not the reason.

(624)	*Ngu*	*hamungĕ*	*máya,*	***ngu***	***ake***	*ngu*	*hambúkhă*	*e-ra.*
	3SG	man	good	3SG	for	3SG	liking	do-1SG.R
	'I like him because he is a good man.'							(elicited, I produced)

(625)	**Ngu*	*hamungĕ*	*máya,*	*ngu*	*hambukhă*	***ngú***	***ake***	*e-ra.*
	3SG	man	good	3SG	liking	3SG	for	do-1SG.R
	*'I like him because he is a good man.'							(elicited, I produced)

Initial independence

The initial *hambukha* 'liking' is rarely independent from the LV *i-* 'do'. One example is shown in (626) below, in the subsection for the property "initial used clearly as noun".

Initial used clearly as noun

In the rare cases when the initial *hambukha* is used as an independent noun, it has meanings of 'liking' and also 'belief' (another, rarer meaning of the LVC is 'believe'). This noun is seen in (626), where it heads a possessive noun phrase that is the subject of a nonverbal predicate headed by the quantifier *amie* 'many'.

(626)	*Ímă*	***hambùkhă***	*amíe*	*ahe.*
	1SG	liking	many	INDEF
	'There are many things that I like / believe in.'			(elicited, consultants produced)

Unexpectedly, despite clearly being a noun in (626) and being able to carry inalienable possession morphology (see the subsection on the property "possessor on initial"), *hambukha* 'liking, belief' cannot be inalienably possessed in the context of (626). In (627), we see that my attempt to add possessive morphology was rejected.

(627) *Ímă **hambukhà-ŏ** amíe ahe.
1SG liking-1SG.POSS many INDEF
'There are many things that I like / believe in.' (elicited, I produced)

Because of this behavior, I encode this property not strictly as "yes", but as "alienable" for this LVC.

Initial modification by adjective when independent from verb

The initial *hambukha* 'liking' is only very rarely seen independently from the verb. While it does behave syntactically like a noun in such cases, modification by adnominals is not attested (not counting modifier insertion, as that is in the context of the LVC, not when the initial is independent).

Initial used clearly as adjective

The initial *hambukha* 'liking' has no attested behavior as an adnominal modifier.

Topic/focus marking of initial

This property was not directly tested for this LVC, but it is completely unattested in the corpus of natural Hurukui speech.

Verb alone in response

The LV cannot occur without the initial in a response to a question. This is shown in (628).

(628) A: *Ima hambukhá-ŏ e-ma?*
1SG liking-1SG.POSS do-2SG.R

B: *Eo, [hambukhá-nă / ***Ø**] e-ra.*
yes liking-2SG.POSS do-1SG.R
A: 'Do you like me?' B: 'Yes, I do.' (elicited, I produced)

Variation in verb selection

This LVC does not allow variation in its LV; it must always be *i-* 'do'. One test for this was using the SS.SIM suffix *-hVhV* to see what elided LV vowels would show up as copy vowels in the suffix. From this test, only the form *hihi* was accepted, indicating an LV root shape of *i-*. For instance, *hambukhă ă-háha* (liking put-SS.SIM) was completely rejected by consultants. Some more rejected attempts at using other LV roots are shown in (629) and (630).

(629) **Na hambukhá-nă **o**-ra.*
2SG liking-2SG.POSS take-1SG.R
*'I like you.' (elicited, I produced)

(630) **Na ima hambukhá-ŏ **a**-ma?*
2SG 1SG liking-1SG.POSS put-2SG.R
*'Do you like me?' (elicited, I produced)

Verb retaining some of its lexical semantics

Because the verb *i-* 'do' is already so semantically empty, it is difficult to say how much it has retained whatever lexical semantics it has in non-light uses. Comparing it with the likely etymologically related *hi-* 'make' (a relationship I often indicate by writing *(h)i-* 'make, do'), however, it is clear that the verb is lighter in the LVC *hambukha i-* 'like', because no physical object is being made/created. In this way, it is clearer that *hambukha i-* 'like' does **not** have this property, whereas for *sáhi-* the LVC meaning is more closely related to the verb's lexical use.

Summary

Table 11.4 shows a summary of the properties that have been shown for this LVC.

Table 11.4: Properties of the LVC *(h)ambukha i-* 'like'

Property	**Subproperty**	**Value(s)**
Stress positions and quantities	3SG.R	primary-none (p-n), none-primary (n-p)
	2SG.IMP	none-primary (n-p)
Vowel raising domain		N/A
Verb elision		yes
Possessor on initial		yes
Object on verb		yes
The possessor/object alternation		yes, double-marking
Initial reduplication		full
Verb reduplication		SS.SIM (2σ)
Modifier insertion		yes
Time adverb insertion		marginal
Anaphora insertion		unattested
Focused argument insertion		no
PP insertion		no
Initial independence		marginal
Initial used clearly as noun		alienable
Initial modification by adjective when independent from verb		no
Initial used clearly as adjective		no
Topic/focus marking of initial		unattested
Verb alone in response		no
Variation in verb selection		no
Verb retaining some of its lexical semantics		no

11.5 Summary

Table 11.5 shows a comparative summary of the properties that have been shown for the LVCs detailed in this chapter.

Table 11.5: Properties of the LVCs detailed in Chapter 11

Property	**Sub-property**	***moko e-* 'speak'**	***sáhi-* 'fix'**	***aipV e-* 'cry'**	***hambukha i-* 'like'**
Stress positions and quantities	3SG.R	p-n	p-n	p-n	p-n, n-p
	2SG.IMP	p-n, s-p	p-p, n-p	n-p, s-p, p-n, p-s	n-p
Vowel raising domain		separate	N/A	N/A	N/A
Verb elision		yes	no	yes	yes
Possessor on initial		no	no	no	yes
Object on verb		yes	yes	no	yes
The possessor/object alternation		no	no	no	yes, double-marking
Initial reduplication		full	except-V, idiosyncratic	idiosyncratic with verb change	full
Verb reduplication		SS.SIM (2σ)	SS.SIM (1σ)	SS.SIM (1σ)	SS.SIM (2σ)
Modifier insertion		yes	no	yes	yes
Time adverb insertion		yes	no	yes	marginal
Anaphora insertion		no	unattested	no	unattested
Focused argument insertion		yes	no	yes	no
PP insertion		yes	no	no	no
Initial independence		yes	no	marginal	marginal
Initial used clearly as noun		yes	N/A	marginal	alienable
Initial modification by adjective when independent from verb		yes	N/A	no	no
Initial used clearly as adjective		no	N/A	no	no
Topic/focus marking of initial	TOPIC	yes	unattested	no	unattested
	FOC	yes	unattested	marginal	unattested
Verb alone in response		no	no	no	no
Variation in verb selection		no	no	no	no
Verb retaining some of its lexical semantics		sometimes	inconclusive	sometimes	no

Some trends can be seen in Table 11.5. First, the primary-none stress pattern in the realis is very common among Hurukui LVCs. That is, the initial retains its primary stress and the LV loses its stress. This ordering is responsible for why some multisyllabic verb stems can still be recognized as historically being LVCs, because they carry antepenultimate stress on the part of the word that was once the penult of their initial.

Another observation is that the initials that behave more like nouns nonetheless do not occur uniformly in all constructions where nominals can occur (e.g., they may be able to be the argument of a predicate but unable to have topic/focus marking or adnominal modifiers). This is one reflection of the fact that word classes are only emergent clusters in terms of occurrence in different slots in different constructions, rather than categories whose members all have the same syntactic distribution (Croft 2001). As LVCs lexicalize and univerbate further, the distributions of their initials become more idiosyncratic, as if by coevolution of the initial and the verb, causing their locations in word class space to scatter.

High collocation and fusion of the LVC does not imply phonological reduction of the LV root, as shown by *sáhi-* 'fix', which is very fused but cannot undergo verb elision to any extent. However, high collocation and fusion do correlate with loss of stress. This can be seen in the fact that the unfused *moko e-* 'speak' and the partially fused *aipV e-* 'cry' can at least sometimes retain secondary stress on their half that lacks primary stress, while the highly fused *sáhi-* 'fix' disallows such a combination of two stresses in its two parts.

The LVCs differ considerably in terms of which kinds of constituents they allow to intervene between the initial and the LV. Even some LVCs that appear to be in the process fusing further, such as *aipV e-* 'cry', still allow modifier insertion, but the class of modifiers is smaller for some of them, some only allowing the most generic intensifiers such as *maisa* 'bad, intense' and *awohe* 'too much'. This class may shrink over time as the collocation strengthens, and in many cases it appears to be creating incipient incorporation of intensifiers into the new complex verb stems.

None of the LVCs detailed in this chapter have the properties of "verb alone in response" or "variation in verb selection". While some LVCs do have verb variation, very few VCs that are

not VOCs allow the initial to be omitted in the response to a question. Even highly compositional VCs made of two clear independent words, such as *mbukura i-* (cold become-) 'become cold' do not allow the initial to be omitted in a response. Ellipsis like that evaluated by "verb alone in response" is also subject to definite null reference of the absent initial if it is a nominal (§7.3.2), so the pragmatics of the discourse context further constrains when the property can apply.

The next chapter will build on the four LVCs detailed here, using them as jumping-off points for further exploration outward into the space defined by the properties.

