

# A Grammar of Ayeri



# A Grammar of Ayeri

DOCUMENTING A FICTIONAL LANGUAGE

*by Carsten Becker*

---

*Benung. The Ayeri Language Resource*

Carsten Becker, 2017.  
Published under CC-BY-SA 4.0 license.  
Last edited: September 13, 2017.

Set in Junicon and Fira Sans with Xe<sub>La</sub>TeX.

Ayeri is a fictional language spoken by fictional people in a fictional setting, and as such is not related to any naturally existing languages. It is thus not to be confused with *Azeri*, a Turkic language spoken in Azerbaijan and its surrounding countries. Ayeri's vocabulary is entirely *a priori*, this means, no real-world languages have been used specifically as sources of vocabulary. Ayeri is also not derived from any specific real-world language family by means of sound changes. Due to Ayeri's sound and spelling aesthetic being inspired by Austronesian languages, however, occasional overlaps with words existing in these may happen, but only accidentally so.

 <https://ayeri.de>  
 <https://github.com/carbeck/ayerigrammar/>  
 <https://creativecommons.org/licenses/by-sa/4.0/>

# Contents

---

<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>vi</b>
<b>Abbreviations</b>	<b>viii</b>
<b>Preface</b>	<b>xi</b>
<b>o Introduction</b>	<b>i</b>
<b>1 Phonology</b>	<b>7</b>
1.1 Phoneme inventory . . . . .	7
1.1.1 Consonants . . . . .	7
1.1.2 Vowels . . . . .	10
1.2 Phonotactics . . . . .	12
1.2.1 Number of syllables per word . . . . .	13
1.2.2 Phonemic makeup of initial syllables . . . . .	16
1.2.3 Phonemic makeup of medial syllables . . . . .	20
1.2.4 Phonemic makeup of final syllables . . . . .	23
1.2.5 Phonemic makeup of single syllables . . . . .	26
1.2.6 Cross-syllable consonant clusters . . . . .	30
1.3 Notes on prosody . . . . .	31
1.3.1 Stress . . . . .	31
1.3.2 Intonation . . . . .	40
<b>2 Writing system</b>	<b>45</b>
2.1 Consonants . . . . .	47
2.2 Vowels . . . . .	48
2.3 Diacritics . . . . .	50
2.3.1 Subscript diacritics . . . . .	50

2.3.2	Prepended diacritics . . . . .	52
2.3.3	Superscript diacritics . . . . .	56
2.4	Numerals . . . . .	56
2.5	Punctuation and abbreviations . . . . .	57
2.6	Styles . . . . .	59
<b>3</b>	<b>Morphological typology</b>	<b>63</b>
3.1	Typology . . . . .	63
3.2	Morphological processes . . . . .	66
3.2.1	Prefixation . . . . .	66
3.2.2	Suffixation . . . . .	69
3.2.3	Reduplication . . . . .	70
3.2.4	Suprasegmental modification . . . . .	71
3.2.5	Clitics . . . . .	71
3.3	Marking strategies . . . . .	102
<b>4</b>	<b>Grammatical categories</b>	<b>107</b>
4.1	Nouns . . . . .	107
4.1.1	Gender . . . . .	107
4.1.2	Number . . . . .	110
4.1.3	Case . . . . .	113
4.1.4	Prefixes on nouns . . . . .	125
4.1.5	Compounding . . . . .	127
4.1.6	Reduplication . . . . .	137
4.1.7	Nominalization . . . . .	139
4.2	Pronouns . . . . .	144
4.2.1	Personal pronouns . . . . .	144
4.2.2	Demonstrative pronouns . . . . .	148
4.2.3	Interrogative pronouns . . . . .	153
4.2.4	Indefinite pronouns . . . . .	155
4.2.5	Relative pronouns . . . . .	163
4.2.6	Reflexives and reciprocals . . . . .	165
4.3	Adjectives . . . . .	167
4.3.1	Comparison . . . . .	168
4.3.2	Negation . . . . .	168
4.3.3	Adjectivization . . . . .	169
4.3.4	Other affixes . . . . .	170
4.4	Adpositions . . . . .	170
4.4.1	Prepositions . . . . .	171
4.4.2	Postpositions . . . . .	176
4.4.3	Adpositions and time . . . . .	178

4.5	Verbs . . . . .	179
4.5.1	Person marking . . . . .	180
4.5.2	Tense . . . . .	192
4.5.3	Aspect . . . . .	198
4.5.4	Mood . . . . .	203
4.5.5	Modals . . . . .	206
4.5.6	Participle . . . . .	209
4.5.7	Other affixes . . . . .	211
4.6	Adverbs . . . . .	214
4.6.1	Attributive adverbs . . . . .	214
4.6.2	Degree and quantity . . . . .	217
4.6.3	Sentence adverbs . . . . .	219
4.6.4	Demonstrative adverbs . . . . .	222
4.7	Numerals . . . . .	223
4.7.1	Cardinal numerals . . . . .	224
4.7.2	Ordinal numerals . . . . .	229
4.7.3	Multiplicative numerals . . . . .	231
4.7.4	Distributive numerals . . . . .	232
4.7.5	Number ranges . . . . .	232
4.8	Conjunctions . . . . .	233
4.8.1	Simple conjunction and disjunction . . . . .	234
4.8.2	Complex conjunction and disjunction . . . . .	235
<b>5</b>	<b>Phrase structures</b>	<b>239</b>
5.1	Noun and determiner phrases . . . . .	241
5.1.1	Noun phrases . . . . .	242
5.1.2	Determiner phrases . . . . .	249
5.1.3	Nominal clitics . . . . .	274
5.2	Adjective and adverb phrases . . . . .	282
5.2.1	Adjective phrases . . . . .	282
5.2.2	Adverb phrases . . . . .	289
5.3	Adpositional phrases . . . . .	298
5.4	Inflectional and verb phrases . . . . .	310
5.4.1	Typological considerations . . . . .	311
5.4.2	Definition of terms . . . . .	317
5.4.3	Tests on subjecthood . . . . .	322
5.4.4	Establishing the constituent structure . . . . .	328
<b>A</b>	<b>Names</b>	<b>329</b>
A.1	Given names . . . . .	329
A.1.1	Masculine names . . . . .	329

A.I.2	Feminine names . . . . .	330
A.I.3	Gender-neutral names . . . . .	330
<b>B</b>	<b>Example Texts</b>	<b>331</b>
B.1	The North Wind and the Sun . . . . .	331
B.2	The Fox and the Rooster . . . . .	333
B.3	Ozymandias . . . . .	335
	<b>Bibliography</b>	<b>341</b>
	<b>Index</b>	<b>351</b>



## List of Figures

---

0.1	First design for an Ayeri script . . . . .	4
0.2	First draft for Tahano Hikamu . . . . .	5
2.1	Box script and Hikamu . . . . .	46
2.2	Tahano Hikamu, ‘book style’ . . . . .	60
2.3	Tahano Hikamu, ‘angular style’ . . . . .	60
2.4	Tahano Hikamu, ‘hand style’ . . . . .	61
2.5	Tahano Hikamu, ‘blackletter style’ . . . . .	61
4.1	The implicational map for indefinite pronoun functions . . . . .	156
4.2	Map of indefinite pronoun functions in Ayeri . . . . .	159
4.3	The syntax and morphology of pronominal affixes . . . . .	189
5.1	F-structure mappings . . . . .	241

## List of Tables

---

1.1	Consonant inventory . . . . .	8
1.2	Vowel inventory . . . . .	10
1.3	Syllabification of inflected verbs . . . . .	13
1.4	Frequency of words with different numbers of syllables . . . . .	14
1.5	Frequency of syllable types per word . . . . .	15
1.6	Frequency of onsets in initial syllables . . . . .	17
1.7	Frequency of nuclei in initial syllables . . . . .	18
1.8	Frequency of codas in initial syllables . . . . .	20
1.9	Frequency of onsets in medial syllables . . . . .	21
1.10	Frequency of nuclei in medial syllables . . . . .	22
1.11	Frequency of codas in medial syllables . . . . .	23
1.12	Frequency of onsets in final syllables . . . . .	24
1.13	Frequency of nuclei in final syllables . . . . .	25
1.14	Frequency of codas in final syllables . . . . .	26
1.15	Frequency of onsets in single syllables . . . . .	27
1.16	Frequency of nuclei in single syllables . . . . .	29
1.17	Frequency of codas in single syllables . . . . .	29
1.18	Frequency of cross-syllable consonant clusters . . . . .	30
1.19	Declension paradigm for Ayeri <i>ṣ̣ niva</i> ‘eye’ . . . . .	32
1.20	Types of heavy syllables . . . . .	33
1.21	Stress patterns for [+ HEAVY, – LONG] in trisyllabic words . . . . .	36
1.22	Stress patterns for [± HEAVY, + LONG] in trisyllabic words . . . . .	37
1.23	Examples of stress patterns in compounds . . . . .	39
2.1	The consonant graphemes . . . . .	47
2.2	Additional consonant graphemes . . . . .	48
2.3	Primary vowel graphemes . . . . .	48
2.4	Secondary vowel graphemes . . . . .	49
2.5	Bottom-attaching diacritics . . . . .	51

2.7	Obligatorily prepended diacritics . . . . .	53
2.8	Allographically prepended diacritics . . . . .	53
2.9	Superscript diacritics . . . . .	57
2.10	The numerals . . . . .	57
2.11	Common punctuation marks . . . . .	58
2.12	Less common punctuation marks . . . . .	59
4.1	Declension paradigm for ሥፋፋ <i>badan</i> ‘father’ . . . . .	108
4.2	Declension paradigm for ሥፋፋ <i>māva</i> ‘mother’ . . . . .	108
4.3	Declension paradigm for ሥፋፋ <i>kirin</i> ‘street’ . . . . .	109
4.4	Declension paradigm for ሥፋፋ <i>pera</i> ‘measure’ . . . . .	109
4.5	Compounds in the Ayeri dictionary . . . . .	128
4.6	Personal pronouns . . . . .	145
4.7	Demonstrative pronouns . . . . .	149
4.8	Token frequencies of attested demonstrative pronouns . . . . .	150
4.9	Interrogative pronouns . . . . .	153
4.10	Declension paradigm for ሥፋፋ <i>sinya</i> ‘who, what’ . . . . .	154
4.11	Indefinite pronouns . . . . .	157
4.12	Relative pronouns . . . . .	162
4.13	Token frequencies of attested complex relative pronouns . . . . .	163
4.14	Prepositions (simple) . . . . .	171
4.15	Prepositions (directional) . . . . .	174
4.16	Postpositions . . . . .	177
4.17	Adpositions with temporal meaning . . . . .	178
4.18	Conjugation paradigm for ሥፋፋ <i>sob-</i> ‘learn, teach’ . . . . .	180
4.19	Conjugation paradigm for ሥፋፋ <i>anl-</i> ‘bring’ . . . . .	181
4.20	Conjugation paradigm for ሥፋፋ <i>no-</i> ‘want’ . . . . .	182
4.21	Conjugation paradigm for ሥፋፋ <i>apa-</i> ‘laugh’ . . . . .	183
4.22	Verb inflection types in Ayeri . . . . .	188
4.23	Modal verbs and particles . . . . .	207
4.24	Adverbial degree and quantifier suffixes . . . . .	218
4.25	Demonstratives relating to adverbial categories . . . . .	223
4.26	Basic cardinal numerals . . . . .	225
4.27	Numerals for factors of 12 . . . . .	225
4.28	Simple fractions from $\frac{1}{2}$ to $\frac{1}{B}$ . . . . .	228
4.29	Basic ordinal numerals . . . . .	229
5.1	Case alternations of ሥፋፋ <i>ling</i> and ሥፋፋ <i>avan</i> . . . . .	304
5.2	Frequency of word order types . . . . .	312

# Abbreviations

---

1	First person	DEIX	Deixis
2	Second person	DET	Determiner
3	Third person	DF	Discourse function
A	Agent	DIM	Diminutive
ABS	Absolutive	DIR	Directional
ACC	Accusative	DIST	Distal
ADJ	Adjunct	DV	Dative voice
ADJ	Adjective	ERG	Ergative
AF	Argument function	EVID	Evidential
AFF	Affirmative	F	Feminine
AGR	Agreement	FOC	Focus
AGTZ	Agentizer	FUT	Future
AN	Animate	GEN	Genitive
ANIM	Animacy	GEND	Gender
ASP	Aspect	GENT	Genitive topic
AT	Agent topic	GF	Grammatical function
AV	Active voice	HAB	Habitative
CASE	Case	HORT	Hortative
CAUS	Causative	IMP	Imperative
CAUT	Causative topic	INAN	Inanimate
COMP	Complement	IND	Indicative
COMP	Comparative	INDEX	Index
COMPAR	Comparison	INDF	Indefinite
COND	Conditional	INS	Instrumental
CONJ	Conjunction	INST	Instrumental topic
DAT	Dative	INT	Intensifier
DATT	Dative topic	IRR	Irrealis
DEF	Definite	ITER	Iterative

IV	Instrumental voice	PRI	Primary
LNK	Linker	PROG	Progressive
LOC	Locative	PRONTYPE	Pronoun type
LOCT	Locative topic	PROX	Proximal
M	Masculine	PRS	Present
MOD	Modality	PRT	Preterite
N	Neuter	PSEM	Prepositional semantics
NEG	Negative	PST	Past
NFUT	Near future	PT	Patient topic
NMLZ	Nominalizer	PTCP	Participle
NOM	Nominative	Q	Question particle
NPST	Near past	REFL	Reflexive
NUM	Number	REL	Relative
OBJ	Object	RFUT	Remote future
OBJ <sub>θ</sub>	Secondary object	RPST	Remote past
OBL	Oblique	SG	Singular
OBL <sub>θ</sub>	Oblique constituent	ST	Strong
OV	Objective voice	SUBJ	Subject
P	Patient	SUBJ	Subjunctive
PCASE	Prepositional case	SUPL	Superlative
PERS	Person	TEL	Telic
PFV	Perfective	TENSE	Tense
PL	Plural	TOP	Topic
POSS	Possessor	TR	Transitive
POSS	Possessive	WK	Weak
PRED	Predicator	XCOMP	Open complement
PREP	Preposition		

\* ungrammatical or undocumented

? questionable

! running counter to expectation

# marked



# Preface

---

This is my latest attempt to write a grammar of Ayeri, a fictional language which I have been developing since December 2003. Getting to work on grammar writing again was triggered by a growing dissatisfaction with not having a central place of documentation, when the first thing people look for on my website is often the grammar, incomplete as well as partially inaccurate and outdated as it may be. In addition to that, there was a seminar on fictional languages at the University of Tübingen, Germany, in the summer semester of 2016 (Buch 2016). Ayeri was one of the languages that was chosen for students to explore and evaluate.

The student group who worked on Ayeri came to the conclusion that its documentation is severely lacking in the description of basic elements and assumptions, since whole chapters of the grammar had been missing to date (Boga et al. 2016: 12).<sup>1</sup> This is to say that previous attempts of writing a full-fledged grammar of Ayeri have been incomplete due to loss of enthusiasm and creeping neglect.

Although the *Ayeri Grammar* has so far been lying dormant for five years, I have written a whole number of blog articles detailing various grammatical issues (Becker 2016a: Blog). These articles have been taken into consideration here. This grammar writing attempt is thus not only a transferral to a different typesetting system, but constitutes an extension to previous formal documentation as well.

I hope that by transferring my previous grammar writing from LibreOffice to L<sup>A</sup>T<sub>E</sub>X, combined with using GitHub as a version control system, maintaining and editing will become faster, more transparent, and more elegant, since L<sup>A</sup>T<sub>E</sub>X operates on plain text files, and version control helps in keeping track of changes over time.

Carsten Becker  
Marburg, July 18, 2016

<sup>1</sup>  *Kutānas-ikan* ‘thanks a lot’ to Bella Boga, Madita Breuninger, Thora Daneyko, and Martina Stama-Kirr for their hard work on making sense of my published materials in spite of information being scattered all over the place, as well as their providing me with the presentation concluding their group work.





## o Introduction

ကပ်ပယံမိဝံသဉာနိဗ္ဗိကံသယသိပယကံ မံပိကံ ခုနိဗ္ဗဒုက။  
*Paronaya adanyāng si ming tabhanoyyāng, edareng voy kotanas.*  
 ‘He who cannot write believes it not to be toil.’  
 — Anonymous<sup>1</sup>

In December 2003, the idea for a new fictional language was born, an idea that turned out to stick with me for over 10 years now.<sup>2</sup> At that time, my seventeen years old self was still fairly new to this whole making-up languages business, read things about linguistics here and there, and was not shy to ask questions about terminology (and, looking at old mails, a little impertinently teenager-like so), for example on the Conlang mailing list and the Zompist BBoard. One thing seemed to catch my interest especially: syntactic alignments other than the NOM/ACC of the few languages I was familiar with, that is, German, English, and French. Apparently this curiosity was big enough for me to grow bored with my second fictional language, Daléian (declared ‘quite complete’ after maybe half a year of work or so), and to start something new from scratch in order to put newly acquired knowledge to test.

I had read about “trigger languages” on Conlang-L and wanted to try my hands on making my own. I cannot remember how long it took me to come up with a first draft of an Ayeri grammar, however, I do remember having been told that a good language cannot be made in a summer. Of course, I still did not really know what I was doing then, even though I thought I had understood things and authoritatively declared “this is how it works” in my first grammar draft when things sometimes really do not work that way. But at least an interest had been whetted.

In order to illustrate the various stages from the beginnings to current Ayerí, I went through some old backups contemporary with the very early days. Here is a

<sup>1</sup> In the original Latin, *Quia qui nescit scribere putat hoc esse nullum laborem*. Scribe's note in Berlin, State Library, Cod. Lat. fol. 270 (see Bluhme 1858: 589).

<sup>2</sup> Most of the text here is taken from the blog article, “Happy 10th anniversary, Ayer!” (Becker 2013) with some slight rephrasings and extensions.

sentence from the oldest existing document related to it, titled “Draft of 8 Ideas for my 3rd Conlang”—the file’s last-changed date is December 14, 2003, though I remember having started work on Ayeri in early December. I added glossing for convenience and according to what I could reconstruct from the notes. This uses vocabulary and grammatical markers just made up on the spot and for illustrative purposes; little of it actually managed to make it into actual work on Ayeri:

- (1) *Ayevhoi      agiaemaesim      coyaielieðamavir      vhaieloyanaiye.*  
 ay-evhoi      agia-ema-esim      coyai-el-i-eðam-avir      vhai-el-o-yaŋa-iye  
 3SG.AN-SUBJ    read-VERB-SUBJ.AN    book-NOUN-AN-INDF-P    bed-NOUN-INAN-on-LOC  
 ‘He reads a book on the bed.’

According to the grammar draft of September 5, 2004, this would have already changed to:

- (2) *Ang      layaiyain      mecoyalei      ling      \*pinamea.*  
 ang      laya-iy-a-in      me-coya-lei      ling      \*pinam-ea  
 A.SUBJ    read-3SG.AN<sub>1</sub>-a<sub>1</sub>-SUBJ    INDF.INAN-book-P.INAN    top.of    bed-LOC  
 ‘He reads a book on the bed.’

A word for ‘bed’—*pinam*—was only (re-)introduced on October 24, 2008, however. In the current state of Ayeri, I would translate the sentence as follows:

- (3) *Ang layaya      koyaley      ling      pinamya.*  
 ang    laya=ya.Ø      koya-ley      ling      pinam-ya  
 AT    read=3SG.M.TOP    book-P.INAN    top.of    bed-LOC  
 ‘He reads a book on a/the bed.’

As you can see, quite a bit of morphology got lost already early on, especially the overt part-of-speech marking (!) and animacy marking on nouns. Also, prepositions were just incorporated into a noun complex as suffixes apparently. Gender was originally only divided into animate and inanimate, but I changed that at some point because only being really familiar with European languages, it felt awkward to me not to be able to explicitly distinguish ‘he’, ‘she’, and ‘it’.

A feature that also got lost is the assignment of thematic vowels in personal pronouns to third-person referents: originally, every third-person referent newly introduced into discourse would be assigned one of /a e i o u/ to disambiguate, and there was even a morpheme to mark that the speaker wanted to dissolve the association. Constituent order was theoretically variable at first, but I preferred SVO/AVP due to familiarity with that. Later on, however, I settled on VSO/VAP. Also, I had no idea about what was called “trigger morphology” on Conlang-L for the longest time—essentially, this referred to the Austronesian, or Philippine,

alignment. I am not claiming that I know all about it now, just that due to reading up on the topic, I have a slightly more informed understanding now. Orthography changed as well over the years, so ⟨c⟩ in the early examples encodes the /k/ sound, not /tʃ/ as it does today; diphthongs were spelled as ⟨Vi⟩ instead of modern ⟨Vy⟩.

What was definitely beneficial for the development of Ayeri was the ever increasing amount of linguistics materials available online and my entering university (to study literature) in 2009, where I learned how to do research and also had a lot of interesting books available at the library.

One of the things people regularly compliment me on is Ayeri's script—note, however, that Tahano Hikamu was not the first one I came up with for Ayeri. Apparently, I had already been fascinated with the look of Javanese/Balinese writing early on;<sup>3</sup> Figure 0.1 shows a draft dated February 9, 2004. However, the letter shapes in this draft looked so confusingly alike that I could never memorize them. About a year later, I came up with the draft in Figure 0.2. What is titled “Another Experimental Script” there is what would later turn into Tahano Hikamu, Ayeri's ‘native’ script. According to the notes in my fictional language ring binder, the script looked much the same as today about a year from then, but things have only been mostly stable since about 2008.

An important date in the history of Ayeri was when I decided to set up an improved website for Ayeri that would include a blog. The idea was that this way, I could more freely write on whatever detail I currently interested me in Ayeri, outside of the constraints of the Grammar. Thus, *Benung. The Ayeri Language Resource* launched on March 1, 2011. Being able to write short articles, however, probably also led to neglecting work on the actual formal reference grammar, which had been lying dormant from January 2011 on. This was always on the premise that I would eventually include the information from blog articles in the grammar. However, juggling such a big document had always felt daunting, so I let laziness take the better part of me eventually as enthusiasm gradually subsided.<sup>4</sup> This renewed attempt at documentation has been started with the intention to right those wrongs.

I hope that by now it should be clear which kind of a fictional language Ayeri is: a personal, artistic language—or *artlang* in community parlance. Thus, my goal in creating Ayeri is not to propose yet another international auxiliary language, like Esperanto. It is also not my goal to make it as logical as possible, like Lojban. Neither is it my goal to engineer it towards certain underlying premises, for example,

<sup>3</sup> Compare, for instance, the charts in Kuipers and McDermott (1996). The Wikipedia articles on either script contain a number of images depicting the scripts in use, both current and historic.

<sup>4</sup> Let me add to my defense, however, that I also worked on my B.A. thesis in 2013 and my M.A. thesis in 2016, which required several months of preparation each and thus left me largely unable to work much on Ayeri.



Figure 0.1: First design for an Ayeri script (February 9, 2004)

to reach a maximal amount of information density, like Ithkuil, or to get by on as few different words as possible, like Toki Pona. It is also not a ‘what-if’ language in the sense of “What could the modern language of Old Irish speakers transplanted to Australia look like?” or “Latin piped through Athabaskan sound changes.”

Ayeri is rather an attempt to create an artificial language for personal enjoyment and intellectual stimulation by creating a feedback loop between reading up on linguistics and actively devising rules for a fictional language accordingly, to see how things work within the frame I created, or to try and see whether certain ideas work together at all when combined, and to better understand why they do or do not. Ayeri will only ever be as perfect as miniature models of things can be, since it has not grown organically from millenia of human interaction, and I cannot and will never know about each and every aspect of language myself, in spite of continued curiosity about these matters. Nor will it be possible for me to replicate all the fascinating twists and irregularities that natural languages normally entail. The ultimate goal in my work on Ayeri is, I suppose, to make it emulate natural languages to at least some degree of depth and complexity.

In writing this grammar, I hope that I will find a good balance between applying linguistic theory to already existing materials and ideas, and going forth to create rules for aspects of the language that have so far been neglected, often due to my

pa ɲ      ɲ = placeholder + vowel carrier      ANOTHER EXPERIMENTAL SCRIPT 03/23/2005

ta ɲa

ka ɲb

ba ɲ

da ɲ

ga ɲ

ma ɲ

na ɲ

ɲa ɲ      -ɲ ɲ

va ɲ

ɲa ɲ < ɲi

ha ɲb

ra ɲ

la ɲ

ɲa ɲ

vivaɲa: ɲ or ɲ or ɲ  
is only used for names

a- ɲ      -in ɲ

e ɲ      -on ɲ

i ɲ

o ɲ      ie = ye ɲ

u ɲ

ɲi ɲ

ɲ = makes vowels long

ɲ = indicates "variable vowel"

1 ɲ

2 ɲ

3 ɲ

4 ɲ

5 ɲ

6 ɲ

7 ɲ

8 ɲ

9 ɲ

10 ɲ

11 ɲ

12 ɲ

13 ɲ

14 ɲ

15 ɲ

16 ɲ

17 ɲ

18 ɲ

19 ɲ

20 ɲ

21 ɲ

22 ɲ

23 ɲ

24 ɲ

25 ɲ

26 ɲ

27 ɲ

28 ɲ

29 ɲ

30 ɲ

31 ɲ

32 ɲ

33 ɲ

34 ɲ

35 ɲ

36 ɲ

37 ɲ

38 ɲ

39 ɲ

40 ɲ

41 ɲ

42 ɲ

43 ɲ

44 ɲ

45 ɲ

46 ɲ

47 ɲ

48 ɲ

49 ɲ

50 ɲ

51 ɲ

52 ɲ

53 ɲ

54 ɲ

55 ɲ

56 ɲ

57 ɲ

58 ɲ

59 ɲ

60 ɲ

61 ɲ

62 ɲ

63 ɲ

64 ɲ

65 ɲ

66 ɲ

67 ɲ

68 ɲ

69 ɲ

70 ɲ

71 ɲ

72 ɲ

73 ɲ

74 ɲ

75 ɲ

76 ɲ

77 ɲ

78 ɲ

79 ɲ

80 ɲ

81 ɲ

82 ɲ

83 ɲ

84 ɲ

85 ɲ

86 ɲ

87 ɲ

88 ɲ

89 ɲ

90 ɲ

91 ɲ

92 ɲ

93 ɲ

94 ɲ

95 ɲ

96 ɲ

97 ɲ

98 ɲ

99 ɲ

100 ɲ

101 ɲ

102 ɲ

103 ɲ

104 ɲ

105 ɲ

106 ɲ

107 ɲ

108 ɲ

109 ɲ

110 ɲ

111 ɲ

112 ɲ

113 ɲ

114 ɲ

115 ɲ

116 ɲ

117 ɲ

118 ɲ

119 ɲ

120 ɲ

121 ɲ

122 ɲ

123 ɲ

124 ɲ

125 ɲ

126 ɲ

127 ɲ

128 ɲ

129 ɲ

130 ɲ

131 ɲ

132 ɲ

133 ɲ

134 ɲ

135 ɲ

136 ɲ

137 ɲ

138 ɲ

139 ɲ

140 ɲ

141 ɲ

142 ɲ

143 ɲ

144 ɲ

145 ɲ

146 ɲ

147 ɲ

148 ɲ

149 ɲ

150 ɲ

151 ɲ

152 ɲ

153 ɲ

154 ɲ

155 ɲ

156 ɲ

157 ɲ

158 ɲ

159 ɲ

160 ɲ

161 ɲ

162 ɲ

163 ɲ

164 ɲ

165 ɲ

166 ɲ

167 ɲ

168 ɲ

169 ɲ

170 ɲ

171 ɲ

172 ɲ

173 ɲ

174 ɲ

175 ɲ

176 ɲ

177 ɲ

178 ɲ

179 ɲ

180 ɲ

181 ɲ

182 ɲ

183 ɲ

184 ɲ

185 ɲ

186 ɲ

187 ɲ

188 ɲ

189 ɲ

190 ɲ

191 ɲ

192 ɲ

193 ɲ

194 ɲ

195 ɲ

196 ɲ

197 ɲ

198 ɲ

199 ɲ

200 ɲ

201 ɲ

202 ɲ

203 ɲ

204 ɲ

205 ɲ

206 ɲ

207 ɲ

208 ɲ

209 ɲ

210 ɲ

211 ɲ

212 ɲ

213 ɲ

214 ɲ

215 ɲ

216 ɲ

217 ɲ

218 ɲ

219 ɲ

220 ɲ

221 ɲ

222 ɲ

223 ɲ

224 ɲ

225 ɲ

226 ɲ

227 ɲ

228 ɲ

229 ɲ

230 ɲ

231 ɲ

232 ɲ

233 ɲ

234 ɲ

235 ɲ

236 ɲ

237 ɲ

238 ɲ

239 ɲ

240 ɲ

241 ɲ

242 ɲ

243 ɲ

244 ɲ

245 ɲ

246 ɲ

247 ɲ

248 ɲ

249 ɲ

250 ɲ

251 ɲ

252 ɲ

253 ɲ

254 ɲ

255 ɲ

256 ɲ

257 ɲ

258 ɲ

259 ɲ

260 ɲ

261 ɲ

262 ɲ

263 ɲ

264 ɲ

265 ɲ

266 ɲ

267 ɲ

268 ɲ

269 ɲ

270 ɲ

271 ɲ

272 ɲ

273 ɲ

274 ɲ

275 ɲ

276 ɲ

277 ɲ

278 ɲ

279 ɲ

280 ɲ

281 ɲ

282 ɲ

283 ɲ

284 ɲ

285 ɲ

286 ɲ

287 ɲ

288 ɲ

289 ɲ

290 ɲ

291 ɲ

292 ɲ

293 ɲ

294 ɲ

295 ɲ

296 ɲ

297 ɲ

298 ɲ

299 ɲ

300 ɲ

301 ɲ

302 ɲ

303 ɲ

304 ɲ

305 ɲ

306 ɲ

307 ɲ

308 ɲ

309 ɲ

310 ɲ

311 ɲ

312 ɲ

313 ɲ

314 ɲ

315 ɲ

316 ɲ

317 ɲ

318 ɲ

319 ɲ

320 ɲ

321 ɲ

322 ɲ

323 ɲ

324 ɲ

325 ɲ

326 ɲ

327 ɲ

328 ɲ

329 ɲ

330 ɲ

331 ɲ

332 ɲ

333 ɲ

334 ɲ

335 ɲ

336 ɲ

337 ɲ

338 ɲ

339 ɲ

340 ɲ

341 ɲ

342 ɲ

343 ɲ

344 ɲ

345 ɲ

346 ɲ

347 ɲ

348 ɲ

349 ɲ

350 ɲ

351 ɲ

352 ɲ

353 ɲ

354 ɲ

355 ɲ

356 ɲ

357 ɲ

358 ɲ

359 ɲ

360 ɲ

361 ɲ

362 ɲ

363 ɲ

364 ɲ

365 ɲ

366 ɲ

367 ɲ

368 ɲ

369 ɲ

370 ɲ

371 ɲ

372 ɲ

373 ɲ

374 ɲ

375 ɲ

376 ɲ

377 ɲ

378 ɲ

379 ɲ

380 ɲ

381 ɲ

382 ɲ

383 ɲ

384 ɲ

385 ɲ

386 ɲ

387 ɲ

388 ɲ

389 ɲ

390 ɲ

391 ɲ

392 ɲ

393 ɲ

394 ɲ

395 ɲ

396 ɲ

397 ɲ

398 ɲ

399 ɲ

400 ɲ

401 ɲ

402 ɲ

403 ɲ

404 ɲ

405 ɲ

406 ɲ

407 ɲ

408 ɲ

409 ɲ

410 ɲ

411 ɲ

412 ɲ

413 ɲ

414 ɲ

415 ɲ

416 ɲ

417 ɲ

418 ɲ

419 ɲ

420 ɲ

421 ɲ

422 ɲ

423 ɲ

424 ɲ

425 ɲ

426 ɲ

427 ɲ

428 ɲ

429 ɲ

430 ɲ

431 ɲ

432 ɲ

433 ɲ

434 ɲ

435 ɲ

436 ɲ

437 ɲ

438 ɲ

439 ɲ

440 ɲ

441 ɲ

442 ɲ

443 ɲ

444 ɲ

445 ɲ

446 ɲ

447 ɲ

448 ɲ

449 ɲ

450 ɲ

451 ɲ

452 ɲ

453 ɲ

454 ɲ

455 ɲ

456 ɲ

457 ɲ

458 ɲ

459 ɲ

460 ɲ

461 ɲ

462 ɲ

463 ɲ

464 ɲ

465 ɲ

466 ɲ

467 ɲ

468 ɲ

469 ɲ

470 ɲ

471 ɲ

472 ɲ

473 ɲ

474 ɲ

475 ɲ

476 ɲ

477 ɲ

478 ɲ

479 ɲ

480 ɲ

481 ɲ

482 ɲ

483 ɲ

484 ɲ

485 ɲ

486 ɲ

487 ɲ

488 ɲ

489 ɲ

490 ɲ

491 ɲ

492 ɲ

493 ɲ

494 ɲ

495 ɲ

496 ɲ

497 ɲ

498 ɲ

499 ɲ

500 ɲ

501 ɲ

502 ɲ

503 ɲ

504 ɲ

505 ɲ

506 ɲ

507 ɲ

508 ɲ

509 ɲ

510 ɲ

511 ɲ

512 ɲ

513 ɲ

514 ɲ

515 ɲ

516 ɲ

517 ɲ

518 ɲ

519 ɲ

520 ɲ

521 ɲ

522 ɲ

523 ɲ

524 ɲ

525 ɲ

526 ɲ

527 ɲ

528 ɲ

529 ɲ

530 ɲ

531 ɲ

532 ɲ

533 ɲ

534 ɲ

535 ɲ

536 ɲ

537 ɲ

538 ɲ

539 ɲ

540 ɲ

541 ɲ

542 ɲ

543 ɲ

544 ɲ

545 ɲ

546 ɲ

547 ɲ

548 ɲ

549 ɲ

550 ɲ

551 ɲ

552 ɲ

553 ɲ

554 ɲ

555 ɲ

556 ɲ

557 ɲ

558 ɲ

559 ɲ

560 ɲ

561 ɲ

562 ɲ

563 ɲ

564 ɲ

565 ɲ

566 ɲ

567 ɲ

568 ɲ

569 ɲ

570 ɲ

571 ɲ

572 ɲ

573 ɲ

574 ɲ

575 ɲ

576 ɲ

577 ɲ

578 ɲ

579 ɲ

580 ɲ

581 ɲ

582 ɲ

583 ɲ

584 ɲ

585 ɲ

586 ɲ

587 ɲ

588 ɲ

589 ɲ

590 ɲ

591 ɲ

592 ɲ

593 ɲ

594 ɲ

595 ɲ

596 ɲ

597 ɲ

598 ɲ

599 ɲ

600 ɲ

601 ɲ

602 ɲ

603 ɲ

604 ɲ

605 ɲ

606 ɲ

607 ɲ

608 ɲ

609 ɲ

610 ɲ

611 ɲ

612 ɲ

613 ɲ

614 ɲ

615 ɲ

616 ɲ

617 ɲ

618 ɲ

619 ɲ

620 ɲ

621 ɲ

622 ɲ

623 ɲ

624 ɲ

625 ɲ

626 ɲ

627 ɲ

628 ɲ

629 ɲ

630 ɲ

631 ɲ

632 ɲ

633 ɲ

634 ɲ

635 ɲ

636 ɲ

637 ɲ

638 ɲ

639 ɲ

640 ɲ

641 ɲ

642 ɲ

643 ɲ

644 ɲ

645 ɲ

646 ɲ

647 ɲ

648 ɲ

649 ɲ

650 ɲ

651 ɲ

652 ɲ

653 ɲ

654 ɲ

655 ɲ

656 ɲ

657 ɲ

658 ɲ

659 ɲ

660 ɲ

661 ɲ

662 ɲ

663 ɲ

664 ɲ

665 ɲ

666 ɲ

667 ɲ

668 ɲ

669 ɲ

670 ɲ

671 ɲ

672 ɲ

673 ɲ

674 ɲ

675 ɲ

676 ɲ

677 ɲ

678 ɲ

679 ɲ

680 ɲ

681 ɲ

682 ɲ

683 ɲ

684 ɲ

685 ɲ

686 ɲ

687 ɲ

688 ɲ

689 ɲ

690 ɲ

691 ɲ

692 ɲ

693 ɲ

694 ɲ

695 ɲ

696 ɲ

697 ɲ

698 ɲ

699 ɲ

700 ɲ

701 ɲ

702 ɲ

703 ɲ

704 ɲ

705 ɲ

706 ɲ

707 ɲ

708 ɲ

709 ɲ

710 ɲ

711 ɲ

712 ɲ

713 ɲ

714 ɲ

715 ɲ

716 ɲ

717 ɲ

718 ɲ

719 ɲ

720 ɲ

721 ɲ

722 ɲ

723 ɲ

724 ɲ

725 ɲ

726 ɲ

727 ɲ

728 ɲ

729 ɲ

730 ɲ

731 ɲ

732 ɲ

733 ɲ

734 ɲ

735 ɲ

736 ɲ

737 ɲ

738 ɲ

739 ɲ

740 ɲ

741 ɲ

742 ɲ

743 ɲ

744 ɲ

745 ɲ

746 ɲ

747 ɲ

748 ɲ

749 ɲ

750 ɲ

751 ɲ

752 ɲ

753 ɲ

754 ɲ

755 ɲ

756 ɲ

757 ɲ

758 ɲ

759 ɲ

760 ɲ

761 ɲ

762 ɲ

763 ɲ

764 ɲ

765 ɲ

766 ɲ

767 ɲ

768 ɲ

769 ɲ

770 ɲ

771 ɲ

772 ɲ

773 ɲ

774 ɲ

775 ɲ

776 ɲ

777 ɲ

778 ɲ

779 ɲ

780 ɲ

781 ɲ

782 ɲ

783 ɲ

784 ɲ

785 ɲ

786 ɲ

787 ɲ

788 ɲ

789 ɲ

790 ɲ

791 ɲ

792 ɲ

793 ɲ

794 ɲ

795 ɲ

796 ɲ

797 ɲ

798 ɲ

799 ɲ

800 ɲ

801 ɲ

802 ɲ

803 ɲ

804 ɲ

805 ɲ

806 ɲ

807 ɲ

808 ɲ

809 ɲ

810 ɲ

811 ɲ

812 ɲ

813 ɲ

814 ɲ

815 ɲ

816 ɲ

817 ɲ

818 ɲ

819 ɲ

820 ɲ

821 ɲ

822 ɲ

823 ɲ

824 ɲ

825 ɲ

826 ɲ

827 ɲ

828 ɲ

829 ɲ

830 ɲ

831 ɲ

832 ɲ

833 ɲ

834 ɲ

835 ɲ

836 ɲ

837 ɲ

838 ɲ

839 ɲ

840 ɲ

841 ɲ

842 ɲ

843 ɲ

844 ɲ

845 ɲ

846 ɲ

847 ɲ

848 ɲ

849 ɲ

850 ɲ

851 ɲ

852 ɲ

853 ɲ

854 ɲ

855 ɲ

856 ɲ

857 ɲ

858 ɲ

859 ɲ

860 ɲ

861 ɲ

862 ɲ

863 ɲ

864 ɲ

865 ɲ

866 ɲ

867 ɲ

868 ɲ

869 ɲ

870 ɲ

871 ɲ

872 ɲ

873 ɲ

874 ɲ

875 ɲ

876 ɲ

877 ɲ

878 ɲ

879 ɲ

880 ɲ

881 ɲ

882 ɲ

883 ɲ

884 ɲ

885 ɲ

886 ɲ

887 ɲ

888 ɲ

889 ɲ

890 ɲ

891 ɲ

892 ɲ

893 ɲ

894 ɲ

895 ɲ

896 ɲ

897 ɲ

898 ɲ

899 ɲ

900 ɲ

901 ɲ

902 ɲ

903 ɲ

904 ɲ

905 ɲ

906 ɲ

907 ɲ

908 ɲ

909 ɲ

910 ɲ

911 ɲ

912 ɲ

913 ɲ

914 ɲ

915 ɲ

916 ɲ

917 ɲ

918 ɲ

919 ɲ

920 ɲ

921 ɲ

922 ɲ

923 ɲ

924 ɲ

925 ɲ

926 ɲ

927 ɲ

928 ɲ

929 ɲ

930 ɲ

931 ɲ

932 ɲ

933 ɲ

934 ɲ

935 ɲ

936 ɲ

937 ɲ

938 ɲ

939 ɲ

940 ɲ

941 ɲ

942 ɲ

943 ɲ

944 ɲ

945 ɲ

946 ɲ

947 ɲ

948 ɲ

949 ɲ

950 ɲ

951 ɲ

952 ɲ

953 ɲ

954 ɲ

955 ɲ

956 ɲ

957 ɲ

958 ɲ

959 ɲ

960 ɲ

961 ɲ

962 ɲ

963 ɲ

964 ɲ

965 ɲ

966 ɲ

967 ɲ

968 ɲ

969 ɲ

970 ɲ

971 ɲ

972 ɲ

973 ɲ

974 ɲ

975 ɲ

976 ɲ

977 ɲ

978 ɲ

979 ɲ

980 ɲ

981 ɲ

982 ɲ

983 ɲ

984 ɲ

985 ɲ

986 ɲ

987 ɲ

988 ɲ

989 ɲ

990 ɲ

991 ɲ

992 ɲ

993 ɲ

994 ɲ

995 ɲ

996 ɲ

997 ɲ

998 ɲ

999 ɲ

1000 ɲ

1001 ɲ

1002 ɲ

1003 ɲ

1004 ɲ

1005 ɲ

1006 ɲ

1007 ɲ

finally discusses how syntactic phrases are built up from words, eventually leading to the formation of complete sentences.