12 Quantitative mapping of LVC space

This chapter describes quantitative methods for observing how the LVCs (and other items such as VOCs and monomorphemic verb roots) are distributed in the space defined by the properties detailed in the previous chapters, and applies these methods to the Hurukui data. One important question is how the VCs cluster together into groups that have similar behavior, or whether they even do so to an appreciable extent at all. Can certain areas of the space be said to have a certain nature, in terms of features **not** directly encoded in the properties, because of the points that inhabit it (e.g., is one area associated with transitivity and high affectedness of the patient)? Questions such as these will be addressed in this quantitative investigation.

12.1 Data encoding

12.1.1 Multicollinearity and logical dependency among features

The methods for getting at these questions are quantitative but take into account domain-specific knowledge about how Hurukui works and how the properties are interrelated. One should not blindly run clustering algorithms on a dataset without understanding the meaning of the data on a qualitative level. Many of the properties have some logical relations to one another. For instance, the properties about what word class the initial is when it is independent from the verb (e.g., "initial clearly used as noun") are dependent on the property of whether the initial can be independent from the verb in the first place. The statistical methods used should not ignore this fact by treating all dimensions as completely orthogonal.

12.1.2 One-hot encoding scheme

The input data matrix was one-hot encoded for the values of properties and their subproperties. This means that a data point (a single VC) gets a value of 1 if it has a certain value of a certain property, and 0 otherwise. For instance, the properties relating to stress placement and quantity have multiple values. An excerpt from the full property table is shown here to illustrate.

Table 12.1: Stress properties of selected LVCs, summary table format

Property	**Sub-property**	***phihi ko-* 'knock'**	***khum ko-* 'sprout'**	***phum ko-* 'explode'**
Stress positions and quantities	3SG.R	p-n	p-n, p-s	p-n, p-s
	2SG.IMP	p-p, n-p	n-p, s-p	unattested

As you can see, the property of "stress positions and quantities" has two subproperties, "3SG.R" and "2SG.IMP", and each of these has multiple possible values: primary-primary (p-p), primary-secondary (p-s), primary-none (p-n), secondary-primary (s-p), and none-primary (n-p). The one-hot encoding treats each property-subproperty-value combination as a separate dimension entirely, and the VCs are given a 1 if they have that value for that property-subproperty combination, and a 0 otherwise. Table 12.2 shows how the information from Table 12.1 was converted into this encoding.

Table 12.2: Stress properties of selected LVCs, one-hot encoding format

Dimension	***phihi ko-* 'knock'**	***khum ko-* 'sprout'**	***phum ko-* 'explode'**
(Stress positions and quantities, 3SG.R, p-p)	0	0	0
(Stress positions and quantities, 3SG.R, p-s)	0	1	1
(Stress positions and quantities, 3SG.R, p-n)	1	1	1
(Stress positions and quantities, 3SG.R, s-p)	0	0	0
(Stress positions and quantities, 3SG.R, n-p)	0	0	0
(Stress positions and quantities, 2SG.IMP, p-p)	1	0	(left blank)
(Stress positions and quantities, 2SG.IMP, p-s)	0	0	(left blank)
(Stress positions and quantities, 2SG.IMP, p-n)	0	0	(left blank)
(Stress positions and quantities, 2SG.IMP, s-p)	0	1	(left blank)
(Stress positions and quantities, 2SG.IMP, n-p)	1	1	(left blank)

As can be seen in the column for *phum ko-* (explode say-) 'explode', under the dimensions for the 2SG.IMP form, cells were left blank in the one-hot encoding when there was no elicitation or corpus data about the corresponding properties. A large proportion of the full feature matrix

was left blank for this reason, and this was dealt with using methods of *imputation*, described in §12.2.

The one-hot encoding scheme also allows for gradience between the numbers 0 and 1 to account for variation in acceptability and production. Each utterance in elicitation sessions was given an acceptability rating on a 5-point scale (completely rejected, mostly rejected, half-and-half, mostly accepted, completely accepted) based on how many consultants stated the utterance's correctness after they reproduced it themselves. This was a crude way of accounting for variation among speakers' grammars, but it was better than marking all utterances with black-and-white acceptability judgments. Another way in which gradience was introduced in the one-hot encoding scheme was when the same property of the same VC was given different judgments in different utterances. These differing values were averaged. For instance, if a VC was tested for verb elision in two different contexts, receiving a judgment of mostly accepted (0.75) for one of them and completely accepted (1) for the other, then its value for that dimension was encoded as 0.875.

12.1.3 N/A values

Often, features do not apply to every VC in the dataset. For instance, the property of the domain of mid vowel raising cannot apply to VCs that do not contain non-final mid vowels in any of their forms. An example of this is *anga a-* 'dawn break', which only contains the vowel *a* and is therefore not subject to raising, in any of its inflections, in either the initial or the LV. In situations like this, I encoded the value as "NA", which was then converted to the number -1 in the one-hot encoding. The reason for doing this rather than leaving it blank was that I did not want the value to be imputed, as it was known, not just unobserved. It had to be placed outside of the interval from 0 to 1 because all of those values represent known degrees from unacceptability to acceptability, while N/A values mean that the test cannot even be done and so acceptability judgments would be meaningless.

The choice of the value -1 was somewhat arbitrary, but it was a reasonable choice given the algorithms used in the subsequent analysis. First, the random forests used for imputation are able to make clean splits between a value outside the [0, 1] interval and values in it. Any value less than zero or greater than 1 would have worked for this purpose. Second, it is closer to 0 than to 1,

because I consider not-applicable values as more similar to absent values than to present values. This is important for the locations of points in the high-dimensional space and the distances between them. However, the distance from 0 that was chosen, being as far away from it as the completely-accepted 1 value is, likely has effects on the distance measures and conclusions of this study. No other values were tested, though, and the effect of the choice of -1 rather than, say, -0.5 or -2, on the conclusions of this study is a direction for future research.

12.1.4 Splitting probability among multiple indistinguishable analyses

Another complication in the data encoding scheme is the treatment of competing indistinguishable analyses. For instance, the fused LVC *maríhi-* 'search for' has SS.SIM form *marihíhi* 'while searching', but it is unclear how many *hV* syllables should be attributed to the SS.SIM suffix *-hV ~ -hVhV* versus the LV root itself. This is because the initial and LV root are not clearly separable in the highly fused stem, so it could be *mariV hi-* or *marihV i-*, and the SS.SIM form *marihíhi* could be parsed as any of:

- *mari-hi-hV*: search-do-SS.SIM
- *marih-i-hV*: search-do-SS.SIM
- *marih-ĭ-hVhV*: search-do-SS.SIM
    - If this parse is correct, the form would have a geminated *h* because the final *h* of the initial would meet with the initial *h* of the suffix, but I consider this still to be a plausible analysis because of the commonness of *h* deletion throughout the language.

As a result, it is unclear whether the SS.SIM form uses the one-syllable or two-syllable suffix, and whether the LV root is elided in the SS.SIM form or not, and this is not answerable from further elicitation. The encoding scheme accounts for this by spreading the probability 1 evenly among the ambiguous options. The LVC *maríhi-* 'search for' therefore was encoded as having values 0.5 for each of (verb reduplication, SS.SIM, 1σ) and (verb reduplication, SS.SIM, 2σ).

12.2 Imputation

Even with lots of targeted elicitation about the properties of these VCs, which was nearly the sole activity of my 2024 field trip, the feature matrix remains sparse. There just is not enough attestation of these combinations of lexical item and property in the natural-speech corpus, and elicitation sessions often had mixed success with being able to apply the tests at all. This could have been due to pragmatic infelicity, my inability to construct a realistic scenario in which one would use rare forms (e.g., inserting arguments marked for contrastive focus between the initial and the LV), simply choosing to skip over certain combinations to save time for more important questions (which I certainly did), or other factors.

As a result of this sparseness in the feature matrix, I needed to fill in missing values somehow in order to see how the VCs are distributed in the space. This is a common problem in data science, and the process of filling in educated guesses for the missing values is called *imputation*. One can *impute* a missing value in a variety of ways, usually using the known values of the feature (and perhaps the other features) from other data points to construct a model that can predict the missing value based on the known values for that data point.

Before imputation, I filtered out any VCs that had less than 10% of the dimensions filled out in the one-hot encoded matrix, and then I filtered out any dimensions that still had data filled out for less than 20% of the VCs. I ended up with 74 VCs and 45 features, a total of 3,330 cells in the matrix. Of these, 2204 (66%) were filled explicitly and the other 1,126 (34%) were missing.

I used the Python package "miceforest" to impute the missing values in my data. The "mice" part of its name refers to *MICE*, which stands for "multiple imputation by chained equations". MICE is a procedure by which multiple imputed versions of the data are created (the "multiple imputation" part) by repeated application of a series of steps for getting the new values (the "chained equations" part). The repeated application of the steps allows the procedure to home in on good guesses for the values by refining the guesses each time based on the current best guesses for the other features. After a certain number of iterations, the imputed dataset copies are finished, and they are then recombined into one final complete data set.

The "forest" part of miceforest's name refers to random forests, a machine learning algorithm that uses ensembles of decision trees to predict unknown values. Decision trees are a robust method for learning complex relationships among many variables. They operate by making a series of splits in the dataset, each split based on the value of a single variable, in order to group the subsets of the data together that have similar values for the features. The splits are chosen based on whatever independent variable (and cutoff value within its range) explains the most variation in the dependent variable, within a given subset that is being split. Because vanilla decision trees have certain drawbacks, random forests train a large number of trees on the same dataset, but each tree is somewhat randomized. This randomization is done by selecting the feature for each split only among a random subset of the features. All the trees in the forest end up looking somewhat different, but still doing a decent job of learning the patterns in the data. The results from all these trees are then pooled together and combined to yield the predictions of the entire forest model.