# 1 Phonology

---

This chapter will present charts depicting the phoneme inventory of Ayeri and describe the various commonly encountered allophones of both consonants and vowels. Following this, a detailed statistical analysis of the words found in a number of translated texts from 2008 to 2016 as well as dictionary entries up to July 2016 will produce insights into Ayeri’s phonotactics. Some notes on stress patterns and intonation will close the chapter.

## 1.1 Phoneme inventory

### 1.1.1 Consonants

At 17 consonants, Ayeri has a “moderately small” inventory, according to Maddieson (2013a). Table 1.1 shows the full chart of consonant phonemes.

Regarding allophony, /tj kj/ and /dj gj/ are usually realized as [tʃ] and [dʒ], respectively, except if a homorganic nasal /n/ or /ŋ/ is preceding: for instance, ႁႏႱ *ankyū* /ʼaŋkju/ ‘really’ is realized as [ʼaŋkju], not as \*[ʼaŋtʃu] or \*[ʼantʃu]. It is important to note, however, that besides this synchronic palatalization process leading to [tʃ] and [dʒ] as *allophones*, there is also a diachronic one in parallel here—or the diachronic process is still ongoing. For example, there is no way to predict whether ႁႏ *cuna* ‘original, initial’, ႏႱ *panca* ‘finally, eventually’, and ႏႱ *vac-* ‘like’, or ႏႱ *jaraŋ* ‘pilgrimage’, ႏႱ *aja-* ‘play’, and ႏႱ *nui-* ‘pour’ have /tj/ or /kj/, /dj/ or /gj/, respectively, unless we consider the clues given by the conservative native spellings of the respective words.<sup>1</sup> We can rather assume two sound changes, (1) tj, kj → tʃ, and (2) dj, gj → dʒ, leading to the *phonemes* /tʃ/ and /dʒ/ in the present-day language.

<sup>1</sup> Actual scribes would typically err in cases where the merger is complete, so this strategy would, in fact, be of limited use in the real world.

Table 1.1: Consonant inventory (divergent orthography in pointed brackets)

	Bilabial		Labiodental		Alveolar		Palatal		Velar		Glottal	
Plosive	p	b			t	d			k	g		
Affricate					tʃ <c>	dʒ <j>						
Nasal		m				n				ŋ <ng>		
Fricative				v	s							h
Tap/Flap						r						
Approximant						l		j <y>				



The plural marker ၵ -ye is commonly contracted to [dʒ] when a case suffix beginning with a vowel follows:<sup>2</sup>

- (1) a. ၵၵၵၵၵၵ nyānyēang → nyānjāng [ˈnjaːndʒaŋ] ‘persons’ (person-PL-A);  
 b. ၵၵၵၵ netuyēas → netujas [neˈtudʒas] ‘brothers’ (brother-PL-P).

The plural marker may also contract before the locative marker ၵ -ya and the dative marker ၵ -yam, basically for dissimilation:<sup>3</sup>

- (2) a. ၵၵၵ nīvayēya → nīvajya [niˈvadʒja] ‘at the eyes’ (eye-PL-LOC);  
 b. ၵၵၵ maviyēyam → mavijyam [maˈvidʒjam] ‘to the sheep’ (sheep-PL-DAT).

Dissimilation of the sequence ၵ -yaya is attested in my translation of Kafka’s short story “Eine kaiserliche Botschaft,” where the relative pronoun ၵ siyaya appears transcribed as *sijya*:

As far as morphophonology is concerned, the relative pronoun complex *sijya* ‘in/at/on which.LOC’ is interesting in so far as it is a contraction of \**siyaya* ‘REL-LOC-LOC’ that I introduced here [...] Since this feature does not occur in previous texts, let’s assume it’s an acceptable variant. (Becker 2012: 12)

The contraction happens “only if both parts are grammatical suffixes” (12), however, so the environments this contraction may appear in are effectively limited to relative pronouns combining locative and locative, or locative and dative marking.

The word ၵ lajāy ‘student’ is special in that it is the only word with ၵ [dʒa] so far. Presumably it is derived from the verb ၵ laya- ‘read’ with the agentive suffix ၵ -maya, except the shortening of the suffix—with or without compensatory lengthening of the final vowel of the modified word stem—was applied irregularly, possibly via \**layāya*. The regular form ၵ layamaya means ‘reader’.

Lastly, /h/ may assimilate to its phonemic environment and is realized as [ç] before front vowels, and as [x] before back vowels in this case:

- (3) a. ၵ tabi [ˈtaçi] ‘favorable’;  
 b. ၵ babo [ˈbaxo] ‘loud’.

While vowels become long when two identical vowels come into succession, consonants do not geminate but are treated like a single consonant:

<sup>2</sup> The customary romanization uses ⟨c⟩ and ⟨j⟩ for allophonic cases of [tʃ] and [dʒ] as well.

<sup>3</sup> ၵ -ea also occurs as an allomorph, so that ၵ -ye + ၵ -ea → ၵ -yēa.

Table 1.2: Vowel inventory (divergent orthography in pointed brackets)

	Front	Center	Back
High	i, i: ⟨ī⟩		u, u: ⟨ū⟩
Mid	e, e: ⟨ē⟩	ə ⟨ə, e⟩	o, o: ⟨ō⟩
Low		a, a: ⟨ā⟩	

- (4) a.  $\text{တပ်တပ်}$  *tavvāng* [ta'va:ŋ] 'you get' (get=2SG.A),  
 b.  $\text{နှိပ်နှိပ်}$  *disyyang* [di'sjaŋ] 'I fasten' (fasten=1SG.A).

With diphthongs, the sequence /V<sub>1</sub>.j/ is treated as though it were /V<sub>2</sub>.j/, so the double /j/ simplifies to just a single /j/; however, the vowel remains lax in spite of being phonetically in an open position now:

- (5)  $\text{မြေပေါ်ပေါ်}$  *tipuyya* [ti'pu.ja] 'on the grass' (grass-LOC).

### 1.1.2 Vowels

Ayeri's vowel system distinguishes five qualities, as shown in Table 1.2; Maddieson (2013c) classifies this as "average." Length, however, is also a factor, and there are five diphthongs as well, as we will see below. At 17 : 5, the consonant–vowel ratio is 4.25, which Maddieson (2013b) again classifies as "average," although Ayeri finds itself at the upper end of the tier.

The lax vowels [ɪ ɛ ɔ ʊ] occur as allophones of their tense counterparts /i e o u/ in closed syllables, for example:

- (6) a.  $\text{မိန့်}$  *ming* [mɪŋ] 'can, be able',  
 b.  $\text{မင်း}$  *enya* ['ɛn.ja] 'everyone',  
 c.  $\text{မင်း}$  *agon* ['a.gɔŋ] 'outer, foreign', and  
 d.  $\text{ပန်းနီ}$  *pakur* ['pa.kʊr] 'ill, sick'.

/ə/ occurs marginally in the tense prefixes  $\text{န}$ : *kə-* 'NPST',  $\text{မ}$ : *mə-* 'PST',  $\text{ပ}$ : *və-* 'RPST', as well as in the prefix  $\text{မ}$ : *mə-* 'some, whichever'. Otherwise, [ə] acts as an allophone of /e/ in final unstressed position, for instance, in the word  $\text{မိန့်}$  *mine* ['minə] 'affair, matter, issue'.

Ayeri also possesses a number of diphthongs, these are: /ai ei ɔi ʊi au/, spelled ⟨ay⟩, ⟨ey⟩, ⟨oy⟩, ⟨uy⟩, and ⟨au⟩. Furthermore, there are long equivalents of the short vowels: /i: e: a: o: u:/; in romanization, long vowels are marked with a macron ⟨̄⟩ over the letter. Long vowels are lexicalized in a few words, for example:

- (7) a. ခဉ်း *nīsa* ‘wanted’, ကုန််း *pasīsa* ‘interesting’;  
 b. ခဉ်း *arēn* ‘anyway, however’, ကုန််း *lēra* ‘whore’;  
 c. ကုန််း *lā* ‘tongue’, ယုန််း *yāng* ‘he’ (he.A);  
 d. ခဉ်း *nōn* ‘will, intention’; and  
 e. ခဉ်း *babūan* ‘barbarian’.<sup>4</sup>

Otherwise, long vowels result from two same vowels next to each other, for instance:

- (8) ခဉ်း *aja-* ‘play’ + ခဉ်း *-an* ‘NMLZ’ → ခဉ်း *ajān* ‘game, play’.

Morphophonologically, long vowels also occur in double-marked relative pronouns where the agreement marker for the relative clause’s head has been omitted, for instance, ခဉ်း *sinā* ‘of which, about which’, as in the following example:

- (9) *Le turayāng taman sinā ang ningay tamala vās.*  
*le tura-yāng [taman-Ø]<sub>1</sub> si-Ø<sub>1</sub>-na ang ning=ay.Ø tamala vās*  
 PT.INAN send=3SG.M.A letter-TOP REL-PT.INAN-GEN AT tell=1SG.TOP yesterday 2SG.P

‘The letter which I told you about yesterday, he sent it.’

This is to disambiguate it from the plain genitive-marked relative pronoun ခဉ်း *sina* ‘which.GEN’:<sup>5</sup>

- (10) *tamanreng ledanena nā sina koronvāng*  
*taman-reng [ledan-ena nā]<sub>1</sub> si-na<sub>1</sub> koron=vāng*  
 letter-A.INAN friend-GEN 1SG.GEN REL-GEN know=2SG.A

‘the letter of my friend which you know’

As pointed out in (7c), the word ကုန််း *lā* ‘tongue’ ends in a long vowel, so the question is what happens when a case suffix beginning with a vowel is appended. To avoid a hiat, a glide /j/ may be inserted, so both of the following renditions are possible:

<sup>4</sup> I have gone years without dictionary entries for /u:/, but it has always seemed slightly odd to me to lack a vowel in that position when all other vowels can be long. Therefore, ခဉ်း *babūan* ‘barbarian’ and its adjective ခဉ်း *babū* ‘barbarian (adj.)’ were coined as ကုန််း *prankaye*—things ‘that you put in specifically to make things fit’, another new coining this decision resulted in.

<sup>5</sup> A variant which combines the allomorphs of the relativizer and the genitive case marker in the opposite way also exists: ခဉ်း *s-* + ခဉ်း *-ena* → ခဉ်း *sena*.

- (11) a. *Aku lāas!*  
           aka-u       lā-as  
           swallow-IMP tongue-P  
           ‘Shut up!’  
       b. *Aku lāyas!*  
           (idem)

With diphthongs—as described above—/ɪ/ coalesces with a following /j/ to /j/, but the initial vowel will not become tense, thus:

- (12) *တိပုယ* *tipuyya* [ti'pu.ja] ‘on the grass’ (grass-LOC).

Moreover, /u/ is commonly realized as [w] when followed by a vowel, for example in *ဘူကာယ* *huākaya* ['wa:kaja] ‘frog’ or *ရူး* *rua-* [rwa] ‘have to, must’. [w] may also be an allophone of /uj/, as in *အသုယ* *adauyi* [a'dawi] ‘then’, *အသုယ* *edaui* [e'dawi] ‘now’, or *အနွယ်* *nekuyi* ['nekwi] ‘eyebrows’. The negative suffix *-oy* is also commonly contracted to [w] before a diphthong:

- (13) *မိမ့်ဝှံ* *mingoyay* → *minguay* [miŋ'wai] ‘I cannot’ (can-NEG=ISG.TOP).

## 1.2 Phonotactics

For the purpose of this statistical analysis, most of the available translations into Ayeri from late 2008 to July 2016 have been used as a text corpus;<sup>6</sup> example sentences from various blog articles have also been added, as well as dictionary entries for all nouns, adjectives, adverbs, pronouns, adpositions, conjunctions, and numerals if they were not prefixes or suffixes.<sup>7</sup> Borrowings have been deleted if they could not reasonably be words in Ayeri. Altogether, the corpus comprises 5,500 words, which

<sup>6</sup> These texts are: A Medieval Neighborhood Dispute (2015), A Message from the Emperor (2012), Article 1 of the Universal Declaration of Human Rights (2011), The Beginning of Tolstoy’s *Anna Karenina* (2014), Conlang Christmas Card Exchange 2008/09 (2009), Conlang Holiday Card Exchange 2010/11 (2011), Conlang Relay 15 (2008), Conlang Relay 17 (2010), Conlang Relay 18 (2011), The First Two Chapters from Saint-Exupéry’s *Le Petit Prince* (2013), The Four Candles (2010), Honey Everlasting (2014), LCC4 Relay (2011), The Lord’s Prayer (2015), The North Wind and the Sun (2016), The Origin of the Wind (2009), Ozymandias (2011), Please Call Stella ... (2008), Psalm 23 (2013), The Scientific Method (2014), The Sheep and the Horses (2012), Sugar Fairies (2011), The Upside-Down Ice Skater (2009). The texts can be accessed from Becker (2016a: Examples).

<sup>7</sup> This section updates and extends a previous analysis of the phonological makeup of dictionary entries (Becker 2010). The previous study had its focus on gathering frequency statistics for word generation, however, we want to know about words generally here.

is a very small figure for such a study, but there are only so many texts available unfortunately. Words may occur more than once.

Among the dictionary entries, verbs have notably been ignored, since verb stems alone do not constitute independent words—they are always inflected in some way, so that they may end in consonants or consonant clusters that independent words cannot end in. This also has repercussions on syllabification and stress, which depend on the inflection of the verb stem:

Table 1.3: Syllabification of inflected verbs

Suffix	<i>ca-</i> ‘love’	<i>gum-</i> ‘work’	<i>babr-</i> ‘mumble’
-ay (1SG)	<i>cáy</i>	<i>gu.máy</i>	<i>ba.bráy</i>
-va (2SG)	<i>cá.va</i>	<i>gúm.va</i>	<i>ba.brá.va</i>
-yam (PTCP)	<i>cá.yam</i>	<i>gúm.yam</i>	<i>bá.bryam</i>

For the purpose of gathering statistics on phonemes, the words from translated texts were converted to IPA first. Fortunately, this is rather easy as Ayerí’s romanization is very straightforward. Syllable breaks have also been inserted semi-automatically.

### 1.2.1 Number of syllables per word

First, let us see how many syllables words commonly have (see Table 1.4). The higher the syllable count, the more likely it is for them to be compounds or inflected words.

Two-syllable words make up the bulk of the sample, which is not surprising since 1,072 entries (55.43%) in the dictionary subsample are disyllabic: most of Ayerí’s roots are disyllabic. Unsurprisingly, most monosyllabic words are function words like the ones cited below. In the following, I will quote a few examples for each number of syllables per word:

- (14) a. ႁႃႏ *ang* (AT),  
           ႃႏ *nay* ‘and’,  
           ႃႏ *rua* ‘must’;  
       b. ႃႏ *datau* ‘normal’,  
           ႃႏ *nasay* ‘near to’;  
       c. ႃႏႃႏ *avanyāng* ‘he sinks’ (sink=3SG.M.A),  
           ႃႏႃႏ *tovaley* ‘a cloak’ (cloak-P.INAN);  
       d. ႃႏႃႏ *binyanveno* (corner.beautiful, a place name),  
           ႃႏႃႏ *mitanena* ‘of the palace’ (palace-GEN);

Table 1.4: Frequency of words with different numbers of syllables (n = 5500)

Segment	Count	Percentage
2 syllables	2277	41.40%
3 syllables	1393	25.33%
1 syllable	1201	21.84%
4 syllables	547	9.95%
5 syllables	74	1.35%
6 syllables	8	0.15%

- e. ဟယ်ပဗာန် *haruyamanas* ‘a beating’ (beat-PTCP-NMLZ-P),  
 နီနွံကိဉ်နီဟန် *sungkorankibas* ‘geography’ (science.map);
- f. ကယ်တမာယဉ် *kaytomayanena* ‘of righteousness’ (righteous-NMLZ-GEN),  
 နီမာယာဉ် *nasimayajang-ben* ‘all followers’ (follow-AGTZ-PL-A=all).

Table 1.5 shows the frequencies of syllable types by position in a word. It is important to note here that phonemes which consist of more than one segment—affricates, diphthongs, and long vowels—have been counted as only one of C (consonant) or V (vowel), respectively. The following subsections will elaborate on which sounds the Cs and Vs correspond to. Moreover, it is important to note that medial syllables have not been further distinguished by position in the word for the sake of this analysis, so anything between the second and the fifth medial syllable is treated the same. It would furthermore be possible to calculate the frequencies of one syllable type following the other, however, no such calculations have been carried out here.

In all positions, CV is the most common syllable type, followed by CVC. With a very big margin, V is the next most common syllable type, which is also most common in initial syllables and least common in monosyllabic words. The cases with only a few attestations are the following:

- (15) a. Initial CVCC:  
 လိန်တန် *linktang* /lin.k.'tan/ ‘they try’ (try=3PL.M.A),<sup>8</sup>  
 လိန်တန် *silvnang* /silv.'nan/ ‘we see’ (see=IPL.A);
- b. Final CCCV:  
 မိဂရိယ *migryo* /mi.grjo/ ‘flourishes’ (flourish-3SG.N),  
 လိနရိယ *subryo* /su.brjo/ ‘ceases’ (cease-3SG.N);
- c. Single V:  
 အို *ay* /a/ ‘I’ (1SG.TOP).

Table 1.5: Frequency of syllable types per word (n = 5500)

Type	Initial		Medial		Final		Single		Total	
CV	2896	67.36%	1974	72.02%	2109	49.06%	578	48.13%	7557	60.26%
CCV	55	1.28%	24	0.88%	46	1.07%	32	2.66%	157	1.25%
CCCV	—	—	—	—	2	0.05%	—	—	2	0.02%
CVC	761	17.70%	610	22.25%	1902	44.24%	298	24.81%	3571	28.48%
CCVC	29	0.67%	10	0.36%	85	1.98%	9	0.75%	133	1.06%
CVCC	2	0.05%	—	—	—	—	—	—	2	0.02%
V	488	11.35%	95	3.47%	67	1.56%	2	0.17%	652	5.20%
VC	68	1.58%	28	1.02%	88	2.05%	282	23.48%	466	3.72%
Total	4299	100.00%	2741	100.00%	4299	100.00%	1201	100.00%	12540	100.00%

The medial and final VC cases may seem like an oddity, but they are mostly due to the previous syllable ending in /ŋ/, with that syllable also containing a lax vowel, which means that this syllable must be closed. An alternative explanation would be to assume that /ŋ/ is ambisyllabic, or actually /n.g ~ ŋ.g/, but realized as [ŋ]. The high number of single-syllable VC is due to ልግ *ang* ‘AT’, which alone appears 255 times in the sample (4.63% of all words, 21.23% of monosyllabic words, 90.43% of monosyllabic VC words).

### 1.2.2 Phonemic makeup of initial syllables

The statistics in the following sections have been gathered from the IPA conversions of translated texts and dictionary entries mentioned above. The transcribed words have been split into syllables and then the collected contents of each position group were written into separate plain text files, one each for:

- all initial syllables of polysyllabic words,
- all medial syllables of polysyllabic words,
- all final syllables of polysyllabic words, and
- all monosyllabic words.

Monosyllabic words are both initial and final syllables at the same time; they have been counted separately for the purpose of this analysis. Onsets, nuclei and codas have been matched by regular expressions; the command line tools *grep*, *sort*, and *uniq* were used to aggregate all occurring variants for each syllable segment as well as their absolute frequencies.<sup>9</sup>

(16) C = (? : t f | d ʒ | [ptkbgdmnŋvshrljw])  
 V = (? : [ae] : ? ɪ | əʊ | [ieaou] : ? | [ɪɛʊə])

As we have seen above (Table 1.5), CCV syllables only make up 1.28% of initial syllables, insofar it is no surprise that consonant clusters all appear at the bottom of Table 1.6. There also seem to be combination patterns in that initial clusters exist for all plosives plus /r/, and almost all bilabials plus /j/, with the exception of /bj/, however, /nj/ is added to the group instead. Combinations with /w/ only occur for /b/, /r/, and /s/, which do not share an obvious connection. Syllables without a consonant filling the onset position are marked with ‘Ø’; these numbers correspond to the VC and VCC rows in Table 1.5.

<sup>8</sup> The verb stem is found in the dictionary as ልግ *linka-*, with a final *-a*, and thus is possibly an entry changed at a later point, or the example from the text (Sugar Fairies) chosen here contains an error.

<sup>9</sup> However, *sort* was unable to handle all IPA characters, so *sed 'y/ɛɪʊə:fʒŋ/EIʊ@:SZN/'* had to be used to compensate by transcribing everything into X-SAMPA.



Table 1.6: Frequency of onsets in initial syllables (n = 4299)

Phoneme	Frequency	Percentage
Ø	556	12.93 %
s	488	11.35 %
t	432	10.05 %
m	418	9.72 %
k	380	8.84 %
n	375	8.72 %
p	334	7.77 %
b	231	5.37 %
d	172	4.00 %
v	164	3.81 %
l	159	3.70 %
r	134	3.12 %
j	126	2.93 %
g	111	2.58 %
h	99	2.30 %
tʃ	30	0.70 %
pr	27	0.63 %
nj	27	0.63 %
kr	8	0.19 %
br	8	0.19 %
tr	6	0.14 %
dʒ	4	0.09 %
gr	3	0.07 %
w	2	0.05 %
sw	1	0.02 %
rw	1	0.02 %
pj	1	0.02 %
mj	1	0.02 %
bw	1	0.02 %

Table 1.7: Frequency of nuclei in initial syllables (n = 4299)

Phoneme	Frequency	Percentage
a	1847	42.96%
i	1011	23.52%
<i>i</i>	802	18.66%
<i>ɪ</i>	209	4.86%
e	705	16.40%
<i>e</i>	523	12.17%
<i>ɛ</i>	164	3.81%
<i>ə</i>	18	0.42%
u	260	6.05%
<i>u</i>	228	5.30%
<i>ʊ</i>	32	0.74%
o	227	5.28%
<i>o</i>	188	4.37%
<i>ɔ</i>	39	0.91%
a:	109	2.54%
aɪ	88	2.05%
eɪ	40	0.93%
e:	4	0.09%
ɔɪ	3	0.07%
ʊɪ	1	0.02%
o:	1	0.02%
i:	1	0.02%
e:ɪ	1	0.02%
aʊ	1	0.02%

Perhaps most striking about the nuclei of initial syllables presented in Table 1.7 is that plain vowels occur most frequently. As mentioned above, lax vowels are counted here as allophones of tense ones as their distribution is complementary and are listed here for the sake of completeness. This is the reason why the plain vowels are presented as grouped with their allophones in this table as well as in subsequent ones. Long vowels and diphthongs find themselves below the 5% threshold, and the words with single occurrences are:

- (17) a. ကွယ်နှံ *kuysān* ‘comparison’,  
 b. ဝဉ် *nōn* ‘will, intention’,  
 c. ဝဉ်န *nīsa* ‘wanted’,<sup>10</sup>  
 d. နှိယှာယ *sēyraya* ‘will overcome’ (FUT-overcome-3SG.M),  
 e. နီဆံ *sautan* ‘cork’.

As the diphthong [e:i] only occurs due to allophony, it should not be counted as a phoneme for the purposes of this analysis. On the other hand, the same could be said for a lot of cases of [a:] included here—this caveat applies to all nouns derived from verbs ending in *-a* with the very common nominalizing suffix နှံ *-an*, as exemplified in (8) above. Similarly, the 18 instances of /ə/ reported here are mostly from tense prefixes also mentioned above, for instance, မသိသိရ *məkoronay* ‘I knew’ (PST-know=1SG.TOP).

Initial-syllable codas (Table 1.8) are far less diverse than consonant onsets: there are only 10 attested segments in comparison to 28 for onsets (not counting empty codas of C(C)V syllables, which constitute the majority by a large margin), and the only two cluster attested are /ŋk/ in the word လိန်တန် *linktang* ‘they try’ (try=3PL.M.A), and /lv/ in the word လိပ်နာ *silvnang* ‘I see’ (see=1PL.A). There only being two incidences of a CC cluster is very probably an effect of the small sample size. Furthermore, the only unvoiced single coda consonants attested are /s/, /h/, /t/, /tʃ/ and /k/, the latter two only once, /h/ twice:

- (18) a. မေ့သိသိရ *mehvāng* ‘you are supposed to’ (be.supposed.to=2SG.A),<sup>11</sup>  
 လိပ်နာ *rohtang* ‘they bite’ (bite=3SG.M.A);  
 b. မိတ် *mutva* ‘you rub’ (rub=2SG.TOP),  
 ကလေး *patlay* ‘cousin’;  
 c. နီနီ *sik-sik* ‘tits’;  
 d. လိပ်နာ *vacvāng* ‘you like’ (like=2SG.A).

<sup>10</sup> ဝဉ်န *nīsa* and ဝဉ် *nōn* are both related to ဝဉ် *no-* ‘want, plan’.

<sup>11</sup> The dictionary entry for the verb is မှ *mya-*, so this may be an instance of my changing a word in the dictionary with the old one staying in the text (The Four Candles).

Table 1.8: Frequency of codas in initial syllables (n = 4299)

Phoneme	Frequency	Percentage
Ø	3441	80.04 %
n	298	6.93 %
ŋ	243	5.65 %
r	129	3.00 %
l	88	2.05 %
m	74	1.72 %
s	20	0.47 %
t	2	0.05 %
h	2	0.05 %
tʃ	1	0.02 %
ŋk	1	0.02 %
lv	1	0.02 %
k	1	0.02 %

### 1.2.3 Phonemic makeup of medial syllables

The onsets of medial syllables (Table 1.9) show properties very similar to those of initial syllables. The order of most common consonants may differ here—for example, the most common onset is /r/, not Ø or /s/—, but there are no restrictions on consonants to appear in this position, with the exception of /ŋ/ for reasons stated above (see section 1.2.1). Regarding initial clusters, there are further attestations for plosive plus /r/ (except for /kr/). As for clusters with /j/, the only one with a bilabial is /bj/, but the set is extended to /sj/ and /kj/. For clusters with /w/, only /sw/ and /kw/ occur here, while attestations for /bw/ and /rw/ as in initial-syllable onsets are lacking. This does not mean that those combinations are not principally possible in this position, however.

As with onset consonants, vowel nuclei of medial syllables (Table 1.10) do not show significant differences compared to those of initial syllables either. /a/ is more common here, and /o/ and /u/ switch places. Instead of /e:ɪ/, there is an attestation of /u:/ (see footnote 4), for which there is more reason to be counted as a phoneme than for /e:ɪ/. The sequences /i:/ and /uɪ/ also only occur once and twice, respectively, namely in the following words:

- (19) a.  *pasisa* ‘interesting’;  
 b.  *puluyley* ‘a mirror’ (mirror-P.INAN),  
 *tipuyya* ‘on the grass’ (grass-LOC).

Table 1.9: Frequency of onsets in medial syllables (n = 2741)

Phoneme	Frequency	Percentage
Ø	123	4.49%
r	343	12.51%
n	260	9.49%
j	233	8.50%
t	222	8.10%
d	213	7.77%
k	189	6.90%
s	170	6.20%
m	169	6.17%
l	149	5.44%
v	148	5.40%
h	147	5.36%
p	119	4.34%
g	92	3.36%
b	89	3.25%
tʃ	20	0.73%
dʒ	15	0.55%
tr	11	0.40%
dr	8	0.29%
pr	7	0.26%
w	6	0.22%
sj	2	0.07%
br	2	0.07%
sw	1	0.04%
kw	1	0.04%
kj	1	0.04%
bj	1	0.04%

Table 1.10: Frequency of nuclei in medial syllables (n = 2741)

Phoneme	Frequency	Percentage
a	1480	53.99%
i	480	17.51%
<i>i</i>	387	14.12%
<i>ɪ</i>	93	3.39%
e	254	9.26%
<i>e</i>	206	7.52%
<i>ɛ</i>	48	1.75%
o	194	7.08%
<i>o</i>	119	4.34%
<i>ɔ</i>	75	2.74%
u	120	4.38%
<i>u</i>	101	3.68%
<i>ʊ</i>	19	0.69%
a:	110	4.01%
aɪ	51	1.86%
ɔɪ	33	1.20%
eɪ	5	0.18%
e:	5	0.18%
aʊ	5	0.18%
ʊɪ	2	0.07%
u:	1	0.04%
i:	1	0.04%

The word in (19a),  $\text{పాసీ}$  *pasīsa* ‘interesting’, is rather transparently constitutes a causative derivation of the verb  $\text{పాసి}$  *pasy-* ‘wonder, be curious, be interested’, essentially meaning ‘making one wonder/curious’—the causative suffix  $\text{-సీ}$  *-isa* can as well be used to derive adjectives with a causative or resultative meaning. Nonetheless it should count as a lexeme in its own right, since it possesses idiomatic meaning.

With medial-syllable codas (Table 1.11) again, sonorants and /s/ make up the largest number of consonants in this position; /t/ and /g/ only occur once each in

- (20) a.  $\text{పాంగిటన్}$  *pangitlan* ‘money change’,<sup>12</sup> and  
 b.  $\text{తెలుగంట}$  *telugtong* ‘they survive’ (survive=3PL.N).

Table 1.11: Frequency of codas in medial syllables (n = 2741)

Phoneme	Frequency	Percentage
Ø	2093	76.36 %
n	313	11.42 %
ŋ	193	7.04 %
r	48	1.75 %
m	39	1.42 %
s	32	1.17 %
l	21	0.77 %
t	1	0.04 %
g	1	0.04 %

As documented in Table 1.5 above, Ayeri very strongly favors CV syllables in medial positions, hence the high count of zero segments here.

#### 1.2.4 Phonemic makeup of final syllables

The onsets of final syllables of polysyllabic words (Table 1.12) show the greatest amount of variety, which is due to Ayeri mostly using suffixes for grammatical purposes. Hence it is no surprise that combinations with /j/ and, indeed, /j/ itself as an onset, are especially common, since /j/ is also what a number of very common suffixes start with, for example the plural marker ʔ -*ye*, the locative marker ʔ -*ya*, the dative and participle marker ʔ -*yam*, as well as third-person animate pronoun agreement suffixes, and the various first-person and third-person animate pronominal clitics. Table 1.3 above shows exemplarily how verbs resyllabify when suffixes are attached. Even though single-segment onsets are strongly preferred, Cr, Cw, and especially C(C)j seem to be generally permissible.<sup>13</sup>

<sup>12</sup> The word for ‘money’ is ʔ pangis, so (20a) is probably a compound, albeit not a fully transparent one. The word for ‘change’ is ʔ tila-, and there seems to be a nominalizing ʔ -an. Ayeri allows noun–verb compounds to have a nominalized verb in the second position in spite of it being the head—noun–noun compounds mostly come in head-initial order. Possibly, what happened at the morpheme borders is that ʔ tilān underwent metathesis to \*ʔ itlān to match the rhyme of ʔ pangis. \*ʔ pangisitlān then underwent irregular haplology (and shortening of the nominalizing suffix) to ʔ pangitlan.

<sup>13</sup> The sequence /sj/ poses difficulty here as there are examples for /Vs.jV/ as well as for /V.sjV/, and I cannot tell for sure if there is a strict rule in operation. It seems that /V.sjV/ is more likely to occur when the second syllable is stressed, whereas /Vs.jV/ is more likely to occur when the first syllable is stressed. Ayeri’s own Tahano Hikamu orthography would not show the difference either, since /sja/ is spelled ʔ either way, and there is no heeding morpheme

Table 1.12: Frequency of onsets in final syllables (n = 4299)

Phoneme	Frequency	Percentage	Phoneme	Frequency	Percentage
Ø	155	3.61%	pr	7	0.16%
j	1101	25.61%	kj	6	0.14%
n	528	12.28%	hj	5	0.12%
r	398	9.26%	bj	5	0.12%
t	268	6.23%	tw	4	0.09%
s	244	5.68%	sw	4	0.09%
l	238	5.54%	sj	4	0.09%
k	199	4.63%	kw	3	0.07%
d	184	4.28%	kr	3	0.07%
m	154	3.58%	br	3	0.07%
v	144	3.35%	vr	2	0.05%
h	128	2.98%	rw	2	0.05%
p	115	2.68%	nw	2	0.05%
g	103	2.40%	tʃj	1	0.02%
dʒ	73	1.70%	rj	1	0.02%
b	73	1.70%	nj	1	0.02%
tʃ	52	1.21%	mw	1	0.02%
vj	26	0.60%	grj	1	0.02%
pj	22	0.51%	dv	1	0.02%
dʒj	17	0.40%	dr	1	0.02%
tr	10	0.23%	brj	1	0.02%
w	9	0.21%			

Nuclei of final syllables (Table 1.13) do not bear striking differences to nuclei in other positions. /a:/ comes out second here due to the common nominalizer *-an*, which lengthens the vowel of verb stems ending in /a/, as demonstrated in (8). /aɪ/ is also fairly common here as it is the topic-marked first-person pronoun/pronominal clitic; for the same reason, /a:ɪ/ occurs a number of times—the vowel-lengthening rule applies here as well, so its status as a phoneme is marginal. All instances of /e:/ in the sample are from the word *arēn* ‘anyway, however’; all evidence for /i:/ is from *sirī* ‘due to which’ (see section 1.1.2). The only evidence for /u:/ in the sample is from *babū* ‘barbarian (adj.)’.

breaks in placing the diacritic. /CsjV/ will be /C.sjV/ in any case, since Ayeri avoids final consonant clusters if possible, see Table 1.5.



Table 1.13: Frequency of nuclei in final syllables (n = 4299)

Phoneme	Frequency	Percentage
a	2408	56.01%
a:	316	7.35%
o	411	9.56%
o	298	6.93%
ɔ	113	2.63%
i	289	6.42%
ɪ	147	3.42%
ɨ	142	3.30%
aɪ	254	5.91%
u	207	4.82%
u	155	3.61%
ʊ	52	1.21%
e	209	4.85%
ɛ	127	2.95%
ə	81	1.88%
ɐ	1	0.02%
ɐɪ	103	2.40%
ɔɪ	42	0.98%
a:ɪ	23	0.54%
ʊɪ	14	0.33%
aʊ	14	0.33%
e:	5	0.12%
i:	3	0.07%
u:	1	0.02%

Table 1.14: Frequency of codas in final syllables (n = 4299)

Phoneme	Frequency	Percentage
Ø	2224	51.73 %
n	899	20.91 %
ŋ	651	15.14 %
s	244	5.68 %
m	225	5.23 %
l	34	0.79 %
r	21	0.49 %
k	1	0.02 %

The list of coda consonants in final syllables (Table 1.14) is very slightly more restrictive than even that of coda consonants in medial syllables (see Table 1.11), since the only non-sonorant attested is /k/, which only occurs in *ṣik-ṣik* ‘tits’ again, which—besides being a vulgar term, thus maybe slightly more dispositioned to allow for deviating phonotactics—looks quite like onomatopoeia, possibly for the sound of sucking.<sup>14</sup>

### 1.2.5 Phonemic makeup of single syllables

Onsets of single syllables (Table 1.15) appear to be the least varied category. Still, none of the basic set of consonant morphemes (see Table 1.1) is missing—the frequency order is just completely different from the other onsets surveyed, not merely a mixture of initial and final syllables. Consonant clusters with /j/, /w/ and /r/ exist here as well. Combinations with /j/ are only present for /m/ and /n/, while /r/ again combines with plosives; /w/ combines with /n/ and /r/ at least, which we have already seen in final-syllable onsets (see Table 1.12). Whereas /mj/ has only occurred once in initial-syllable onsets so far (see Table 1.6), it occurs a few more times here, all in the word *ṣmya* ‘be supposed to’, which is very commonly used as an unconjugatable modal particle.

A consonant onset that can only be found in monosyllables is /ŋ/,<sup>15</sup> in *ṣṅas* ‘almost’, a quantifier suffix that has managed to sneak in due to being marked as an adverb in the dictionary, since it can modify a verb:

<sup>14</sup> Kroonen (2013: 489–490) identifies PGmc *\*sūgan-*, *\*sūkan-* ‘to suck’ as an iterative of PGmc *\*sukkōn-*, *\*sugōn-* ‘to suck’ and reconstructs PIE *\*souk-neh<sub>2</sub>-*. However, he does not say anything about the Germanic word being onomatopoeic in origin.