The miceforest package uses random forests as the model by which each variable is imputed based on the values of all the other variables. This is the "equation" that is chained, i.e., applied repeatedly for a number of iterations, to get a single imputed copy of the dataset. Finally, all the multiple imputed copies are combined and the final dataset is returned with all empty values now filled with good guesses.

12.3 Principal component analysis

Once a full dataset of VCs with values for all the features (either observed or imputed) was available, the next step was to place the data points (one for each VC) into a high-dimensional space that reflected their similarities and differences in behavior. The naïve way to do this would be to just treat each of the features in the matrix as a separate dimension, with all dimensions orthogonal to each other, and measure the Euclidean distance between the points.

However, because of the many logical dependencies and high correlations among the variables (a problem called *multicollinearity*), this approach would result in inflated measurements of distance due to what essentially amounts to double counting of some underlying trait by multiple variables that reflect that trait in similar ways. For instance, the

properties of "initial used clearly as noun" and "initial modified by adjective when independent" are highly correlated because nouns in Hurukui are, of course, often found with adnominal modifiers. The correlation is not perfect, but it is large enough that the presence of both of these dimensions does not contribute very much more information to the characterization of LVCs than would be present if only one of the dimensions were used. Thus, while they are not equivalent, they should be treated more like lines that are nearly parallel (together contributing only slightly more distance compared to if only one of them was used), rather than as lines that are perpendicular (carrying as much relative weight as do dimensions that are completely unrelated).

To mitigate this multicollinearity problem and get a more accurate view of the LVC space that is not distorted by the over-counting of multiple very similar dimensions, I used a technique of *dimensionality reduction* called *principal component analysis* (PCA). PCA creates a new coordinate system, in which the first coordinate is the direction in the space along which the data varies the most, the second coordinate is the variance-maximizing direction in the subspace orthogonal to the first coordinate, and so on. The number of these coordinates (called *principal components*, here abbreviated *PCs*) can be chosen directly or based on how many is needed to explain some majority amount of the variation (e.g., 95%). This amounts to projecting the data into a smaller number of dimensions (unless the number of components is the same as the original number of dimensions) and rotating it within that subspace so that multicollinearity is minimized.

The PCA algorithm cannot create more components than there are dimensions in the data, nor more than there are data points. I chose to maximize the number of components so that I could measure a weighted distance in the resulting space, described in §12.4.1.

Because PCA chooses each component based on maximizing the variance in the data that can be accounted for ("explained") by that component, it is sensitive to the scale of each dimension. That is, one could artificially inflate the variance explained by a given dimension by simply multiplying the values by a large number along that dimension. While the scales used in my data are very limited because of the one-hot encoding scheme (having a minimum of -1 and a

maximum of 1 for every feature), the problem is still worth mitigating because the distributions of the data points are not the same for each feature. For instance, many features have no N/A values and so they are confined instead to the interval [0, 1], and some features have a distribution that is more skewed toward 1, 0, or N/A than others.

The way to deal with this and make the PCA results as "fair" as possible is to *center* and *scale* the data along each dimension. Centering subtracts the mean of each dimension from its values so that its mean becomes zero. Scaling multiplies the values of each dimension by a constant so that its standard deviation becomes one. None of this changes the correlation coefficients between the dimensions, so it is still necessary to use PCA to reduce multicollinearity, and PCA will still be able to pick out directions in which the data appears most stretched out (due to an underlying correlation between the dimensions that combine to create that line).

PCA identifies components that are essentially weighted averages (a.k.a. *linear combinations*) of the dimensions in the dataset. For instance, a component could be a line defined by the expression:

0.15 * (initial independence) - 0.09 * (modifier insertion) - 0.24 * (PP insertion) + ...

The coefficients in the above expression are what matters, and they are what PCA optimizes over. Table 12.3 shows the first ten coefficients of each of the first three principal components found in this study, sorted in descending order of absolute value to reflect the amount of impact each feature has on the component. It is worth mentioning how much variance each of these components explained in the data: 15.66% for PC1, 11.91% for PC2, and 6.30% for PC3. Because no small number of components explains the vast majority of the variance, we can conclude that the data really is high-dimensional; it takes more than just a few underlying features to characterize the VCs.

Table 12.3: Most heavily weighted features of the first three principal components

PC1		PC2		PC3	
feature	**weight**	**feature**	**weight**	**feature**	**weight**
initial independent	+0.3200	possessor-object alternation (any)	+0.3156	stress in 3SG.R is primary-none	+0.3214
initial modified by adjective	+0.3158	possessor-object alternation (double-marked)	+0.2980	verb root reduplication (any)	+0.3115
initial is a noun	+0.3083	verb root elision in SS.SIM form	-0.2833	stress in 2SG.IMP is none-primary	+0.2749
time adverb insertion	+0.2672	initial is an adjective	+0.2767	stress is 3SG.R is secondary-primary	+0.2645
PP insertion	+0.2666	initial reduplication (any)	-0.2481	verb root reduplication (full)	+0.2553
focused argument insertion	+0.2648	possessor marking on initial	+0.2458	verb root reduplication (SS.SIM, 2 syllables)	+0.2435
verb alone in response	+0.2541	verb retaining some lexical semantics	-0.2454	possessor-object alternation (alternating)	+0.2331
verb retaining some lexical semantics	+0.2459	time adverb insertion	-0.2284	modifier insertion	+0.2265
initial reduplication (except-V)	-0.2283	focused argument insertion	-0.2276	possessor-object alternation (any)	+0.2165
verb root elision	-0.1891	initial reduplication (except-V)	-0.2101	verb root elision in SS.SIM form	+0.2122

The components are a bit hard to interpret just from looking at the weights like this. We can see a few interesting things, though. The first component cares a lot about the behavior of the initial as a separate word and insertion of material between the initial and the LV. The second component seems to capture information about possessor marking on the initial, but it also contains higher weights for some seemingly unrelated dimensions such as verb root elision. The

third component is more about morphophonology: the stress placements and quantities in the realis versus the imperative, and the forms of the verb root when it is reduplicated.

More insight can be gained from looking not at the composition of the components themselves, but where the VCs are found along each component. Table 12.4 shows, for each of the first three principal components, the VCs at its two extreme ends as well as those most in the middle. The signs of the values do not matter; i.e., there is no inherent meaning of a VC being at the low end vs the high end of a component. What matters is that the two ends are opposite poles, and where the VCs are relative to each other, but the negatives and positives could be flipped with no impact on the interpretation.

Table 12.4: Locations of the most extreme and middlemost VCs on the first three PCs

PC1		**PC2**		**PC3**	
VC	**value**	**VC**	**value**	**VC**	**value**
sáhi- 'fix'	-4.481	*moko e-* 'speak'	-3.461	*ekha-* 'descend'	-6.404
ekha- 'descend'	-4.085	*sono ekha-* 'bathe'	-3.366	*epu a-* 'be enough'	-3.636
inóngi- 'do what'	-3.702	*nama a-* 'sit'	-3.217	*ukumi-* 'grab'	-3.462
hupúri- 'show'	-3.190	*kuru hi-* 'make bird blind'	-3.021	*óka-* 'be hungry'	-3.227
hióki- 'drop'	-3.129	*wo hoka-* 'work'	-2.939	*pe we-* 'vanish'	-3.123
...		...		...	
khenge pi- 'ask'	-0.404	*khaníngi-* 'do thus'	-0.396	*hióki-* 'drop'	+0.505
hambukha i- 'like'	-0.384	*erére-* 'pet'	-0.378	*ngáhi-* 'tell'	+0.566
pe we- 'vanish'	-0.268	*hóhi-* 'rub'	-0.324	*ipe e-* 'call name'	+0.580
phíhi- 'disembowel'	+0.010	*hióki-* 'drop'	-0.316	*inóngi-* 'do what'	+0.599
...		...		...	
hauyu e- 'wind blow'	+4.717	*íre-* 'beg'	+5.442	*eme i-* 'look around'	+1.645
wo hoka- 'work'	+5.100	*pokhona pi-* 'help'	+5.479	*ura i-* 'fear'	+2.753
hipa re- 'defecate'	+5.115	*mbukhu i-* 'be wet'	+5.652	*sono ekha-* 'bathe'	+2.771
opopo i- 'eat'	+5.674	*khunukha pa-* 'smell'	+6.590	*hambukha i-* 'like'	+3.190
sono ekha- 'bathe'	+5.915	*khane a-* 'feel sorry'	+7.823	*íre-* 'beg'	+3.446

The first component clearly has to do with the amount of syntactic univerbation, with lower values being more fused and higher values being less fused. The presence of the root *ekha-* 'descend', known to be completely monomorphemic and included as a control data point, at the opposite end of the scale from *sono ekha-* (water descend-) 'bathe', which includes *ekha-*, makes this interpretation of PC1 very clear. The ranking also shows that *sáhi-* 'fix' was a good choice for the prototype of a completely fused LVC, as detailed in §11.2. This matches the insight from PC1's coefficients that it is mainly about the behavior of the initial as a separate word and insertion of material between the initial and the LV, as we should expect highly fused VCs not to allow these while unfused VCs may allow them.