<sup>15</sup> At least according to the analysis chosen here, see section 1.2.1 for an explanation.

Table 1.15: Frequency of onsets in single syllables (n = 1201)

Phoneme	Frequency	Percentage
Ø	284	23.65%
n	231	19.23%
s	147	12.24%
j	144	11.99%
k	51	4.25%
v	48	4.00%
m	46	3.83%
l	44	3.66%
t	41	3.41%
d	33	2.75%
r	26	2.16%
h	23	1.92%
mj	16	1.33%
p	13	1.08%
tʃ	9	0.75%
g	9	0.75%
nj	8	0.67%
rw	7	0.58%
b	7	0.58%
pr	5	0.42%
dʒ	3	0.25%
tr	2	0.17%
nw	1	0.08%
ŋ	1	0.08%
kr	1	0.08%
br	1	0.08%

- (21) *Apayeng-ngas.*  
 apa=yeng=ngas  
 laugh=3SG.F.A=almost  
 ‘She almost laughed.’

Here, ၵာန *-ngas* modifies the verb complex like any other adverb:

- (22) *Apayeng baho.*  
 apa=yeng baho  
 laugh=3SG.F.A loudly  
 ‘She laughs loudly.’

However, whereas ခပ် *baho* ‘loud’ is treated as a separate unit in terms of intonation, ၵာန *-ngas* is unstressed and binds to what it follows:

- (23) a. ၵာန-ၵာန ၵာန *Apayeng-ngas*. [apa'jɛŋas];  
 b. ၵာန-ၵာန ခပ် *Apayeng baho*. [apa'jɛŋ 'baxo].

As with onset consonants of monosyllabic words, nuclei of this syllable type are the least diverse group again (Table 1.16). One segment that is notably absent is /au/, and the marginally phonemic /e:/ is not present either. By having /a/, /ai/, /a:/ at the top, monosyllabic words behave similar to final syllables of polysyllabic words (see Table 1.13), however, the order of the most common vowels bears more similarities to that of initial and medial syllables (see Tables 1.7 and 1.10). The very uncommon /o:/ features twice in this group, namely in two instances of the word ခပ် *nōn* ‘will, intention’.<sup>16</sup>

Like the other syllable segments of monosyllabic words, coda consonants (Table 1.17) as well show the lowest degree of variety among all the coda consonants of the various syllable classes discussed so far. The order is basically the same as that of final-syllable codas (see Table 1.14), though /ŋ/ supersedes /n/ and there is some attestation of final /h/. As noted above, the prevalence of /ŋ/ is due to the agent-topic marker ၵာ *ang* (see section 1.2.1). /h/ only occurs in the interjections ခပ် *ah!* and ခပ် *āh!*, so its status as an actual phoneme in this position is marginal at best.

<sup>16</sup> Ayeri used to have ခပ် *-on* as a nominalizer beside ခပ် *-an*, however, it was not very productive and has long fallen out of use. ခပ် *nōn* is thus, in fact, originally a nominalization of ခပ် *no-* ‘want, plan’.

Table 1.16: Frequency of nuclei in single syllables (n = 1201)

Phoneme	Frequency	Percentage
a	568	47.29 %
aɪ	171	14.24 %
a:	140	11.66 %
i	113	9.41 %
<i>i</i>	65	5.41 %
<i>ɪ</i>	48	4.00 %
e	104	8.66 %
<i>ɛ</i>	65	5.41 %
<i>e</i>	34	2.83 %
<i>ə</i>	5	0.42 %
o	45	3.75 %
<i>ɔ</i>	30	2.50 %
<i>o</i>	15	1.25 %
u	20	1.67 %
a:I	14	1.17 %
ɔɪ	10	0.83 %
i:	6	0.50 %
eɪ	5	0.42 %
ʊɪ	3	0.25 %
o:	2	0.17 %

Table 1.17: Frequency of codas in single syllables (n = 1201)

Phoneme	Frequency	Percentage
Ø	612	50.96 %
ŋ	377	31.39 %
n	105	8.74 %
s	58	4.83 %
m	36	3.00 %
l	6	0.50 %
h	4	0.33 %
r	3	0.25 %

Table 1.18: Frequency of cross-syllable consonant clusters (n = 1270)

Interval [%]	Consonant cluster
0.00 ... 0.09	g.t, h.t, h.v, k.s, l.n, lv.n, m.bj, m.d, m.dʒ, m.l, m.n, m.pr, m.r, n.dv, n.g, <sup>17</sup> n.h, n.w, ŋ.dʒj, ŋ.kw, ŋ.m, ŋ.n, ŋ.rj, ŋ.t, ŋk.t, r.b, r.dʒ, r.g, r.l, r.m, r.sj, r.tʃ, r.v, s.dʒ, s.h, s.l, s.n, s.p, s.v, t.v, tʃ.v (0.08%).
0.10 ... 0.24	l.bj, m.br, m.t, n.s, ŋ.b, ŋ.h, ŋ.p, ŋ.w, r.dʒj, r.pj, s.dʒj, s.m, t.l (0.16%); l.dʒ, l.p, m.k, n.sj, ŋ.dʒ, ŋ.g, ŋ.s, r.pr (0.24%).
0.25 ... 0.49	m.v, r.s, s.r (0.31%); n.r, s.t (0.39%); m.pj, n.dʒj, r.d (0.47%).
0.50 ... 0.74	ŋ.kj, ŋ.v, r.k, r.n (0.55%); l.b, l.t, ŋ.r (0.71%).
0.75 ... 1.00	r.p, r.t (0.87%); l.vj (0.94%).
1.0 ... 2.4	m.j (1.18%); ŋ.l (1.34%); n.tʃ (1.50%); n.dʒ (2.13%); n.v (2.28%); l.j (2.36%).
2.5 ... 4.9	m.p (2.52%); s.j (2.60%); n.l (2.91%); l.v (3.15%); m.b (3.23%); ŋ.k (3.78%).
5 ... 9	n.t (5.28%); n.d (6.85%); ŋ.j (7.32%); r.j (8.98%).
10+	n.j (25.35%).

### 1.2.6 Cross-syllable consonant clusters

Since a table detailing every combination with its absolute and relative frequency would be too large here, Table 1.18 gives the attested combinations ordered by brackets. As can be expected, bilabials cluster mostly with bilabials (83/112 purely bilabial CC combinations = 74.11%), alveolars with alveolars (317/948 = 33.44%), and velars with velars (59/207 = 28.51%). However, at least for alveolars and velars, the score is even higher with /j/: 52.64% and 44.93%, respectively. /j/ is also the most common second consonant overall, at 47.8% of all consonant clusters; /n.j/ is the most common cluster at a total of 25.35%. Alveolars provide the highest variety of both first and second consonants, with 6 different phonemes making up 74.65% of C<sub>1</sub>, and 8 different phonemes making up 28.74% of C<sub>2</sub>.

Labiodentals and glottals occur least frequently, on the other hand: There is only one cluster with /v/ as a first consonant, namely, /lv.n/ (0.08%). For /h/, there are two, which are /h.v/ and /h.t/ (0.16%). Altogether, however, there are 97 combinations in /v/ (7.64%)—most commonly /l.v/ (3.15%) and /n.v/ (2.28%)—while there are only 4 in /h/ (0.31%): /n.h/, /s.h/, and twice /ŋ.h/.

<sup>17</sup> This is from གཙུག་གཙུག་ *gan-gan* ‘grandchild’ being given as /gan.gan/ in the dictionary, which is the intended pronunciation, however. Assimilation to [ŋ] does not operate in this case.

At 924 first consonants (72.76%), the nasals /m/, /n/, and /ŋ/ make up the largest group going by manner of articulation, followed by the tap /r/, which appears 175 times (13.78%) as the first consonant. For second consonants, approximants constitute the largest group at 669 combinations (52.68%), followed by 387 pairs with plosives second (30.47%).

## 1.3 Notes on prosody

### 1.3.1 Stress

Ayeri uses dynamic accent, that is, stress is based on differences in the loudness of syllables, among others.<sup>18</sup> Which syllable is stressed depends on a mixture of which position in a word a syllable occupies and the phonemic shape of it. In fact, English, which also has phonemic stress in pairs such as *record* /'rɛkərd/ (noun) and /rɪ'kɔrd/ (verb) that Ayeri lacks, does a similar thing (examples adapted from Halle 1998: 552):

- (24) *admire* /æd'maɪər/ — *admirable* /'ædmərəbl/ [English]  
*carnivore* /'karnɪvɔr/ — *carnivorous* /kar'nɪvərəs/  
*ignore* /ɪg'nɔr/ — *ignorant* /'ɪgnərənt/

Stress does not stay at fixed intervals in these words and they even change their sound structure a little, but there are a number of variables which can nonetheless be formally described and applied here (564–565).<sup>19</sup> To demonstrate how word stress moves around in Ayeri, the complete declension paradigm for *ꠄ niva* ‘eye’ is presented in Table 1.19.

It may appear that in the table above, stress is always on the penultimate syllable, which is indeed the case for most forms quoted there, but compare the superficially unmarked form *ꠄ niva*, which is disyllabic with stress on the first (= penultimate) syllable, to the agent and patient singular forms, *ꠄꠤꠤ nivaṅ* and *ꠄꠤꠤ nivaś*, respectively. These are also disyllabic, however, they are stressed on the second (= ultimate) syllable. Similarly, compare the agent and patient plural forms to each other: the agent plural form *ꠄꠤꠤ nivaṅ* is trisyllabic and has its main

<sup>18</sup> For a discussion of terms, see Kager (2007), for instance.

<sup>19</sup> Halle (1998) takes a generativist approach rather than a more modern Optimality-Theory based one like Kager (2007) does, who only deals with fixed-stress systems in this introductory article, though Halle’s article is still informative. Simplifying a lot, English essentially tries to construct trochaic feet from the right edge of the word. If the last syllable’s vowel is not light, it is skipped and stress moves to the antepenultimate syllable; this process is recursive for words with multiple feet, although some suffixes introduce irregularities in rule application.

<sup>20</sup> Final-syllable stress is possible as well, also in the plural.

Table 1.19: Declension paradigm for Ayeri ຂໍ້ *niva* ‘eye’

	Singular		Plural	
TOP	<i>ní.va</i>	‘the eye’	<i>ni.vá.ye</i>	‘the eyes’
A	<i>ni.vǎ́ng</i>	‘eye’	<i>ni.va.jǎ́ng</i>	‘eyes’
P	<i>ni.vǎ́s</i>	‘eye’	<i>ni.vá.jas</i>	‘eyes’
DAT	<i>ni.vá.yam</i> <sup>20</sup>	‘to the eye’	<i>ni.vá.jyam</i>	‘to the eyes’
GEN	<i>ni.vá.na</i>	‘of the eye’	<i>ni.va.yé.na</i>	‘of the eyes’
LOC	<i>ni.vá.ya</i>	‘at the eye’	<i>ni.vá.jya</i>	‘at the eyes’
CAUS	<i>ni.va.í.sa</i>	‘due to the eye’	<i>ni.va.jí.sa</i>	‘due to the eyes’
INS	<i>ni.vá.ri</i>	‘with the eye’	<i>ni.va.yé.ri</i>	‘with the eyes’

stress on the third (= ultimate) syllable, while the equally trisyllabic patient plural form ຂໍ້ຖັງ *nivajas* is stressed on the second (= penultimate) syllable again.

It should have become clear that even though the basic form ຂໍ້ *niva* has first-syllable stress, *ni* will not necessarily carry stress across the whole paradigm. It should also have become clear that the basic algorithm to determine stressed syllables in Ayeri has something to do with counting syllables from the right edge of a word, although some complications need to be factored in. The following sections will try to describe these formally.

#### Analysis of stress patterns in disyllabic words

The basic foot in Ayeri is a trochee, and for the most part it does not matter whether the syllable is open or closed, or whether there are complex onsets or codas, or no onsets or codas at all:<sup>21</sup>

- (25) a.     $\acute{x}$          $x$   
           *ha* - *ri* ‘pithy, striking’
- b.     $\acute{x}$          $x$   
           *sa*    - *yan* ‘hole, cave’  
           *sem* - *ba* ‘comb’
- c.     $\acute{x}$          $x$   
           *bri* - *ha* ‘grace’  
           *ba* - *brya* ‘(he) mumbles’  
           *a*    - *gu* ‘chicken’

<sup>21</sup> In the following, a syllable will be marked by  $\langle x \rangle$  and receives an acute accent ( $\acute{\phantom{x}}$ ) when carrying primary stress, a grave accent ( $\grave{\phantom{x}}$ ) when carrying secondary stress, and no accent when unstressed. Feet are marked by horizontal lines ( $\langle \rangle$ ).



It can be deduced from words with more than two syllables that stress assignment is trochaic. Stress assignment furthermore runs from right to left, so that in a word with more than two syllables, the last two syllables form a full foot:

- (26) a.    x    |     $\acute{x}$     x  
           *ba* - *ha* - *lan* 'target, goal'  
           *jar* - *ma* - *ya* 'pilgrim'
- b.     $\acute{x}$     x    |     $\acute{x}$     x  
           *ho* - *ra* - *ma* - *ya* 'sinner'  
           *ya* - *ma* - *na* - *ti* 'causer'

In the case of (26b), the stressed syllables of the first foot bear secondary stress while those of the second foot bear primary stress. Complications, then, come in the form of syllables ending in /ŋ/, containing a long vowel, or containing a diphthong, or a combination of those features. Ayeri does not have syllables that contain a diphthong and also end in /ŋ/, though, since consonant codas after a diphthong are largely avoided.<sup>22</sup> Since the presence or the absence of a certain element that is suspected to have an effect on stress assignment is a yes–no decision, we can make a matrix of binary features:

Table 1.20: Types of heavy syllables

	[+ DIPH, – ŋ]	[– DIPH, + ŋ]	[– DIPH, – ŋ]
[+ LONG]	++	++	++
[– LONG]	+	+	–

The feature matrix above (Table 1.20) shows the various kinds of syllable types that we will now see have a manipulative effect on trochaic stress assignment. These syllable types can be considered 'heavy' in that they attract stress and thus modify the regular assignment of stress to every other syllable from the right edge of a word. For the time being, we will only test their effects on disyllabic words as the most common type. As will be shown in the following example, heavy syllables in ultimate positions attract stress while quasi-regular results are produced when they are in penultimate position and the ultimate syllable is not heavy:

- (27) a.    x         $\acute{x}$   
           *ma* - *tay* 'summer, wet season'  
           *pa* - *dang* 'mind; heart, mood'  
           *ka* - *nāy* 'I marry' (marry=1SG.TOP)  
           *bra* - *syāng* 'he bathes' (bathe=3SG.M.A)  
           *na* - *rān* 'word; speech'

<sup>22</sup> It might thus be possible to alternatively analyze diphthongs in /ɪ/ as /Vj/ sequences, essentially.

- Unfortunately, there are no disyllabic examples for the feature sets [+LONG, -DIPH, +η] and [+LONG, +DIPH, -η] in the first syllable (syllables of the type /C(C(j))V:η/ or /C(C(j))Vɪη/). If there were, they would group with (27b).

(28) a. x             $\acute{x}$   
*bay* - *hāy* 'I govern' (govern=ISG.TOP)  
*say* - *lyāng* 'he sails' (sail=3SG.M.A)  
*kay* - *vān* 'container'

b. x             $\acute{x}$   
*kong* - *āyn* 'we enter' (enter=IPL.TOP)  
*keng* - *vāng* 'you notice' (notice=2SG.A)  
*lang* - *-vā* 'in the most tiresome way' (tiresome=SUPL)

We can see here that these words have primary stress invariably on the last/long syllable in spite of a heavy syllable preceding in the examples in (28b). The question then is, however, what happens if we invert this order. This is more problematic than it sounds, however, as initial [+LONG, +DIPH, -η] and [+LONG, -DIPH, +η], as well as final [-LONG, +DIPH, +η] do not occur, thus, there will only be one possible combination here—the reverse pattern of  $\text{ḥaṭṭa}$  *lang-vā* ‘in the most tiresome way’ from (28b) above, also compare with (27):

- There is only one pattern possible here, which is very little to make a point, however, other words following this syllable pattern, like  $\text{နာရေ}$  *nāreng* ‘rather’, for example, behave in the same way. A long syllable has precedence over other kinds of heavy syllables, so  $\text{နာ}$  *-nang* does not take away stress from  $\text{ငါး}$  *cā-* as one might expect from the examples in (27a). Another question is what happens if we pit

elements from the [ $\pm$  LONG] rows against another feature combination of the same row. As above, we will start with the [ $-$  LONG] row:

- (30) a.    x         $\acute{x}$   
           *bay* - *tang* ‘blood’  
       b.    x         $\acute{x}$   
           *pang* - *lay* ‘goddess’

In the case of examples for [ $+$  LONG] pattern combinations, we need to keep in mind again that initial [ $+$  LONG,  $+$  DIPH,  $-$   $\eta$ ] and [ $+$  LONG,  $-$  DIPH,  $+$   $\eta$ ] are not attested, so again, there will only be one possible combination of two syllables with a long vowel:

- (31)     $\acute{x}$          $\acute{x}$   
           *mā* - *sāy* ‘I traveled’ (PST-travel=ISG.TOP)

Combining two long syllables with each other will result in both being stressed, which is otherwise avoided in Ayeri, as we will see later. Moreover, the following patterns emerge if we combine each pattern with itself; the combinatorial restrictions mentioned above apply again, of course:

- (32) a.    x         $\acute{x}$   
           *kay* - *vay* ‘without’  
           *dang* - *reng* ‘bell’ (bell-A.INAN)  
       b.     $\acute{x}$          $\acute{x}$   
           *bā* - *mā* ‘parents, mom-and-dad’

As demonstrated in (28), the last heavy syllable will receive primary stress, except if two long syllables collide, in which case the first long syllable will receive secondary stress.

To summarize the above findings:

1. Ayeri assigns trochaic stress from the right edge of a word. A foot thus consists of two syllables, of which the first is stressed.
2. Syllables ending in / $\eta$ / or ones containing a diphthong are considered heavy. They attract stress and take it away from a preceding stressed syllable if the following syllable is not stressed already.
3. Syllables containing a long vowel are considered superheavy and override both light and heavy syllables in attracting stress, since long vowels cannot be unstressed.

Table 1.21: Stress patterns for [+ HEAVY, – LONG] in trisyllabic words

–H –H +H	<i>prantanley</i>	x   x ́	‘question’ (question–P.INAN)
–H +H –H	<i>sarayya</i>	x   ́ x	‘(he) bows’ (bow–3SG.M)
+H –H –H	<i>taykondam</i>	x   ́ x	‘break (n.)’
–H +H +H	<i>ralanghay</i>	x   x ́	‘thumbnail’
+H –H +H	<i>kaybunay</i>	́   x ́	‘by the way’
+H +H –H	<i>maykongas</i>	x   ́ x	‘harbor’ (harbor–P)
+H +H +H	<i>saylingyang</i>	́   x ́	‘I progress’ (progress=1SG.A)

4. Primary stress is assigned to the last stressable syllable, or otherwise the last heavy syllable. In the rare case of two long/superheavy syllables after another, the first syllable receives secondary stress and reduces in duration.

As we will see in the in the next section, however, another rule needs to be added to this set:

5. Secondary stress is assigned to syllables that are eligible for word stress but which are not in the final foot.

#### Analysis of stress patterns in trisyllabic words

So far, we have only considered all the possible combinations of two heavy and light syllables. Doing the same for all combinations of three and more syllables would be possible, though the list of examples were to become even longer. Since the feature pair [ $\pm$  DIPH,  $\pm$   $\eta$ ] behaves the same way throughout and both features are in complementary distribution, we need not test iterations of them separately, but can subsume them under the label [ $\pm$  HEAVY]. The parameters that need testing, then, are [ $\pm$  HEAVY] in combination with [ $\pm$  LONG]. There are 4 possible outcomes for these two features, which in the case of three syllables leads us to  $(2 \times 2)^3 = 64$  theoretically possible combinations. For this reason, I want to point out just a few cases, since the general rules sketched out above still apply.

First, let us look at [+ HEAVY, – LONG] combined with [– HEAVY, – LONG] in all positions (Table 1.21). Finding words that fit the respective permutations is not too much of a problem, especially in cases where there is only one heavy syllable. It becomes clear from Table 1.21 that the rules stated at the end of the previous section (p. 35) also hold in the case of trisyllabic words whose syllables alternate short syllables based on the [ $\pm$  HEAVY] feature: ကုဉ်ဒ်ဒ် *prantanley*, ကုဉ်ဒ်ဒ် *ralanghay*, ကုဉ်ဒ်ဒ် *kaybunay*, and ကုဉ်ဒ်ဒ် *saylingyang* receive final-syllable stress since this is their

Table 1.22: Stress patterns for [ $\pm$  HEAVY, + LONG] in trisyllabic words

-L -L +L	<i>peraysān</i>	x   x ǰ	‘paste’
-L +L -L	<i>raypānya</i>	x   ǰ x	‘at the stop’ (stop-LOC)
+L -L -L	<i>nōneri</i>	ǰ   ǰ x	‘deliberate, intentional’
-L +L +L	<i>mə-cān-cān</i>	x   ǰ ǰ	‘whatever fling’ (whatever=fling)
+L -L +L	<i>sānisān</i>	ǰ   x ǰ	‘copula; clutch (n.)’
+L +L -L	<i>lērayon</i>	ǰ   ǰ x	‘manwhore’
+L +L +L	—	—	—

last heavy syllable. The first syllables of *တယ်ကံ* *taykondam* and *မယ်ကွန်* *maykongas*, on the other hand, lose the secondary stress they would normally be assigned as two stressed syllables after another are normally avoided; the requirement of long syllables to not be unstressed does not come into effect here. *တယ်ကံ* *taykondam* is also an example for the rule that even if a syllable is not heavy, the last syllable that can be assigned stress will receive primary stress.

Carrying out the same analysis as above and moving the feature [+ LONG] through the various positions, we receive the results depicted in Table 1.22.<sup>23</sup> Since long syllables override stress of both light and heavy syllables as pointed out above (p. 35), the example words in this chart contain both of these syllable types. It was not too hard finding examples for all the slots in this case either, except that words with two long syllables in succession are rather rare. Still, only the case of three long syllables must remain unattested.<sup>24</sup>

Again, we can see that long syllables attract stress, in that the final syllables of *တယ်ကံ* *peraysān* and *နီနီကံ* *sānisān* are stressed even though the penultimate syllable is heavy on the virtue of containing a diphthong. As it is in an unstressed position and there is no requirement for the syllable to be stressed, the first syllable of *တယ်ကံ* *raypānya* loses stress adjacent to the stressed long penultimate syllable. *နီနီ* *nōneri* and *လီရယ်* *lērayon* display a secondary-stressed and a primary-stressed syllable next to each other, in the first case due to the rule that long syllables must not be unstressed and *နီနီ* *-eri* forming a valid disyllabic foot that receives regular trochaic stress, and in the second case due to two long syllables next to each other, of which

<sup>23</sup> For more precision, modifications will be made to the symbols given in footnote 21: let a double acute (ˆˆ) denote superheavy syllables with primary stress, and a double grave (˘˘) denote superheavy syllables with secondary stress.

<sup>24</sup> It would be possible to construct a word with three long syllables if the habitative suffix *-asa* did not delete the vowel at the end of the verb stem if there is one. *မယ်ကံ* *māsāy* ‘I traveled’ (PST-travel=ISG.TOP) would then become *\*မယ်ကံ* *\*māsāsāy* ‘I used to travel’ (PST-travel-HAB=ISG.TOP) instead of the actual form *မယ်ကံ* *māsasāy*; the verb stem is *မယ်* *asa-* ‘travel’.

the first—again—must not be unstressed. မဲးဝဲးဝဲးဝဲး *mə-cān-cān* is interesting insofar as the long-syllable stress rules operate on the second foot regularly, while the marker မဲး *mə-* ‘whatever’ is by default unstressed, so even if it is added to stressed monosyllabic words like ခဲး *hin* ‘box’, the foot of မဲးခဲး *mə-hin* ‘which box so ever’ would appear iambic, even though the syllable ခဲး *hin* itself is not heavy.

A further exception is formed by monosyllabic quantifying clitics like ခဲး *-ben* ‘all, every’ or ဘဲး *-ngas* ‘almost’. These are unstressed also when following an unstressed syllable of their host if they contain a short vowel.

### Stress in compounds

Ayeri has a number of lexicalized compound nouns that are treated as one word morphologically (33a). This is in contrast to compounds that are not as established terms, or formed ad hoc (33b):

- (33) a. *Ang bengay kardangirayya ya Litareng.*  
 ang beng=ay.Ø kardang.iray-ya ya Litareng  
 AT attend=1SG.TOP school.high-LOC LOC Litareng  
 ‘I attend university in Litareng.’
- b. *Ang pasyye Pila sungkoranyam kibas.*  
 ang pasy-ye Ø Pila sungkoran-yam kibas  
 AT be.interested.in-3SG.F TOP Pila science-DAT map  
 ‘Pila is interested in geography.’

For purposes of surveying stress patterns, we will only deal with the kind in (33a), though it may be noted that when not being overtly inflected, the second kind of compound will be treated as a word as well: နီးခဲးဝဲးဝဲး *sungkorankibas* ‘geography’. Another kind of indivisible compound as in (33a) is one formed from reduplication, for instance, နီးခဲးနီးခဲး *kusang-kusang* ‘model’, from နီးခဲး *kusang* ‘double’. Table 1.23 gives several examples along with their stress patterns. As a reference for the various rules in operation, consider the list above (p. 35).

The first word, ခဲးဝဲးဝဲး *apan-apan*, is not very noteworthy but I included it nonetheless as a reference for regular stress assignment to light syllables. The word decomposes into two feet, each of them has trochaic stress, which does not change when reduplicated. Per rule, the first syllable of the word receives secondary stress while the penultimate syllable bears primary stress. နီးခဲးနီးခဲး *kusang-kusang* is following the normal rules as well in that a heavy syllable takes stress from a light one. This does not change in reduplication.

မဲးဝဲးဝဲးဝဲး *depangcāti* is noteworthy since it follows the same stress pattern as ခဲးဝဲးဝဲး *apan-apan* in spite of consisting of one foot with a heavy second syllable (မဲးဝဲး *de-pang*) and another with a superheavy first syllable (ဝဲးဝဲး *cāti*). To avoid a clash, stress

Table 1.23: Examples of stress patterns in compounds

Word	Pattern	Translation	Constituents					
			Word	Pattern	Translation	Word	Pattern	Translation
<i>apan-apan</i>	ˈx x   ˈx x	‘extensive’	<i>apan</i>	ˈx x	‘wide’	<i>apan</i>	ˈx x	‘wide’
<i>depangcāti</i>	ˈx x   ˈx x	‘cuckold’	<i>depang</i>	x ˈx	‘fool’	<i>cāti</i>	ˈx x	‘lover’
<i>kusang-kusang</i>	x ˈx   x ˈx	‘model’	<i>kusang</i>	x ˈx	‘double’	<i>kusang</i>	x ˈx	‘double’
<i>latunkema</i>	ˈx x   ˈx x	‘tiger’	<i>latun</i>	ˈx x	‘lion’	<i>kema</i>	ˈx x	‘stripe’
<i>malingkaron</i>	ˈx x   ˈx x	‘coast, seashore’	<i>maling</i>	x ˈx	‘shore’	<i>karon</i>	ˈx x	‘water’
<i>māvaganeng</i>	ˈx x   x ˈx	‘mother’s siblings’	<i>māva</i>	ˈx x	‘mother’	<i>ganengan</i>	x   ˈx x	‘siblings’
<i>pikunanding</i>	ˈx x   x ˈx	‘mustache’	<i>piku</i>	ˈx x	‘beard’	<i>nanding</i>	x ˈx	‘lip’
<i>sapayyila</i>	ˈx x   ˈx x	‘limbs’	<i>sapay</i>	x ˈx	‘hand’	<i>yila</i>	ˈx x	‘foot’

is not shifted to the heavy syllable in  $\text{depang}$ , since it is not strictly necessary for it to be stressed:

$$(34) \quad (*\text{depang} \rightarrow \text{depang}) + \text{cāti} \rightarrow \left\{ \begin{array}{l} * \text{depang} + \text{cāti} \\ \text{depang} + \text{cāti} \end{array} \right\}$$

In the case of  $\text{malingkaron}$  again the rule operates that prohibits two stressed non-long syllables after another. Thus, even if the first component  $\text{maling}$  contains a heavy syllable, stress will not move there. In  $\text{latunkema}$ , the syllable /tun/ is assimilated to [tʊŋ] before the /k/ onset of the next syllable. For one, however, this does not make it heavy, and second, even if it did, stress would stay on the first syllable of the word for the same reason as in  $\text{malingkaron}$ . The same rule of stress hiatus avoidance operates in  $\text{sapayyila}$ .

Besides the shortening of the second component of the compound,  $\text{māvaganeng}$  retains the stress pattern of its constituents. Since /ma:/ is not in a final foot, it receives secondary stress. Moreover,  $\text{māvaganeng}$  and  $\text{pikunanding}$  both show that it is acceptable for two light syllables to follow each other.

### 1.3.2 Intonation

Peterson (2015: 66) writes that if “you’re creating a language on your own and you’re the only speaker, intonation is usually not high on the list of features to focus on, but intonational flavoring is well worth it (read: crucial) when it comes to making an authentic language.” Indeed, this has so far been a rather neglected topic in my work on Ayeri. Even though I made a handful of recordings in the past, I have never considered intonation much. Yet, of course, the spoken words in those recordings do not sound like robot speech either, so there must be intonational patterns that I have been subconsciously applying.<sup>25</sup>

Since intonation contours are notoriously difficult to display in print, I will give very approximate graphs of pitch in the respective examples for each surveyed pattern. Certainly there will be other patterns as well which would require more detailed gradations, but for the time being, I will only try to briefly describe those that are most prominent.

#### *Declarative statements*

Declarative statements have a gradually falling pitch contour based around an average pitch height, not deviating considerably on both ends:

<sup>25</sup> Whenever this happens, the fallback is likely to be a mixture of German and English, since those are the languages I am most familiar with.



(35)



*Ang gibayo Pintemis minganeri-ben yona.*  
ang giha-yo Ø Pintemis mingan-eri=hen yona  
AT blow-3SG.N TOP North Wind ability-INS=all 3SG.N.GEN.

‘The North Wind blew with all of his might.’

#### Yes-no questions

Since Ayeri does not use a particle or word order to mark closed questions as such, intonation is used to mark the difference from a declarative statement. To achieve a strong contrast, questions exhibit gradually rising intonation:

(36)



*Ang gibayo Pintemis minganeri-ben yona?*  
ang giha-yo Ø Pintemis mingan-eri=hen yona  
AT blow-3SG.N TOP North Wind ability-INS=all 3SG.N.GEN.

‘Did the North Wind blow with all of his might?’

#### ‘Wh-’ questions

Unlike English, Ayeri marks open questions with an in-situ question word. Open questions are thus marked by the question word causing a sharp rise and fall in the overall contour of the question. The first half of the clause has the rising contour of a question, the second half has gradually falling pitch.

(37)



*Ang engyo mico sinyá luga toya sam?*  
ang eng-yo mico sinyá-Ø luga toya sam  
AT be.more-3SG.N strong who-TOP among 3PL.N.LOC two

‘Who was the stronger of the two?’

#### Lists

List statements have the general gradual downward slope of declarative statements, but the individual items can nonetheless be marked by a pitch rise on the primary accent of each item.

(38)



*Le vacyeng seygo, disu, betay nay vasra.*  
 le vac=yeng seygo-Ø disu-Ø betay-Ø nay vasra-Ø  
 PT.INAN like=3SG.F.A apple-TOP banana-Ø berry-Ø and nut-Ø

‘She likes apples, bananas, berries and nuts.’

### Complement and relative clauses

Complement clauses are characterized by the short spike at the end of the preceding main clause followed by a short break which together signal the beginning of a new syntactic unit within the context of the current sentence, which is broadly similar to list statements. Otherwise, statements with complement clauses as well bear the overall downward-sloping contour of declarative statements if included in such.

(39)



*Ang manga rantong, engyo mico sinyāng.*  
 ang manga ran=tong eng-yo mico sinyang  
 AT PROG argue=3PL.N.A be.more=3SG.N strong who-A

‘They were arguing who is stronger.’

Relative clauses, on the other hand, do not receive special prosodic marking, but are treated the same as other basic sentence types. They display a continuous downward slope if part of a declarative statement, or a continuous upward slope if part of a question:

(40) a.



*Lugaya asāyāng si sitang-naykonyāng kong tova.*  
 luga-ya asāya-ang si sitang=naykon=yāng kong tova-ya  
 pass=3SG.M traveler-A REL self=wrap=3SG.M.A inside cloak-LOC

‘A traveler passed who had wrapped himself into a cloak.’

b.



*Adareng asāyās si le ninyāng tova?*  
 ada-reng asāya-as si le nin=yāng tova-Ø  
 that-A.INAN traveler-P REL PT.INAN wear=3SG.M.A coat-TOP

‘Is that the traveler who wore the coat?’

*Contrast*

Ayeri uses a kind of topic system for highlighting constituents in a clause by morphosyntactic means, but this is still different from emphasis on semantic grounds, for example when the speaker wants to highlight a semantic difference in the same syntactic position, as in the following example, which presents a possible answer to the question posed in (4ob):

(4t)



<i>Adareng</i>	<i>asāyās</i>	<i>si</i>	<i>le</i>	<i>nin-yāng</i>	<i>kegan.</i>
ada-reng	asāya-as	si	le	nin=yāng	kegan-Ø
that-A.INAN	traveler-P	REL	PT.INAN	wear=3SG.M.A	hat-TOP

‘It is the traveler who wore the *hat*.’

We can see here a spike towards the end of the utterance where the word <sup>ḥḥ</sup>*kegan* ‘hat’ is placed. This word receives extra stress for contrast with <sup>ḥḥ</sup>*tova* ‘coat’, which is what the other person had asked about.



## 2 Writing system

In the previous chapter, example words were given in Ayerī's script, ၵၵၵၵၵ ၵၵၵၵ *Tabano Hikamu*, wherever possible. Thus, it seems advisable to include a description of Ayerī's native writing system here as well. Literally, ၵၵၵၵၵ ၵၵၵၵ *Tabano Hikamu* means 'Round Script' (script round), which is an old formation based on the word ၵၵၵၵ *taban-* 'write' that stuck. The current word for 'script' is ၵၵၵၵ *tabanan* 'writing'. Tahano Hikamu was originally named thus because of an earlier draft for a script that never made it very far beyond the drawing board and which was a lot more angular and boxy, see Figure 2.1—Tahano Hikamu was a lot more bubbly in comparison, especially early on (Figure 0.2).<sup>1</sup>

As we have seen in the previous chapter, Ayeri's prosody strongly emphasizes the syllable as a unit. Thus, it is not a surprise that Ayeri's native script, Tahano Hikamu, is an alphasyllabary similar to the Brāhmī alphabets of India and Southeast Asia (Salomon 1996; Court 1996). Scripts like these are

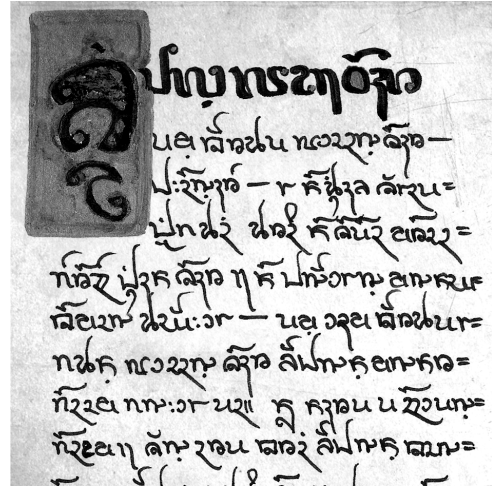
based on the unit of the graphic “syllable” [...], which by definition always ends with a vowel (type V, CV, CCV, etc.). Syllables consisting of a vowel only (usually at the beginning of a word or sentence) are written with the *full* or *initial vowel signs* [...]. But when, as is much more frequently the case, the syllable consists of a consonant followed by a vowel, the vowel is indicated by a diacritic sign attached to the basic sign for the consonant [...]. (Salomon 1996: 376)

For Tahano Hikamu the definition that a syllable consisting only of a vowel is written with an initial vowel sign is only true under certain circumstances, as we will see below. Moreover, Brāhmī scripts are often characterized by conjuncts of clustered consonants which may become quite large and sometimes behave in an idiosyncratic way. Consonant conjuncts like Devanāgarī त्व ⟨tva⟩ from त ⟨ta⟩ + व ⟨va⟩ or idiosyncratic conjuncts like क्ष ⟨kṣa⟩ for क ⟨ka⟩ + ष ⟨ṣa⟩ are not known in Tahano Hikamu, however. Subscript notation for consonant clusters and special diacritics marking coda consonants like in Javanese (Kuipers and McDermott 1996: 478–479) are equally unknown to Tahano Hikamu. This does not mean, however,

<sup>1</sup> Unfortunately, there is no documentation of the Box script surviving that I know of.



(a) Old and aborted draft: Box script



(b) Ayeri's native script: Tahano Hikamu

Figure 2.1: Box script and Hikamu

that final consonants are simply omitted in writing, since closed syllables are reasonably common enough in Ayeri to warrant indicating them. Thus, there is “a special mark to eliminate the vowel of the previous syllable, thereby leaving a consonant in a syllable-final position” (Kuipers and McDermott 1996: 476). That is, a diacritic exists which marks the absence of an inherent vowel, rendering the syllable consonant-only.

Another difference from Brāhmī-family scripts is that vowel length and diphthongs in [i] are indicated by dedicated diacritics, so the long vowels are not doubled versions of their short counterparts. Like in Kharoṣṭhī—another historically important ancient script of India—initial vowels are not represented by unique graphemes but they are all written like post-consonantal vowel diacritics (Salomon 1996: 377), though in Tahano Hikamu with a character without an inherent sound value. For this reason, the character is indicated in the table below as  $\emptyset$  /Ø/; its native name is  $\text{ṛṇṇ}$  *ranyan* ‘nothing’.<sup>2</sup> Similar to a number of Brāhmī scripts, Tahano Hikamu puts diacritics not only below or above consonant bases, but also before them. This, however, is not limited to vowel graphemes as in Devanāgarī ि ⟨i⟩ or Javanese ꦺ ⟨e, é/è⟩ (Kuipers and McDermott 1996: 478).<sup>3</sup>

<sup>2</sup> I will give the native names of graphemes here, but will refer to them by their English names for clarity in the running text.

<sup>3</sup> Kuipers and McDermott (1996) do not say, but it seems that both might be related, since they are both functionally the only prepended vowel diacritics and both represent a high front sound; this is just a guess, however.