The second component can be most succinctly characterized as capturing information about different kinds of argument structure. The low end of the scale consists of intransitive predicates (counting VCs that have clear objects, such as *kuru hi-* 'make bird blind', as "intransitive" here because another object cannot be added). The high end consists mostly of predicates that have some experiencer, beneficiary, or maleficiary, which is marked as a possessor on the initial, an object on the verb, or both. In the middle we have transitive or ambitransitive predicates which can take object marking on the verb root.

The third component seems to be about the amount of morphophonological univerbation, because it is sensitive to whether the initial can be stressed. Some of the VCs included are known to just be monomorphemic roots, such as *ekha-* 'descend' and *ukumi-* 'grab', which I included as controls. For these, I treated them as if they were like the fused LVCs, e.g., *sáhi-* 'fix' or *óme-* 'light up', for the purposes of locating the "initial" and the "verb root". Thus, I treated *ekha-* 'descend' as if it were *ekhV a-*, and *ukumi-* 'grab' as if it were *ukumV i-*, for the purposes of encoding these properties. These both occur in the lowest reaches of PC3, while some VCs that are more clearly still two words synchronically occur at its high end, e.g., *sono ekha-* (water descend-) 'bathe' and *ura i-* (fear do-) 'fear'. However, the presence of some unexpected interlopers on both ends of the scale, such as *epu a-* (enough put-) 'be enough' on the more-univerbated end and *íre-* 'beg' on the less-univerbated end, is puzzling. It could be due to deficient data, aberrant values for some

features predicted during the imputation process. The signal could also be real, a sign that the morphophonemic behavior of the VCs with regard to stress and verb root reduplication is rather fluid and the groups have begun to bleed together in some ways.

The following figures show the space of VCs projected into two-dimensional planes. Figure 12.1 shows PC1 and PC2, Figure 12.2 shows PC1 and PC2, and Figure 12.3 shows PC2 and PC3.

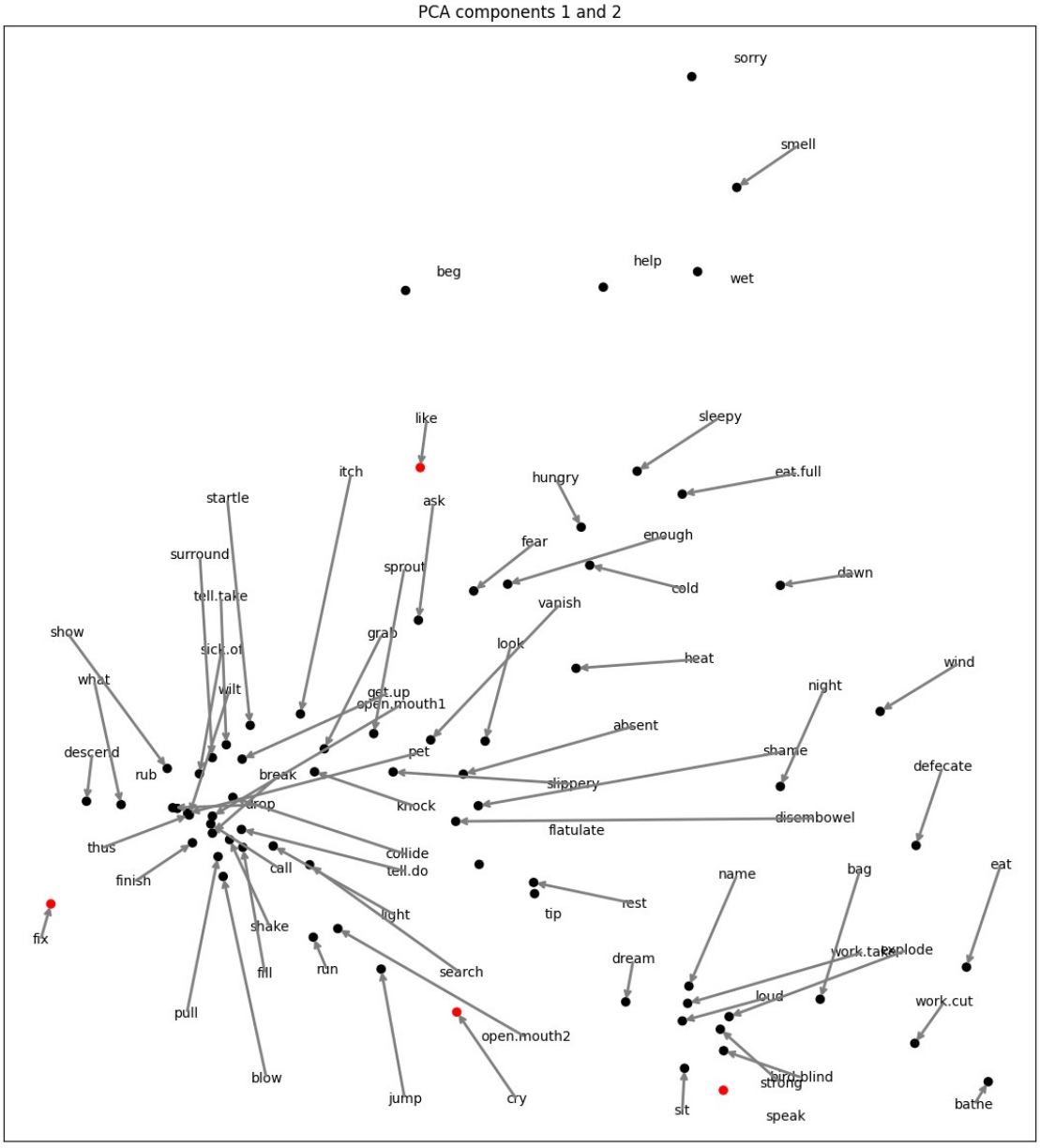


Figure 12.1: VCs with PC1 (morphosyntactic univerbation) on the *x*-axis and PC2 (argument structure) on the *y*-axis

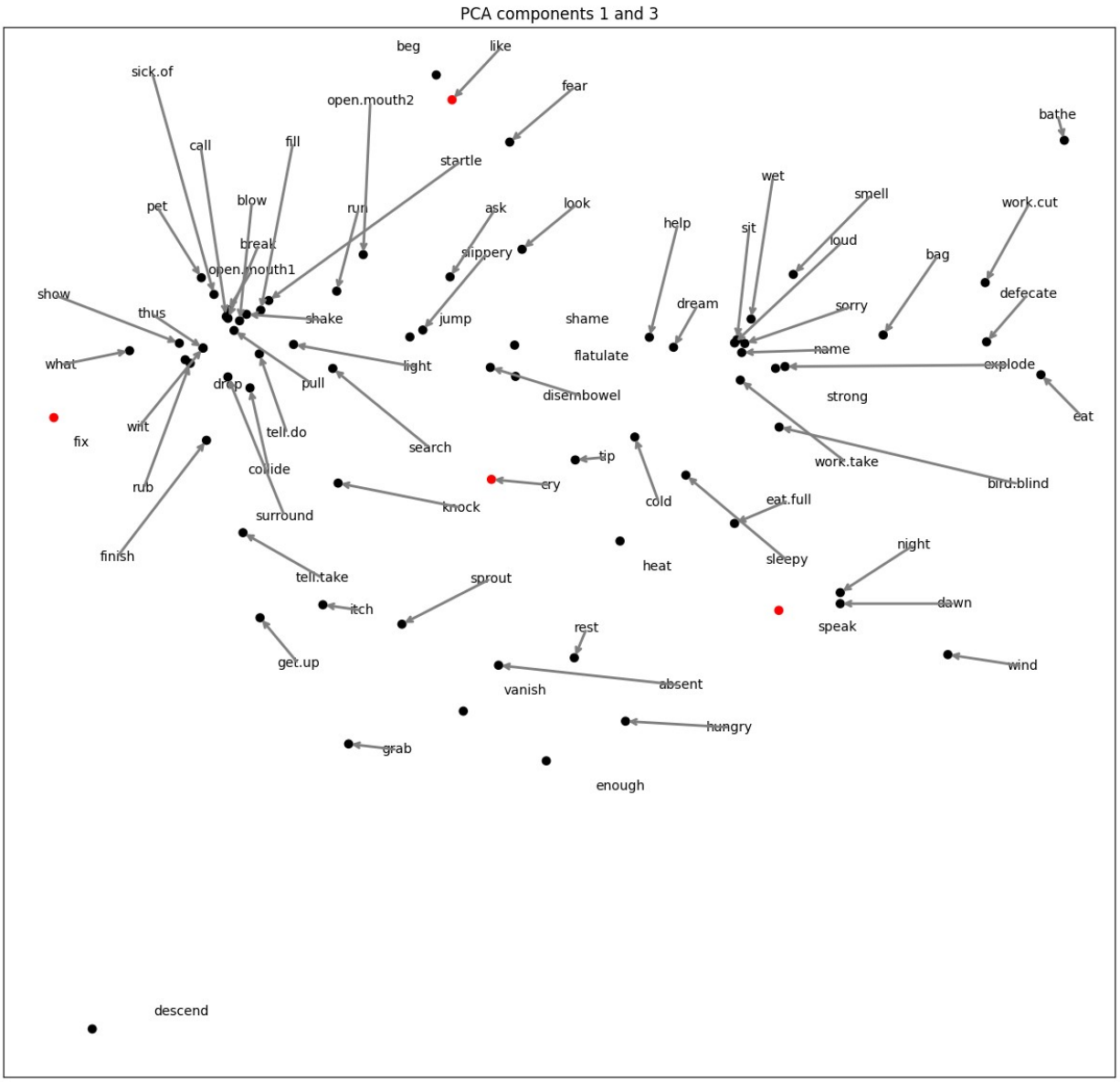


Figure 12.2: VCs with PC1 (morphosyntactic univerbation) on the *x*-axis and PC3 (morphophonemic univerbation) on the *y*-axis