Table 2.1: The consonant graphemes

/pa/	/ta/	/ka/	/ba/	/da/	/ga/
ᵑ	ᵑ	ᵑ	ᵑ	ᵑ	ᵑ
/ma/	/na/	/ɲa/	/va/	/sa/	/ha/
ᵑ	ᵑ	ᵑ	ᵑ	ᵑ	ᵑ
/ra/	/la/	/ja/	/Ø/		
ᵑ	ᵑ	ᵑ	ᵑ		

## 2.1 Consonants

Tahano Hikamu is mainly based on consonant bases that are modified by diacritics. Since the vowel /a/ is so highly frequent in Ayeri, it is also the vowel that is *inherent* to every consonant grapheme if not further modified by vowel diacritics. Consonant letters are simply referred to as *pa*, *ta*, *ka*, ... Table 2.1 displays all the main consonants. The customary collation is—similar to the IPA table—roughly grouping the letters according to their sound value by anteriority (front → back) and sonority (low → high). The script is monocameral, that is, there is no distinction between capital letters and minuscule letters as in the Latin, Greek, Cyrillic, Georgian, and Armenian alphabet. It is also written in lines from left to right.

ᵑ, which in Ayeri has no sound value but is used as a base for initial vowels, may also serve as the character for /ʔa/. What is, moreover, interesting about ᵑ ⟨nga⟩ is that even though before, /ɲ/ was treated strictly as a coda consonant in the previous chapter, it is in fact treated as an onset consonant in writing if a vowel is following:

- (1) ᵑ + ᵑ  
 /pa/      /ɲis/  
 ᵑ pangis /paɲ.is/ ‘money’

Tahano Hikamu knows a few ligatures. First of all, when two ᵑ ⟨na⟩ are in succession within a word, they will form a ligature ᵑ ⟨nana⟩:

- (2) ᵑ + ᵑ → ᵑ  
 /na/      /na/      /nana/

This is distinct from conjuncts like in Devanāgarī et al., though, since the unmodified sound value will still be /nana/, not \*/nna/, so the inherent vowel of each ᵑ

Table 2.2: Additional consonant graphemes

/fa/	/wa/	/tsa/	/za/	/ʃa/	/ʒa/
ṃ	ṃ	ṃ	ṃ	ṃ	ṃ
/ɕa/	/ksa/	/kwa/	/xa/	/ɣa/	
ṃ	ṃ	ṃ	ṃ	ṃ	

Table 2.3: Primary vowel graphemes

	/i/	/e/	/a/	/o/	/u/	/ə/	/au/
Diaritics	ṃ	ṃ	(ṃ)	ṃ	ṃ	ṃ	ṃ
Independent	ṃ	ṃ	ṃ	ṃ	ṃ	ṃ	ṃ

⟨na⟩ is not deleted, and each 𑌕 ⟨na⟩ retains the ability to be modified by diacritics. Tahano Hikamu also has a few ligatures of the kind you would find in Brāhmī scripts. The difference is that they are not productive, but fossilized.

- (3) a. 𑌕 ⟨kwa⟩ ← 𑌕 ⟨ka⟩ + 𑌕 ⟨va⟩,  
 b. 𑌕 ⟨tsa⟩ ← 𑌕 ⟨ta⟩ + 𑌕 ⟨sa⟩, and  
 c. 𑌕 ⟨ksa⟩ ← 𑌕 ⟨ka⟩ + 𑌕 ⟨sa⟩.

These conjunct letters are, however, not normally employed by Ayeri. Table 2.2 shows all additional consonants, added to write other languages. Individual languages may adapt the sound values slightly to fit their own purposes.

## 2.2 Vowels

As mentioned above, vowels are written as diacritics that are added to consonants. In principle, every consonant has two slots for vowels, a primary one atop it, and a secondary one below it. Vowels added to consonants in the primary slot delete their inherent /a/:

- (4) 𑌕 → 𑌕  
 /pa/      /pe/

Table 2.3 gives the primary vowel signs. Of the vowel signs given there, only 𑌕 ⟨ə⟩ is not used in Ayeri. 𑌕 ⟨au⟩ is the only diphthong for which a dedicated



Table 2.4: Secondary vowel graphemes

/i/	/e/	/a/	/o/	/u/	/ə/	/au/

grapheme exists, even though its occurrence is rather limited. The independent vowel graphemes are used at the beginning of words or inside words when there is no other way to spell the vowel, which is occasionally the case for secondary vowels. Secondary vowels are vowels that are not parts of diphthongs (even though another language might use them to spell diphthongs that are not covered by default), but follow the vowel of a syllable directly. They are attached underneath a consonant base, for example:

- (5)  $\text{ɰ} \rightarrow \text{ɰ}^{\text{◌}} \rightarrow \text{ɰ}^{\text{◌}}$   
       /ja/        /je/        /jea/

In fact, the principle that every consonant base with its diacritics represents one syllable is slightly violated here, which is also the reason why secondary vowels very occasionally need to be spelled as independent vowels, for example when the secondary vowel is long, as in the word  $\text{ရူၤ}$  *ruān* ‘duty’:

- (6)  $\text{ၣ်} \rightarrow \text{ၣ်}^{\text{◌}} \quad (\text{ၣ်}^{\text{◌}})$   
       /ru/        /rwa:/        /ru:a/

Example (6) uses a diacritic,  $\text{◌}$ , to indicate length. If  $\text{◌}$  is put directly under  $\text{ၣ်}$  (*ru* (the  $\text{◌}$  diacritic moves down where it is not in the way), the syllable will incorrectly spell /ru:a/ instead of the intended /rua:/. This is because diacritics modify consonants and primary vowels, but there is no way to modify a secondary vowel directly. Table 2.4 gives a list of secondary vowels corresponding to that of primary vowels above. The vowels as well are just referred to by their sound value; ‘primary’ and ‘secondary’, ‘superscript’ and ‘subscript’ or ‘upper’ and ‘lower’ may be chosen to disambiguate their positions; the native names may use  $\text{ၣ်}$  *iray* ‘high’ and  $\text{ၣ်}$  *eyra* ‘low’ to disambiguate, so  $\text{ၣ်}$  *e iray* denotes the superscript ⟨e⟩ diacritic while  $\text{ၣ်}$  *e eyra* denotes its subscript counterpart.

As a further exception, those consonant bases with an ascender ( $\text{ၣ်}$  ⟨ka⟩,  $\text{ၣ်}$  ⟨da⟩,  $\text{ၣ်}$  /ça/) move the primary vowel to the secondary slot below the consonant by default while indicating the vacancy of the primary slot at the top with a dot. This is done to avoid crossing the ascender of the consonant with a vowel diacritic:

$$(7) \quad \begin{array}{ccccc} \text{ḱ} & \rightarrow & \text{ḱ} & \rightarrow & \text{ḱ} \\ /ka/ & & /ka.i/ & & /ki/ \end{array}$$

If the primary vowel slot were not silenced by the  $\circ$  diacritic, it could reasonably be assumed that the consonant is not losing its inherent /a/ and the vowel below the consonant indicates a secondary vowel, spelling /CaV/. If, however, a secondary vowel is *actually* added, primary and secondary vowels will be assigned the regular primary and secondary slots, respectively, again (8a). This condition also holds true for subscript diacritics (8b).

$$(8) \quad \begin{array}{lcl} \text{a.} & \begin{array}{ccc} \text{ḱ} & \rightarrow & \text{ḱ} \\ /ki/ & & /ki.e/ \end{array} \\ \text{b.} & \begin{array}{ccc} \text{ḱ} & \rightarrow & \text{ḱ} \\ /ki/ & & /ki:/ \end{array} \end{array}$$

The order of secondary vowels and subscript diacritics is iconic insofar as it follows the order of sounds in the syllable. Thus, secondary vowels appear below the consonant-doubling diacritic,  $\text{ḱ}$ , while they appear above the syllable-final homorganic nasal diacritic,  $\text{ḱ}$ :

$$(9) \quad \begin{array}{lcl} \text{a.} & \begin{array}{ccc} \text{ḱ} & \rightarrow & \text{ḱ} \\ /ppa/ & \rightarrow & /ppea/ \end{array} \\ \text{b.} & \begin{array}{ccc} \text{ḱ} & \rightarrow & \text{ḱ} \\ /peN/ & \rightarrow & /peaN/ \end{array} \end{array}$$

## 2.3 Diacritics

We have already encountered a few diacritics, though Tahano Hikamu comes with a lot more, some of which undergo non-trivial positioning and repositioning rules. As vowels are primarily expressed as superscripts, diacritics are primarily realized as subscripts, so in the following I will first describe subscript diacritics; then prepended diacritics, which Ayeri also has a number of, both as graphemes in their own right and as allographs of other subscript diacritics; and then, lastly, superscript diacritics.

### 2.3.1 Subscript diacritics

Table 2.5 shows the bottom-attaching diacritics. The ‘large diacritics’ ( $\text{ḱ}$  through  $\text{ḱ}$ ) cause the secondary slot of consonants to move down below the diacritic. ‘Small

Table 2.5: Bottom-attaching diacritics

	Native name	Function	Example
◌̄	တုပာဆိ <i>tupasati</i> ‘long-maker’	Lengthens the primary vowel of the syllable	၎ <i>pa</i> → ပျ <i>pā</i>
◌့	ယုရော <i>ya eyra</i> ‘low ya’	⟨ya⟩ following another consonant, also across syllables. Marks palatalization of တ <i>ta</i> , သ <i>da</i> , ခ <i>ka</i> , ဂ <i>ga</i> and ယ <i>ya</i> in Ayeri.	အာ <i>ara</i> → အျာ <i>arya</i> ; တာ <i>ta</i> → တျာ <i>ca</i>
◌့	ရိဂယ <i>ringaya</i> ‘raiser’	Palatalizes a consonant (not used in Ayeri)	တာ <i>ta</i> → တျာ <i>/tʰa/</i> , <i>/tʃa/</i>
◌့	အိလငယ <i>ulangaya</i> ‘breather’	Aspiration or frication of a consonant (not used in Ayeri)	တာ <i>ta</i> → တျာ <i>/tʰa/</i> , <i>/θa/</i>
◌့	ရပယုရော <i>raypāya eyra</i> ‘low stopper’	Glottal stop coda or glottalization of a consonant (consonant letters with ascenders; not used in Ayeri)	ခ <i>ka</i> → ချ <i>/kaʔ/</i> ; သ <i>da</i> → သျ <i>/dʰa/</i>
◌့	ဂိယ <i>gondaya</i> ‘extinguisher’	Deletes the inherent <i>/a/</i> of a consonant, e.g. in consonant clusters or closed syllables	ပာ <i>para</i> → ပာ <i>pra</i> , ပာ <i>par</i>
◌့	ဝိနာ <i>vināti</i> ‘nasalizer’	Indicates a homorganic nasal or nasalizes the vowel, depending on the language	ပာ <i>pada</i> → ပာ <i>panda</i> <i>/panda/</i> or <i>/pāda/</i>
◌့	ကွေအိယ <i>kusangisāti</i> ‘duplicator’	Indicates a geminated or otherwise double consonant	ပာ <i>pala</i> → ပာ <i>palla</i>

diacritics' (◌◌ through ◌◌) can attach in this place as well as secondary vowels, as does the homorganic nasal diacritic ◌◌ in this diacritic-fraught example:

- (10)  $\text{တၢ်နီၤ} + \text{ပုၤလၢ} \rightarrow \text{တၢ်နီၤပုၤလၢ}$   
 /tʃa:n/      /pu'lɔ/      /tʃa:mpu'lɔ/  
 တၢ်နီၤပုၤလၢ *cāmpuluy* 'heterosexual'

It also needs to be noted that diacritics like ◌◌ are applied progressively to words as a whole, not stopping at morpheme and syllable boundaries, so even though *toryeng* 'she sleeps' may be composed of တၢ်: *tor-* 'sleep' + ပုၤ *-yeng* (=3SG.F.A) and syllabifies as /tor.'jeŋ/, the spelling is not \*တၢ်ပုၤ as one might expect, but တၢ်ပုၤ.

Even though the primary position for small diacritics is underneath consonants, the diacritic deleting the inherent vowel, ◌◌, very commonly also appears after a consonant letter at the end of words:

- (11)  $\text{ပ} \quad \text{နီၤပုၤလၢ} \quad \text{ပၢၤ} \quad \text{နီၤပုၤလၢ}$   
*ya nimreng      pangan narānyena.*  
*ya nim-reng      pangan-Ø narān-ye-na*  
 LOCT appear=3SG.INAN.A end-TOP word-PL-GEN  
 'It appears at the end of words.'

This strategy is advantageous in that Tahano Hikamu leaves very little space between individual words: ပနီၤပုၤလၢပၢၤနီၤပုၤလၢ. With the dot after the consonant, word boundaries are more visible.

### 2.3.2 Prepended diacritics

Example (10) leads us directly to the next class of diacritics—those that are prepended to the consonant letter, either because they are simply placed there or because of allography. Let us first list those diacritics that appear in front of consonants obligatorily (Table 2.7).

As Table 2.7 shows, the only obligatorily prepended diacritic that Ayeri uses is the one that marks diphthongs, ◌◌. It needs to be noted here that ◌◌ changes into ◌◌ (ya) proper when a vowel follows, but stays ◌◌ when a ◌◌ (ya) follows:

- (12) a.  $\text{ပုၤ} \text{ baday 'hero'} \rightarrow \text{ပုၤပုၤ} (*\text{ပုၤပုၤ}) \text{ badayang 'the hero' (hero-A);}$   
 b.  $\text{ပုၤ} \text{ tipuy 'grass'} \rightarrow \text{ပုၤပုၤ} (*\text{ပုၤပုၤ}) \text{ tipuyya 'in the grass' (grass-LOC).}$

<sup>4</sup> In a Tahano Hikamu orthography I devised for English once, ◌◌ was used for /ə/, as in the NURSE vowel in American English: ◌◌ nurse.

Table 2.7: Obligatorily prepended diacritics

Native name	Function	Example
၂ <sup>○</sup> ကုလ်ဃံကု <i>lentankusang</i> 'double-sound'	Marks a diphthong with /ɪ/	ပံ <i>pe</i> → ပံ၂ <i>pey</i>
၃ <sup>○</sup> တီလံဃပ <i>tilamaya</i> 'changer'	Marks raised vowels (i.e. umlaut; not used in Ayeri)	ပံ <i>po</i> → ပံ၃ <i>/pø/</i>
၃ <sup>○</sup> ဟိယံဃပ <i>hiyamaya</i> 'roller'	Marks retroflex consonants (not used in Ayeri) <sup>4</sup>	တံ <i>ta</i> → တံ၃ <i>/tʰa/</i>

Table 2.8: Allographically prepended diacritics

Native name	Function	Example
၁ <sup>○</sup> တုပာသိဗာရ် <i>tupasati marin</i> 'anterior long-maker'	Lengthens the primary vowel of the syllable	ဟံ <i>syā</i> → ဟံ၁ <i>syā̄</i> , နံ <i>na</i> → နံ၁ <i>nā</i>
၂ <sup>○</sup> ယာရ် <i>ya marin</i> 'anterior ya'	⟨ya⟩ following another consonant, also across syllables.	နံ <i>na</i> → နံ၂ <i>nya</i>
ကံ၁ပာရ် <i>ringaya marin</i> 'anterior raiser'	Also used as an allograph for the palatalization proper diacritic.	ဟံ <i>/sʰa/</i> → ဟံ၁ <i>/sʰi/</i>
၁ <sup>○</sup> ဝီလံဂံ၁ပာရ် <i>ulangaya marin</i> 'anterior breather'	(Pre-)Aspiration or frication of a consonant (not used in Ayeri)	ကံ <i>nga</i> → ကံ၁ <i>/ŋʰa/</i> ; တံ <i>ta</i> → တံ၁ <i>/tʰa/</i>

Besides  $\text{ᳵ}$ , there are also a number of diacritics that are also obligatorily prepended to consonants, but do so as context-sensitive allographs (Table 2.8). The selection of the variant diacritics is not random or up to the aesthetic eye of the writer (even though the device itself is certainly a matter of aesthetics), but it is governed by rules. The prepended forms listed in Table 2.8 are thus triggered

1. when there is no stem or bowl for the regular subscript diacritic to attach to, which is the case for  $\text{ᳵ}$  ⟨na⟩,  $\text{ᳶ}$  ⟨nga⟩,  $\text{᳷}$  ⟨va⟩, and  $\text{᳸}$  ⟨wa⟩:

$$\begin{array}{ll}
 (13) \quad \text{a. } \text{ᳵ} \rightarrow \text{ᳵᳵ} & \text{c. } \text{᳷} \rightarrow \text{᳷᳷} \\
 \quad \quad \quad /na/ & \quad \quad \quad /va/ \\
 \quad \quad \quad /na:/ & \quad \quad \quad /va:/ \\
 \quad \quad \quad \text{b. } \text{ᳶ} \rightarrow \text{ᳶᳶ} & \text{d. } \text{᳸} \rightarrow \text{᳸᳸} \\
 \quad \quad \quad /ŋa/ & \quad \quad \quad /wa/ \\
 \quad \quad \quad /ŋa:/ & \quad \quad \quad /wa:/
 \end{array}$$

2. when a large subscript diacritic would be added after another large subscript diacritic—this position can only be occupied once, so further large subscripts are prepended:

$$\begin{array}{ccccccc}
 (14) & & + \text{ᳵ} & & + \text{ᳶ} & & + \text{᳷} & & + \text{᳸} \\
 \text{ᳵ} & \rightarrow & \text{ᳵᳵ} & \rightarrow & 2\text{ᳵᳵ} & \rightarrow & 2\text{ᳵᳵᳶ} & \rightarrow & 22\text{ᳵᳵᳶ᳷} \\
 /ta/ & & /tʰa/ & & /tʰja/ & & /tʰji/ & & /tʰji:/
 \end{array}$$

The order of diacritics follows the logic of the respective language's phoneme inventory, so if there are, for example, retroflex consonants and both dental and retroflex consonants can be aspirated, retroflexion would be marked first, then aspiration. If there is a palatalization contrast on top of this, the diacritic would be added after aspiration.

When adding large diacritics to stemless consonants, they are prepended from the beginning, as we saw in (13), and just like in (14), this principle continues:

$$\begin{array}{ccccccc}
 (15) & & + \text{ᳶ} & & + \text{᳸} & & + \text{ᳵᳵ} \\
 \text{ᳵ} & \rightarrow & 2\text{ᳵ} & \rightarrow & 22\text{ᳵ} & \rightarrow & ᳵᳵ22\text{ᳵ} \\
 /na/ & & /nja/ & & /nja:/ & & /nja:ɿ/
 \end{array}$$

3. with consonants directly following  $\text{ᳵ}$  ⟨na⟩, to avoid a clash with its swash:

$$(16) \quad \underset{/na/}{\text{Ⴌ}} + \underset{/pa:/}{\text{Ⴌ}} \rightarrow \underset{/napa:/}{\text{ႬႬ}} \quad (*\text{ႬႬ})$$

An exception to this exception occurs, however, when the consonant is not directly following. In this case, no reordering happens, only Ⴌ ⟨na⟩ *may* reduce its swash in size to accommodate the following prepended diacritic:<sup>5</sup>

$$(17) \quad \underset{/na/}{\text{Ⴌ}} + \underset{/pai/}{\text{ႬႬ}} \rightarrow \underset{/napai/}{\text{ႬႬႬ}} \quad (^\text{?}\text{ႬႬႬ})$$

4. in other cases where a clash of subscript diacritics needs to be avoided:

$$(18) \quad \underset{/di/}{\text{ႬႬ}} + \underset{/pa:/}{\text{Ⴌ}} \rightarrow \underset{/dipa:/}{\text{ႬႬႬ}} \quad (*\text{ႬႬႬ})$$

Alternatively, the following solution is permissible:

$$(19) \quad \underset{/di/}{\text{ႬႬ}} + \underset{/pa:/}{\text{Ⴌ}} \rightarrow \underset{/dipa:/}{\text{ႬႬႬ}}$$

When two long syllables follow each other, as in *bāmā* ‘mom-and-dad’, one of the length diacritics should definitely be pulled to the front:

$$(20) \quad \begin{array}{l} \underset{/ba:/}{\text{ႬႬ}} + \underset{/ma:/}{\text{ႬႬ}} \rightarrow \underset{/ba:ma:/}{\text{ႬႬႬႬ}} \quad (^\text{?}\text{ႬႬႬႬ}) \\ \text{or:} \quad \underset{/ba:/}{\text{ႬႬ}} + \underset{/ma:/}{\text{ႬႬ}} \rightarrow \underset{/ba:ma:/}{\text{ႬႬႬႬ}} \end{array}$$

Generally, prepended diacritics apply only to a single consonant grapheme, not a whole consonant cluster as such. Thus, for instance, in words like *pray* ‘smooth’ ႬႬ appears before Ⴌ ⟨ra⟩, not before Ⴌ ⟨pa⟩, since Ⴌ ⟨ra⟩ is the closest consonant before the syllable nucleus which we are modifying by adding the ႬႬ. Since in the case of *pray* the inherent vowel of Ⴌ ⟨pa⟩ is silent, it receives a diacritic Ⴌ to mark this fact:

<sup>5</sup> The font I am using here is designed so that the reduced combination looks nicer, but if unreduced, Ⴌ ⟨na⟩’s swash is not so long as to cross the descender of ႬႬ either in this particular case.

(21)     $\text{ṛṛ}$              $(^*\text{ṛṛ})$   
              /praɪ/

What (21) shows is that essentially, /praɪ/ is split into /p/ + /raɪ/ for purposes of spelling, rather than /pr/ + /aɪ/. If necessary, it is also possible this way to distinguish, for instance,  $\text{ṛṛ}$  /tʃa/ from  $\text{ṛṛ}$  /tʃa/. It would be up to the respective language's orthography to decide whether either combination spells /tʃa/ or whether the  $\text{ṛ}$  diacritic is needed on both consonants—that is,  $\text{ṛṛ}$ —to spell the retroflex affricate.

### 2.3.3 Superscript diacritics

Ayeri's standard position for diacritics is below consonants, but sometimes it is nicer to put them on top, especially for the letter  $\text{ṛ}$  (na) due to its swash, as well as for  $\text{ṛ}$  (va) since the space below its flag is empty otherwise, thus not providing much of a visual connection. The only diacritic that is normally attaching to the top of consonants is that for the glottal stop—we have already encountered its subscript allograph earlier. Since Ayeri's phoneme inventory does not possess a phonemic glottal stop or glottalization, this diacritic is not used in Ayeri. The list of superscript diacritics is given in Table 2.9.

At times, it may be necessary to attach both a superscript diacritic and a vowel sign above a consonant. In this case, the consonant-modifying diacritic is placed first and the vowel diacritic on top of it—this is exactly equivalent to the rule exemplified for subscript diacritics in (9).

(22) a.     $\text{ṛ}$      $\rightarrow$      $\text{ṛ}$   
              /vva/    $\rightarrow$    /vve/  
       b.     $\text{ṛ}$      $\rightarrow$      $\text{ṛ}$   
              /vva/    $\rightarrow$    /vvaN/

## 2.4 Numerals

Ayeri uses a duodecimal number system, that is, a system based on the powers of 12, which is a typological rarity.<sup>6</sup> There is a digit for zero, so the system is positional, like the Hindu–Arabic digits used by the Latin alphabet. The numerals for the numbers from 1 to 12 are shown in Table 2.10.

<sup>6</sup> And one possibly overrepresented by fictional languages due to its rarity in natural languages.



Table 2.9: Superscript diacritics

Native name	Function	Example
◌် ꨀꨣꨩꨣꨩꨣꨩ <i>gondaya ling</i> ‘upper extingisher’	Deletes inherent /a/ of consonant, e.g. in consonant clusters or closed syllables	ṽṽ <i>vara</i> → ṽṽ <i>vra</i>
◌် ꨀꨣꨩꨣꨩꨣꨩ <i>vināti ling</i> ‘upper nasalizer’	Indicates a homorganic nasal or nasalizes the vowel, depending on language/context	ṽṽ <i>naka</i> → ṽṽ <i>nanka</i> /naṅka/ or /nāka/
◌် ꨀꨣꨩꨣꨩꨣꨩ <i>kusangisāti ling</i> ‘upper duplicator’	Indicates a geminated or otherwise double consonant	ṽṽ <i>pana</i> → ṽṽ <i>panna</i>
◌် ꨀꨣꨩꨣꨩꨣꨩ <i>raypāya stopper</i>	Glottal stop coda or glottalization of a consonant (not used in Ayeri)	ṽṽ <i>ta</i> → ṽṽ /taʔ/; ṽṽ <i>sa</i> → ṽṽ /sʔa/

Table 2.10: The numerals

1	2	3	4	5	6
ṽṽ	ṽṽ	ṽṽ	ṽṽ	ṽṽ	ṽṽ
7	8	9	A	B	10
ṽṽ	ṽṽ	ṽṽ	ṽṽ	ṽṽ	ṽṽ

## 2.5 Punctuation and abbreviations

Tahano Hikamu’s system of manipulating the sound of syllables is very sophisticated, so it comes as no surprise that it is also host of a large number of punctuation marks. Table 2.11 lists the ones commonly encountered, Table 2.12 the ones not so commonly encountered.

◌ (◌) does not look very much like a dot or a point, but it is derived from a sign that looks like two circles stacked on top of each other, similar to : ◌ (see Figure 0.2). There is no mark for a comma as such, so : ◌ or – ◌ cannot be used in this way. Instead of a comma, a wide word space is used to separate syntactic units. A long dash — ◌ is also sometimes found at the end of paragraphs or texts to mark their end. The strong exclamation mark ◌ may appear in its exclamatory

	Native name	Function	Example
॥	ṣṣ <i>dan</i> ‘dot’	Full stop	ṣṣṣṣṣṣ <i>Sarayāṅg.</i> ‘He left.’
:	ṣṣṣṣ <i>dan-dan</i> ‘little dot’	A separator for small things, like clitics and abbreviations; divides the constituents of reduplication	ṣṣṣṣṣṣ <i>ada-nanga</i> ‘this house’; ṣṣṣṣ ṣṣṣṣ <i>pd</i> ‘ṣ hrs’; ṣṣṣṣṣṣ <i>dan-dan</i> ‘dot-dot, little dot’
—	ṣṣṣṣṣ <i>puntān</i> ‘dash’	General sign for a longer pause, equivalent to a dash, colon, semicolon, brackets	ṣṣṣṣṣṣ <i>Yan – saru!</i> ‘Yan – go!’
ṣ	ṣṣṣṣṣṣ <i>damprantan</i> ‘question point’	Marks questions	ṣṣṣṣṣ <i>Manisu?</i> ‘Hello?’
ṣ	ṣṣṣṣṣ <i>dambahān</i> ‘shouting point’	Marks exclamations; strong exclamations may be marked by the ṣ variant.	ṣṣṣṣṣ <i>Manisu!</i> ‘Hello!’; ṣṣṣ <i>Yi!</i> ‘Urgh!’

Regarding the less common marks, some of these seem like all to bland copies of modern punctuation in the Latin alphabet, especially the brackets and the decimal point. Still, however, they may serve their purpose sometimes, and the brackets may maybe come with the redeeming notion that they push off the text around the inclusion rather than encapsulating the inclusion within it, so the visual effect is slightly different. The name brackets are interesting and maybe useful insofar as many names in Ayeri are derived from common words, for example, ၼံၼ် ၼ်း *Ajān*, a male name, is literally ‘play, game’, relating to a playful character; ၼံၼ် ၼ်း *Migoray*, a female name, literally means ‘flower’. The name brackets, then, make it unmistakably clear that a proper noun is intended rather than a common noun. The line-breaker serves the purpose of marking the continuation of a clause at the end of a line either generally or where there would be ambiguity with a comma, which, as described above, is a large blank that would otherwise be invisible at the end of a line.

Two very common abbreviations are symbolic in nature, like the ampersand (&) in the Latin alphabet, and incidentally, they correspond to it in that the very com-

Table 2.12: Less common punctuation marks

Native name	Function	Example
◌ <sup>◌</sup> ◌ <sub>◌</sub> မ်းဝ်း <i>danarān</i> ‘speaking point’	Quotation marks	ၵုၵ်းမၵ်းဝ်း <i>Narayāng</i> “ <i>Manisu!</i> ” ‘He says, “Hello!”’
◌◌    မ်းဝ်း <i>dankayvo</i> ‘beside-point’	Bracketing of text	ၵုၵ်းမၵ်းဝ်း <i>bahis (larau)</i> ‘a (nice) day’
◌ <sub>◌</sub> မ်းဝ်း <i>dangaran</i> ‘name-point’	Explicitly marks a name as such. For the closing name parenthesis, ◌ <sub>◌</sub> can be found as well.	ၵုၵ်းမၵ်းဝ်း <i>Ajān Savati</i> ; ၵုၵ်းမၵ်းဝ်း <i>Pila Lay</i> <i>Maran</i>
◌    မ်းဝ်း <i>dansinday</i> ‘number-point’	Marks (duo)decimal fractions	ၵုၵ်းမၵ်းဝ်း <i>17.4582</i> ‘19.37482’
◌    မ်းဝ်း <i>adrumaya</i> ‘breaker’	Marks line breaks within a phrase	

mon small word *ၵုၵ်း* *nay* ‘and’ may be abbreviated as *ၵု*. Based on this, its reduplicated form *ၵုၵု* *naynay* ‘furthermore, also’ may be abbreviated as *ၵု*.

## 2.6 Styles

Over the course of the years since Tahano Hikamu’s inception, I have liked to experiment with different styles of writing, that is, I tried applying a number of different writing styles to the script to change its look and feel while still staying true to the overall character shapes and the system behind the script. The example text I will be using to illustrate the different styles in the following is an Ayeri translation of the first article of the United Nations *Universal declaration of human rights* (Becker 2011a):

*Sa vesayon keynam-ikan tiganeri nay kaytanyeri sino nay kamo.*  
*Ri toraytos tenuban nay iprang, nay ang mya rankyon sitanyās ku-netu.*

[All human beings are born free and equal in dignity and rights.  
They are endowed with reason and conscience and should act towards each other in  
a spirit of brotherhood.] (United Nations 1948: Article 1)

The examples above are all using a style I call ‘book’ style since it comes close to printed letters, or also what might be conceivable as being written with quills or

nibs on parchment or paper—of course, pen and paper is also what I used to make up the letters in the first place, without second thought about the limitations of the supposed original writing utensils. The ‘book’ style letters are what I consider the canonical form. Figure 2.2 shows the above article in this letter style.



Figure 2.2: Tahano Hikamu, 'book style'

As described above, I have long found the look of the Javanese script<sup>7</sup> rather interesting and thus I tried applying the general aesthetics of what I had seen of it to Tahano Hikamu at some point. As mentioned above as well, there are no subscript letters in Ayeri, and the number of large swirling diacritics is also rather low, so there is still definitely a difference in appearance. The ‘angular’ style is also the one that is comparable in function to our bold face or italic style letters, since it is used in captions or to highlight special text within running text. This letter style (ꦲꦶꦂꦶꦏꦸ *hinya* ‘angular’) is displayed in Figure 2.3.



Figure 2.3: Tahano Hikamu, ‘angular style’

The greatest difference to the ‘book’ style is that many of the main strokes double to become a thick and a parallel thin line and the shape of 𐌺 ⟨na⟩ changes to have its swirl straightened into a simple descending line. The vowel carrier 𐌰 changes to a flattened *O*-like circle, and the bottom curl in 𐌹 ⟨ta⟩ changes to a wedge. While the right side of the 𐌱 ⟨sa⟩ character in the ‘book style’ consists of two strokes—a flag and a downwards bow, both independently attached to the main stem—they connect here to form an *R*-like shape.

Neatly reproducing the shapes of either the ‘book’ style or the ‘angular’ style by hand goes rather slowly, so I was wondering what daily handwriting could look like. Of course, this presupposes pen and paper again; Salomon (1996: 377) mentions that inscriptions of Brāhmī and related scripts have been found on copper plates and plates made of other metals, besides stone. Metal plates can be inscribed with metal styluses and should allow similar shapes as modern pens. Wax tablets—a staple in European literacy until the use of paper became widespread—should as

<sup>7</sup> For examples, see Everson (2008), or *Wikipedia*.

well allow for relative freedom of stroke direction, so character shapes are probably not implausible even without assuming that pen and paper are (widely) available. Figure 2.4 shows what Tahano Hikamu might look like quickly jotted down by hand.



Figure 2.4: Tahano Hikamu, ‘hand style’

Many letter shapes become simplified, specifically ခ (ba), ခ (ga), ခ (ka), ခ (na), ခ (nga), the vowel carrier ခ, and the vowel ဝ (i). Not shown here is the vowel length diacritic, ခ, which is simplified to a reverse C shape. The abbreviation ခ *nay* ‘and’ is used throughout, though in a shape that is more similar to its ‘angular’ form ခ. ခ (na) is also taken from the ‘angular’ style ခ, which opens the possibility that this is actually the basic shape rather than the ‘book’ style’s ခ, or both are different developments from a shared ancestor.

Most recently, I also wondered what Tahano Hikamu might look like if it were adapted to European blackletter style with its characteristic broken arches. This, of course, constitutes a sharp contrast to Ayer’s usual look and feel, which made the experiment all the more interesting, though decidedly non-‘canonic’. Figure 2.5 shows what our example passage might have looked like at a time when Gothic book hands flourished.



Figure 2.5: Tahano Hikamu, ‘blackletter style’

The letter shapes from the ‘book’ style stay largely intact here, though all curves are broken up into at least two strokes, and strokes from the bottom right to the top left, which push a quill in a way that may cause ink to splatter from scratching, are avoided completely. The characters that differ most are ခ (ga), ခ (ra), ခ (nga), and the vowel carrier ခ. ခ (na) again appears in the ‘angular’ shape, though without its descender word-internally and in the abbreviation ခ *nay*. ခ (ta) comes with a horizontal stroke instead of a curl at the bottom; ခ (sa) gains a descender, as does ခ (ra). Not shown here either are changes to the ‘large’ diacritics.



## 3 Morphological typology

---

The first chapter dealt with the smallest constituent parts of words—speech sounds, which ones there are, and how they assemble into valid words. Consequently, the following two chapters will be about the next step up from this: morphemes, the atoms of meaning. First, we will have a more general look at which kinds of morphemes there are, and then look at them more closely by part of speech: what is their distribution, and how are morphemes put together to form inflected words? This chapter on morphological typology will first deal with general questions about Ayeri’s degree of synthesis, and then will try to answer questions about the kinds of functions the various morpheme classes carry out in the language.

### 3.1 Typology

For the largest part, Ayeri is an *agglutinative* language. Comrie (1989) says of agglutinating languages that in these, typically,

a word may consist of more than one morpheme, but the boundaries between morphemes in the word are always clear-cut; moreover, a given morpheme has at least a reasonably invariant shape, so that the identification of morphemes in terms of their phonetic shape is also straightforward. [...] As is suggested by the term agglutinating (cf. Latin *gluten* ‘glue’), it is as if the various affixes were just glued on one after the other (or one before the other, with prefixes). (43–44)

In Ayeri, root morphemes are modified by affixes for the purposes of inflection and derivation, and these affixes, in the form of suffixes more specifically, can be stacked, especially on verbs. Indeed, they vary little, so that they are always easily recognizable. Suffixation in Ayeri is especially prominent on verbs:

- (1) *Le kondasayāng hemaye pruyya nay napayya kayvay.*  
le kond-asa=yāng hema-ye-Ø pruy-ya nay napay-ya kayvay  
PT.INAN eat-HAB=3SG.M.A egg-PL-TOP salt-LOC and pepper-LOC without

‘He always eats his eggs without salt and pepper.’

The verb root  $\text{ႤႬ}$ : *kond-* ‘eat’ is inflected here for a habitual action with the suffix  $\text{ႤႬ}$  *-asa*, and also carries a pronominal clitic,  $\text{ႤႬ}$  *-yāng*, marking a third person singular masculine agent. With the notable exception of pronouns and related pronominal clitics, affixes tend to encode a single grammatical function. Verbs are not the only part of speech that can inflect; nouns, adjectives, and the relativizing conjunction can as well:

- (2) a. *Ang mətabanay tamanyeley yeyam.*  
 ang mə-tahan=ay.Ø taman-ye-ley yeyam.  
 AT PST-write=1SG.TOP letter-PL-P.INAN 3SG.F.DAT  
 ‘I wrote letters to her.’
- b. *Ang koronya Kaman apyanas palay-eng.*  
 ang koron-ya Kaman apyan-as palay-eng  
 AT know-3SG.M Kaman joke-P funny-COMP  
 ‘Kaman knows a funnier joke.’
- c. *Le turayāng taman sinā ang ningay tamala vās.*  
 le tura=yāng taman-Ø si-Ø-na ang ning=ay.Ø tamala vās  
 PT.INAN send=3SG.M.A letter-TOP REL-PT.INAN-GEN AT tell=1SG.TOP yesterday 2SG.P  
 ‘The letter which I told you about yesterday, he sent it.’

The principle of not conflating several grammatical functions into a single suffix can be observed in (2a) regarding the word  $\text{ႤႬ}$  *tamanyeley* ‘letters’, in which the plural marker  $\text{ႤႬ}$  *-ye* is distinct from the inanimate-patient case marker  $\text{ႤႬ}$  *-ley* (the latter, however, conflates animacy and case). Strictly speaking, the pronoun  $\text{ႤႬ}$  *yeyam* ‘to her’ is also composed, namely of the third person feminine base form  $\text{ႤႬ}$  *ye* and the dative case marker  $\text{ႤႬ}$  *yam*. Example (2c) is one we have already encountered before (p. 11). Here, the relative pronoun,  $\text{ႤႬ}$  *sinā* ‘of/about which’ is inflected for genitive case, and stress on the usually unstressed last syllable suprasegmentally marks that this form is contracted from  $\text{ႤႬ}$  *sileyena* (*si-ley-ena*, REL-P.INAN-GEN).

So far, we have concentrated on suffixes, but there are a number of prefixes as well; (2a) exhibits the past prefix  $\text{ႤႬ}$  *mə-* (which is actually redundant in this case). There are also demonstrative prefixes on nouns, however. In the following example, the prefix  $\text{ႤႬ}$  *eda-* ‘this-’ joins the noun  $\text{ႤႬ}$  *peham* ‘carpet’ to indicate a specific carpet.

- (3) *Le no intoyyang eda-peham.*  
 le no int-oy=yang eda=peham-Ø  
 PT.INAN want buy-NEG=1SG.A this=carpet-TOP  
 ‘I do not want to buy this carpet.’



Besides prefixes and suffixes, Ayeri also possesses at least one element in both the verb cluster and cooccurring with adpositions which straddles the border between inflection and a function word. This is the case with the clitic marker  $\text{ᄇᆞᆯ}$  *manga*, which is treated as an independent word, but can modify verbs and adpositions—heads of verb phrases (VPs) and prepositional phrases (PPs), respectively—is unstressed and appears at the margin of its modification target:

- (4) a. *Ang manga yavaya ayon bariley.*  
 ang manga yava-ya ayon-Ø bari-ley  
 AT PROG roast-3SG.M man-TOP meat-P.INAN  
 ‘The man is roasting meat.’
- b. *Ya mətapyyāng maritay misley manga luga bari.*  
 ya mə-tapy=yāng maritay mis-ley manga luga bari-Ø  
 LOCT PST-put=3SG.M.A before spit-P.INAN DIR between meat-TOP  
 ‘The meat, he had put a spit through it before.’

In (4a),  $\text{ᄇᆞᆯ}$  *manga* modifies the verb  $\text{ᄇᆞᆯ}$  *yava-* ‘roast’ and indicates that this is a temporarily ongoing action, like the English progressive, except not as strongly grammaticalized.<sup>1</sup> In (4b),  $\text{ᄇᆞᆯ}$  *manga* modifies the preposition, on the other hand, to indicate that it is directional:  $\text{ᄇᆞᆯ}$  *luga* by itself means ‘among, between’, while its directional form  $\text{ᄇᆞᆯ}$   $\text{ᄇᆞᆯ}$  *manga luga* means ‘through; during, for’.

As we have seen in the examples above, person suffixes on verbs are single morphemes that encode more than one property, for example  $\text{ᄇᆞᆯ}$  *-yeng* encodes the person features third person, feminine, singular, and agent. Personal pronouns, of which the person clitics on verbs are an instance, are the main case of fusion among agglutination in Ayeri, although some of the forms, like  $\text{ᄇᆞᆯ}$  *yeyam* ‘to her’ above, can be decomposed into root and suffix without problem.<sup>2</sup>

Perpendicular to the axis isolation–agglutination runs the axis analytic–synthetic. On the latter axis, Ayeri scores mostly as *synthetic*, since it prefers compactness over spreading a construction over several words, though it does not incorporate object noun phrases (NPs) and it is not possible to form ‘sentence-words’ either, so

<sup>1</sup> I suppose, a better parallel is the so-called *rheinische Verlaufsform* ‘Riparian progressive’ (*sein* ‘be’ + *am/beim* ‘at the’ + infinitive) in German, a construction common in the colloquial language which parallels the English progressive construction and is not yet fully grammaticalized (Eisenberg et al. 2016: 435). Speakers will thus accept both *Er lernt gerade*, literally ‘He studies right now’, and *Er ist am lernen* ‘He is studying’.

<sup>2</sup> Originally, Ayeri’s personal pronouns were indeed agglutinative as well, so  $\text{ᄇᆞᆯ}$  *yeng* ‘she’ used to be  $\text{ᄇᆞᆯ}$  *iyegang* (*iy-e-ang*, 3SG-F-A). This also gives an explanation to Boga et al. (2016)’s observation that Ayeri’s plural pronouns are formed “[v]ielleicht sogar zu regelmäßig” ([15]; ‘possibly in an even too regular way’).

it is not going so far as to be polysynthetic (Comrie 1989: 45–46). It is nonetheless theoretically possible, due to suffixation being a prominent pattern, to form foot-long words like

- (5) *da-mətabasongoyyang-ikan*  
 da=mə-taha-asa-ong-oy=yang=ikan  
 such=PST-have-HAB-IRR-NEG=ISG.A=much  
 ‘I would not much used to have had such’

Cases of analytic morphology are compound prepositions as we have seen with *ᄇᆞᆯᆫᆯᆯᆫ manga luga* ‘through’ in (4b), but verbs as well show analytic structures not only with the progressive marker, but also with modals:

- (6) *Ming saboyyang dabas.*  
 ming saha-oy=yang dabas  
 can come-NEG=ISG.A today  
 ‘I can’t come today.’

Most of the information the VP contains in this example is marked on the content verb, *ᄇᆞᆯᆫᆯᆯᆫ saba-* ‘come’, except for ability, which is expressed by the particle *ᄇᆞᆯᆫᆯᆯᆫ ming* ‘can’. *ᄇᆞᆯᆫᆯᆯᆫ ming* is an uninflected form of the verb expressing ability and we might count them as auxiliary verbs in that the full semantic content of the VP is spread out over two verb forms, one major, one minor—this probably should not be understood as a serial verb construction, however (Aikhenvald 2006).<sup>3</sup> As we will see later (section 3.2.5), though, these modal particles behave more like clitics than function words. Consider, on the other hand, the following example in which *ᄇᆞᆯᆫᆯᆯᆫ ming* is inflected like a regular verb:

- (7) *Da-mingya ang Diyan.*  
 da=ming-ya ang Diyan.  
 so=can-3SG.M A Diyan  
 ‘Diyan can (do it).’

## 3.2 Morphological processes

### 3.2.1 Prefixation

Prefixes in Ayeri apply mainly to verbs, but nouns, pronouns, adjectives and conjunctions as well can appear with them, some of which are likely clitics; reasons

<sup>3</sup> *ᄇᆞᆯᆫᆯᆯᆫ manga* has, in fact, a verbal counterpart *ᄇᆞᆯᆫᆯᆯᆫ manga-* ‘move; remove’ as well, which presumably served as the origin of both the progressive and the directional marker.

for their being clitics will be discussed below in section 3.2.5. With verbs, prefixes that are most certainly ‘true’ prefixes—that is, bound morphemes that have been semantically bleached by grammaticalization to the point where they only express grammatical functions (Lehmann 2015: 157 ff.) and which subcategorize words rather than phrases (Klavans 1985: 117), with a rather high obligation to be marked on every conjunct in coordination (Spencer and Luís 2012: 139)—are the tense prefixes marking both three degrees of past and future tense, for example:

- (8) *Ang səsarāyn ya Makapetang.*  
 ang sə-sara=ayn.Ø ya Makapetang  
 AT FUT-go=IPL.TOP LOC Makapetang  
 ‘We will go to Makapetang.’

Here, the prefix ꠘ: sə- marks future tense on the verb, ꠘꠘ: sara- ‘go’. The other tense prefixes are ꠘꠘ: kə- (NPST), ꠘ: mə- (PST), ꠘꠘ: və- (RPST), and ꠘ: pa- (NFUT) and ꠘꠘ: ni- (RFUT). Besides this set of prefixes, there are also a number of proclitics that can appear with verbs, though not exclusively. These are the anaphora ꠘꠘ: da- ‘thus, so, such’ and the reflexive marker ꠘꠘꠘ: sitang- ‘self’:

- (9) *Da-mingya ang Diyan.*  
 da=ming-ya ang Diyan.  
 so=can=3SG.M A Diyan  
 ‘Diyan can (do it).’

- (10) *Sitang-kecāng.*  
 sitang=ket=yāng  
 self=wash=3SG.M.A  
 ‘He washes himself.’

ꠘꠘꠘ: sitang- can also be used as a preverbal particle in situations where the agent is also the instrument, so both of the following two sentences are equivalent in meaning:

- (11) a. *Sa apicāng nanga ikan sitang-yari.*  
 sa apit=yāng nanga ikan sitang=yari  
 PT clean=3SG.A house complete self=3SG.M.INS  
 ‘He cleaned the whole house by himself.’  
 b. *Sa sitang-apicāng nanga ikan.*  
 sa sitang=apit=yāng nanga ikan  
 PT self-clean=3SG.A house complete  
 (idem)

Example (11a) shows the more common application of *sitang-*, that is, as a reflexive modifier of pronouns. The prefix *da-* can as well be used with noun phrases and is part of the demonstrative set of prefixes (which behave, in fact, like proclitics), *da-* ‘such’, *eda-* ‘this’, and *ada-* ‘that’:

- (12) *eda-* / *ada-* / *da-ganang*  
*eda=* / *ada=* / *da=gan-ang*  
 this= / that= / such.a=child-A  
 ‘this/that/such a child’

The demonstrative prefixes are also used to form the demonstrative pronouns *edanya* ‘this one’, *adanya* ‘that one’ and *danya* ‘such one’. A special case in this regard is the postposition *da-nārya* ‘in spite of, despite’ where *da-* combines with the conjunction *nārya* ‘but, although, except’. Originally, *dikapisa* ‘respective’ is derived from *da-* + *ikapisa* ‘bound, dependent’, which is an example of a combination with an adjective. There is also a fixed adverbial expression using one of these prefixes, *eda-tadayyam* ‘for the time being, for now’ (this=time-DAT).

Last but not least, the prefix *ku-* ‘like, as though’ (also a proclitic) can be used with both adjectives and nouns, as well as complement phrases:

- (13) a. *ku-koyaya*  
           *ku=koya-ya*  
           like=book-LOC  
           ‘like in a book’  
       b. *ku-prasi*  
           *ku=prasi*  
           like=sour  
           ‘as though (it were) sour’  
       c. *ku-adareng*       *turavangas*  
           *ku=ada-reng*     *turavang-as*  
           like=that-A.INAN problem-P  
           ‘as though that were a problem’

An example of a set-phrase adverbial consisting of *ku-* and a verb is *ku-nasya* ‘as follows’, *nasya-* meaning ‘follow’. What is curious here is that this fossilized form is lacking person marking and is just extended with an epenthetic *-a* since *-sy* is not a permissible coda. The expected form would be *\*ku-nasyareng* (like-follow=3SG.INAN.A).

Following Klavans (1985), who suggests that clitics best be defined as “affixation at the phrasal level” (117), a very common kind of prefix to the verb *phrase* are the

topic markers. They are counted as parts of the VP but do not interact with it regarding stress assignment (they are always unstressed) while always being in an initial position, preceding any other preverbal elements:

- (14) a. *Ang tabanya tamanley.*  
 ang tahan-ya taman-ley  
 AT write-3SG.M letter-P.INAN  
 ‘He writes a letter.’  
 b. *Ang mətabanya tamanley.* ‘He wrote a letter.’  
 c. *Ang manga mətabanya tamanley.* ‘He was writing a letter.’  
 d. *Ang manga no mətabanya tamanley.* ‘He was wanting to write a letter.’

### 3.2.2 Suffixation

As a largely agglutinative language, most grammatical marking in Ayeri is done by means of suffixes. These occur mainly with nouns and verbs, however, some basic quantifiers take the shape of suffixes as well, but behave more like enclitics. Quantifiers may modify content words almost regardless of their part of speech—noun, verb, adjective or adverb. The most pervasive examples of suffixation are certainly those of case marking on nouns and of person marking on verbs, for example:

- (15) *Sa pəharuyang va manga miday tangya vana suyareri, vimyon!*  
 sa pə-haru=yang va.Ø manga miday tang-ya vana suyar-eri vimyon  
 PT NFUT-beat=1SG.A 2SG.TOP DIR around ears-LOC 2SG.GEN ladle-INS monkey  
 ‘I’ll beat you around your ears with a ladle, you monkey!’

This example shows marking of 𐌵𐌹𐌶𐌹 *tang* ‘ears’ with the locative case suffix 𐌶𐌹𐌶𐌹 -*ya* and the marking of 𐌸𐌹𐌶𐌹 *suyar* ‘ladle’ with the instrumental case suffix 𐌸𐌹𐌶𐌹𐌶𐌹 -*eri*; the previous examples already provide instances of the exceedingly common markers for agent and patient case, 𐌵𐌹𐌶𐌹 -*ang* and 𐌵𐌹𐌶𐌹 -*as*, respectively. Besides case, nouns can also be marked for plural with the suffix 𐌶𐌹𐌶𐌹 -*ye*, and verb roots may be extended by the mood and aspect markers 𐌵𐌹𐌶𐌹 -*ong* (IRR), 𐌵𐌹𐌶𐌹 -*asa* (HAB) and 𐌵𐌹𐌶𐌹 -*oy* (NEG), the last of which is the most frequently occurring one. The mood suffixes can also be stacked, leading to the long word in (5) above. Person marking on verbs is realized as agreement suffixes or of clitic personal pronouns depending on whether an agent NP proper is present or not for the verb to agree with. In (15), a cliticized agent pronoun 𐌵𐌹𐌶𐌹 -*yāng* ‘he’ (3SG.M.A) appears.

As mentioned above, quantifiers appear as enclitics on almost any type of content word, like on the adverb 𐌵𐌹𐌶𐌹 *para* ‘fast’ in the following example:

- (16) *Tigalyeng para-ma.*  
 tical=yeng para=ma  
 swim=3SG.F.A fast=enough  
 ‘She swims fast enough.’

### 3.2.3 Reduplication

There are two patterns of reduplication for verbs, one with complete reduplication of the imperative form to create a hortative statement (17a), and one with partial reduplication as a way to express that an action takes place again, that is, partial reduplication expresses a iterative, as it were (17b). The imperative iterative, then, has a hortative function as well (17c):

- (17) a. *naru-naru*  
 naru~nara-u  
 HORT~speak-IMP  
 ‘let us speak’  
 b. *na-narayeng*  
 na~nara=yeng  
 ITER~speak=3SG.F.A  
 ‘she speaks again’  
 c. *na-naru*  
 na~nara-u  
 ITER~speak-IMP  
 ‘let us speak again’

With nouns, full reduplication is used to create a diminutive form (18a), though some reduplications are also lexicalized and may use roots from other parts of speech as well to form nouns, for instance, the words in (18b–d). There are also a number of adjectives for which there exists a lexical reduplication with an intensifying meaning; (19) lists a few examples. This, however, is not a productive derivation strategy.

- (18) a. *veney* ‘dog’ → *veney-veney* ‘little dog, doggie’  
 b. *gan* ‘child’ → *gan-gan* ‘grandchild’  
 c. *kusang* ‘double (adj.)’ → *kusang-kusang* ‘model’  
 d. *veh-* ‘build’ → *veha-veha* ‘tinkering’  
 (19) a. *apan* ‘wide’ → *apan-apan* ‘extensive’  
 b. *kebay* ‘alone’ → *kebay-kebay* ‘all alone’  
 c. *pisu* ‘tired’ → *pisu-pisu* ‘exhausting’

### 3.2.4 Suprasegmental modification

As written above (section 1.1.2), case agreement on a complex-marked relative pronoun can drop out under certain circumstances and is replaced by compensatory stress on the secondary case marker, which lengthens the syllable's nucleus vowel:

- (20) ... *tamanley sinā* (\**sina*) *ang ningay tamala vās*  
 ... [taman-ley]<sub>1</sub> si-Ø<sub>1</sub>-na (\*si-na<sub>1</sub>) ang ning=ay.Ø tamala vās  
 ... letter-P.INAN REL-PT.INAN-GEN (\*REL-GEN) AT tell=1SG.TOP yesterday 2SG.P  
 '... the letter which (\*whose) I told you about yesterday'

This can be reinterpreted so that vowel length/stress itself is what signifies the agreement of the relativizer with the preceding NP. Which grammatical role the relativizer's head figures as an agreement controller is essentially underspecified, hence I will gloss it as -AGR in the following example instead of as full -P.INAN:

- (21) /si.lei.'ena/ → /si.'na(:)/  
 /si-lei-ena/ → /si-'-na(:)/  
 REL-P.INAN-GEN REL-AGR-GEN

Since  $\text{na}$  as a light syllable cannot be stressed in word-final position under normal circumstances, it has to lengthen to  $\text{nā}$ .

### 3.2.5 Clitics

I have been using the term 'clitic' above and claimed that the one or the other morpheme in Ayeri is a clitic. Clitics, however, cannot be easily defined in a formal way, as it appears (Spencer and Luís 2012: 126). Based on Spencer and Luís (2012), with recourse to Zwicky and Pullum (1983), some typical characteristic traits are:

- Clitics behave in part like function words and in part like affixes, but in any case they are not free morphemes (Spencer and Luís 2012: 38, 42).
- Clitics tend to be phonologically weak items (39).
- Clitics prominently—and importantly—tend to attach 'promiscuously' to surrounding words. That is, unlike inflection, they are not limited to connect to a certain part of speech or to align with their host in semantics (40, 108–109).
- Clitics tend to appear in a second position, whether that is after a word, or an intonational or syntactic phrase (41).
- Clitics tend to be templatic and to cluster, especially if they encode inflection-like information (41, 47–48).
- Clitics have none of the freedom of ordering found in independent words and phrases (43).
- Positions of 'special' clitics tend to not be available to free words (44).

- Clitics tend to be functional morphemes, and to realize a single morphosyntactic property (Spencer and Luís 2012: 67, 179).
- There are no paradigmatic gaps (108–109).
- There tends to be no morphophonemic alteration like vowel harmony, stress shift or sandhi between a clitic and its host (108–109).
- There tends to be no idiosyncratic change in meaning when a clitic and a clitic host come together, unlike there may be with inflection (108, 110).
- Similar to affixes, clitics and their host tend to be treated as a syntactic unit, that is, lexical integrity prevents that word material can be put in between a clitic and its host (108, 110).
- Clitics usually get joined to a host word after inflection (108, 110).
- Affixes tend to go on every word in a conjunct (narrow scope), while clitics have a tendency to treat a conjunct as a unit to attach to (wide scope; 139, 196 ff.).

However, Spencer and Luís (2012) point out many counterexamples to this list in order to highlight that the border between clitics and affixes is often fuzzy. Given this fuzziness, it comes as no surprise that according to their assessment, there is a lot of miscategorization in individual grammars as a result (107). Another consequence of this lack of a clear delineation between clitics and affixes is that, since not all of the traits described above are always present, making a checklist and summing up the tally is only of limited value. The traits listed above are thus sufficient conditions only, not necessary ones. In the following, I want to elaborate on the classification of various prefixes, suffixes, and particles as clitics.<sup>4</sup>

#### *Preposed particles and prefixes*

What should be rather unproblematic with regards to classification as clitics in Ayeri are the preverbal particles, that is, the topic marker, one or several modal particles, the progressive marker, and also the emphatic affirmative and negative discourse particles. All of these particles essentially have functional rather than lexical content, and are usually unstressed. They come in a cluster with a fixed order, and they appear in a position where no ordinary word material could go,

<sup>4</sup> The following discussion incorporates most of the content of a blog article I previously wrote on this topic, Becker (2017), with some additions and corrections. Since clitics sit at the junction of morphology and syntax, it will be necessary at times to deal with topics roughly which will be elaborated on later in more detail.



since Ayeri is strictly verb-initial.<sup>5</sup> In conjuncts it is also unnecessary to mark every verb with one or several preverbal particles:

- (22) a. *Ang kece            nay dayungisaye   māva            yanjas   yena.*  
 ang ket-ye        nay dayungisa-ye   māva-Ø        yan-ye-as   yena  
 AT wash-3SG.F and dress-3SG.F mother-TOP boy-PL-P 3SG.F.GEN  
 ‘The mother washes and dresses her boys.’
- b. *Manga sahaya        rangya        nay nedraya   ang Tikim.*  
 manga saha-ya        rang-ya        nay nedra-ya   ang Tikim  
 PROG come-3SG.M home-LOC and sit-3SG.M A Tikim  
 ‘Tikim is coming home and sitting down.’
- c. *Ang mya            ming sidegongya            nay la-lataya*  
 ang mya            ming sideg-ong=ya.Ø        nay la~lata=ya.Ø  
 AT be.supposed.to can repair-IRR=3SG.M.TOP and ITER~sell=3SG.M.TOP  
*ajamyeley.*  
 ajam-ye-ley  
 toy-PL-P.INAN  
 ‘He should be able to repair and resell the toys.’

In (22a), therefore, the agent-topic marker *ang* only occurs before *kece* ‘(she) washes’, and the conjoined verb *dayungisaye* ‘(she) dresses’ is also within its scope. Repeating the marker as well before the latter verb could either be considered ungrammatical because there is only one topic there—*māva* ‘mother’—or the sentence could be interpreted as having two conjoined clauses with different subjects: ‘[She washes] and [mother dresses] her boys.’ The latter outcome has *māva* as the topic only of *dayungisaye*, while *kece*’s topic is the person marking on the verb—a pro-drop subject, essentially.<sup>6</sup>

In (22b), then, the progressive marker *manga* equally has scope over both verb conjuncts, *sahaya* ‘(he) comes’ and *nedraya* ‘(he) sits’ in what is presumably a case of extended/distributed exponence. This is to say that functionally contiguous information can sometimes be split over several words, so that the functional annotation of each verb in (22b) can be represented in the fashion of the (incomplete) f-structure matrix (see Bresnan et al. 2016; Butt and King 2015) shown in (23), which is an attempt to represent the phrase *ang manga saha-* ‘is coming’ formally. *manga* is treated there as being part of things the verb inflects for, that is, progressive aspect, in spite of appearing superficially as a

<sup>5</sup> The translation of ‘Ozymandias’ in section B.3 deviates from this rule in *ḡamāng sam kāryo nay taryankay bengyon adābalya*. ‘Two big and torsoless legs stand in that desert’ by having the subject NP *ḡamāng sam* ... ‘two legs ...’ precede the verb *bengyon* ‘(they) stand’. This is non-standard syntax in a poetic text.

<sup>6</sup> This claim is further investigated below, p. 89 ff.; also compare section 4.5.1.

function word. The topic marker ၵၵၵ *ang* does not reflect a morphological property of the verb in the way the progressive marker does, but announces the case and—for agents and patients—the animacy value of the topicalized noun phrase (NP), so the f-structure in (23) lists this information under the TOP relation.<sup>7</sup>

(23)	<table> <tr> <td>PRED</td><td>'come &lt;((↑ SUBJ), (↑ OBL<sub>loc</sub>))&gt;'</td></tr> <tr> <td>ASP</td><td>PROG</td></tr> <tr> <td>TOP</td><td> <table> <tr> <td>CASE</td><td>A</td></tr> <tr> <td>ANIM</td><td>+</td></tr> <tr> <td>...</td><td>...</td></tr> </table> </td></tr> <tr> <td>...</td><td>...</td></tr> </table>	PRED	'come <((↑ SUBJ), (↑ OBL <sub>loc</sub> ))>'	ASP	PROG	TOP	<table> <tr> <td>CASE</td><td>A</td></tr> <tr> <td>ANIM</td><td>+</td></tr> <tr> <td>...</td><td>...</td></tr> </table>	CASE	A	ANIM	+	...	...	...	...
PRED	'come <((↑ SUBJ), (↑ OBL <sub>loc</sub> ))>'														
ASP	PROG														
TOP	<table> <tr> <td>CASE</td><td>A</td></tr> <tr> <td>ANIM</td><td>+</td></tr> <tr> <td>...</td><td>...</td></tr> </table>	CASE	A	ANIM	+	...	...								
CASE	A														
ANIM	+														
...	...														
...	...														

Modal particles, exemplified in (22c), are probably slightly less typical as clitics since it seems feasible for them to be stressed for contrast. What is not possible, however, is to front either ၵၵ *mya* 'be supposed to' or ၵၵၵ *ming* 'can', and the verb itself also cannot precede the particles, which is demonstrated in (24). It is also not possible to coordinate any of the elements in the preverbal particle cluster with ၵၵ *nay* 'and', as shown in (25).

- (24) a. \**mya ang ming sidegongya*  
 b. \**ming ang mya sidegongya*  
 c. \**sidegongya ang mya ming*
- (25) a. \**ang nay mya ming sidegongya*  
 b. \**ang mya nay ming sidegongya*  
 c. \**ang mya ming nay sidegongya*

It needs to be pointed out that unlike verbs, modal particles in Ayeri resist inflection, so in (22c) the irrealis suffix ၵၵၵ *-ong* is realized on the verb ၵၵၵ *sidegongya* '(he) would repair' instead of on one or both of the modal particles as \**ၵၵၵၵ* \**mingong* and \**ၵၵၵၵ* *myong*, respectively. The combination of ၵၵ *mya* 'be supposed to' with an irrealis-marked verb together indicates that the speaker thinks the action denoted by the verb *should* be carried out. The marking on the verb may then be interpreted as distributing to the constituent parts of whole verb complex. The same goes for negation: only the verb can be negated, but not the modal particle. Possibly, it would be useful in this case to abstract the modal particles as a

<sup>7</sup> In the chart, angular brackets group grammatical functions together. Since the verb is basically the head of the clause, the first PRED (predicator) lists the verb with its argument structure (a-structure). In the case of (23), (↑ SUBJ) and (↑ OBL<sub>loc</sub>) indicate that the verb, 'come', governs two arguments: one syntactic subject and one oblique argument in the form of a locative adverbial, which have been omitted for brevity in the chart.

feature [MODALITY] as listed by ParGram (2009–2016: Feature Table) for purposes of functional representation. At least superficially, it looks as though Ayeri acts differently from English here in that the content verb is possibly not a complement of the modal element. This assumption is supported by the fact that in Ayeri, the verb inflects, not the modal particle. Furthermore, modal particles cannot be modified by adverbs in the way regular verbs can:

- (26) a. *Ming tigalye ban nilay ang Diya.*  
 ming tigal-ye ban nilay ang Diya  
 can swim-3SG.F good probably A Diya  
 ‘Diya can probably swim well.’
- b. \**Ming nilay tigalye ban ang Diya.*  
 ming nilay tigal-ye ban ang Diya  
 can probably swim-3SG.F well A Diya

Combinations of topic particle and modal particle, as well as modal particle and verb, can likewise not be interrupted by parenthetical material like *naratang* ‘they say’, so that:

- (27) a. *Naratang, ang ming tigalye ban Diya kodanya.*  
 nara=tang ang ming tigal-ye ban Ø Diya kodan-ya  
 say=3PL.M.A AT can swim-3SG.F well TOP Diya lake-LOC  
 ‘They say Diya can swim well in a lake.’
- b. \**Ang naratang, ming tigalye ban Diya kodanya.*
- c. \**Ang ming, naratang, tigalye ban Diya kodanya.*
- d. ?*Ang ming tigalye, naratang, ban Diya kodanya.*
- e. *Ang ming tigalye ban, naratang, Diya kodanya.*
- f. *Ang ming tigalye ban Diya, naratang, kodanya.*
- g. *Ang ming tigalye ban Diya kodanya, naratang.*

Besides verbs, nouns as well have preposed modifiers. This is the case with proper nouns specifically, where the name is preceded by a case particle instead of receiving a case-marking suffix like generic nouns do. This case marker is phonologically weak in that its phonological make-up is similar to that of affixes, and unstressed, with the exception of the causative case marker *sā*, which bears at least secondary stress since it contains a long vowel. We already saw case particles preceding names in (22b) and (26) above: *ang Tikim* and *ang Diya*; *ang* marks the proper-noun NPs as agents in both cases. The case marker is missing when the NP is topicalized, as indicated in (27), where the agent NP appears as just *Diya*, not *ang Diya*. While case suffixes have narrow scope as shown in (28a) and thus need to be repeated on every NP in a conjunct, preposed case

(28) a. *Toryon veneyang nay badanang.*  
tor-yon veneyang nay badan-ang  
sleep-3PL.N dog-A and father-A  
  
'The dog and father are (both) sleeping.'

b. \**Toryon veney nay badanang.*  
tor-yon veney\_ nay badan-ang  
sleep-3PL.N dog\_ and father-A

c. *Sa sobisayan ang Niva nay Mico narānye.*  
sa sobisa-yan ang Niva nay \_ Mico narān-ye-Ø  
PT study-3PL.M A Niva and \_ Mico language-PL-TOP  
  
'Languages is what Niva and Mico study.'

d. *Sa sobisayan ang Niva nay ang Mico narānye.*  
sa sobisa-yan ang Niva nay ang Mico narān-ye-Ø  
PT study-3PL.M A Niva and A Mico language-PL-TOP  
  
'Languages is what Niva and Mico (each) study.'

Taking the above characteristics into account—inability to insert word material, special positioning, and wide scope—one may argue that the preposed case markers are clitics. It should be noted furthermore that a single NP cannot be marked for two grammatical functions at the same time, so that case markers cannot be coordinated, as is attempted in (29a) below with \**sa nay s̄a Sopan*:

- (29) a. \**Ang delacan* *sa nay sã Sopan.*  
ang delak=yan.Ø *sa nay sã Sopan*  
AT suffer.from=3PL.M.TOP P and CAUS Sopan  
*Intended: ‘They suffer from and due to Sopan.’*
- b. *Ang delacan* *sa Sopan, nay yasa.*  
ang delak=yan.Ø *sa Sopan nay yasa*  
AT suffer.from=3PL.M.TOP P Sopan and 3SG.M.CAUS  
*‘They suffer from Sopan, and due to him.’*

The case markers of proper nouns are necessarily proclitics rather than enclitics to preceding word material, since it is possible for them to begin utterances, where it is not possible to lean to the left, but only to the right. This is the case in equative sentences such as the one in (30a). In these cases as well, it is not possible for parenthetical material to be placed between the case marker and its target of modification, as in (30b); the particle and its head cohere closely and behave essentially like a unit.

- (30) a. *Ang Misan lajāyas puti.*  
 ang Misan lajāy-as puti  
 A Misan student-P zealous  
 ‘Misan is a zealous student.’
- b. \**Ang paronyang, Misan lajāyas puti.*  
 ang paron=yang Misan lajāy-as puti  
 A believe=1SG.A Misan student-P zealous

The fact that case particles attach always to a proper noun very specifically makes them unlike ‘typical’ clitics, since according to Spencer and Luís (2012), a typical and important feature of clitics is their ‘promiscuous’ attachment, as described initially. This puts case particles closer to affixes—just like the suffixed case markers. On the other hand, as previously pointed out, clitics do not have to exhibit all traits often associated with them in order to be counted as such. Yet more typical of function words, on the other hand, is the fact that there is no morphophonemic interaction between a case particle and the word it modifies. Thus, for instance, there is no form /sa:dʒa:n/ resulting from the combination of ꨀ *sa* (P) with ꨀꨀꨀꨀ *Ajān*. This overlap in form between affix and function word is typical of clitics, according to the traits excerpted from Spencer and Luís (2012) above.

As discussed previously, ꨀꨀ *manga* may not only modify verbs, but also adpositions—which in the case of prepositions are often very transparently derived from nouns. ꨀꨀ *manga* in combination with an adposition indicates that there is motion along a path. The directional marker ꨀꨀ *manga* is thus a functional morpheme and it always appears before the adposition itself. Adpositions do not otherwise inflect, but ꨀꨀ *manga*, due to its functional nature, could reasonably be construed as inflection, in spite of appearing as a function word, just as its (related) verbal counterpart. This double nature makes it a good candidate for a clitic. Applying a shuffling or coordination test here to figure out whether ꨀꨀ *manga* is an adjunct is moot, since there is nothing else which can appear in this position—the position ꨀꨀ *manga* appears in is thus syntactically privileged; ꨀꨀ *manga* can be said to exhibit special syntax, which is further evidence for it being a clitic. With regards to the distinction between special and simple clitics (Zwicky 1977), it ought to be classified as the former, since even though it may be derived from the verb ꨀꨀ: *move*, this verb does not constitute the particle’s associated full form:

- (31) a. *Ang saraya Ajān manga kong nangaya.*  
 ang sara-ya Ajān manga kong nanga-ya  
 AT go-3SG.M Ajān DIR inside house-LOC  
 ‘Ajān goes into the house.’

- ‘Ajān goes to move inside the house.’

(32) a. *Ang saraya Ajān, narayāng, manga kong nangaya.*  
 ang sara-ya Ajān nara=yāng manga kong nanga-ya  
 AT go-3SG.M Ajān say=3SG.M.A DIR inside house-LOC

b. \*Ang saraya Ajān manga, **narayāṅ**, kong nangaya.

Also, when testing *ṁṁ manga*'s behavior in terms of distribution over coordinated NPs, we can see in (33b) that there is no problem in condensing the sentence given in (33a) to the extent that *ṁṁ manga* governs two adpositions in coordination—*ṁṁṁṁ miday* 'around' and *ṁṁṁṁ kong* 'inside'—sharing the same adpositional object, *ṁṁṁṁ nanga* 'house'.

- (33) a. *Ang saraya Ajān manga miday nangaya nay manga kong nangaya.*  
ang sara-ya Ajān manga miday nanga-ya nay manga kong nanga-ya  
AT go-3SG.M Ajān DIR around house-LOC and DIR inside house-LOC

b. *Ang saraya Ajān manga miday nay kong nangaya.*  
 ang sara-ya Ajān manga miday nay kong nanga-ya  
 AT go-3SG.M Ajān DIR around and inside house-LOC

For all intents and purposes, thus, *en* *manga* behaves syntactically like a typical clitic in that it has wide scope over conjuncts, coheres tightly with its target of modification, is located in a syntactically privileged position, and unites properties of both function words and inflection.

From this discussion of prenominal (and one pre-adpositional) particles, let us return to verbs again for a moment. Besides the preverbal particles discussed above, there is also what is spelled as a prefix on the verb which appears to be a little odd as such in that it can have wide scope over conjoined verbs. This is the prefix 𐀓 *da-* often meaning ‘so, thus’, displayed in (34).

- (34) *Ang da-pinyaya nay hisaya Yan sa Pila.*  
 ang da=pinya-ya nay hisa-ya Ø Yan sa Pila  
 AT so=ask-3SG.M and beg-3SG.M TOP Yan P Pila

‘Yan asks and begs Pila to (do so).’

𐀓 *da-*, where it is not used for presentative purposes,<sup>8</sup> is a functional morpheme in that it basically acts as an anaphora for a complementizer phrase (CP) the speaker chooses to drop. Thus, it does not mark any of the intrinsic morphological categories of the verb (tense, aspect, mood, modality, finiteness), just like the topic marker refers to a syntactic relation the verb subcategorizes but none of its proper categories of inflection. As an anaphora, 𐀓 *da-* cannot stand alone, though it is possible to use a full demonstrative form 𐀓𐀕𐀗𐀓 *danya* ‘such one’ in its place:

- (35) *Ang pinyaya nay hisaya Yan sa Pila danyaley.*  
 ang pinya-ya nay hisa-ya Ø Yan sa Pila danya-ley  
 AT ask-3SG.M and beg-3SG.M TOP Yan P Pila such.one-P.INAN

‘Yan asks and begs Pila such.’

Unlike the preverbal particles, 𐀓 *da-* can be associated with a full form, though it still displays special syntax in that unlike English *-n’t* or *’ll*, for instance, it does not occur in the same place as the full form. Note also how 𐀓 *da-* is appended to the right of tense prefixes, which *do* express a property of the verb, as shown in (36).

- (36) a. *Ang da-məpinyaya sa Pila.*  
 ang da=mə-pinya=ya.Ø sa Pila  
 AT so=PST=ask=3SG.M.TOP P Pila  
 ‘He asked Pila to.’  
 b. *Ang da-məpinyaya nay məhisaya Yan sa Pila.*  
 ang da=mə-pinya-ya nay mə-hisa-ya Ø Yan sa Pila  
 AT so=PST=ask-3SG.M and PST-beg-3SG.M TOP Yan P Pila

‘Yan asked and begged Pila to.’

<sup>8</sup> Although this use is probably related to the anaphoric use.

The verb form in (36) becomes ungrammatical with the order of its prefixes reversed, so it is not acceptable to say: မာ်ဒါပိယာ *məda-pinyaya*, although note that pre- and suffixes proper also have a fixed order in Ayeri, so this alone is probably not enough evidence to claim that မာ် *da-* is not possibly a prefix. Furthermore, while the tense prefixes undergo crasis, this is not the case with မာ် *da-*:

- (37) a. *Māmangreng.*  
           mə-amang=reng  
           PST-happen=3SG.INAN.A  
           ‘It happened.’  
       b. \**Məamangreng.*
- (38) a. *Da-amangreng.*  
           da=amang=reng  
           thus=happen=3SG.INAN.A  
           ‘It happens thus.’  
       b. \**Dāmangreng.*

Besides the characteristic of not seeking out certain parts of speech, the မာ် *da-* prefix satisfies the criteria of being a phonologically reduced form of an otherwise free functional morpheme, and it occurs in a place where normal syntax would not put its corresponding full form. It has wide scope over conjuncts, is attached outside of inflection for proper categories of the verb, and doesn’t interact with its host with regards to morphophonemics. In addition to these more typical traits of clitics, there is also no way to place words between မာ် *da-* and the verb stem:

- (39) *Da, naratang, amangreng.*  
       da  nara=tang  amang=reng  
       thus say=3PL.M.A  happen=3SG.INAN.A

The prefix မိတာ် *sitang-* ‘self’ behaves in the same way as မာ် *da-*, since it also abbreviates a reflexive NP, for instance, မိတာ်ယေ *sitang-yes* ‘herself’ where ‘herself’ as a patient is coreferential with the agent of the clause. One might assume that reflexivity is a verbal category of inflection in Ayeri, although, on the other hand, Ayeri also does not have any verbs which appear as grammatically reflexive to indicate anticausativity like in Romance languages. The reflexive marking in Ayeri is thus semantically motivated, not functionally.

- (40) a. *Adruara  biratayreng.*  
           adru-ara  biratay-reng  
           break-3SG.INAN  pot-A.INAN  
           ‘The pot broke.’



- b. \**Sitang-adruara biratayreng.*  
 sitang=adru-ara biratay-reng  
 self=break-3SG.INAN pot-A.INAN

*Intended:* ‘The pot broke.’ (an unspecified force broke it)

- (41) a. *Le pot s’est cassé.* [French]  
 le pot se=est cassé  
 the pot self=be.3SG.PRS broken

‘The pot broke.’ (an unspecified force broke it)

- b. *Le pot est cassé.*  
 le pot est cassé  
 the pot be.3SG.PRS broken

‘The pot is broken.’

Ayeri has a tendency to reuse prefixes with different parts of speech, and thus 𐏃 *da-* is also used with nouns, forming part of the series of deictic prefixes, 𐏃 *da-* ‘such (a)’, 𐏃 *eda-* ‘this’, 𐏃 *ada-* ‘that’. The prefix in all these cases represents a grammatical function, is unstressed, and may have wide scope over conjoined NPs, unless an individuating interpretation is intended, as in (42b). These traits are typical of clitics, as we have seen, though (43) shows that unlike with verbs, the deictic prefixes do undergo crasis here, which is a trait more typically associated with affixes.

- (42) a. *Sinyāng eda-ledanas nay viretāyās tondayena-ben?*  
 sinyā-ang eda=ledan-as nay viretāya-as tonday-ena=hen  
 who-A this=friend-P and supporter-P art-GEN=all  
 ‘Who is this friend and supporter of all arts?’
- b. *Sinyāng eda-ledanas nay eda-viretāyās tondayena-ben?*  
 sinyā-ang eda=ledan-as nay eda=viretāya-as tonday-ena=hen  
 sinyā-A eda=ledan-P nay eda=viretāya-P tonday-GEN=hen  
 ‘Who is/are this friend and this supporter of all arts?’

- (43) *Sa ming nelnang edāyon.*  
 sa ming nel=nang eda=ayon-Ø  
 P can help=IPL.A this=man-TOP

‘This man, we can help him.’

The deictic prefixes also cannot be used with all types of NPs, only with those headed by generic and proper nouns; the picky nature of the deictic prefixes also makes them more typical of affixes than of clitics. The preverbal particles, on the other hand, also only occur with verbs, and it was nonetheless argued for them to be

classified as clitics above due to the presence of other traits which make the particle under scrutiny clitic-like.