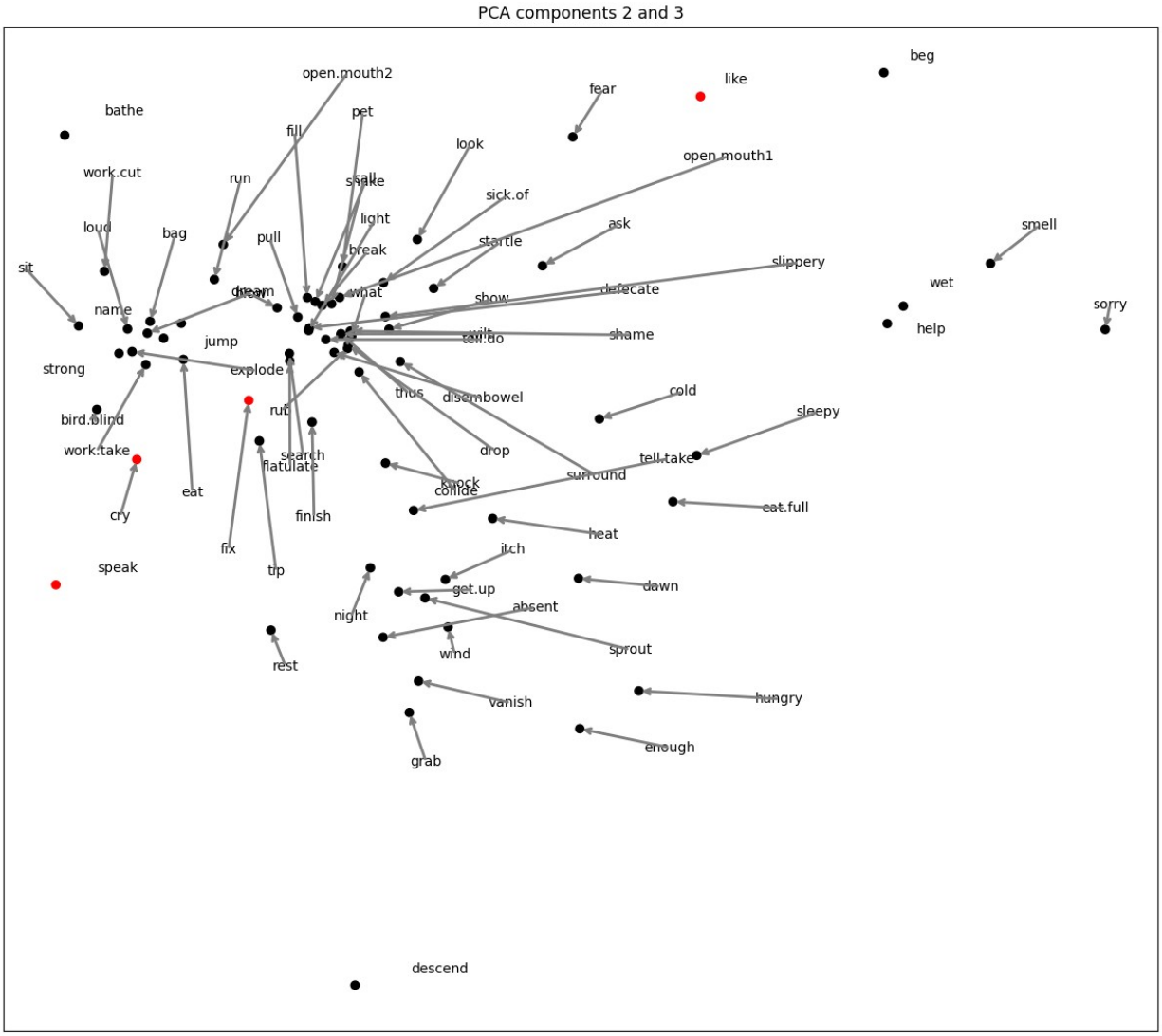


Figure 12.3: VCs with PC2 (argument structure) on the *x*-axis and PC3 (morphophonemic univerbation) on the *y*-axis

The next section describes the methods used for further analysis of how the VCs are distributed in the space.

12.4 Methods using distance measurements

12.4.1 Weighted distance metric

Beyond projecting the VCs' data points onto one-dimensional lines (as in Table 12.4) or two-dimensional planes (as in Figures 12.1, 12.2, and 12.3), this study aimed to observe the overall grouping of points in the higher-dimensional space. To do this, I needed a distance metric that could measure how (dis-)similar the VCs are, without falling prey to the pitfalls of a naive approach that does not take into account the high degree of correlation between some pairs of features. A transformation of the data into the coordinate system defined by the PCs is a good start, as these have dealt with much of the multicollinearity problem already. One could then simply take a Euclidean distance in that space, but that would weight each of the PC dimensions equally. I chose not to do this because I wanted to assign higher weight to differences along dimensions that mattered more for characterizing the similarities and differences between the VCs. I therefore modified the Euclidean distance in the PCA space by weighting the dimensions according to the amount of variance they explained in the original dataset. The Euclidean distance function is:

$$d(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$

whereas the weighted distance function is:

$$d(x, y) = \sqrt{\sum_i w_i (x_i - y_i)^2}$$

where x and y are points in the space with coordinates on dimension i given by x_i and y_i respectively, and w_i is the weight assigned to dimension i. Because PC1 explained 15.66% of the variance, its weight was 0.1566. Similarly, PC2 explained 11.91% of the variance and was weighted 0.1191, and so on for all of the principal components. Table 12.5 shows an excerpt of the distance matrix calculated using this metric.

Table 12.5: Sample distances between VCs in variance-explained-weighted PCA space

	áhi- **'open mouth'**	***áipe*** **'cry'**	***akhawáe-*** **'be loud'**	***anga a-*** **'dawn break'**	***áuni-*** **'finish'**
áhi- **'open mouth'**	0.000	1.866	2.862	3.025	1.276
áipe **'cry'**	1.866	0.000	2.188	2.667	1.868
akhawáe- **'be loud'**	2.862	2.188	0.000	2.577	2.788
anga a- **'dawn break'**	3.025	2.667	2.577	0.000	3.006
áuni- **'finish'**	1.276	1.868	2.788	3.006	0.000

12.4.2 Multi-dimensional scaling

This PCA-defined space is still high-dimensional (with 45 PCs because I used the maximum possible given that the dataset had 76 data points and 45 features after filtering out rows and columns that were too incomplete). Because of this, visualizing the relative positions of the points is difficult. However, there are good methods for getting a qualitative impression of the relative closeness of points in such a space. I used *multi-dimensional scaling* (MDS), which distorts the space in a nonlinear way so as to place points close together with high probability if they are close together in the actual space. MDS output can be visualized in either two or three dimensions, and I used two-dimensional plots for this study. Figure 12.4 shows the result.