As mentioned initially, Spencer and Luís (2012) give numerous counterexamples to the catalog of traits typically associated with clitics. One of these counterexamples is what they call ‘suspended affixation’. This phenomenon occurs in Turkish, for instance, where the plural suffix *-ler* and subsequent suffixes can be left out in coordination (44a), as well as case markers (44b), and adverbials with case-like functions (44c):

(44) Turkish (199):

- a. *bütün kitap(...) ve defter-ler-imiz*  
all book and notebook-PL-IPL.POSS  
‘all our books and notebooks’
- b. *Vapur hem Napoli(...) hem Venedik’-e uğruyormuş*  
boat and Naples and Venice-LOC stops.EVID  
‘Apparently the boat stops at both Naples and Venice’
- c. *öğretmen-ler(...) ve öğrenci-ler-le*  
teacher-PL and student-PL-WITH  
‘with (the) students and (the) teachers’

Spencer and Luís (2012) note that, in “the nominal domain especially, wide scope inflection is widespread in the languages of Eurasia, becoming more prominent from west to east,” and that wide scope affixation “can be found with inflectional and derivational morphology in a number of languages, and it is often a symptom of recent and not quite complete morphologization” (200). They report further that Wälchli (2005) finds that suspended affixation is especially common with ‘natural coordination’, that is, the combination of items very frequently occurring in pairs like *knife and fork* or *mother and father*, as opposed to cases of occasional coordination (Spencer and Luís 2012: 200). Whether this is also true for Ayeri as of now would require a separate survey.<sup>9</sup> Ayeri is not (intended to be) of Eurasian stock, though since there is evidence for this phenomenon, it should at least be considered.

Given the evidence from Turkish, the categorization of deictic prefixes as *either* affixes *or* clitics is unclear, especially since the diagnostic of scope is devalued by the Turkish examples. On the other hand, suffixes on nouns do not behave this way in Ayeri, as demonstrated in (45)—they rather behave like typical affixes in that they mandatorily occur on each conjunct. The question is, thus, whether an exception should be made for prefixes on nouns. We may as well assume that they are clitics.

<sup>9</sup> Or rather, devising supplemental rules.

- (45) a. *sobayajang*      *nay lajāyjang*  
           *sobaya-ye-ang* *nay lajāy-ye-ang*  
           teacher-PL-A    and student-PL-A  
           ‘(the) teachers and (the) students’
- b. \**sobayaye*      *nay lajāyjang*  
           *sobaya-ye*    *nay lajāy-ye-ang*  
           teacher-PL    and student-PL-A
- c. \**sobaya*      *nay lajāyjang*  
           *sobaya*      *nay lajāy-ye-ang*  
           teacher      and student-PL-A

From a functional point of view, the exact nature of the deictic prefixes should not matter either way—ParGram (2009–2016: Feature Table) also cites a [DEIXIS] feature with PROXIMAL and DISTAL as its values, which fits *eda-* ‘this’ and *ada-* ‘that’ just fine. At present it is unclear, however, how to represent ‘such (a)’ in this respect, since it is clearly deictic, but neither proximal nor distal. In this case, it should be possible to use [DEIX this/that/such] as well, hence:

- (46) a. *edāyon*  
           *eda=ayon*  
           this=man  
           ‘this man’
- b.  $\begin{bmatrix} \text{PRED} & \text{‘man’} \\ \text{DEIX} & \text{this} \end{bmatrix}$

As described above, proper nouns are case marked by clitic case markers preceding the noun. In fact, these markers must be located somewhere at the left periphery of the NP, so the deictic prefixes stand in between the case marker and the proper noun itself, which is unproblematic for lexical integrity, since the deictic prefixes are not free morphemes. And even if they were part of inflection, the case markers, as clitics, would be on the outside—the order DEICTIC PREFIX – CASE MARKER – NOUN is ungrammatical. An example of this is given in (47).

- (47) a. *Ang koronay*      *sa eda-Kagan.*  
           *ang koron=ay.Ø*    *sa eda=Kagan*  
           AT    know=1SG.TOP    P    this=Kagan  
           ‘I know this Kagan.’
- b. \**Ang koronay*      *eda-Kaganas.*  
           *ang koron=ay.Ø*    *eda=Kagan-as*  
           AT    know=1SG.TOP    this=Kagan-P

- c. \**Ang koronay eda-sa Kagan.*  
 ang koron=ay.Ø eda=sā Kagan  
 AT know=ISG.TOP this=P Kagan

The question now is, what happens to coordinated proper nouns? Since the suffixed case markers on common nouns have the distributional properties of affixes, they occur on every conjunct, the deictic prefixes, however, only occur on the first unless an individuating reading is intended, as shown in (41). For proper nouns it ought to be possible for both a case marker and a deictic prefix to have scope over coordinated proper nouns, as in (48a). Yet, however, this seems slightly odd-sounding, so the strategy in (48b) is preferable, which avoids the problem altogether by making the names an adjunct to the demonstrative  $\text{da-}$  *edanya* ‘this/these one(s)’.<sup>10</sup> The example in (48c) is unproblematic and here as well indicates that the two persons are referred to individually and not as a group.

- (48) a. *Ang koronay sa eda-Kagan nay Ijān.*  
 ang koron=ay.Ø sa eda=Kagan nay Ijān  
 AT know=ISG.TOP P this=Kagan and Ijān  
 ‘I know these Kagan and Ijān.’
- b. *Ang koronay edanyās, Kagan nay Ijān.*  
 ang koron=ay.Ø edanya-as Kagan nay Ijān  
 AT know=ISG.TOP this.one-P Kagan and Ijān  
 ‘I know these, Kagan and Ijān.’
- c. *Ang koronay sa eda-Kagan nay eda-Ijān.*  
 ang koron=ay.Ø sa eda=Kagan nay eda=Ijān  
 AT know=ISG.TOP P this=Kagan and this=Ijān  
 ‘I know this Kagan and this Ijān.’

Of the deictic prefixes,  $\text{da-}$  is not only available to verbs and nouns, but also to adjectives. Like with verbs, it is short for  $\text{da-}$  *danya* ‘such one’ in this case, as demonstrated in (49a). The resulting meaning is ‘the ADJECTIVE one’;  $\text{da-}$  essentially acts as a nominalizer, at least to the extent that the compound of  $\text{da-}$  and an adjective inherits the distributional properties of  $\text{da-}$  *danya* as a demonstrative pronoun. Thus, it can be case- and topic- marked, as shown by (49bc). It can also be modified by another adjective, as in (49c). On the other hand, it cannot be reduplicated for diminution, and can also not be pluralized. Since adjectives follow their heads, the original order of DEMONSTRATIVE – ADJECTIVE remains intact.  $\text{da-}$

<sup>10</sup> Honestly, it is these cases where you wish it were possible to just ask a speaker of your language for their judgement instead of relying on your own intuition—which will most certainly be tainted by interference from your native language—in order not to carry over all too familiar patterns into your creation.

*da-* is thus similar in distribution to English simple clitics such as *'ll*, which occurs in the same place as its full form, the future tense auxiliary *will*.

- (49) a. *Le noyang danyaley tuvo.*  
 le no=yang danya-Ø tuvo  
 PT.INAN want=ISG.A such.one-TOP red  
 'The red one I want.'
- b. *Ang noay da-tuvoley.*  
 ang no=ay.Ø da-tuvo-ley  
 AT want=ISG.TOP one-red-P.INAN  
 'I want the red one.'
- c. *Le noyang da-tuvo kivo.*  
 le no=yang da-tuvo-Ø kivo  
 PT.INAN want=ISG.A one=red-TOP small  
 'The little red one I want.'

The prefix, again, coheres tightly in that no additional material can be inserted. Like with nouns above, inflecting each form in a group of coordinated adjectives results in an individuating reading in (50a). It should be possible for the prefix to take wide scope as in (50b), though it seems better to me to instead rephrase the coordinated adjective as a relative clause like in (49c), for instance, besides using the full form  $\text{ᄃᆞᆫ}$  *danya* + adjectives. Since case is obligatorily marked on every conjunct, (50d) is not grammatical.

- (50) a. *Ang noay da-tuvoley nay da-lenoley.*  
 ang no=ay.Ø da=tuvo-ley nay da=leno-ley  
 AT want=ISG.TOP one=red-P.INAN and one=blue-P.INAN  
 'I want the blue one and the red one.'
- b. <sup>?</sup>*Ang noay da-tuvoley nay lenoley.*  
 ang no=ay.Ø da=tuvo-ley nay leno-ley  
 AT want=ISG.TOP one=red-P.INAN and blue-P.INAN  
 'I want the red and blue one.'
- c. *Ang noay adaley si tuvo nay leno.*  
 ang no=ay.Ø ada-ley si tuvo nay leno  
 AT want=ISG.TOP that-P.INAN REL red and blue  
 'I want that which is red and blue.'
- d. <sup>\*</sup>*Ang noay da-tuvo nay lenoley.*  
 ang no=ay.Ø da=tuvo nay leno-ley  
 AT want=ISG.TOP one=red and blue-P.INAN

Possessive pronouns like  $\mathfrak{z}$  *nā* ‘my’,  $\mathfrak{r}$  *vana* ‘your’, etc. behave the same way when derived from their usual role as modifiers to free-standing anaphoras ( $\mathfrak{d}$  *da-nā* ‘mine’,  $\mathfrak{d}$  *da-vana* ‘yours’, etc.), except they cannot themselves be modified by adjectives in the way  $\mathfrak{d}$  *da-tuvo* ‘the red one’ is in (49c). Taking all of the examples above into account,  $\mathfrak{d}$  *da-* with adjectives and possessive pronouns seems to be most like a simple clitic according to Zwicky (1977)’s definition, compared to the other contexts it can appear in:

Cases where a free morpheme, when unaccented, may be phonologically subordinated to a neighboring word. Cliticization of this sort is usually associated with stylistic conditions, as in the casual speech cliticization of object pronouns in English; there are both formal *full* pronouns and casual *reduced* pronouns. (5)

Typical of a simple clitic as well, the distribution of  $\mathfrak{d}$  *da-* is restricted by grammatical context, as pointed out regarding example (48b). Unlike in English, which Zwicky (1977) gives examples of, the condition in Ayeri is likely not merely phonological in this case. The nature of the condition, however, is not predetermined in Spencer and Luís (2012), when they elaborate on Zwicky (1977)’s definition that

we may therefore need to define simple clitics along the lines of Halpern (1998), namely, as clitics that may be positioned in a subset of the positions within which the full forms are found, rather than as clitics that have the same distribution as their full-form counterparts as in Zwicky (1977). Under this broader definition, we capture the fact that simple clitics differ from special clitics in that they can appear in some of the positions that are occupied by their corresponding full forms, while special clitics never can. (Spencer and Luís 2012: 44)

Besides deictic prefixes, nouns may also receive a prefix expressing likeness,  $\mathfrak{k}$  *ku-*. This prefix is also applicable to adjectives, and is maybe more adverbial in terms of semantics than purely functional morphemes like  $\mathfrak{d}$  *da-*. In contrast to  $\mathfrak{d}$  *da-*,  $\mathfrak{k}$  *ku-* has no full-form equivalent. Some examples of it leaning on nouns are given in (51). Like the deictic prefixes,  $\mathfrak{k}$  *ku-* appears in a position which is restricted to functional morphemes. Any other modifiers which appear as free words or phrases (adjectives, relative clauses, nominal adjuncts) follow nouns and cannot appear in the position of  $\mathfrak{k}$  *ku-*. Slightly untypical of a clitic, again, it is not fully ‘promiscuous’ regarding its phonological host in that it requires a nominal, adjectival or phrasal host.

- (51) a. *Ang misya Amān ku-depangas.*  
 ang mis-ya Ø Amān ku=depang-as  
 AT act-3SG.M TOP Amān like=fool-P  
 ‘Amān acts like a fool.’
- b. *Ang misya Amān ku-depangas nay karayās.*  
 ang mis-ya Ø Amān ku-depang-as nay karaya-as  
 AT act-3SG.M TOP Amān like=fool-P and coward-P  
 ‘Amān acts like a fool and a coward.’

- c. *Ang misya Amān ku-depangas nay ku-karayās.*  
ang mis-ya Ø Amān ku=depang-as nay ku-karaya-as  
AT act-3SG.M TOP Amān like=fool-P and like=coward-P  
‘Amān acts like a fool and like a coward.’
- d. *Ang misya Amān ku-ada-depangas.*  
ang mis-ya Ø Amān ku=ada=depang-as  
AT act-3SG.M TOP Amān like=that=fool-P  
‘Amān acts like that fool.’
- e. *Ang misya Amān ku-ada-depangas nay ada-karayās.*  
ang mis-ya Ø Amān ku=ada=depang-as nay ada=karayās  
AT act-3SG.M TOP Amān like=that=fool-P and that=coward-P  
‘Amān acts like that fool and that coward.’
- f. \**Ang misya Amān ada=ku=depangas.*  
ang mis-ya Ø Amān ada=ku=depang-as  
AT act-3SG.M TOP Amān that=like=fool-P

Generally, ꠘꠞꠦ ꠕꠤ- fulfills the function of the preposition *like* in English in (51). However, if it were a preposition in Ayeri, it should trigger the locative on its dependent. In the examples above, however, the NP ꠘꠞꠦ ꠕꠤ- modifies takes the patient case, like predicative NPs are otherwise wont to do. Moreover, while prepositions like ꠘꠞꠦ ꠕꠤ ꠈꠞꠦ ꠕꠤꠇ ꠈꠞꠦ ꠕꠤꠇ ‘inside’ in (52) are free morphemes in Ayeri, ꠘꠞꠦ ꠕꠤ- is bound, which becomes apparent by introducing a parenthetical remark in (53).

- (52) a. *Ang yomāy, suprareng, kong sayanya.*  
ang yoma=ay.Ø surpa=reng kong sayan-ya  
AT be=1SG.TOP seem=3SG.INAN.A inside cave-LOC  
‘I am, it seems, inside a cave.’  
b. *Ang yomāy kong, suprareng, sayanya.*
- (53) a. *Ang misya Amān, suprareng, ku-depangas.*  
ang mis-ya Ø Amān surpa=reng ku=depang-as  
AT acts=3SG.M TOP Amān seem=3SG.INAN.A like=fool-P  
‘Amān acts, it seems, like a fool.’  
b. \**Ang misya Amān ku, suprareng, depangas.*

Examples (51ab) show that similar to the deictic prefixes,  $\text{ku-}$  precedes its target of modification and can have wide scope over coordinated NPs. As (51c) shows, narrow scope is possible as well, and in this case, again, each conjunct is to be interpreted separately instead of  $\text{ku-}$  modifying both conjuncts collectively. As illustrated in (51d),  $\text{ku-}$  even precedes  $\text{ada-}$  as a deictic prefix, for instance, if they appear together. Reversing the order of the prefixes is not possible, as is

shown in (51f). As (51e) shows,  $\frac{3}{2}$ : *ku-* may also have scope over two individuated noun phrase conjuncts. Besides nouns,  $\frac{3}{2}$ : *ku-* is also applicable to pronouns, which makes (54) possible, for example.

- (54) a. *Ang silvye Pada ku=yes.*  
 ang silv-ye Ø Pada ku=yes  
 AT look-3SG.F TOP Pada like=3SG.F.P  
 ‘Pada looks like her.’
- b. *Sa silvye ang Pada ku=ye.*  
 sa silv-ye ang Pada ku=ye  
 PT look-3SG.F A Pada like=3SG.F.TOP  
 ‘Like *her* Pada looks.’

With proper nouns, the same distributional properties as with generic nouns apply, except that  $\frac{3}{2}$ : *ku-* appears, rather idiosyncratically, as a suffix at the right edge of an NP—or at the right edge of the first NP conjunct—if the NP is preceded by a case marker, as shown in (55).

- (55) a. *Ang lentava sa Tagāti diyan-ku.*  
 ang lenta=va.Ø sa Tagāti diyan=ku  
 AT sound=2.TOP P Tagāti worthy=like  
 ‘You sound like Mr. Tagāti.’
- b. *Ang lentava sa Tagāti diyan-ku nay diranas yana.*  
 ang lenta=va.Ø sa Tagāti diyan=ku nay diran-as yana  
 AT sound=2.TOP P Tagāti worthy=like and uncle-P 3SG.M.GEN  
 ‘You sound like Mr. Tagāti and his uncle.’
- c. *Sa lentavāng ku-Tagāti diyan.*  
 sa lenta=vāng ku=Tagāti diyan  
 PT sound=2.A like=Tagāti worthy  
 ‘Like Mr. Tagāti you sound.’

With adjectives, however, there are no idiosyncrasies to this degree.  $\frac{3}{2}$ : *ku-* appears only as a prefix here, as with generic nouns:

- (56) a. *Surpya ku-suta ang Maran.*  
 surp-ya ku=suta ang Maran  
 seem-3SG.M like=busy A Maran  
 ‘Maran seems like he’s busy.’
- b. *Surpya ku-suta nay baras ang Maran.*  
 surp-ya ku=suta nay baras ang Maran  
 seem-3SG.M like=busy and gruff A Maran  
 ‘Maran seems like he’s busy and gruff.’



As (56b) shows, ꠘꠞ ꠘꠞ- again can have wide scope over conjuncts. What further distinguishes ꠘꠞ ꠘꠞ- from a prefix here is that it does not undergo crasis if the adjective begins with an /u/, hence ꠘꠞꠘꠞ ꠘꠞ-ubo /ku'ubo/ ‘like bitter’, not \*ꠘꠞꠘꠞ \*kūbo /'ku:bo/. Again, the position ꠘꠞ ꠘꠞ- appears in is special in that whatever modifies adjectives usually trails after them.

Besides attaching to words, ꠘꠞ ꠘꠞ- is furthermore able to subordinate infinite CPs. Since ꠘꠞ ꠘꠞ- leans on a whole phrase in (56), which affixes (at least in Ayeri) otherwise cannot do, its status as a clitic should be unmistakable in this context.

- (57) *Silvyeng ku-tabayam misungas.*  
 silv=yeng ku=taha-yam misung-as  
 look=3SG.F.A like=have-PTCP secret-P  
 ‘She looks as though having a secret.’

### Suffixes

Besides a number of prefixes and particles occurring before lexical heads which are likely clitics, Ayeri also has a number of morphemes trailing lexical heads as suffixes which do not seem quite like typical inflection. These are, for one, a part of the person suffixes on the verb. Especially tricky in this regard is maybe that “a pronominal affix or incorporated pronominal is effectively a clitic masquerading as an affix. Therefore, if there are pronominal affixes then they should behave exactly like clitics with respect to crucial aspects of morphosyntax” (Spencer and Luís 2012: 144; also compare Corbett 2006: 101). Spencer and Luís (2012) then proceed to give examples from Breton and Irish where the person marking on the verb is in complementary distribution with full NPs:

- (58) Breton (Spencer and Luís 2012: 145; from Borsley et al. 2007):
- a. *Bremañ e lennont al levrioù*  
 now PRT read.PRS.3PL the books  
 ‘Now they are reading the books’
  - b. *Bremañ e lenn ar vugale al levrioù*  
 now PRT read.PRS.3SG the children the books  
 ‘Now the children are reading the books’
  - c. \**Bremañ e lennont ar vugale al levrioù*  
 now PRT read.PRS.3PL the children the books

(59) Irish (Spencer and Luís 2012: 145; from McCloskey and Hale 1984):

- a. *Chuirfinn (\*mé) isteach ar an phost sin*  
 put.COND.ISG (I) in on the job that  
 ‘I would apply for that job’
- b. *Chuirfeadh sibh isteach ar an phost sin*  
 put.COND.3SG you in on the job that  
 ‘You would apply for that job’
- c. *Chuirfeadh Eoghan isteach ar an phost sin*  
 put.COND.3SG Owen in on the job that  
 ‘Owen would apply for that job’

What we can see in (58) is that, according to Spencer and Luís (2012), the verb shows no number marking, defaulting to the singular form, in non-negative clauses if the subject of the verb is overt as either a full noun or a pronoun: plural marking on the verb and a full subject cannot coincide in this case, which is why (58c) is marked ungrammatical. In (59a) we can see that there is no need for an explicit first-person pronoun, since that function is already expressed by person marking on the verb—person inflection on the verb seems to be in complementary distribution with full subject pronouns at least for some parts of the paradigm. In (59b) we have an overt second-person subject pronoun, but in this case, the verb does not agree with it and instead defaults to the third-person form, a clear case of which is given in (59c).

While in Ayeri, there is no defaulting to a certain person in the presence of an overt subject NP as such, there is still the effect of complementary distribution between a pronominal suffix in the absence of an overt subject NP, and a functionally impoverished as well as phonologically reduced form in its presence:

- (60) a. *Suta ang Niyas.*  
 suta ang Niyas  
 busy A Niyas  
 ‘Niyas is busy.’
- b. *Yāng suta.*  
 yāng suta  
 3SG.M.A busy  
 ‘He is busy.’
- (61) a. *Lampya ang Niyas.*  
 lamp-ya ang Niyas  
 walk-3SG.M A Niyas  
 ‘Niyas walks.’



Example (64) is an attempt to conceptualize formally that functionally, the inflection takes the role of the subject relation. Thus, (62a) is ungrammatical in that the pronominal suffix  $\text{ḡḡ}$  *-yāng* on the verb is redundant in the presence of a full NP which expresses the same features except that the subject NP's [PRED] value is  $\text{ḡḡ}$  *Niyas*, not 'pro'. In effect, what is attempted in (62), is to fill a grammatical relation with essentially the same content in two places, which is redundant. Assuming a 'pro' value for the [PRED] feature of  $\text{ḡḡ}$  *-yāng* 'he' is LFG's way to model the fact that this pronominal suffix functions as a pro-form available for predication, like a pronoun. Pronouns and full NPs necessarily exclude each other, however.

Example (65) shows, then, the annotations for  $\text{ḡḡ}$  *lampya* 'walks' as agreeing with an overt NP. Here, the suffix does not have a [PRED] feature—it is not available for predication, so that a full NP is allowed as a controller of agreement in the clause, with the person suffix is its person-inflection agreement target. The agreement suffix  $\text{ḡḡ}$  *-ya* thus reflects that the subject NP needs to have a certain set of person features. The NP which controls verb agreement (in canonical cases the agent NP) needs to match these features in order to establish an agreement relationship. By constraining ( $=_c$ ) the subject's predicator to not be a pro-form in (65) it should also be possible to rule out cases like in (62b), where person agreement is triggered by a pronominal NP, which is ungrammatical, since Ayeri basically supplants person agreement with a pronominal suffix in those cases, as we have seen. If  $\text{ḡḡ}$  *-yāng* were a simple inflectional affix, one of the two examples in (62) should be grammatical.

(65)	$\text{ḡḡ}$ <i>-ya</i> <sub>I</sub>	$V_{\text{infl}}$	(↑ SUBJ PRED)	$=_c$	$\neg$ 'pro'
			(↑ SUBJ PERS)	=	3
			(↑ SUBJ NUM)	=	SG
			(↑ SUBJ GEND)	=	M
			(↑ SUBJ ANIM)	=	+

The behavior of the pronominal person marking on the verb is thus rather complex, and decidedly unlike inflection in that what looks like an affix on the verb is also an argument of it, like a pronoun, as displayed in (63). Another layer of complexity is added by the fact that such an incorporated pronoun is also eligible for topicalization. As we have seen above, topic marking on nouns is realized by suppressing the realization of the overt case marker, whether it is a proclitic or a suffix. The topic- marked forms of pronouns are also underspecified for case, and they happen to be the same as those of the person-agreement suffixes, as exemplified by  $\text{ḡḡ}$  *-ya* in (61a). Thus, a topic-marked pronominal suffix on the verb will look exactly like ordinary agreement with a full NP, except that there is no full NP to agree with—hence the subscript numbers in (65) and (66) to distinguish between both kinds of  $\text{ḡḡ}$  *-ya*.<sup>12</sup>

<sup>12</sup> Assuming that the person suffix on the verb always co-indexes the topic and that one there-

(66)	ᵛᵛ ᵛᵛᵛ	V <sub>infl</sub>	(↑ SUBJ PRED)	=	'pro'
			(↑ SUBJ PERS)	=	3
			(↑ SUBJ NUM)	=	SG
			(↑ SUBJ GEND)	=	M
			(↑ SUBJ ANIM)	=	+
			(↑ SUBJ CASE)	=	∅ ⇒ (↑ TOP) = ↓

Comparing the feature list in (66) with that in (63) and (65), we see that (66) is basically the same as (63), except that either the [CASE] feature is absent, or that the suffix is underspecified for case. In absence of an NP to agree with, it follows from this definitional lack that the person marking on the verb itself is to be identified as constituting the topic, and the correspondent of the preverbal topic marker. In the following case, the preverbal topic marker defines that the topic is an animate agent; this information is united with the functional annotations in (66).

(67)	ᵛᵛᵛ ang	Cl	(↑ TOP CASE)	=	A
			(↑ TOP ANIM)	=	+

Instances of other case-unmarked nouns can be ruled out as being also part of the topic relation on the grounds of cohesion and functional uniqueness: if the topic is defined as an agent and it cannot be assumed from context that the case-unmarked noun in question is also part of the agent NP, discard it as a candidate. Besides, every core  $\theta$ -role (agent, patient, recipient) can only be assigned once, so if the role specified by the topic marker is already assigned, another NP in the same clause cannot also be assigned the same role. This gets more difficult with non-core roles, though it may be assumed that oblique arguments are less likely to be topicalized.

Possibly confusing with regards to the status of the pronominal suffixes as clitics is that “a pronominal affix or incorporated pronominal is effectively a clitic masquerading as an affix” (Spencer and Luís 2012: 144). While the pronominal suffixes in Ayeri behave in a special way regarding syntax, they lack wide scope, which is typical of affixes (apart from the examples from Turkish in (44)). Unlike Breton (58) or Irish (59), Ayeri’s pronominal affixes do not default to some form, and verbs cannot be unmarked either, that is, verbs always have to be inflected in some way, mostly for phonotactic reasons. Thus, in coordination, every conjunct has to be inflected for person features:

fore does not need to distinguish a person-agreement suffix from a homophonous topicalized pronominal suffix is a premature conclusion. In fact, the agreement suffix always agrees with the subject of the verb, which is most often the agent NP but may, in absence of an agent NP, also be the patient NP. The topic may consist of any NP, also oblique ones. The verb generally does not agree with non-agent or non-patient NPs, however.

- (68) a. *Nedrayāng nay layayāng.*  
 nedra=yāng nay laya-yāng  
 sit=3SG.M.A and read=3SG.M.A  
 ‘He is sits and reads.’
- b. \**Nedrayāng nay laya.*  
 nedra=yāng nay laya  
 sit=3SG.M.A and read
- c. \**Nedra nay layayāng.*  
 nedra nay laya=yāng  
 sit and read=3SG.M.A

In the case of 𐌵𐌳𐌹: *nedra*- ‘sit’ and 𐌵𐌳𐌹: *laya* ‘read’ in (68), leaving off the person marking would theoretically generate valid words, since \*𐌵𐌳𐌹: \**nedra* and \*𐌵𐌳𐌹: \**laya* satisfy phonotactic constraints (see section 1.2). However, Ayeri also has a great number of verb stems which end in a consonant cluster, such as 𐌵𐌳𐌹: *anl*- ‘bring’ or 𐌵𐌳𐌹: *tapy*- ‘set’, which do not form valid words as bare stems since words cannot end in CC. What would be possible instead is that one conjunct might carry the full pronominal suffix as a ‘strong’ form and the other one might only partially co-index the required features by using the less specific corresponding agreement marker as a ‘weak’ form. Differential marking of this kind, though, is simply not established.

After briefly delving into the realm of syntax, let us return to morphology for the second group of suffixes which need clarification. While Ayeri has quantifiers which are independent words, there are also a number of very common ‘little’ quantifiers and degree adverbs which are customarily spelled as suffixes, for instance, 𐌵𐌳𐌹: *-ikan* ‘much, many; very’, 𐌵𐌳𐌹: *kay* ‘few; a little’, 𐌵𐌳𐌹: *nama* ‘just, only’, and 𐌵𐌳𐌹: *-nyama* ‘even’. All of these are adverbial in meaning and while they are comparatively light in their semantics compared to regular content words, they do not particularly resemble functional morphemes either. A natural language which also contains suffixed quantifiers is West Greenlandic (Bittner 1995). According to Bittner (1995)’s terminology, 𐌵𐌳𐌹: *-ikan* in Ayeri as modifying a noun would be a D-quantifier, since it forms “a constituent [with] a projection of N” (59). This is in contrast to A-quantifiers, which are defined as forming “a constituent with some projection of V” (59). That is, A-quantifiers are adverbs like *almost* (𐌵𐌳𐌹: *-ngas* in Ayeri), *mostly*, or *never*, which modify verbs, while D-quantifiers are words like *most*, *some*, or *every*, which modify nouns. Ayeri makes no distinction between A- and D-quantifiers with regards to their being treated as suffixes, however, so that one may find suffixed quantifiers in both groups, sometimes even to the extent that the same quantifier may modify both nouns or verbs. Example (69) gives two instances of suffixed quantifiers from West Greenlandic for comparison with Ayeri in (70).

(69) West Greenlandic (adapted from Bittner 1995):

- a. *qaatuur-tuaanna-ngajap-p-a-a*  
break-always-almost-IND-TR-3SG<sub>1</sub>.3SG<sub>2</sub>  
'he almost always breaks it' (60)
- b. *qaqutigu-rujussuaq*  
rarely-very  
'very rarely' (63)

- (70) a. *ang adruya                      tadayen-ngas      adaley*  
ang adru=ya.Ø                      tadayen=ngas      ada-ley  
AT   break=3SG.M.TOP   always=almost   that-P.INAN  
'he almost always breaks it'
- b. *kora-ikan*  
kora=ikan  
rarely=very  
'very rarely'

As we can see in (69a), West Greenlandic incorporates the quantifier suffixes into the verb, while Ayerí—not a polysynthetic language—proceeds more freely in (70a) in that *tadayen* 'always, every time' is an adverb and as such a free morpheme which is, however, in turn modified by a suffixed quantifier. Since orthography may be treacherous, let us first try to establish whether *-ngas* 'almost' and *-ikan* 'many, much, very' and their like are free morphemes or not. As discussed initially regarding the preverbal particles, it is possible to reorder free morphemes, while clitics—as bound morphemes—cannot move around. Adverbs and adjectives are, if they optionally add additional information to a lexical head, adjuncts, and according to Carnie (2013) it is possible for adjuncts to switch places within the same syntactic domain. Adjuncts can also be coordinated with other adjuncts in the same syntactic domain. Furthermore, it is possible to replace X' nodes with pro-forms, like *one* in English.

- (71) a. *kipisānye-ikan              bino              kāryo*  
kipisān-ye=ikan              bino              kāryo  
painting-PL=many   colorful   big  
'many big colorful paintings'
- b. *kipisānye-ikan kāryo bino*  
'many colorful big paintings'
- c. <sup>1</sup>*kipisānye bino-ikan kāryo*  
'very colorful big paintings'
- d. <sup>1</sup>*kipisānye bino kāryo-ikan*  
'very big colorful paintings'

- (72) a. *kipisānye-ikan bino nay kāryo*  
           *kipisān-ye=ikan bino nay kāryo*  
           painting-PL=many colorful and big  
           ‘many big and colorful paintings’
- b. \**kipisānye-ikan nay bino kāryo*  
           ‘many and colorful big paintings’
- c. <sup>1</sup>*kipisānye bino-ikan nay kāryo*  
           ‘big and very colorful paintings’

As (71cd) shows, moving *ṣṣṣ* -*ikan* ‘many, much, very’ into different positions results not necessarily in ungrammatical expressions, but in ones with meanings different from what was intended, since *ṣṣṣ* -*ikan*’s scope changes from the noun to the adjective it is appended to. On the other hand, comparing (71a) and (b), it is possible for *ṣṣṣ* *kāryo* ‘big’ and *ṣṣṣ* *bino* ‘colorful’ to switch places with no ill effects. Example (72b) demonstrates that placing a coordinating conjunction between *ṣṣṣ* -*ikan* and *ṣṣṣ* *bino* does not work. The coordination in (72c), on the other hand, is not a problem—not because it is possible to coordinate *ṣṣṣ* -*ikan* and *ṣṣṣ* *kāryo*, but because *ṣṣṣ* *bino-ikan* ‘very colorful’ is considered one syntactic unit which is coordinated with *ṣṣṣ* *kāryo*. Thus, in (71b), we have actually been trying to coordinate *ṣṣṣ* *kipisānye-ikan* ‘many paintings’ with *ṣṣṣ* *bino* ‘colorful’, which does not work, since it is not possible to coordinate a lexical head with an adjunct supposed to modify it, because they are of different syntactic categories. In this regard it is worth mentioning that Ayeri’s quantifier suffixes are rather not complements either, since they are not required to satisfy their head’s argument structure.

One might argue that in (71) and (72) we tried to compare apples to oranges in that *ṣṣṣ* -*ikan* ‘many, much, very’ and *ṣṣṣ* *bino* ‘colorful’ are of different categories, since they do not appear to operate on the same levels. So instead, let us look at possibilities of word order change and coordination between different quantifiers to ensure that we actually stay on the same level. With this there is the problem, however, that it seems strange to modify the same lexical head with multiple different quantifiers, so this test does not really seem feasible to produce grammatical results. Also, with regards to coordination of quantifiers, it is maybe more natural to oppose them with *ṣṣṣ* *soyang* ‘or’ than to coordinate them; the grammatical structure of two categorially identical elements connected by a grammatical conjunction (even if the meaning is disjunctive) remains the same in either case.

- (73) a. \**keynam-ikan-kay*  
           *keynam=ikan-kay*  
           people=many-few  
           ‘few many people’



- b. <sup>?</sup>*keynam-ikan soyang kay*  
 keynam=ikan soyang kay  
 people=many or few  
 ‘few or many people’

In example (73a) we see that it is indeed not possible to combine multiple quantifiers to jointly modify a head in the way it is possible for multiple adjectives to modify the same head as in (71a), for instance. The example of quantifier disjunction in (73b) is also odd unless we permit a reading where *keynam* ‘people’ has been suppressed in the second disjunct to avoid repetition, although in the corresponding case of (74b) below, *da-kay* ‘few ones’ would be preferable.

- (74) a. <sup>?</sup>*keynam[-ikan soyang -kay]*  
 ‘[few or many] people’  
 b. [*keynam<sub>i</sub>-ikan*] *soyang* [<sub>i</sub>-*kay*]  
 ‘[few <sub>i</sub>] or [many people<sub>i</sub>]’

Both tests, moving *keynam-ikan* ‘many, much, very’ into other positions and coordination, have failed so far, and we have evidence that *keynam-ikan* forms a syntactic unit with its head, which points to it being a bound morpheme similar to an affix in spite of its adverb-like meaning. As with free words, it is also possible to replace a quantifier’s head with a pro-form, as mentioned above in the comment on (74b), and shown in more detail in (75). With quantifier suffixes there seems to be an overlap between word-like and affix-like properties, which is typical of clitics.

- (75) a. *Ang vacyan keynam-ikan seygoley.*  
 ang vac-yan keynam-Ø=ikan seygo-ley.  
 AT like-3PL.M people-TOP=many apple-P.INAN  
 ‘Many people like apples.’  
 b. *Ang vacyan danya-ikan seygoley.*  
 ang vac-yan danya-Ø=ikan seygo-ley.  
 AT like-3PL.M such.one-TOP=many apple-P.INAN  
 ‘Many of them like apples.’  
 c. *Ang vacyan da-ikan seygoley.*  
 ang vac-yan da=ikan-Ø seygo-ley.  
 AT like-3PL.M one=many-TOP apple-P.INAN  
 ‘Many (of them) like apples.’

Somewhat untypical of affixes, it seems to be possible to modify suffixed quantifiers with adverbs like *ekeng* ‘too’ and *kagan* ‘far too’, as (76) shows. This suggests that at least in this context, *keynam-ikan* ‘many, much, very’ may actually be

(76) *Ang vacyan keynam-ikan kagan disuley.*  
 ang vac-yan keynam-Ø=ikan kagan disu-ley  
 AT like-3PL.M people-TOP=many far.too disu-P.INAN  
 ‘Far too many people like bananas.’

(77) a. *Ang vacyan, narayang, keynam-ikan kagan disuley.*  
 ang vac-yan nara=yang keynam-Ø=ikan kagan disu-ley  
 AT like-3PL.M say=1SG.A people-TOP=many far.too disu-P.INAN  
 'Far too many people, I say, like bananas.'

b. \**Ang vacyan keynam, narayang, ikan kagan disuley.*

c. \**Ang vacyan keynam-ikan, narayang, kagan disuley.*

d. *Ang vacyan keynam-ikan kagan, narayang, disuley.*

(78) *keynam-ikan patu*  
 keynam=ikan patu  
 people=many surprisingly  
 'surprisingly many people'

The question here as well is whether *ṇṣṣ patu* refers to just ṇṣṣ-*ikan* or to ṇṣṣ-*keynam-ikan*. Replacing ṇṣṣ *keynam* with a pronoun produces a grammatical outcome (79a); doing so with ṇṣṣ-*keynam-ikan*, however, does not (79b). Replacing just ṇṣṣ-*ikan* at good last produces a very questionable expression as well, however (79c).

- (79) a. *keynam-ikan patu*  
           keynam=ikan patu  
           people=many surprisingly  
           ‘surprisingly many people’
- b. \**danyāng patu*  
       danya-ang patu  
       such.one-A surprisingly  
       ‘surprisingly ones’
- c. <sup>??</sup>*keynam da-patu*  
       keynam da=patu  
       people so=surprisingly  
       ‘surprisingly so people’

Another interesting distributional property of suffixed quantifiers in Ayeri is that in spite of their being suffixed, to verbs for instance, they can form arguments of the verb, similar to pronominal suffixes. Thus, with verbs like *ᵐᵒᵐ kond-* ‘eat’, *ᵐᵒᵐ -ma* ‘enough’ appears suffixed to the verb instead of as a predicative adverb. Incidentally, the examples in (80) also show that a quantifier attaches after pronominal suffixes, which we have already established as being clitics. An inflectional affix would not normally appear in post-clitic position, which is further evidence to the hypothesis that quantifier suffixes in Ayeri are clitics.

- (80) a. *Kondanang=ma.*  
           kond=nang=ma  
           eat=IPL.A=enough  
           ‘We ate enough.’
- b. *Ang tangay-ikan vana.*  
      ang tang=ay.Ø=ikan vana  
      AT hear=ISG.TOP=much 2.GEN  
      ‘I’ve heard much about you.’

Since Ayeri possesses a zero copula, equative phrases which treat quantifier suffixes as predicative adverbs pose a difficulty in that quantifier suffixes cannot stand alone like predicatives normally would. Thus, in a similar fashion to the behavior of *ᵐᵒᵐ -ma* ‘enough’ in (80a), the predicative *ᵐᵒᵐ -ma* in (81b) cliticizes to the only available word: the subject, *ᵐᵒᵐ adareng* ‘that’.

- (81) a. *Adareng edaya.*  
          ada-reng edaya  
          that-A.INAN here  
          ‘It is here.’