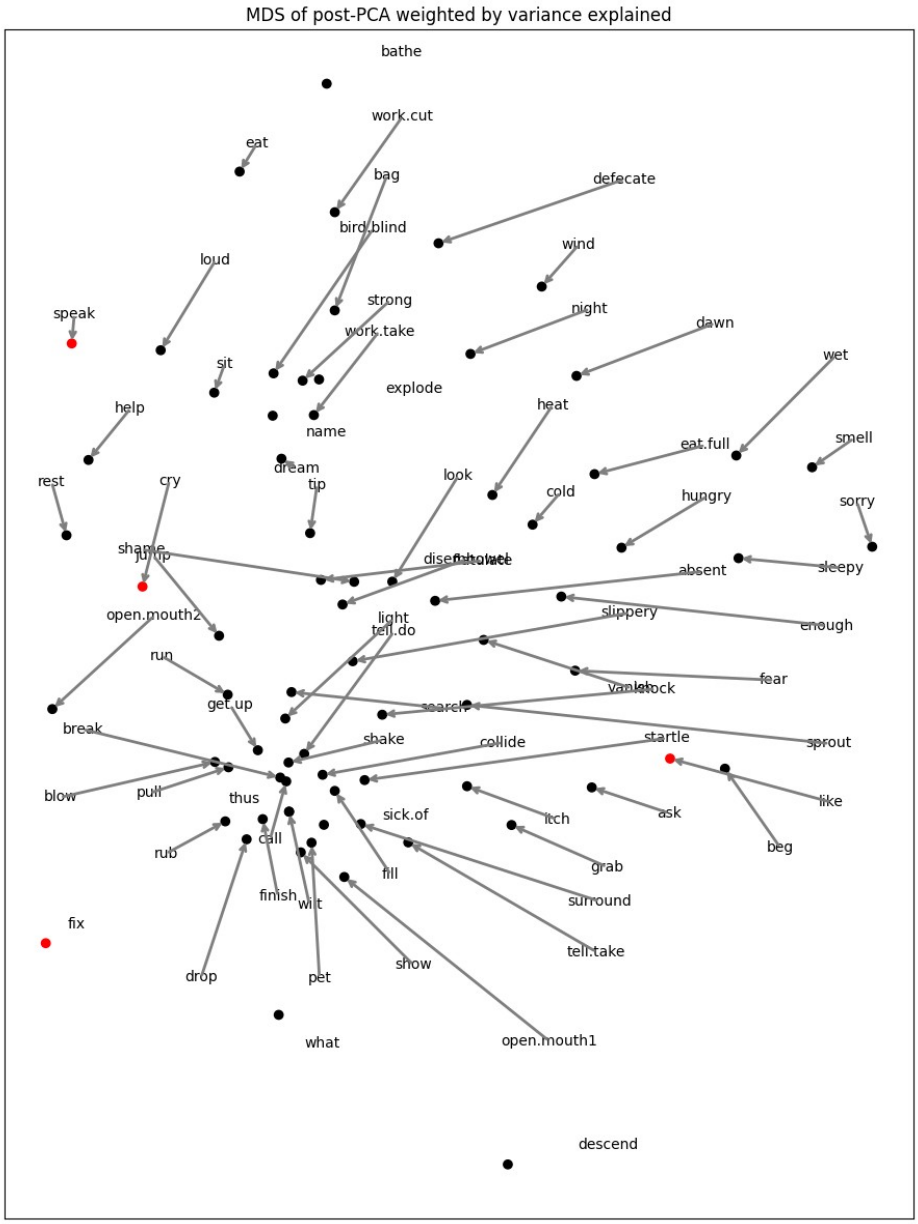


Figure 12.4: MDS projection of the VCs using the variance-explained-weighted distance metric

In an MDS plot, because of the nonlinear and stochastic nature of the transformation algorithm, the *x* and *y* axes are meaningless. What matters is the relative proximity or distance between pairs of points. With this is mind, in Figure 12.4 we can see some meaningful structure. The points for *ekha-* 'descend' and *sono ekha-* (water descend-) 'bathe' are good references for nearly perfect one-word and two-word VCs, respectively, and they are found right at the bottom

and the top of the plot, respectively. Other VCs found in the lower part of the plot are highly fused, such as *sáhi-* 'fix' and *inóngi-* 'do what'. There is a large cluster about a third of the way from the bottom to the top of the plot, containing various mostly fused VCs, such as *hukV e-* (blow take-) 'blow', *yakhV a-* (get.up put-) 'get up', and *ongV ya-* (startle go.up-) 'startle'.

In the PCA and MDS figures, no two points are in exactly the same place. Part of the reason for this is imputation. If the observed vectors of two points are not exactly the same, then the model is likely to put them in slightly different places because the predictors used for filling in the missing values are different for the two points. If I had restricted the space based only on what values I had observed, looking at some set of LVCs only regarding properties that are defined for all of them, then there would be some collisions. I also filtered out LVCs with too few properties observed, and I filtered out properties with too few LVCs observed.

Generalizations are gradient. People have subconscious estimates for the probability that a certain form can be used successfully with a certain function in a certain context, rather than a hard-and-fast rule for the form-meaning mapping. Nonetheless, that probability could be very close to 0 or 1. Additionally, infrequent items could be observed with only one value for this question (all successful, or all unsuccessful) by sheer chance due to the low sample size. By this mechanism, two very infrequent LVCs could be observed to have the same properties, even though the generalizations producing them are not exactly the same. The diagrams presented in this chapter would fail to account for this, because I removed the highly infrequent LVCs, as well as due to the imputation.

12.5 Clusters within the space

The more fused LVCs pattern together nicely, largely occupying a single region of the space. The most fused group, LVCs involving the LV *i-* 'do', is the one with the tightest clustering. It is also remarkable that LVCs that would seem to have rather different diachronic sources, such as *khuhV ya-* (jump go.up-) 'jump' and *honV yo-* (fill come-) 'fill up', nonetheless have arrived in the same behavioral region. This suggests a diachronic explanation in terms of convergence: as VCs have gotten more fused in Hurukui, they have gravitated toward the same kinds of fusional properties. There are many ways for something to be two words, but few ways for something to be

one word. There may be a sort of diachronic sink in the area of the space where this cluster lives. This raises questions about the dynamics within the space. How stable is the cluster? Does anything ever leave it once it has entered? How stable are other areas of the space? Are there other sinks that simply have fewer members and thus are less noticeable on the map? Are there areas that repel diachronic trajectories, i.e., combinations of behaviors that are very hard for any lexical item to reach for some reason? Finally, if the mostly fused VC cluster is such a powerful sink, why then are some of the most clearly one-word items (*ekha-* 'descend', *sáhi-* 'fix', and *inóngi-* 'do what') not a part of it, instead being found even farther to the lower edge of the map? Many questions such as these arise and must await datasets of significantly larger size and higher granularity in terms of teasing apart competing analyses of the language. The measurements and plots shown in this chapter may be only the first glimpse into an enormous multifaceted ocean of variation among lexical items in Hurukui, an ecosystem into which I hope to be able to peer more deeply over time.

12.6 Discussion

The approach that I use in this study, while allowing insight into how lexical items cluster together on the language's own terms rather than based on notions of similarity imposed by the analyst, has some notable drawbacks. It could very well be that some other properties that were ignored in my construction of the VC space would have been highly informative, and thus that the plots we have seen here are only crude projections of a space that captures the rich structure of the system. For instance, perhaps encoding the argument structure of the VC directly into the dataset using the one-hot encoding would have obviated the need for PCA to find it indirectly using PC2. This would have split argument structure into multiple features because of its multiple possible values, allowing the PCA to then more directly find correlations between certain types of argument structure and other types of behaviors accounted for by the properties. Conversely, it could be that some features that I included in the dataset should have been omitted because they still impose too much top-down analyst intuition on the language's structures. Questions of feature selection such as these are a large topic in data science, and their resolution in the context

of Hurukui LVCs would be a fruitful area for further research combining the disciplines of minority language documentation and quantitative research.

12.6.1 Discussion: Light verb constructions and similar constructions

Lightness is gradient. LVs are the same as Vs in some ways and different in other ways. This gradates as they coevolve with their initials and also depends on how much they remain connected in the speaker's mind to a lexical verb, and how much the initial remains connected to some other entity. If one or both of these parts is replaced elsewhere in the lexicon, it becomes hapax-type in its LVC and then is let loose more in its new life. The debate about whether LVs exist needs gradience the same way as the debate about irrealis does.

Distinguishing macroscopic types (e.g. verb+object, verb+adjunct) is easier in the source cxs rather than in the complex predicates (re: debate over SVC vs AVC vs LVC). The sources leave some traces in the new items' behavioral properties. As these evolve and start interacting with other nascent members of the new system(s), they become intertwined with behaviors of those too. Questions about trying to separate these subtypes of complex predicates also need clusterism rather than either synchronic criteria alone (Seiss 2009) or diachronic source alone (Anderson 2006).

In my view, LVs and LVCs are not dead ends, against what has been argued by some authors (Butt & Lahiri 2002; Butt 2003; Butt 2010). Nor is there just one cline (Bowern 2008), nor are the items necessarily going to finish the trajectory. New gravitational forces can start to emerge as things are in flux and can pull clusters off into new lives.

Because the Hurukui system is still morphologizing, different lexical items are at different stages of univerbation. It is not necessarily the case that all of the LVCs are at different stages on the same diachronic path. Some of them may stop in the intermediate space between unfused and completely-fused, then evolve differently, becoming something else. There are multiple source constructions (object+verb, adverb+verb, ideophone+verb, etc.), which may be the primary reason for this varied behavior.

12.6.1.1 Relationships of the LVCs within speakers' mental grammars

The extent to which speakers' metalinguistically aware analyses would align with my analysis as a linguist remains a very open question and a direction for much future research. That is, to what extent do speakers perceive the behaviors of, say, the fused LVCs as being similar enough to warrant treating them as a group of the same kind of thing? To what extent could speakers consciously and analytically perceive the multi-word LVCs as something different from VOCs? It is likely that speakers' analyses would differ from mine somewhat. My analyses themselves have also evolved over time as my understanding of the language has changed. The history of my analyses would also show different ways of grouping LVC-like structures together based on similarity.

I highly doubt that there is a large schema of "LVC" from which all of the LVC lexical items inherit properties and are children of in the speakers' mental grammars. Many of the highly fused LVCs may be perceived as very different from the unfused ones like *moko e-* (speech take-) 'speak', because the former behave mostly as one word which has a memorized irregular stress pattern. I suspect that if I showed the speakers my list of LVCs (at least the ones that fit my definition reasonably well), they would say that some of them are two words and others are one word. I have already had some discussions like this with them during elicitation sessions about LVCs, and analysis of this metalinguistic commentary in my recordings could shed light on the question of how speakers would characterize these LVC lexical items.

It may be possible to measure the extent to which the two things are still talking to each other in the mental grammar, based on how much they're able to transfer properties to each other. For instance, if one LVC lexical item starts taking a new modifier, or occurring in a new construction, do others follow suit? And if so, is that evidence for how much those are connected in the mind of the speaker as if they should behave the same? I am not familiar with literature on questions such as these, and it would be a fruitful direction for future research on the Hurukui LVC system.