(83) a. *Ang tahisayan koyās kāryo nay kībasley-ikan.*  
 ang tahisa=yan.Ø koya-as kāryo nay kīhas=ley=ikan  
 AT own=3PL.M.TOP book-P big and map-P.INAN=many  
 ‘They own a big book and many maps.’  
*Not:* ‘They own many big books and maps.’

b. *Veneyang alingo patu nay para-ven.*  
 vney-ang alingo patu nay para=ven  
 dog-A clever surprisingly and quick=pretty  
 ‘The dog is surprisingly clever and pretty quick.’  
*Not:* ‘The dog is surprisingly pretty clever and quick.’

The interpretations marked as erroneous in (83) can be correctly achieved by ordinarily placing the adjective after the compounds so that the adjective itself has scope over both conjuncts. This is demonstrated in (84) and (85). Again, an unambiguous and individuating interpretation can be achieved by placing the quantifier suffix on each conjunct.

- (84) a. *Ang tabisayan koyajas nay kīhasyeley kāryo.*  
 ang tahisa=yan.Ø koya-ye-as nay kīhas-ye-ley kāryo  
 AT own=3PL.M.TOP book-PL-P and map-PL-P.INAN big  
 ‘They own big books and maps.’

- b. *Ang tabisayan koyās nay kīhasley-ikan kāryo.*  
 ang tahisa=yan.Ø koya-as nay kīhas-ley=ikan kāryo  
 AT own=3PL.M.TOP book-PL-P and map-P.INAN=many big  
 ‘They own many big books and maps.’

- (85) a. *Veneyang alingo nay para patu.*  
 veney-ang alingo nay para patu  
 dog-A clever and quick surprisingly  
 ‘The dog is surprisingly clever and quick.’

- b. *Veneyang alingo nay para-ven patu.*  
 veney-ang alingo nay para=ven patu  
 dog-A clever and quick=pretty surprisingly  
 ‘The dog is surprisingly pretty clever and quick.’

A slight problem is posed by the fact that the comparative suffixes on adjectives, *-eng* (COMP) and *-vā* (SUPL) are obviously derived from their quantifier counterparts meaning ‘rather’ and ‘most’. The question here is, whether they act as clitics as well, or whether grammaticalization has stripped them of some of the clitic-like properties of quantifier suffixes. Consider, for instance example (86):

- (86) a. *Ang koronya Kaman apyanas palay nay ban-eng.*  
 ang koron-ya Kaman apyan-as palay nay ban-eng  
 AT know-3SG.M Kaman joke-P funny and good-COMP  
 ‘Kaman knows a rather funny and good joke.’  
 or: ‘Kaman knows a funnier and better joke.’
- b. *Ang koronya Kaman apyanas palay-eng nay ban-eng.*  
 ang koron-ya Kaman apyan-as palay-eng nay ban-eng  
 AT know-3SG.M Kaman joke-P funny-COMP and good-COMP  
 ‘Kaman knows a funnier and better joke.’  
 or: ‘Kaman knows a rather funny and rather good joke.’

- What all of the above examples in (86) show is that in principle, both an interpretation of ᄃᆫᆯ -eng as a quantifier suffix and as a comparative suffix are possible, so that it is not easy to clearly distinguish between an inflectional affix and a clitic here. A clear distinction also cannot be made on phonological grounds in that even in the reading as a quantifier clitic, ᄃᆫᆯ -eng (COMP) and ᄃᆫᆯ -vā (SUPL) are stressed (compare Spencer and Luis 2012: 90–92). With ᄃᆫᆯ -eng it is possible at least to test whether it undergoes crasis if appended to an adjective stem ending in ᄃᆫᆯ -e. As we have seen, however, this is not a fully reliable metric either in that deictic prefixes show clitic-like behavior but may still phonologically meld with the stem they attach to. The only adjective ending in ᄃᆫᆯ -e currently listed in the dictionary is ᄃᆫᆯ *nake* ‘large, tall’:

- As illustrated by (87b), the purely comparative variant of *ᄒᆞᆫ* -eng should be able to be affected by crisis of two alike vowels. Since adjectives in *ᄒᆞᆫ* -e are exceedingly rare, though, this observation should not matter much in effect.

With regards to the dichotomy head-dependent marking, Ayeri is rather thoroughly dependent marking, albeit with the exception of agreement morphology on the verb. Dependent marking is exhibited, for instance, in the expression of possessive relationships, where the dependent is marked for genitive case:

- (88) a. *dema na Tuvo*  
 dema **na** Tuvo  
 aunt GEN Tuvo  
 ‘Tuvo’s aunt’



- b. *kasu bariri nā*  
 kasu bari-ri **nā**  
 basket meat-INS 1SG.GEN  
 ‘my basket of meat’



In (88a), <sup>ᑭᑭᑭ</sup> *Tuvo* is grammatically in possession of her <sup>ᑭᑭᑭ</sup> *dema* ‘aunt’; the possessee forms the head of the phrase while it is modified by the possessor, which receives the marking. In (88b), <sup>ᑭᑭᑭ</sup> *kasu* ‘basket’ forms the head and thus also the possessee while <sup>ᑭᑭᑭ</sup> *nā* ‘my’ serves as the dependent possessor; the genitive case is, then again, marked on the dependent. A further example of dependent marking is the locative case, which is marked on the prepositional object while the preposition itself, as the head of the PP, does not receive marking:

- (89) *agonan minkayya*  
 agonan minkay-**ya**  
 outside village-LOC  
 ‘outside of the village’



The relativizer, likewise, may agree in case with the NP in the matrix clause to which it links the relative clause. This typically happens mainly in formal language and—in terms of linear succession of words at the surface level of the clause—if the relativizer cannot be immediately adjacent to the NP which the relative clause modifies, for example, when an adjective or a possessive pronoun is following the noun:

- (90) *sangalas kivo sas ...*  
*sangal-as kivo s-as ...*  
 room-P small REL-P ...  
 ‘the small room which ...’



The only instance of head-marking there is in Ayeri is person-marking on the verb, which manifests when the NP following the verb (agent or patient) is not pronominal and thus there is no pronoun to cliticize to the verb stem, but the verb still receives a suffix that indicates a relation with, usually, the agent NP:

- (91) *Malya ang Amān.*  
*mal-ya ang Amān*  
 sing-3SG.M A Amān  
 ‘Amān sings.’



Sentences containing more than one NP also have topic marking on the verb, so that the verb agrees with one of the NPs in topicality. This may be the NP it has person agreement with or any other NP. The topicalized NP as a dependent of the verb is, in turn, zero-marked, so that the marking relationship for topics is bilateral and thus mixed:

- (92) *Sa manya ang Ajān Pila.*  
*Sa man-ya ang Amān Ø Pila*  
 PT greet-3SG.M A Ajān TOP Pila  
 ‘Pila, Ajān greets her.’



In this example, the verb exhibits canonic agreement with the agent, *Ajān*, in person, gender, and number. It is additionally marked for a patient topic, *Pila*,



*Pila*, and thus serves as an agreement target for two different controller NPs. The agreement relation is of a different kind for each of the two NPs, however.



## 4 Grammatical categories

---

While the previous chapter was about general mechanisms of marking in Ayeri, this chapter will dive into the various parts of speech in order to define their morphology with a closer look. I will begin with nouns as the main carriers of meaning, then deal with other parts of speech that regularly feature in or in combination with the noun phrase—pronouns, adjectives, and adpositions. Following this, there will be a discussion of verbs and adverbs before moving on to numerals and conjunctions.

### 4.1 Nouns

Nouns in Ayeri have *gender* and *number* as their inherent grammatical properties. Besides common nouns, there are, of course, also proper nouns (that is, names) and nominalizations. Nouns, as the heads of NPs, are also assigned *case* by the verb, which is a third grammatical property they display. For an illustration of the declension paradigms, compare Tables 4.1–4.4.

#### 4.1.1 Gender

Grammatical gender in Ayeri consists of two tiers which are subdivided into four classes based on a mixture of semantic and ontological properties:



Table 4.1: Declension paradigm for  $\text{ᩈᩣ᩠ᩅ}$  *badan* ‘father’ (animate; consonantal root)

	Singular		Plural	
TOP	<i>badan</i>	‘the father’	<i>badanye</i>	‘the fathers’
A	<i>badanang</i>	‘father’	<i>badanjang</i>	‘fathers’
P	<i>badanas</i>	‘father’ (obj.)	<i>badanjas</i>	‘fathers’ (obj.)
DAT	<i>badanyam</i>	‘to the father’	<i>badanjyam</i>	‘to the fathers’
GEN	<i>badanena</i>	‘of the father’	<i>badanyena</i>	‘of the fathers’
LOC	<i>badanya</i>	‘at the father’	<i>badanjya</i>	‘at the fathers’
CAUS	<i>badanisa</i>	‘due to the father’	<i>badanjisa</i>	‘due to the fathers’
INS	<i>badaneri</i>	‘with the father’	<i>badanyeri</i>	‘with the fathers’

Table 4.2: Declension paradigm for  $\text{ᩈᩣ᩠ᩅ}$  *māva* ‘mother’ (animate; vocalic root)

	Singular		Plural	
TOP	<i>māva</i>	‘the mother’	<i>māvaye</i>	‘the mothers’
A	<i>māvāng</i>	‘mother’	<i>māvajang</i>	‘mothers’
P	<i>māvās</i>	‘mother’ (obj.)	<i>māvajas</i>	‘mothers’ (obj.)
DAT	<i>māvayam</i>	‘to the mother’	<i>māvajyam</i>	‘to the mothers’
GEN	<i>māvana</i>	‘of the mother’	<i>māvayena</i>	‘of the mothers’
LOC	<i>māvaya</i>	‘at the mother’	<i>māvaiya</i>	‘at the mothers’
CAUS	<i>māvaisa</i>	‘due to the mother’	<i>māvajisa</i>	‘due to the mothers’
INS	<i>māvai</i>	‘with the mother’	<i>māvayeri</i>	‘with the mothers’

The animate gender refers, broadly speaking, to entities that are considered alive or are closely associated with living things, such as events, concepts, or activities executed or connected to them. The ‘masculine’ and ‘feminine’ subcategories are applied to humans, animals whose sex is known (for example on behalf of breeding them or keeping them as pets), and gods—basically anything that shows sexual dimorphism or is assumed to be an exponent of it as well as nouns referring to such entities in a functional way, for instance,  $\text{ᩈᩣ᩠ᩅ}$  *badan* ‘father’ and  $\text{ᩈᩣ᩠ᩅ}$  *māva* ‘mother’. The remainder falls into the ‘neuter’ category—plants, for instance, body parts, or animals whose sex is unknown. The ‘inanimate’ category typically contains materials and things, such as tools. Furthermore, animals and plants change their category to inanimate as well if they serve as food. There are exceptions to either group, where elements appear in them for no obviously discernable reason. In order to illustrate,

Table 4.3: Declension paradigm for *kirin* ‘street’ (inanimate; consonantal root)

	Singular		Plural	
TOP	<i>kirin</i>	‘the street’	<i>kirinye</i>	‘the streets’
A	<i>kirinreng</i>	‘street’	<i>kirinyereng</i>	‘streets’
P	<i>kirinley</i>	‘street’ (obj.)	<i>kirinyeley</i>	‘streets’ (obj.)
DAT	<i>kirinyam</i>	‘to the street’	<i>kirinjyam</i>	‘to the streets’
GEN	<i>kirinena</i>	‘of the street’	<i>kirinyena</i>	‘of the streets’
LOC	<i>kirinya</i>	‘at the street’	<i>kirinjya</i>	‘at the streets’
CAUS	<i>kirinisa</i>	‘due to the street’	<i>kirinjisa</i>	‘due to the streets’
INS	<i>kirineri</i>	‘with the street’	<i>kirinyeri</i>	‘with the streets’

Table 4.4: Declension paradigm for *pera* ‘measure’ (inanimate; vocalic root)

	Singular		Plural	
TOP	<i>pera</i>	‘the measure’	<i>peraye</i>	‘the measures’
A	<i>perareng</i>	‘measure’	<i>perayereng</i>	‘measures’
P	<i>peraley</i>	‘measure’ (obj.)	<i>perayeley</i>	‘measures’ (obj.)
DAT	<i>perayam</i>	‘to the measure’	<i>perajyam</i>	‘to the measures’
GEN	<i>perana</i>	‘of the measure’	<i>perayena</i>	‘of the measures’
LOC	<i>peraya</i>	‘at the measure’	<i>perajya</i>	‘at the measures’
CAUS	<i>peraisa</i>	‘due to the measure’	<i>perajisa</i>	‘due to the measures’
INS	<i>perari</i>	‘with the measure’	<i>perayeri</i>	‘with the measures’

here are a few examples for each category:

- (2) a. Animate masculine:  
*badan* ‘father’, *netu* ‘brother’, *aguyan* ‘rooster’, *Ajān*, *Latun*;
- b. Animate feminine:  
*māva* ‘mother’, *kina* ‘sister’, *aguway* ‘hen’, *Maha*, *Trānay*;
- c. Animate neuter:  
*adang* ‘palm tree’, *binu* ‘color’, *ikam* ‘deer’, *kadān* ‘harvest’, *cān* ‘love’, *nanga* ‘house’, *tampu* ‘luck’, *yila* ‘foot’;
- d. Inanimate:  
*ahal* ‘sand’, *bema* ‘egg’, *kaban* ‘spear’, *melung* ‘yogurt’, *nusān* ‘damage’, *payutān* ‘mathematics’.

Ayeri only distinguishes singular and plural in nouns, which receive plural marking; verbs, then, agree with agent NPs in number in the canonical case. Ordinarily, nouns in Ayeri are countable, however, there is also a group of uncountable nouns as well as a (small) group of nouns which are always plural. As above, I will list a few words from each group for illustration:

## (4) a. Countable nouns:

၁၂၁၂ *ajam* ‘toy’ — ၁၂၁၂၂ *ajamye* ‘toys’,  
 ၂၁၂ *devo* ‘head’ — ၂၁၂၂ *devoye* ‘heads’,  
 ၁၂၁၂ *inun* ‘fish’ — ၁၂၁၂၂ *inunye* ‘fish’ (pl.),  
 ၁၂၁၂ *netu* ‘brother’ — ၁၂၁၂၂ *netuye* ‘brothers’;

## b. Uncountable nouns:

၁၂၁၂ *abal* ‘sand’, ၁၂၁၂ *bakay* ‘stuff’, ၁၂၁၂ *gabān* ‘hope’, ၁၂၁၂ *mingan* ‘ability’;

## c. Plurale tantum nouns:

၁၂၁၂ *burang* ‘lifestock, cattle’,<sup>1</sup> ၁၂၁၂ *ganengan* ‘siblings’, ၁၂၁၂ *keynam* ‘people’, ၁၂၁၂ *tang* ‘ears’.

Most concrete things that exist as discrete entities are countable, also, for instance, animals and livestock—fish, deer, sheep etc. are thus countable, unlike in English; pants, pliers, scissors, glasses, etc. are by default singular as well. Uncountable, on the other hand, are materials in general or abstract concepts. There are also a number of nouns which are plural by default, most notably entities which often occur in groups, but there is as well the odd word for which there seems to be no reason to be included in this group, for instance, ၁၂၁၂ *mino* ‘paint’, and ၁၂၁၂ *gimbay* ‘sorrows’. A few body parts are also plurale tantum nouns, especially those which occur in pairs (၁၂၁၂ *niva* ‘eye’ is a notable exception).

As demonstrated in (4a), the noun plural marker is ၁၂ *-ye*, which in native orthography also occurs in the variant ၁၂ or ၁၂. As described above (section 1.1.1, p. 7), the plural marker may also be reduced to [dʒ] ⟨-j⟩ before case suffixes that begin with /j/ or with a vowel other than /e/, like ၁၂၁၂ *-ang* (A) or ၁၂၁၂ *-yam* (DAT):

(5) a. ၁၂၁၂၁၂ *diranang* (uncle-A) + ၁၂ *-ye* (PL)

→ ၁၂၁၂၁၂ *diranjang* (uncle-PL-A),

b. ၁၂၁၂ *diranena* (uncle-GEN) + ၁၂ *-ye* (PL)

→ ၁၂၁၂၁၂ *diranyena* (uncle-PL-GEN),

c. ၁၂၁၂၁၂ *diranyam* (uncle-DAT) + ၁၂ *-ye* (PL)

→ ၁၂၁၂၁၂ *diranjyam* (uncle-PL-DAT).

For plurale tantum, to express a singular entity, it is always possible to use a genitive phrase like —၁၂၁၂၁၂ ...-*ena men* ‘one of ...’ (...-GEN one), for instance:

<sup>1</sup> Specifically in this meaning; ၁၂၁၂ *burang* can also simply mean ‘animal’, in which case there is a plural form ၁၂၁၂၁၂ *burangye* ‘animals’.

- (6) a. *Nupayon tangang nā.*  
 nupa-yon tang-ang nā  
 hurt-3PL.N ears-A ISG.GEN  
 ‘My ears hurt.’
- b. *Na nupareng tang nā men.*  
 na nupa=reng tang-Ø nā men  
 GENT hurt=3SG.INAN.A ears-TOP ISG.GEN one  
 ‘Of my ears, one is hurting.’

Number in nouns can also be manipulated by quantifiers which attach to declined nouns as suffixes—or rather, enclitics. In this case, when plurality is indicated by the quantifier, the noun is not additionally marked for number; the verb, however, keeps agreeing in number:

- (7) a. *Ajayon ganjang kivo.*  
 aja-yon gan-ye-ang kivo  
 play-3SG.N child-PL-A small  
 ‘The small children are playing.’
- b. *Ajayon ganang-ikan kivo.*  
 aja-yon gan-ang=ikan kivo.  
 play-3SG.N child-A=many small  
 ‘Many small children are playing.’

Likewise, when nouns are modified by numerals, plurality is not normally marked again on the noun. In example (8a), we see a plural noun, *ṣṣ nanga* ‘house’, and in (8b) the same phrase is repeated again with plurality implied by the use of a numeral, *ṣṣ sam* ‘two’; the plural noun itself appears unmarked in its singular form in this case.

- (8) a. *Ang no vehya sitang-yām nangajas veno nay biro.*  
 ang no veh=ya.Ø sitang=yām nanga-ye-as veno nay hiro  
 AT want build=3SG.M.TOP self=3SG.M.DAT house-PL-P pretty and new  
 ‘He wants to build himself pretty new houses.’
- b. *Ang no vehya sitang-yām nangās sam veno nay biro.*  
 ang no veh=ya.Ø sitang=yām nanga-as sam veno nay hiro  
 AT want build=3SG.M.TOP self=3SG.M.DAT house-P two pretty and new  
 ‘He wants to build himself two pretty new houses.’

An exception to this is the use of words for the numeral powers, like *ṣṣ lan* ‘dozen’, *ṣṣ menang* ‘gross’, *ṣṣ samang* ‘myriad’, etc. in an unspecified way like ‘dozens of people’. In this case, to convey that the numeral is not to be understood



as a precise value, the modified noun will appear in the plural—even if it is a plurale tantum like ꠘꠘꠘ *keynam* ‘people’:

- (9) *Bengyon keynamjang menang.*  
 beng-yon keynam-ye-ang menang  
 attend-3SG.N people-PL-A gross  
 ‘Hundreds of people attended.’

As we have seen in various examples above, proper nouns in Ayeri do not receive inflection for case by suffixes as common nouns do, and for the purpose of number they are treated as uncountable in Ayeri—they resist inflection by suffixation, marking their special status.<sup>2</sup> However, they can still be modified by quantifiers and quantifying suffixes; verb agreement as well can be used to indicate plurality:

- (10) a. *Sabayan cabo ekeng ang Yan.*  
 saha-yan cabo ekeng ang Yan  
 come-3PL.M late too A Yan  
 ‘The Yans are coming too late.’  
 b. *Ang apateng sa Yan-ikan.*  
 ang apa=teng sa Yan=ikan  
 AT laugh=3PL.F.A P Yan=all  
 ‘They laughed at (all) the Yans.’

#### 4.1.3 Case

As demonstrated in the declension tables at the beginning of this section (Tables 4.1–4.4), Ayeri’s NPs are marked for case, which is governed by the verb. Since Ayeri uses a split alignment system with some additional complications, it is not very straightforward, in my opinion, to use the classical labels of nominative (S/A) and accusative (O), or of absolutive (S/P) and ergative (O) for the first two core roles. Hence, I will be using the terms ‘agent’ and ‘patient’, which I hope brings about some more clarity, especially when discussing the mentioned complications later on.

##### Agent

What I call ‘agent’ here is, to quote Fillmore (2003 [1968]), “the case of the typically animate perceived instigator of the action identified by the verb” (46). Fillmore

<sup>2</sup> Many common names in Ayeri are derived from regular words in the language, so the language needs to have a way to distinguish between regular use and use as a name. For instance, the name ꠘꠘ *Yan* also means ‘boy, son’ as a common noun.

(2003 [1968]) himself qualifies this definition, however, in that the “escape qualification ‘typically’ expresses my awareness that contexts which I will say require agents are sometimes occupied by ‘inanimate’ nouns like *robot* or ‘human institution’ nouns like *nation*” (Fillmore 2003 [1968]: 46, footnote 31). Payne (1997) summarizes on prototypical agents with regards to their topicality that a “less technical way of expressing this fact is to say that people identify with and like to talk about things that act, move, control events, and have power” (151).

Agents in Ayeri frequently embody the properties quoted by both Fillmore (2003 [1968]) and Payne (1997) in this regard, including Fillmore (2003 [1968])’s caveat. However, importantly, ‘agent’ in Ayeri is a macrorole that may be applied to, for instance, instruments, experiencers, and less typical actors as well, specifically, in absence of more prototypical candidates for agenthood in a sentence. It thus comes very close to a nominative, except that it does not need to be locus of the sentence’s topic—although agents very typically are topics, as Payne (1997: 151) goes on to note.<sup>3</sup> Thus, the first NP after the verb in all of the following examples is treated as an agent; the agent is marked by the suffix :ᐃᑦ *-ang* for animate referents and the suffix ᐃᑦᑦ *-reng* for inanimate referents; names and verbal topic agreement are marked by the clitic case markers ᐃᑦ *ang* and ᐃᑦᑦ *eng*, respectively:

- (11) a. *Ang tinkaya Yan kunangley.*  
 ang tink-a-ya Ø Yan kunang-ley  
 AT open-3SG.M TOP Yan door-P.INAN  
 ‘Yan opens the door.’
- b. *Le tinkaya ayonang kunang.*  
 le tink-a-ya ayon-ang kunang-Ø  
 PT open-3SG.M man-A door-TOP  
 ‘The door is opened by a/the man’,  
 or: ‘The door, a/the man opens it.’
- c. *Eng tinkāra tinkay kunangley.*  
 eng tink-a-ara tinkay-Ø kunang-ley  
 AT.INAN open-3SG.INAN key-TOP door-P.INAN  
 ‘The key opens the door.’
- d. *Tinkāra kunangreng.*  
 tink-a-ara kunang-reng  
 open-3SG.INAN door-A.INAN  
 ‘The door opens.’

<sup>3</sup> This is the main reason I spoke of ‘complications’ above: Ayeri’s notion of ‘subject’ is somewhat problematic due to topicalization, which is why I try to avoid complicating terminology by using ‘nominative’ for agent topics and ‘ergative’ for agent non-topics, and ‘accusative’ for patient non-topics and ‘absolutive’ for patient topics, respectively.

- e. *Sā tinkaya ang Yan kunangley yan.*  
 sā tinka-ya ang Yan kunang-ley yan.Ø  
 CAUT open-3SG.M A Yan door-P.INAN 3SG.M.TOP  
 ‘They make Yan open a/the door’,  
 or: ‘Because of them, Yan opens the door.’

In predicative constructions, the constituent which a quality is assigned to or about which a judgement is made is also assigned the agent case:

- (12) a. *Tado tinkayreng.*  
 tado tinkay-reng  
 old key-A.INAN  
 ‘The key is old.’  
 b. *Ang Yan nimpayās ban.*  
 ang Yan nimpaya-as ban  
 A Yan runner-P good  
 ‘Yan is a good runner.’

With regards to constituents’ roles in ditransitive verb frames, donors are represented by agents in Ayeri as well, since they are the origin of whatever is conceptually passed on to the recipient party:

- (13) *Le ilya ang Yan tinkay yam Cānlay.*  
 le il-ya ang Yan tinkay-Ø yam Cānlay  
 PT give-3SG.M A Yan key-TOP DAT Cānlay  
 ‘The key, Yan gives it to Cānlay.’

#### Patient

Patients are less of a definitional problem than agents, since in transitive sentences, they are very typically undergoers, that is, the constituent that is acted on, affected, or produced by the action expressed by the verb. The patient case is thus the one assigned by default to direct objects—but also to predicative nominals. In ditransitive sentences, the theme is represented by the patient. Animate patients are marked by 𐀓𐀮 -as, inanimate ones by 𐀓𐀮 -ley; for names and verbal topic agreement, the markers are 𐀓 sa and 𐀓 le, respectively:

- (14) a. *Ang silvye Briha sa Taryan.*  
 ang silv-ye Ø Briha sa Taryan  
 AT see-3SG.F TOP Briha P Taryan  
 ‘Briha sees Taryan.’

- b. *Sa manye ang Briha Taryan.*  
 sa man-ye ang Briha Ø Taryan  
 PT greet-3SG.F A Briha TOP Taryan  
 ‘Taryan is greeted by Briha’,  
 or: ‘Taryan, Briha greets him.’

- (15) a. *Ang rimaye Briha kunangley.*  
 ang rima-ye Ø Briha kunang-ley  
 AT close-3SG.F TOP Briha door-P.INAN  
 ‘Briha closes a/the door.’

- b. *Le rimaye ang Briha kunang.*  
 le rima-ye ang Briha kunang-Ø  
 PT.INAN close-3SG.F A Briha door-TOP  
 ‘The door is closed by Briha’,  
 or: ‘The door, Briha closes it.’

- (16) *Ang ilya Taryan koyaley yam Kandan.*  
 ang il-ya Ø Taryan koya-ley yam Kandan  
 AT give-3SG.M TOP Taryan book-P.INAN DAT Kandan  
 ‘Taryan gives Kandan a book.’

As the translations of the examples above show, topicalizing the patient can be used to create an effect similar to English’s passive voice, except that the patient will not become marked by the agent case for logical reasons—this is a notable difference from the nominative. Even if the agent NP is omitted, the patient NP will not be changed to the agent case, since that would reverse the direction of action:

- (17) *Manya sa Taryan. ≠ Manya ang Taryan.*  
 man-ya sa Taryan Man-ya ang Taryan  
 greet-3SG.M P Taryan greet-3SG.M A Taryan  
 ‘Taryan is greeted.’ ≠ ‘Taryan greets.’

This example shows that the case of the NP will not change, however, the verb will: it now agrees with the next argument in line, the patient NP. It will not do so, however, if the order of arguments is just scrambled, as exemplified by (18). This is to say that the verb does not simply agree with whichever NP follows it, even if it can be assumed that verb agreement in Ayeri developed along similar lines in-world, which will become especially apparent in the discussion of pronouns.<sup>4</sup>

<sup>4</sup> Mismatches in agreement in connection to scrambling such as exemplified by (18b) are to be expected, however, since the brain can only handle so much information between the controller and the target of an agreement relationship. Corbett (2006), notes that with regards

- (18) a. *Sa manye Taryan ang Briha.*  
 sa man-ye Ø Taryan ang Briha  
 PT greet-3SG.F TOP Taryan A Briha  
↑  
person agreement  
 ‘Taryan is greeted by Briha’,  
 or: ‘Taryan, Briha greets him.’
- b. \**Sa manya Taryan ang Briha.*  
 sa man-ya Ø Taryan ang Briha  
 PT greet-3SG.M TOP Taryan A Briha  
↑  
\*person agreement

Besides being the default case for direct objects, the patient case is also assigned to predicative nominals, by analogy with transitive sentences and in spite of the likening nature of the construction:

- (19) *Ang Yan nimpayās ban.*  
 ang Yan nimpaya-as ban  
 A Yan runner-P good  
 ‘Yan is a good runner.’

#### Dative

The most typical use of the dative is for the recipient NP in a ditransitive clause; as such, it may be a recipient proper or the entity to whose benefit the action is carried out. A number of transitive verbs also use the dative for their object, for example, when it is the target of address. The dative can furthermore be used to mark movement toward a place. The case suffix for datives is ꠘꠗ -*yam* for both animate and inanimate entities. Names and verbal topic agreement are marked equally by ꠘꠗ *yam*. Verbs do not exhibit person agreement with dative NPs, since experiencers are treated as agents.

- (20) a. *Ang ilya Taryan koyaley ayonyam.*  
 ang il-ya Ø Taryan koya-ley ayon-yam  
 AT give-3SG.M TOP Taryan book-P.INAN man-DAT  
 ‘Taryan gives a book to the man.’

to agreement in NP conjuncts, “distant agreement is rare, and that agreement with the nearest noun phrase or agreement with all (resolution) is much more common” (62). If there were an extensive corpus of texts written by Ayeri speakers, it might be interesting to gather statistics on the number of words between target and controller in relation to the prevalence of agreement mismatches.

- b. *Ang ilya                      Taryan koyaley              yam Kandan.*  
 ang il-ya              Ø    Taryan koya-ley              yam Kandan  
 AT   give-3SG.M   TOP   Taryan   book-P.INAN   DAT   Kandan  
 ‘Taryan gives Kandan a book.’
- c. *Yam ilya                      ang Taryan koyaley              ayon.*  
 yam il-ya                      ang Taryan koya-ley              ayon-Ø  
 DATT   give-3SG.M   A   Taryan   book-P.INAN   man-TOP  
 ‘The man is given a book by Taryan’,  
 or: ‘The man, Taryan gives him a book.’

The three examples in (20) show the regular use of the dative as the case the recipient of the theme appears in. What distinguishes Ayeri from a pure split-S language is that all constituents can serve as topics, not just agents and patients with regards to their function as syntactic subjects. Thus, it is also possible for dative NPs to appear as topics—person agreement is unaffected by this, though. The following example shows the addressee of a speech act in the dative case; the message is treated as the theme which is passed on:

- (21) *Ang ningye              māva              ninganas              ganyam              yena.*  
 ang ning-ye              māva-Ø              ningan-as              gan-yam              yena  
 AT   tell-3SG.F   mother-TOP   story-P   child-DAT   3SG.F.GEN  
 ‘The mother tells her child a story.’

As mentioned above, the dative can also take on an allative meaning insofar as it marks the target of a motion, as displayed in (22a). As an extension of this means, the adpositional object may as well appear in the dative, since Ayeri cannot distinguish, for instance, ‘up’ from ‘to the top of’ with just the preposition, in this case *ling* ‘on top of’. With the adpositional object in the locative case (see below), the phrase in (22b) would imply that the man were literally going to the top of the temple, that is, possibly ending up on its roof.

- (22) a. *Ang nimpye              lay              māvayam              yena.*  
 ang nimp-ye              lay-Ø              māva-yam              yena  
 AT   run-3SG.F   girl-TOP   mother-DAT   3SG.F.GEN  
 ‘The girl runs to her mother.’
- b. *Ang saraya              ayon              manga ling              natrangyam.*  
 ang sara-ya              ayon-Ø              manga ling              natrang-yam  
 AT   go-3SG.M   man-TOP   DIR   top   temple-DAT  
 ‘The man goes up to the temple.’

*Genitive*

The genitive is used to mark possessors; attributive genitives follow the possessee. It can also be used for ablative meanings, that is, to mark the place from which a motion originates, in analogy to the dative's allative use. The genitive is marked on common nouns with the suffix : $\text{ṛ}$  -*na*. If a noun stem ends in a consonant, the marker becomes : $\text{ṛ}$  -*ena*, compare Tables 4.1–4.4 above. Names and verbal topic agreement are marked by  $\text{ṛ}$  *na*. There is no animacy distinction in the genitive case.

- (23) a. *Pakur ledanang netuna nā.*  
 pakur ledan-ang netu-na nā  
 sick friend-A brother-GEN 1SG.GEN  
 'My brother's friend is sick.'
- b. *Kopo dilengyereng ajānena.*  
 kopo dileng-ye-reng ajān-ena  
 difficult rule-PL-A.INAN game-GEN  
 'The rules of the game are difficult.'
- c. *Ang nakasyo tamo ibangya na Niyas.*  
 ang nakas-yo tamo-Ø ibang-ya na Niyas  
 AT grow-3SG.N wheat-TOP field-LOC GEN Niyas  
 'There is wheat growing on Niyas's field.'
- d. *Na nakasyo tamoang ibangya Niyas.*  
 na nakas-yo tamo-ang ibang-ya Ø Niyas  
 GENT grow-3SG.N wheat-A field-LOC TOP Niyas  
 'Regarding Niyas, there is wheat growing on his field.'

Futhermore, Ayeri does not make a distinction between alienable and inalienable possession, so that typically inalienable things such as body parts, relatives and family members, or personal items and tools are all treated as described above. Consider the following example for illustration:

- (24) *Ang puntaye māva nā mitrangas yena sembari yena.*  
 ang punta-ye māva-Ø nā mitrang-as yena semba-ri yena  
 AT brush-3SG.F mother-TOP 1SG.GEN hair-P 3SG.F.GEN comb-INS 3SG.F.GEN  
 'My mother is brushing her hair with her comb.'

The above examples show the regular use of the genitive as a marker of possession. The following examples, on the other hand, show the genitive in its ablative function:

- (25) *Ang sabaya Vetayan rimanena.*  
 ang saha-ya Ø Vetayan riman-ena  
 AT come-3SG.M TOP Vetayan city-GEN  
 ‘Vetayan comes from the city.’

### Locative

The locative marks basic locations, often the default that is associated with a verb. It is also the case in which adpositional objects normally appear, besides the special cases using the dative and the genitive mentioned above. Common nouns are marked by  $\text{ya}$  *-ya*;<sup>5</sup> names and verbal topic agreement use the marker  $\text{u}$  *ya*. There is no difference made between animate and inanimate referents in the locative.

- (26) a. *Ang nedraya paray binya.*  
 ang nedra-ya paray-Ø hin-ya  
 AT sit-3SG.M cat-TOP box-LOC  
 ‘The cat sits in the box.’  
 b. *Ang naraya Ajān ya Kaman.*  
 ang nara-ya Ø Ajān ya Kaman  
 AT speak-3SG.M TOP Ajān LOC Kaman  
 ‘Ajān speaks to Kaman.’  
 c. *Ya mica ang Kaman Visamhinang.*  
 ya mit-ya ang Kaman Ø Visamhinang  
 LOCT live-3SG.M A Kaman TOP Visamhinang  
 ‘Kaman lives in Visamhinang’,  
 or: ‘Visamhinang is where Kaman lives.’

The example sentences in (26) show locative NPs that are not further specified by adpositions so that the correct interpretation may be dependent on context and the experience of the addressee. Example (26a) is an instance of this circumstance, insofar as experience tells that cats like to sit inside boxes, so further specifying the position with the preposition  $\text{ႬႬႬ}$  *kong* ‘inside’ would be emphasizing that the cat is not sitting just anywhere, but really *inside* the box as opposed to on top of it, for instance. The following example has the cat sitting on top of the box:

- (27) *Ang nedraya paray ling binya.*  
 ang nedra-ya paray-Ø ling hin-ya  
 AT sit-3SG.M cat-TOP on.top box-LOC  
 ‘The cat sits on the box.’

<sup>5</sup> Older texts still exhibit an allomorph  $\text{ႬႬႬ}$  *-ea*, used especially in combination with the plural suffix  $\text{ႬႬႬ}$  *ye*, giving  $\text{ႬႬႬ}$  *-yēa*. The modern language uses  $\text{ႬႬႬ}$  *-fya*.



Ayeri also has a number of postpositions, which does not change that the adposition object is marked for locative case, however:

- (28) *Ang mican edaya **tenyanya** tan pesan.*  
 ang mit-yan edaya tenyan-ya tan pesan  
 AT live-3PL.M here death-LOC 3PL.M.GEN until  
 ‘They lived here until their death.’

#### Causative

The causative marks the cause or causer of an action, the instigator or the reason on behalf of which an agent is acting. It is thus similar to the agent case, though it does not replace it in Ayeri; verbs do not exhibit person agreement with causers even though their action logically supersedes or precedes that of the agent in the embedded event. Dixon (2000) writes that a “causer refers to someone or something (which can be an event or state) that initiates or controls the activity. This is the defining property of the syntactic–semantic function A (transitive subject)” (30). According to Comrie (1989: 176), the causee—the agent of the event controlled by the causer—normally takes the highest place in the hierarchy of syntactic constituents that is not already filled, in this case, by the causer. This observation, however, is complicated by Ayeri’s more or less semantics-based case marking as well as topicalization. In the following, I will give examples of nominal marking for cause as before; a discussion of the morphosyntax of Ayeri’s morphological causative constructions will be deferred to the section on valency-increasing operations.

Causers or causes are marked by *isa* for common nouns; names and verbal topic agreement use the marker *sā*. As stated above, verbs do not agree with causers even though they have agent-like semantics. There is no animacy distinction in the marking of causers.

- (29) a. *Ang rua sarāyn seyaranisa.*  
 ang rua sara=ayn.Ø seyaran-isa  
 AT must leave=IPL.TOP rain-CAUS  
 ‘We had to leave due to the rain.’  
 b. *Ang yomāy edaya sā Apican.*  
 ang yoma=ay.Ø edaya sā Apican  
 AT be=ISG.TOP here CAUS Apican  
 ‘I am here because of Apican.’

- c. *Sā nimpvāng hakasley yan.*  
*sā nimp=vāng hakas-ley yan.Ø*  
 CAUT run=2SG.A mile-P.INAN 3PL.M.TOP  
 ‘You run a mile because of them’,  
 or: ‘Due to them, you run a mile’,  
 or: ‘They make you run a mile.’

Regarding the typological oddities mentioned above, example (29c) shows what happens in Ayeri with regards to the marking of causers. Essentially, the causer topic was grammaticalized to express a causation relationship.

#### Instrumental

The instrumental marks the means by which an action is carried out by an agent. This can be a tool as well as an animate being by whose help the action is brought about. The instrumental thus, in effect, marks secondary agents (but not causees); verbs, however, never show person agreement with instrumental NPs. Common nouns are marked by *-ri* when ending in a vowel and with *-eri* when ending in a consonant; names and verbal topic agreement are marked by *ri*. With nouns ending in *-e*, as well as the plural marker *-ye*, there is variation regarding whether *-ri* or *-eri* is used, so that in the case of the plural marker both *-yeri* and *-yēri* occur. In passive-like constructions, it is not grammatical to reintroduce the agent as an instrumental; the agent simply remains in the clause in this case, though as a non-topic constituent.