A perfect brain, with infinite memory and infinite computational power, if given an infinite amount of experience with everything, would not need to use categorization or mental

models to predict things, because it would already know all the properties of all things. The entity that is the whole universe is the closest thing to this that exists. The universe has information about everything, but it cannot tell us anything unless we do experiments on it. Any mind or computational system that does not already have information about literally everything has to do some amount of imperfect modeling in order to make predictions.

Language speakers are an example of imperfect brains in this sense, as they have a limited amount of exposure and have to make generalizations and predictions about how a linguistic form can be used and what someone meant by using it in some construction in some context. Speakers have a mental model of how the language works based on incomplete information. They also each have different amounts of exposure to different parts of the language, so certain form-meaning mappings are more entrenched for some speakers than for others.

A highly open question for future research is the extent to which it is useful for speakers to make a certain categorization in their mental grammars, for the purposes of producing and understanding communication successfully. This may be able to be answered quantitatively, although I do not yet know what methods could be used for this. One could try measuring how much predictive power a certain category has for making predictions about unknown properties of forms in that category. One could devise some score value for how informative classifying a novel form as an LVC would be about how it behaves and is distributed regarding the LVC properties.

The purpose of mental categories is to act as a starting point for knowing roughly where something is among the clusters in the feature space. For instance, when asked where they are from, most people will provide a region rather than an exact street address. This allows the listener to make some predictions with a small amount of information transfer. Categories are extremely useful mentally as that kind of first-order approximation. Speakers are doing this somehow in their mental grammars in order to use and comprehend infrequent and novel forms. I hypothesize that treating the fused and unfused LVCs as separate groups would give a lot more predictive power than treating them as the same category, because the space of all LVCs in Hurukui is extremely heterogeneous. There are probably small subgroups where people do treat

all member LVC lexical items as behaving basically the same. Within one of those groups, a speaker could treat all member LVC lexical items the same as they learn new ones, and then refine information about specific lexical items over time as they are exposed to more of the language. However, it remains unclear to what extent speakers actually use such categorizations in their production and perception of Hurukui form-function mappings. A thorough corpus study would be required to shed light on these questions.

Even if speakers do use categorization of subgroups of LVCs, the fact that the individual lexical items have different forms and different meanings still allows them to be pried apart in the feature space. That is, just because two or more lexical items occupy the same point in the feature space at some time does not mean that they must always have the same properties. Because the different lexical items occur in different contexts and co-texts (Catford 1965), they can begin to behave differently, even if speakers started from a generalization that treated them as behaving entirely the same. Usage-based theory emphasizes dynamical processes by which this can happen. For instance, one of the LVCs could become frequent in a certain scenario where the other is not often used, then its verb could become defective because certain inflections are rarely or never used in that scenario. Then the other LVC would retain properties that the one had lost, and the two would diverge in feature space. So the points in feature space are never glued together, even if they start off the same because speakers are making a categorical generalization at some point in time.

12.6.2 Discussion: Categorization and gradience

The Hurukui LVCs show gradient and clustering patterns. Subsective gradience within my definition of LVC. Intersective gradience between LVC and other types like VOC and monomorphemic verb root. Clusterism is a good compromise between lumping and splitting. It has synchronic gradience and diachronic flux built into it. You don't have to create arbitrary cutoffs for when one thing splits into multiple things. Instead, you decide how to carve up the space based on what helps your understanding best, guided by the natural clustering that the data shows.

You have to discretize to be able to measure/observe anything. But as you get more data about what you've treated as one thing, you can go more granular and start looking at how its subparts behave if considered as separate things. For instance, whether to treat a group of formally and semantically similar LVC tokens as one lexical item or multiple. Another example is whether to split a property into more fine-grained ones.

There is the question of which of the groups of points in the high-dimensional space of Hurukui (L)VCs correspond to constructions. That is, which (L)VC lexical items can be analyzed as inheriting from the same parent construction? Because of the gradience of whether a form-meaning mapping is a separate construction or not (Ungerer 2023), the decision to group points together to answer this question inherently involves some arbitrary cutoffs and choices by the analyst. I mostly dodge this question for the purposes of this study. One may ask whether the LVC schema is one entity in the mental grammar or many, and the answer to this question in my clusterist framework of categorization is that it depends on what analytical purpose the grouping(s) would be used for. There is no objective answer to the question of how many constructions are involved here.

References

Aarts, Bas. 2004a. Modelling linguistic gradience. *Studies in Language* 28(1). 1–49. https://doi.org/10.1075/sl.28.1.02aar.

Aarts, Bas. 2004b. Conceptions of gradience in the history of linguistics. *Language Sciences* 26(4). 343–389. https://doi.org/10.1016/j.langsci.2003.07.001.

Aikhenvald, Alexandra Y. 2007. Typological distinctions in word-formation. In Timothy Shopen (ed.), *Language typology and syntactic description III: Grammatical Categories and the Lexicon*, vol. 3, 1–65. 2nd ed. Cambridge, UK; New York: Cambridge University Press.

Anderson, Gregory D. S. 2006. *Auxiliary verb constructions* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.

Barðdal, Jóhanna, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.). 2015. *Diachronic Construction Grammar* (Constructional Approaches to Language). Vol. 18. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/cal.18.

Bisang, Walter. 2010. Grammaticalization in Chinese: A construction-based account. In Elizabeth Closs Traugott & Graeme Trousdale (eds.), *Gradience, Gradualness and Grammaticalization* (Typological Studies in Language), vol. 90, 245–277. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.90.13bis.

Bowern, Claire. 2008. The diachrony of complex predicates. *Diachronica* 25(2). 161–185. https://doi.org/10.1075/dia.25.2.03bow.

Bromley, H. Myron. 1981. *A Grammar of Lower Grand Valley Dani* (Pacific Linguistics C). Vol. 63.

Bruening, Benjamin. 2016. Light Verbs are Just Regular Verbs. *University of Pennsylvania Working Papers in Linguistics* 22(1). 50–60.

Butt, Miriam. 2003. The Light Verb Jungle. Presented at the Workshop on Multi-Verb Constructions, Trondheim.

Butt, Miriam. 2010. The light verb jungle: still hacking away. In Mengistu Amberber, Brett Baker & Mark Harvey (eds.), *Complex Predicates*, 48–78. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511712234.004.

Butt, Miriam & Aditi Lahiri. 2002. Historical Stability vs. Historical Change. University of Konstanz, ms. https://ling.sprachwiss.uni-konstanz.de/pages/home/butt/main/papers/stability.pdf. (17 April, 2024).

Bybee, Joan L. & Clay Beckner. 2012. Usage-Based Theory. In Bernd Heine & Heiko Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 827–856. 1st edn. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199544004.013.0032.

Catford, J. C. 1965. *A Linguistic Theory of Translation*. Vol. 31. London: Oxford University Press.

Cleary-Kemp, Jessica. 2015. *Serial Verb Constructions Revisited: A Case Study from Koro*. University of California, Berkeley.

Corbett, Greville G. 2007. Canonical Typology, Suppletion, and Possible Words.

Croft, William. 2001. *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William A. 2007. Beyond Aristotle and gradience: A reply to Aarts. *Studies in Language* 31(2). 409–430. https://doi.org/10.1075/sl.31.2.05cro.

Darwin, Charles. 1859. *The Origin of Species: By Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: Murray. https://doi.org/10.1017/CBO9780511694295.
De Smet, Hendrik. 2010a. Grammatical interference: Subject marker for and the phrasal verb particles out and forth. In Elizabeth Closs Traugott & Graeme Trousdale (eds.), *Gradience, Gradualness and Grammaticalization* (Typological Studies in Language), vol. 90, 75–104. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.90.06des.
De Smet, Hendrik. 2010b. English -ing-clauses and their problems: The structure of grammatical categories. *Linguistics* 48(6). https://doi.org/10.1515/ling.2010.038.
DeLancey, Scott. 2001. Lectures on Functional Syntax.
Denison, David. 2010. Category change in English with and without structural change. In Elizabeth Closs Traugott & Graeme Trousdale (eds.), *Gradience, Gradualness and Grammaticalization* (Typological Studies in Language), vol. 90, 105–128. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.90.07den.
deVries, Lourens. 2010. From clause conjoining to clause chaining in Dumut languages of New Guinea. *Studies in Language* 34(2). 327–349. https://doi.org/10.1075/sl.34.2.04vri.
Diessel, Holger. 2019. *The Grammar Network: How Linguistic Structure Is Shaped by Language Use*. 1st edn. Cambridge University Press. https://doi.org/10.1017/9781108671040.
Family, Neiloufar. 2014. *Semantic spaces of Persian light verbs* (Brill's Studies in South and Southwest Asian Languages volume 6). Leiden ; Boston: Brill.
Fedden, Sebastian. 2012. Switch-reference and temporal reference in Mian. In Volker Gast & Holger Diessel (eds.), *Clause Linkage in Cross-Linguistic Perspective*. Berlin, Boston: DE GRUYTER. https://doi.org/10.1515/9783110280692.393.
Francis, Elaine J. 2021. *Gradient Acceptability and Linguistic Theory*. 1st edn. Oxford University PressOxford. https://doi.org/10.1093/oso/9780192898944.001.0001.
Fried, Mirjam. 2009. Construction Grammar as a tool for diachronic analysis. *Constructions and Frames* 1(2). 261–290. https://doi.org/10.1075/cf.1.2.04fri.
Givón, T. 2000. Internal reconstruction: As method, as theory. In Spike Gildea (ed.), *Typological Studies in Language*, vol. 43, 107. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.43.05giv.
Givón, Talmy. 2001. *Syntax Volume 1*. Vol. 1. 2 vols. Amsterdam; Philadelphia: Benjamins.
Goldberg, Adele. 2005. *Constructions at Work*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199268511.001.0001.
Goldberg, Adele E. 1995. *Constructions: a construction grammar approach to argument structure* (Cognitive Theory of Language and Culture). 5. pr. Chicago: The Univ. of Chicago Press.
Haiman, John. 1978. Conditionals Are Topics. *Language* 54(3). 564. https://doi.org/10.2307/412787.
Haiman, John. 1983. Iconic and Economic Motivation. *Language* 59(4). 781. https://doi.org/10.2307/413373.
Harris, Alice C. 2008. Light verbs as classifiers in Udi. *Diachronica* 25(2). 213–241. https://doi.org/10.1075/dia.25.2.05har.

Haspelmath, Martin. 2011. The gradual coalescence into 'words' in grammaticalization. In Bernd Heine & Heiko Narrog (eds.), *The Oxford Handbook of Grammaticalization*, 342–355. 1st edn. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199586783.013.0027.
Haspelmath, Martin. 2024. Inflection and derivation as traditional comparative concepts. *Linguistics* 62(1). 43–77. https://doi.org/10.1515/ling-2022-0086.
Hilpert, Martin. 2010. What can synchronic gradience tell us about reanalysis?: Verb-first conditionals in written German and Swedish. In Elizabeth Closs Traugott & Graeme Trousdale (eds.), *Gradience, Gradualness and Grammaticalization* (Typological Studies in Language), vol. 90, 181–201. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.90.10hil.
Hilpert, Martin. 2021. *Ten Lectures on Diachronic Construction Grammar*. BRILL. https://doi.org/10.1163/9789004446793.
Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticization: Opposite or orthogonal? In Walter Bisang, Nikolaus Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? a look from its fringes and its components* (Trends in Linguistics 158), 21–42. Berlin ; New York: Mouton de Gruyter.
Hjørland, Birger. 2017. Classification. *Knowledge Organization* 44(2). 97–128.
Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56(2). 251. https://doi.org/10.2307/413757.
Jespersen, Otto. 1954. *A Modern English Grammar - On Historical Principles - Part VI - Morphology*. S.l.: Routledge.
Jones, Wesley Kuhron. 2023. Hurukui Language Documentation Project. Collection HK1 at catalog.paradisec.org.au [Open Access]: PARADISEC. https://doi.org/https://dx.doi.org/10.26278/3S4S-QF78.
Jones, Wesley Kuhron. in prep. Measuring the consistency of the "irrealis" domain. SocArXiv. https://doi.org/10.31235/osf.io/3md6u.
Jones, Wesley Kuhron. submitted. How pragmatic implicature can thwart a common grammaticalization pathway: From location to completion in Hurukui. SocArXiv. https://doi.org/10.31235/osf.io/x7fuw.
Keenan, Edward J. 1976. Towards a universal definition of subject. Academic Press.
Kuteva, Tania. 2001. *Auxiliation: an enquiry into the nature of grammaticalization* (Oxford Linguistics). 1. publ. new in paperback. Oxford: Oxford University Press.
Langacker, Ronald W. 2006. On the continuous debate about discreteness. *Cognitive Linguistics* 17(1). https://doi.org/10.1515/COG.2006.003.
Levshina, Natalia, Savithry Namboodiripad, Marc Allassonnière-Tang, Mathew Kramer, Luigi Talamo, Annemarie Verkerk, Sasha Wilmoth, et al. 2023. Why we need a gradient approach to word order. *Linguistics* 61(4). 825–883. https://doi.org/10.1515/ling-2021-0098.
McClelland, James L & Joan Bybee. 2007. Gradience of Gradience: A reply to Jackendoff. *The Linguistic Review* 24(4). https://doi.org/10.1515/TLR.2007.019.
McElhanon, K.A. 1975. Isolates: Morobe District: Wasembo (or Gusap). In S.A. Wurm (ed.), *New Guinea area languages and language study, Vol. 1, Papuan languages and the New Guinea*

linguistic scene. The Australian National University: Pacific Linguistics. http://sealang.net/archives/pl/pdf/PL-C38.897.pdf. (23 January, 2020).
McElhanon, Kenneth & Sigkepe Sogum. 1976. English-Wasembo Vocabulary. *Data Papers on Papua New Guinea Languages*. Ukarumpa: Summer Institute of Linguistics 113–119.
Megerdoomian, Karine. 2012. The status of the nominal in Persian complex predicates. *Natural Language & Linguistic Theory* 30(1). 179–216. https://doi.org/10.1007/s11049-011-9146-0.
Minter, Paul. 2009. *Iyo grammar sketch* (Data Papers on Papua New Guinea Languages v. 56). Ukarumpa, Papua New Guinea: SIL-PNG Academic Publications.
Mithun, Marianne. 1999. The Languages of Native North America. Cambridge University Press.
Nash, Léa & Pollet Samvelian. 2016. Introduction: Approaches to Complex Predicates. In Léa Nash & Pollet Samvelian (eds.), *Approaches to complex predicates* (Syntax & Semantics), vol. 41, 1–19. Leiden ; Boston: Brill.
Nolan, Brian. 2014. Complex predicates and light verb constructions in Modern Irish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics* 27(1). 140–167. https://doi.org/10.1075/resla.27.1.06nol.
Pagin, Peter & Dag Westerståhl. 2010. Compositionality I: Definitions and Variants. *Philosophy Compass* 5(3). 250–264. https://doi.org/10.1111/j.1747-9991.2009.00228.x.
Pawley, Andrew. 2012. On the Argument Structure of Complex Predicates in Kalam, a Language of the Trans New Guinea Family. *Annual Meeting of the Berkeley Linguistics Society* 32(2). 83. https://doi.org/10.3765/bls.v32i2.3494.
Pawley, Andrew & Harald Hammarström. 2018. The Trans New Guinea family. In Bill Palmer (ed.), *The Languages and Linguistics of the New Guinea area: A Comprehensive Guide* (The World of Linguistics volume 4), 21–195. Berlin; Boston: De Gruyter Mouton.
Pompei, Anna & Valentina Piunno. 2023. Light Verb Constructions in Romance languages: An attempt to explain systematic irregularity. In Anna Pompei, Lunella Mereu & Valentina Piunno (eds.), *Light Verb Constructions as Complex Verbs*, 99–148. De Gruyter. https://doi.org/10.1515/9783110747997-005.
Reesink, Ger P. 1993. "Inner speech" in Papuan languages. *Languages and Linguistics in Melanesia* 24. 217–225.
Roberts, John R. 1988. Switch-reference in Papuan languages: A syntactic or Extrasyntactic device? *Australian Journal of Linguistics* 8(1). 75–117. https://doi.org/10.1080/07268608808599392.
Roberts, John R. 1997. Switch-reference in Papua New Guinea: A Preliminary Survey. *Papers in Papuan Linguistics* 3. 101–241.
Rosch, Eleanor. 1978. Principles of Categorization. In *Readings in Cognitive Science*, 312–322. Elsevier. https://doi.org/10.1016/B978-1-4832-1446-7.50028-5.
Sadock, Jerrold. 2017. The Subjectivity of the Notion of Polysynthesis. In Michael Fortescue, Marianne Mithun & Nicholas Evans (eds.), *The Oxford Handbook of Polysynthesis*, 99–114. 1st edn. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199683208.013.7.
Seiss, Melanie. 2009. On the difference between auxiliaries, serial verbs and light verbs. In *Proceedings of the LFG09 Conference*, 19. CSLI Publications.

Sorace, Antonella & Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115(11). 1497–1524. https://doi.org/10.1016/j.lingua.2004.07.002.
Taylor, John R. 1998. Syntactic Constructions as Prototype Categories. In *The New Psychology of Language*. Routledge.
Timberlake, Alan. 2007. Aspect, tense, mood. In Timothy Shopen (ed.), *Language typology and syntactic description III: Grammatical Categories and the Lexicon*, vol. 3, 280–333. 2nd ed. Cambridge, UK; New York: Cambridge University Press.
Traugott, Elizabeth Closs & Graeme Trousdale. 2010. Gradience, gradualness and grammaticalization: How do they intersect? In Elizabeth Closs Traugott & Graeme Trousdale (eds.), *Gradience, Gradualness and Grammaticalization* (Typological Studies in Language), vol. 90, 19–44. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.90.04tra.
Ungerer, Tobias. 2023. A gradient notion of constructionhood. *Constructions*. Constructions 15. 2023: Special Issue "35 Years of Constructions" (Editors: Lotte Sommerer&Stefan Hartmann). https://doi.org/10.24338/CONS-543.
van Gijn, Rik. 2016. Switch reference: An overview. In Rik van Gijn & Jeremy Hammond (eds.), *Typological Studies in Language*, vol. 114, 1–54. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.114.01van.
Wells, James D. 2012. *Effective Theories in Physics: From Planetary Orbits to Elementary Particle Masses* (SpringerBriefs in Physics). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34892-1.
Wittgenstein, Ludwig. 1953. *Philosophical investigations*. Oxford: Basil Blackwell.
Z'graggen, John A. 1975. The languages of the Madang District, Papua New Guinea. *Pacific Linguistics* B-41. 160.
Zingler, Tim. 2020. *Wordhood issues: Typology and grammaticalization*. University of New Mexico.