- (30) a. *Ang visye Pila seygoley tibangeri yena.*  
*ang vis-ye Ø Pila seygo-ley tihang-eri yena.*  
 AT cut-3SG.F TOP Pila apple-P.INAN knife-INS 3SG.F.GEN  
 ‘Pila cuts an apple with her knife.’
- b. *Ang liboyya-ma badan nibanyeri (nibanyēri).*  
*ang liha-oy-ya=ma badan-Ø nihan-ye-ri (nihan-ye-eri)*  
 AT earn-NEG-3SG.M=enough father-TOP nihan-PL-INS (nihan-PL-INS)  
 ‘Father did not earn enough with his fruits.’
- c. *Ang lingya Mindan mebiras ri Kadijān.*  
*ang ling-ya Ø Mindan mehir-as ri Kadijān.*  
 AT climb.up-3SG.M TOP Mindan tree-P INS Kadijān  
 ‘Mindan climbs a tree with Kadijān’s help.’
- d. *Ri tavya gino ang Kan nimpur.*  
*ri tav-ya gino ang Kan nimpur-Ø*  
 INST become-3SG.M drunk A Kan wine-TOP  
 ‘Kan becomes drunk on the wine’,  
 or: ‘The wine, Kan becomes drunk on it.’

The instrumental may also be used for comitative meanings where the instrumental NP acts as a nominal adjunct describing an attribute of its antecedent head, for example:

- (31) *Ang pegayo sinyā kasuley bari nā?*  
 ang pega-yo sinyā-Ø kasu-ley bari-ri nā  
 AT steal-3SG.N who-TOP basket-P.INAN meat-INS ISG.GEN

‘Who stole my basket of meat?’

In this case, *ari* *bari* is marked as an instrumental since it is an attribute of sorts to *kasu*: the instrumental NP describes what its antecedent contains or entails more specifically: it is a basket *with* meat in it. Note, however, that this comitative use of the instrumental is different from mere accompaniment. Thus, it is not possible to say

- (32) \**Ang sahaya Ajān ri Pila.*  
 ang saha-ya Ø Ajān ri Pila  
 AT come-3SG.M TOP Ajān INS Pila

to express ‘Ajān comes (together) with Pila’. The sentence in (32) would instead imply that Pila helps Ajān to come, for example, because he has a sprained ankle and thus needs support to get around. To express accompaniment, instead, the preposition *kayvo* ‘with, along, beside’ has to be used; the prepositional object appears in the locative case:

- (33) *Ang sahaya Ajān kayvo ya Pila.*  
 ang saha-ya Ø Ajān kayvo ya Pila  
 AT come-3SG.M TOP Ajān with LOC Pila

‘Ajān comes (together) with Pila.’

Theoretically, it should be possible as well to use the instrumental together with prepositions for some kind of prolocative meaning. The adposition would indicate the place *by way of* a motion is happening:

- (34) *Ang pukay manga luga labaneri.*  
 ang puk=ay.Ø manga luga lahan-eri  
 AT jump=1SG.TOP DIR top fence-INS

‘I jump over the fence.’

This use of the instrumental is unattested in previous translations into Ayeri, however, but could be considered a stylistic alternative—in the case of the example above, to a construction with the word for ‘over’, *eyrarya*:

- (37) a. ... *nay ang mya rankyon sitanyaās ku-netu.*  
 ... *nay ang mya rank=yon.Ø sitanya-as ku=netu*  
 ... *and AT be.supposed.to treat=3PL.N.TOP each.other-P like=brother*  
 ‘... and they shall treat each other like brothers.’<sup>6</sup> (Becker 2011a)
- b. ... *ang nunaya ku-vipin* ...  
 ... *ang nuna=ya.Ø ku=vipin* ...  
 ... *AT fly=3SG.M.TOP like=bird* ...  
 ‘... he (would) fly like a bird ...’ (Becker 2012: 14)

Strikingly, in example (37a),  $\text{netu}$  ‘brother’ in  $\text{ku-netu}$  ‘like brothers’ is not even inflected for plural; likewise,  $\text{ku-vipin}$  ‘like a bird’ in (37b) is not inflected for case. Both of these NPs form adverbials of manner, and this apparently prevented case marking in both cases. The instrumental as the case used to mark the how—means or helpers by which an action is brought about—would be expectable here.

Nouns may also be unmarked if they act as modifiers in a compound and the head is marked for the NP’s case and number, for instance:

- (38) *ralanyeri mapang*  
 ralan-ye-ri mapang  
 nail-PL-INS finger  
 ‘with the fingernails’

And lastly and probably most importantly, nouns appear superficially unmarked if topicalized, since the topic marker is  $-\emptyset$  if viewed systematically:

- (39) *Saru-nama, ang nupoyya veney aruno vās.*  
 sar-u=nama ang nupa-oy-ya veney- $\emptyset$  aruno vās  
 go-IMP=just AT hurt-NEG-3SG.M dog-TOP brown 2SG.P  
 ‘Just go, the brown dog won’t hurt you.’

#### 4.1.4 Prefixes on nouns

All of the nominal morphology we have so far dealt with in this section was suffixing. As mentioned in the previous section already (p. 67), however, there are also a number of prefixes that can be applied to nouns. I have just given two examples of the prefix  $\text{ku-}$  ‘like, as though’ above, but  $\text{ku-}$  applies not only to nouns, but can be combined with other parts of speech as well. As discussed in section 3.2.5 (p. 86 ff.), it behaves in the way of a special clitic in Zwicky (1977)’s terminology, since no corresponding full form exists in its place. To cite from the Ayeri translation of Kafka’s short story “Eine kaiserliche Botschaft” again:

- (40) ... *saylingyāng kovaro naynay, ku-ranyāng palung.*  
 ... sayling=yāng kovaro naynay ku=ranya-ang palung  
 ... progress=3SG.M.A easy also like=nobody-A else  
 ‘... he also got on easily, like nobody else.’ (Becker 2012: 12)

<sup>6</sup> The original English text this was translated from has “and should act towards one another in a spirit of brotherhood” (United Nations 1948: Article 1).

(41) a. *Ang lentava sa Tagāti diyan-ku.*  
ang lenta=va.Ø sa Tagāti diyan=ku  
AT sound=2.TOP P Tagāti worthy=like  
‘You sound like Mr. Tagāti.’

b. *Ang lentava sa Tagāti diyan-ku nay diranas yana.*  
ang lenta=va.Ø sa Tagāti diyan=ku nay diran-as yana  
AT sound=2.TOP P Tagāti worthy=like and uncle-P 3SG.M.GEN  
‘You sound like Mr. Tagāti and his uncle.’

c. *Sa lentavāng ku-Tagāti diyan.*  
sa lenta=vāng ku=Tagāti diyan  
PT sound=2.A like=Tagāti worthy  
‘Like Mr. Tagāti you sound.’

(42) a. *da-nanga kāryo*  
           *da=nanga kāryo*  
           such=house big  
           ‘such a big house’

      b. *edāyon nake*  
           eda=ayon nake  
           this=man tall  
           ‘this tall man’

For the purpose of gaining at least a little insight into which types of compounds Ayer allows—besides endocentric compounds—I conducted a small (non-exhaustive) survey based on 130 compounds from the Ayer dictionary (Becker

Table 4.5: Compounds in the Ayeri dictionary (Becker 2016a) and their classification (n = 130)

Type	Harmonic		Disharmonic		Total	
Endocentric (N + N)	67	51.54%	2	1.54%	69	53.08%
Endocentric (N + Adj)	18	13.85%	4	3.08%	22	16.92%
Synthetic (V + N)	16	12.31%	4	3.08%	20	15.38%
Coordinative (N + N)	9	6.92%	—		9	6.92%
Exocentric (N + N)	1	0.77%	3	2.31%	4	3.08%
Unclear	6	4.62%	—		6	4.62%
Total	117	90.00%	13	10.00%	130	100%

2016a: Dictionary); Table 4.5 shows the various compound classes and the number of words for each. ‘Harmonic’ and ‘disharmonic’, respectively, refer to the order of elements; the order is ‘harmonic’ if it is following the normal constituent order of the language and ‘disharmonic’ if it is at odds with it (Gaeta 2008).

Unsurprisingly, the largest number of compound nouns in the sample were endocentric compounds of the regular kind, which means that, just like genitive attributes follow nouns, noun compounds are headed left. Especially compounds with adjectives are interesting insofar as this is also the normal order for free adjectives, so to illustrate, some tests will be necessary to show that these adjectives form a unit with the head noun and are unable to undergo comparison, for instance. Synthetic compounds exist in Ayeri and produce nouns. These are compounds in which “the modifying element in the compound is (usually) interpreted as an argument of the verb from which the head is derived” (Bauer 2001: 701). There are also a number of coordinative compounds; this group, however, is lexicalized and not productive. Exocentric compounds constitute the minority of the sample. In the following I will give examples for each type.

It needs to be noted that unlike Germanic languages, Ayeri does not allow compounds of arbitrary length to be strung together, like in the following ridiculous but no less real example from (former) German legislation (see, for instance, *Süd-deutsche Zeitung* 2013):

- (44) *Rindfleischetikettierungsüberwachungsaufgabenübertragungsgesetz* [German]  
 Rind-Fleisch-Etikettierung-s-Überwachung-s-Aufgabe-n-Übertragung-s-Gesetz  
 cow-meat-labeling-LNK-surveillance-LNK-duty-LNK-delegation-LNK-law  
 ‘Law on the delegation of duties in the surveillance of beef labeling’

In stark contrast, Ayeri allows only two elements in compounds. Furthermore, this section on compounds is located within the section on nouns because Ayeri



almost only possesses compounds involving nouns, and the majority of these also results in a noun.

#### Endocentric compounds

To start with the largest group, endocentric/*tatpuruṣa* compounds, the bulk of these compounds combines two nouns, one of which is the head which is modified by a dependent noun. As Ayeri exhibits a rather strict head-first word order, it comes as no surprise, following Gaeta (2008), that most of these compounds follow this order strictly: the second noun modifies the first, which is opposite of how English, for instance, typically operates:

- (45) a.  $\text{betaynimpur}$  'grape' ←  $\text{betay}$  'berry' +  $\text{nimpur}$  'wine'  
 b.  $\text{karirayan}$  'vertigo' ←  $\text{kar}$  'fear' +  $\text{irayan}$  'height'<sup>7</sup>  
 c.  $\text{pikunanding}$  'mustache' ←  $\text{piku}$  'beard' +  $\text{nanding}$  'lips'  
 d.  $\text{tapayperin}$  'sunblind' ←  $\text{tapay}$  'screen' +  $\text{perin}$  'sun'

The example words in (45) show that the relationships between the modifier and the head are various: a grape is a berry *used* to make wine from (compare Bauer 2001: 702); vertigo is the fear *of* height; a mustache is a beard *located* over the lips (702); and a sunblind is a screen *against* the sun. Bauer (2001) mentions that "there may be special morphophonemic processes which apply between the elements of compounds", such as "phonological merger[s] between the elements of the compound" (695). This also occasionally happens in Ayeri, as the next few example words show:

- (46) a.  $\text{avararan}$  'wetland'  
 ←  $\text{avan}$  'ground' +  $\text{raro}$  'wet' +  $\text{-an}$  (NMLZ)  
 b.  $\text{mehimitrang}$  'fiber tree'  
 ←  $\text{mehir}$  'tree' +  $\text{mitrang}$  'hair, fiber'  
 c.  $\text{ningampinam}$  'bedtime story'  
 ←  $\text{ningan}$  'story' +  $\text{pinam}$  'bed'  
 d.  $\text{padilamican}$  'gravitational force'  
 ←  $\text{padilan}$  'attraction' +  $\text{mican}$  'force, power'

There is a modicum of alteration happening in all of the heads of the example words in (46), mostly nasals assimilating to the stop or nasal which the modifier begins with ( $/n/ + /p/ \rightarrow /mp/$ ,  $/n/ + /m/ \rightarrow /m/$ ), though  $\text{avararan}$  and  $\text{mehimitrang}$  even delete whole coda segments. Bauer (2001: 703) notes that very commonly, genitive and plural markers may form linking elements,

<sup>7</sup>  $\text{irayan}$ , however, is a transparent nominalization of  $\text{iray}$  'high'.

though he also gives examples of languages which allow other case markers on the modifying element in languages with head-right order; individual languages may allow even more case inflection. However, this appears not to happen in Ayeri. The only element that comes up time and again in between the two halves of compounds is the nominalizer *ဆ်* *-an*, which signifies that the head is being formed by a nominalized root, such as in *ပုၤလၢမိၤဆ်* *padilamican*, where *ပုၤလၢမိၤ* *padilan* ‘attraction’ is a nominalization of *ပုၤလၢမိၤ* *padil-* ‘attract’, or in *နီၤဂၢၢ်ဆ်* *ningampinam*, where *နီၤဂၢၢ်* *ningan* ‘story’ is derived from the verb *နီၤဂၢၢ်* *ning-* ‘tell’. However, since Ayeri is head-first and possessive phrases are dependent marking, genitive or other case marking would be expected on the second element, not the first. Case marking on a compound, however, does not inflect just the modifier, but the whole NP:

- (47) *Ang ningya sipikanena koyababisena.*  
 ang ning-ya sipik-an-ena koyabahis-ena  
 AT talk-3SG.M.TOP keep-NMLZ-GEN book.day-GEN  
 ‘He talks about keeping a journal.’

*နီၤပုၤလၢမိၤဆ်* *koyababisena* in this example is not to be interpreted as ‘book of day(s)’ but as ‘of a day-book’. Inflection between the parts of a compound can happen nonetheless, though. In compounds which are formed *ad hoc* or which are otherwise transparent in their composition (‘loose’ compounds), inflection often is deferred to the noun head noun instead of the edge of the compound as a whole; the modifier is treated as an adjunct in this case, and stays uninflected:

- (48) *Sa trayeng tipin ralanyeri mapang yena.*  
 sa tra=yeng tipin-Ø ralan-ye-ri mapang yena  
 PT scratch=3SG.F.A itch-TOP nail-PL-INS finger 3SG.F.GEN  
 ‘The itch, she scratches it with her fingernails.’

Besides noun modifiers, there are also compounds where the modifier is an adjective. In classical Sanskrit terminology, this type is called *karmadhāraya* (Bauer 2001: 698–699).<sup>8</sup> Examples in Ayeri include:

- (49) a. *နီၤကၢၤဂီၤကၢၤ* *kandangiray* ‘university’ ← *နီၤကၢၤဂီၤ* *kandang* ‘school’ + *နီၤကၢၤ* *iray* ‘high’  
 b. *မၢၤရၢၤဆၢၤ* *marashari* ‘witticism’ ← *မၢၤရၢၤ* *maras* ‘phrase’ + *ဆၢၤ* *hari* ‘pithy’  
 c. *နီၤလၢၤဆ်* *silvanikan* ‘overview’ ← *နီၤလၢၤ* *silvan* ‘view’ + *ဆ်* *ikan* ‘whole’  
 d. *နီၤပုၤလၢမိၤဆ်* *vipimakārya* ‘crow’ ← *နီၤပုၤလၢမိၤ* *vipin* ‘bird’ + *ဆ်* *makārya* ‘black’

<sup>8</sup> Bauer (2001) also mentions that appositional compounds like *maid-servant*, *woman doctor* and *fighter-bomber* are counted in this category (699). Ayeri, however, does not possess such formations in particular.

In all of these cases, the adjective forms a unified lexeme with the head noun, hence it is not comparable, for example:

- |      |   |   |
|------|---|---|
| (50) | COMPARATIVE   | SUPERLATIVE   |
| a.   | * <i>kardangiray-eng</i><br>kardang-iray=eng<br>school-[high=COMP]<br>'higher-school' | <i>kardangiray-vā</i><br>kardang-iray=vā<br>school-*[high=SUPL]<br>'highest-school' |
| b.   | * <i>marashari-eng</i><br>maras-hari=eng<br>phrase-[pithy=COMP]<br>'pithier-phrase'   | <i>marashari-vā</i><br>maras-hari=vā<br>phrase-*[pithy=SUPL]<br>'pithiest-phrase'   |

In fact, it is possible to form *ကံသင်္ဂရာယ-ဝါ* *kardangiray-vā* and *မာရသရိယ-ဝါ* *marasari-vā*, but they mean 'most universities' and 'most witticisms', that is, *-vā* here does not mark the adjectival part as a superlative form; the suffix modifies the noun-adjective compound as a whole: [school-high]=most, [phrase-pithy]=most. *-eng* 'rather' as a quantifier does not combine with nouns, which is why the first examples in (50ab) are both ungrammatical *per se*.

Since the meaning composed from noun-adjective compounds is often idiomatic, they also cannot be divided as shown above in (48), since a *ကံသင်္ဂရာယ* *kardangiray* 'university' is not a *ကံသင်္ဂ* *kardang* 'school' which is *သင်္ဂ* *iray* 'high' in the literal sense, but a school of the highest tier. *ကံသင်္ဂရာယသင်္ဂ* *kardangena iray* (school-GEN high), then, can only be interpreted in the literal sense, 'of the high school', but not as 'of the university', which thus can only be *ကံသင်္ဂရာယသင်္ဂ* *kardangirayena*.

In the sample, there were also a few compounds I categorized as noun-noun combinations, which look as though they violate head-first order. All of these involve *သိတာ* *sitang* 'self' as a modifier:

- (51) a. *သိတာဂုဏ်သံ* *sitanglentan* 'vowel'  
← *သိတာ* *sitang* 'self' + *ဂုဏ်သံ* *lentan* 'sound'
- b. *သိတာပါဝါသံ* *sitangparonān* 'self-confidence'  
← *သိတာ* *sitang* 'self' + *ပါဝါသံ* *paronān* 'faith'
- c. *သိတာမေတ္တသံ* *sitangtenyan* 'suicide'  
← *သိတာ* *sitang* 'self' + *မေတ္တသံ* *tenyan* 'death'

*သိတာ* *sitang* does not exist as a noun by itself in Ayeri, the word for 'self' is its nominalization *သိတာသင်္ဂ* *sitangan*. Nonetheless, it looks like it could have plausibly been a noun once. However, this noun may have been grammaticalized into a reflexive morpheme of a more general kind, which in turn birthed the form *သိတာသင်္ဂ* *sitangan*

as a renovation.<sup>9</sup> The reflexive *sitang* is used—as we have seen in the previous chapter—as a prefix, so there are two ways to interpret these formations: first, *sitang* may be the reflexive prefix here and thus the compound follows the normal syntactic order; or second, the order of elements is reversed and thus may reflect an earlier stage of Ayeri where *sitang* was still a noun and modifiers could still appear in front of their heads, at least optionally so (Gaeta 2008: 133–137).

There are a number of genuinely reversed endocentric compounds as well, however, in which the modifier comes first and the head last. Since there are only a few of these to date, I will give all of them in the following example:

- (52) a. *baripata* ‘ground meat’  
 ← *bari* ‘meat’ + *pata* ‘mash’  
 b. *kayvolentan* ‘consonant’  
 ← *kayvo* ‘with’ + *lentan* ‘sound’  
 c. *māvaganeng* ‘mother’s siblings’  
 ← *māva* ‘mother’ + *ganeng* ‘siblings’  
 d. *matinanding* ‘labia’  
 ← *matikan* ‘hot’ + *nanding* ‘lips’  
 e. *muyavirang* ‘brass’  
 ← *muya* ‘false’ + *avirang* ‘gold’  
 f. *tonisaytang* ‘self-assured’  
 ← *tonisa* ‘assured’ + *sitangan* ‘self’

Given the discussion of *sitang* above, one word among the examples above that is not too clear is *tonisaytang*, which appears to contain a deviant form of either *sitang* or *sitangan*, which is preceded by the adjective *tonisa* ‘assured, ascertained’.

All of the previously mentioned compounds involving nominal elements formed nouns, though, there are also a few denominal compounds in the sample. This process is not productive, however, and interestingly, only noun–adjective combinations appear in this group:

<sup>9</sup> A little bit of language history would certainly simplify things here and lend them credence. Let us simply assume that *sitang* used to be a noun meaning something like ‘self’ at a previous stage of Ayeri and was repurposed as a reflexive prefix. Lehmann (2015) quotes a few examples of what he calls ‘autophoric’ nouns that came to be used as reflexive pronouns in their respective language: “Typical examples are Sanskrit *tan* ‘body, person’ and *ātmán* ‘breath, soul’, Buginese *elena* ‘body’, Okinawan *dūna* ‘body’, !Xu *l’esi* ‘body’, Basque *burua* ‘head’, Abkhaz *a-xə* ‘the head’. In their respective languages, all these nouns are translation equivalents of English *self*” (45–46). Thus, it would not be out of line at all to assume such a grammaticalization path for Ayeri as well.

- (53) a. ၶီၵၢၵ်း *mirampaluy* ‘otherwise’  
 ← ၶီၵ်း *miran* ‘way’ + ၵၢၵ်း *palung* ‘different’
- b. ပါဒဃ *padabanya* ‘insane’  
 ← ပါး *padang* ‘mind’ + ဃ *banaya* ‘sick’
- c. တဲနရိး *tenkarisa-* ‘be frightened to death’  
 ← တဲ *ten* ‘life’ + ရိး *karisa* ‘frightened’

ၶီၵၢၵ်း *mirampaluy* is an adverb, the modifier probably a mangling of ၵၢၵ်း *palung*; ပါဒဃ *padabanya* is an adjective meaning ‘insane’ rather than the expected ‘insanity’ (instead: ပါဒဃ *padabanyān*); and တဲနရိး *tenkarisa-* acts as a verb, possibly from conversion or reinterpretation, since the suffix -*isa* also forms morphological causatives of a number of verbs. Besides these irregularities, there is also at least one noun compound which uses a postposition as an adjectival modifier:

- (54) ၶီၵ်း *silvankayvay* ‘blindness’ ← ၶီၵ်း *silvan* ‘sight’ + ကယ *kayvay* ‘without’

This compound must be derived from the phrase ၶီၵ်း *silvanya kayvay* ‘without sight’ (see-NMLZ-LOC without), though here as well, the word roots are simply juxtaposed, as noted above is the common way to form compounds in Ayeri.

#### Synthetic compounds

According to Bauer (2001), (semi-)synthetic compounds, or verbal(-nexus) compounds, are compounds that have “been variously defined as being based on word-groups or syntactic constructions (Botha 1984: 2), or as compounds whose head elements are derived from verbs (Lieber 1994: 3607)” (Bauer 2001: 701). Examples of this type in English would include *truck-driver*, *peace-keeping*, and *home-made*. He mentions also that synthetic compounds have been mainly discussed with regards to Germanic languages, but that according to Lieber (1994: 3608), the phenomenon is much more widespread. Ayeri possesses compounds like this as well, and the regular case again follows the constituent order, here that of verbs and nouns: Ayeri is a VO language, and thus the verb as the head of the compound is usually found on the left side with its nominal modifier following it (compare Gaeta 2008: 129–133):

- (55) a. ၶီၵ်း *anlāgonan* ‘pronunciation’  
 ← ၶီၵ်း *anl-* ‘bring’ + ၶီၵ်း *agonan* ‘outside’
- b. ဂုၵ်း *napakaron* ‘acid’  
 ← ဂုၵ်း *nap-* ‘burn’ + ၶီၵ်း *karon* ‘water’
- c. ဂုၵ်း *napaperin* ‘sunburn’  
 ← ဂုၵ်း *nap-* ‘burn’ + ၶီၵ်း *perin* ‘sun’
- d. ၶီၵ်း *telbasasān* ‘waysign’  
 ← ၶီၵ်း *telba-* ‘show’ + ၶီၵ်း *sasān* ‘way’

The relations between the verb and the noun are various again, that is, the nominal modifier is not simply the direct object of the verb: to pronounce something means to bring it *to* the outside; a sunburn is a burn *caused by* the sun; and a waysign shows the way (မန္တရ် *sasān* is the object here). Even though ကရုဏ် *karon* may serve as an agent (or a causer) of the burning effect of acid (similarly for နာပာပရိ *napaperin* ‘sunburn’), the verb-first order is justified here as well, since verbs always go first in Ayeri sentences, and any other NPs, whether actor or undergoer, are following.

Just as with endocentric compounds, there are a number of seeming exceptions to the verb-first order of synthetic compounds as well. These are just as far and few between, however, and whether they should all be counted as noun–verb combinations is also questionable as they appear to all be formed with nominalized verbs. The verbal element may thus be only indirectly verbal for the purposes of compounding. If interpreted as noun–noun combinations, the nominal first element would reasonably form the head again for some of the below example words.

- (56) a. မာရိပုဏ္ဏယံ *maripuntayam* ‘spread’  
 ← မာရ် *marin* ‘surface’ + ပုဏ္ဏ *punta-* ‘stroke’ + ယံ *-yam* (DAT)
- b. မာရ်နီယံ *sasanlekān* ‘labyrinth’  
 ← မာရ် *sasān* ‘way’ + နီယံ *leka-* ‘guess’ + ယံ *-an* (NMLZ)
- c. မာရ်နီယံ *selangnunān* ‘plane’  
 ← မာရ် *selang* ‘machine’ + နီယံ *nuna-* ‘fly’ + ယံ *-an* (NMLZ)
- d. မာရ်နီယံ *sinturān* ‘radio’  
 ← မာရ် *sinto* ‘wave’ + နီယံ *tura-* ‘send’ + ယံ *-an* (NMLZ)

မာရ်ပုဏ္ဏယံ *maripuntayam* is special in that it contains the dative suffix ယံ *-yam* which is lexicalized as part of the word: something made or intended for spreading on a surface. A few more such verbal derivations can be found, though not compounds, among others:

- (57) a. မာရ်ဂရီယံ *grenyam* ‘extremity’ ← မာရ် *gren-* ‘reach out’  
 b. မာရ်လူယံ *lugayam* ‘password’ ← မာရ် *luga-* ‘go through’  
 c. မာရ်ယံ *sabayam* ‘future’ ← မာရ် *saba-* ‘come’

There is also မာရ်ပုဏ္ဏ *maripunta-* ‘spread over’ as the verb corresponding to မာရ်ပုဏ္ဏယံ *maripuntayam*, though its meaning is less specific. Here as well, however, the verbal part is last instead of first. For the other example words (56b–d), an interpretation of the second part as a deverbal noun is possible: a labyrinth as a way or path which requires guessing, a plane a machine for flight, and radio as a sending of waves. In the latter case, မာရ်နီယံ *sinturān*, however, the head is still on the wrong side even if one interprets all of the above examples as noun–noun compounds with a deverbal element.

*Coordinative compounds*

Coordinative compounds are a very small group among the sample drawn from the dictionary, and not a very productive one. Bauer (2001) defines this class as having “two or more words in a coordinate relationship, such that the entity denoted is the totality of the entities denoted by each of the elements” (699). He cautions that they are very easy to confuse with appositional (also *karmadhāraya*) compounds in that both types of compound allow inserting an *and* between both elements. The following nominal coordinative compounds are included in the dictionary sample:

- (58) a.  $\text{bāmā}$  ‘mom-and-dad’  
 ←  $\text{bā(bā)}$  ‘dad’ +  $\text{mā(mā)}$  ‘mom’  
 b.  $\text{pruynapay}$  ‘seasoning’  
 ←  $\text{pruy}$  ‘salt’ +  $\text{napay}$  ‘pepper’  
 c.  $\text{sapayyila}$  ‘hands-and-feet’  
 ←  $\text{sapay}$  ‘hand’ +  $\text{yila}$  ‘foot’  
 d.  $\text{simileno}$  ‘horizon’  
 ←  $\text{simil}$  ‘country’ +  $\text{leno}$  ‘sky’  
 e.  $\text{sitemrugon}$  ‘thunderstorm’  
 ←  $\text{sitem}$  ‘lightning’ +  $\text{rugon}$  ‘thunder’  
 f.  $\text{vekamdekey}$  ‘dishes’  
 ←  $\text{vekam}$  ‘plate’ +  $\text{dekey}$  ‘fork’

None of the two elements recognizably forms the head in these examples, but both elements are typical exponents of the thing the compound signifies. Bauer (2001: 699) mentions that coordinative adjective compounds are rare, or at least rarely documented in the grammars he surveyed. In the sample I took, only the following compound is included, which forms a noun from the combination of two adjectives, insofar it is relevant to this section even though the component parts are not nouns:

- (59)  $\text{makagisu}$  ‘twilight’ ←  $\text{maka}$  ‘light’ +  $\text{gisu}$  ‘dark’

The sample also includes the following two words, however, which are neither made up from nouns, nor do they form a noun in combination. Instead, they are technically verbs combining to form directional adverbs and have been exceptionally included here for completeness:

- (60) a.  $\text{mangasaba}$  ‘towards’ ←  $\text{manga-}$  ‘move’ +  $\text{saba-}$  ‘come’  
 b.  $\text{mangasara}$  ‘away’ ←  $\text{manga-}$  ‘move’ +  $\text{sara-}$  ‘go’

*Exocentric compounds*

In exocentric compounds, the modifier is not a hyponym of its head (Bauer 2001: 700), which means that the modifier is not describing a property that more closely determines its head. So while a *dog kennel* is a type of kennel made for dogs, the head of an *egghead* is neither for eggs, nor containing eggs, nor made of eggs; instead, it refers to an egg-shaped skull metaphorically. And while a *bluecollar* may wear a blue shirt professionally, the referent it signifies is not a type of collar, but the relationship is metonymical in that the blue collar is part of the guise of the signified entity as a whole. The sample from the Ayeri dictionary contains a few compounds of this kind as well, though again, it is not a very productive group:

- (61) a. ၼံၼ်ၼ်ၼ် *avanyonang* 'artery'  
           ← ၼံၼ် *avan* 'bottom, down' + ၼ်ၼ် *yonang* 'stream'
- b. ၼံၼ်ၼ် *baytandevo* 'headache'  
           ← ၼံၼ် *baytang* 'blood' + ၼ်ၼ် *devo* 'head'
- c. ၼံၼ်ၼ် *linyonang* 'vein'  
           ← ၼံၼ် *ling* 'top, up' + ၼ်ၼ် *yonang* 'steam'
- d. ၼံၼ်ၼ် *sindaynanga* 'address'  
           ← ၼံၼ် *sinday* 'number' + ၼ်ၼ် *nanga* 'house'

What is striking here is that only one out of four examples shows the expected head-left order: ၼံၼ်ၼ် *sindaynanga*. The other three examples all have the head component on the right side, preceded by a modifier. However, what all of these have in common, is that they are only metaphorically or metonymically describing the thing they signify: veins and arteries are not literally streams going up or down (they are a kind of stream flowing in different directions, however, so these are probably on the borderline between exocentric and endocentric); a headache is related to the head, but has not directly to do with being made of or containing blood (the rationale behind this being a superstition that you have too much blood in your head, which is said to cause the pain); and a house number may be part of an address, but is in a *pars pro toto* relationship to it.

*A few mysterious cases*

The following words from my sample were either undeterminable as to their composition due to parts of the word not being clear regarding one of their constituent parts, either because I tweaked the constituent so much as to not be readily recognizable anymore, or because I forgot to make an entry in the dictionary, or even deleted or changed that. The words in question are the following:



- (62) a. ခါဇာန်ဗဲ *batangiman* ‘mosquito’ ← ခုဇာန် *baytang* ‘blood’ + ?  
 b. ခဲင်ဇာန် *kirinalang* ‘avenue’ ← ခဲင် *kirin* ‘street’ + ?  
 c. ခဲင်ဇာန် *ningambakar* ‘telltale’ ← ခဲင် *ningan* ‘story’ + ?  
 d. ဂဗျန် *ragayesuy* ‘grid’ ← ဂဗျန် *ragan* ‘line’ + ?  
 e. မဲဒုဗဲ *teraymino* ‘melancholic’ ← ? + မဲဒု *mino* ‘happy’  
 f. ဖဲဒုဗဲ *vetaysano* ‘fare’ ← ? + မဲဒု *sasān* (earlier မဲဒု *sasano*) ‘way’

For all of the components represented by a question mark, there is no corresponding dictionary entry. At least in ခုဇာန်ဗဲ *baytangiman*, the \*မဲဒု *\*iman* part looks as though it could be a noun due to the -*an* nominalizer suffix. \*မဲဒု *\*teray* in မဲဒုဗဲ *teraymino* might also be an adjective supposed to mean ‘sad’ (which would make it an adjectival coordinative compound), although the dictionary entry for that is မဲဒု *giday*. Even though parts of all these words are unclear, they all seem to follow the correct syntactic order, judging by those parts that are identifiable. And even in the case of ဖဲဒုဗဲ *vetaysano*, which is missing the first part, it can be reasonably assumed that the identifiable part, \*မဲဒု *\*sano*, is the modifier, and \*မဲဒု *vetay* may have once been intended to mean ‘money’ or ‘fee’ or something along these lines.

With the exception of ခဲင်ဇာန် *ningambakar*, all of the mystery words were entered into the dictionary in 2006. Digging through old archives and translations, I could determine at least that \*ခဲင် *\*bakar* was once intended to mean ‘lie’, and \*မဲဒု *\*teray* was indeed meant to mean ‘sad’.

#### 4.1.6 Reduplication

Wiltshire and Marantz (2000) write that it has been suggested that reduplication serves an iconic function, “with the repetition of phonological material indicating a repetition or intensity in the semantics” (561), so with regards to nouns it mainly serves to indicate plurality of various kinds. However, they find that in fact, reduplication serves all kinds of functions, also ones without iconic meanings, and mention Agta, an Austronesian language of the Philippines, which uses reduplication to form diminutives (Healey 1960: 6–9). As we have seen in section 3.2.3 above, so does Ayeri, and it is justified in doing so since there is real-world evidence for this use of reduplication. Examples for diminutive reduplication in Ayeri include:

- (63) a. လိမ္မ *limu* ‘shirt’ → လိမ္မလိမ္မ *limu-limu* ‘little shirt’  
 b. ဘာ *nanga* ‘house’ → ဘာဘာ *nanga-nanga* ‘little house’  
 c. မာပ *sapay* ‘hand’ → မာပမာပ *sapay-sapay* ‘little hand’  
 d. ဖဲဒု *veney* ‘dog’ → ဖဲဒုဖဲဒု *veney-veney* ‘little dog’

Diminutive reduplication involves full stem reduplication in Ayeri. Besides the productive use of reduplication for diminutive marking, there are a number of diminutive formations which have been lexicalized, such as in the following examples:

- (64) a. ၼ်ႇ ၼ်ႇ *agu* ‘chicken’ → ၼ်ႇၼ်ႇ *agu-agu* ‘chick’  
 b. ၼ်ႇ *gan* ‘child’ → ၼ်ႇၼ်ႇ *gan-gan* ‘grandchild’  
 c. ၼ်ႇ *pasing* ‘tube’ → ၼ်ႇၼ်ႇ *pasing-pasing* ‘straw’  
 d. ၼ်ႇ *poyu* ‘cheek; bacon’ → ၼ်ႇၼ်ႇ *poyu-poyu* ‘butt’

There are also at least two documented cases where the reduplicated root is not a noun, but the reduplication results in a noun:

- (65) a. ၼ်ႇ *kusang* ‘double (adj.)’ → ၼ်ႇၼ်ႇ *kusang-kusang* ‘model’  
 b. ၼ်ႇ *veh-* ‘build’ → ၼ်ႇၼ်ႇ *veha-veha* ‘tinkering’

Reduplicated nouns behave like regular nouns with regards to inflection, that is, they receive prefixes and suffixes just like the simplexes from which they are derived:

- (66) *Puco mino veney-veneyang.*  
 puk-yo mino veney~veney-ang  
 jump-3SG.N happily DIM~dog-A

‘The little dog is jumping happily.’

In this example, the reduplicated noun ၼ်ႇၼ်ႇ *veney-veney* is marked as an agent in that the agent suffix ၼ်ႇ *-ang* is appended to the noun as a unit *after* reduplicating the noun stem. In other words, the following formation in which the root is reduplicated along with its declension suffix is ungrammatical for the purpose of forming a diminutive:

- (67) \**ၼ်ႇၼ်ႇ ၼ်ႇ *veneyang-veneyang**

Likewise, the reduplicated form is not treated in the way an endocentric compound would be, so case and plural marking cannot be appended to the first element:

- (68) \**ၼ်ႇၼ်ႇ ၼ်ႇ *veneyang veney**

While ordinary nouns undergo full reduplication to form a diminutive, in compounds, only the head is reduplicated, unless the compound is strongly lexicalized or has an idiomatic meaning going beyond that of its components. The following example shows the simple case of a transparent endocentric compound:

- (69) *Ya yomayo mehir-mehirang seygo veno kay pang nanga nana.*  
 ya yoma-yo mehir~mehir-ang seygo veno kay pang nanga-Ø nana  
 LOCT be-3SG.N DIM~tree-A apple pretty three back house-TOP IPL.GEN

‘There are three pretty little apple trees behind our house.’

In this example, being endearing or otherwise small is treated as a property of the head, မ်းမိာ် *mehir* ‘tree’, not of the whole compound မ်းမိာ်နီၤ *mehirseygo* ‘apple tree’, or the dependent, နီၤ *seygo* ‘apple’—after all, an apple tree which is small is rather a small tree with apples on it than a tree with small apples. The avoidance of the fully reduplicated form မ်းမိာ်နီၤမ်းမိာ်နီၤ *mehirseygo-mehirseygo* is probably related to the notion of economy of expression.

#### 4.1.7 Nominalization

Some accidental ways of deriving nouns have been mentioned above, for instance, some reduplicated non-nominal roots like နီၤနီၤ *kusang* ‘double’ or ဖ်းဃုး *veha-* ‘build’ may form nouns. However, Ayeri also has some dedicated morphology to derive nouns from other parts of speech. The most common and highly productive way to derive a noun, is the suffix -an. The examples in (70) illustrate some derivations from verbs, and (71) shows derivations from adjectives to nouns. As နီၤနီၤ *kuban* ‘oar’ shows, the nominalization may have an idiomatic meaning.

- (70) a. ခာၣ်နီၤ *balang-* ‘search (v.)’ → ခာၣ်နီၤနီၤ *balangan* ‘search (n.)’  
 b. နီၤနီၤ *kub-* ‘row’ → နီၤနီၤနီၤ *kuban* ‘oar’  
 c. ဂိၣ်း *rig-* ‘draw’ → ဂိၣ်းနီၤ *rigan* ‘drawing’  
 d. ဖ်းဃုး *veh-* ‘build’ → ဖ်းဃုးနီၤ *vehan* ‘building’
- (71) a. မ်းမိာ် *apitu* ‘pure’ → မ်းမိာ်နီၤ *apituan* ‘purity’  
 b. ဂိၣ်း *gira* ‘urgent’ → ဂိၣ်းနီၤ *girān* ‘hurry’  
 c. ပါၣ်း *pakis* ‘serious’ → ပါၣ်းနီၤ *pakisan* ‘seriousness’  
 d. ဝါၣ် *vapa* ‘skillful’ → ဝါၣ်နီၤ *vapan* ‘skill’

Occasionally, it may even happen that a noun is derived from a noun with a related but sometimes more basic meaning using the nominalizer -an. This process, however, is not productive, so compared to deverbalization and deadjectivization, examples of this derivation strategy are few.

- (72) a. မ်းမိာ် *ajam* ‘toy’ → မ်းမိာ်နီၤ *ajaman* ‘game’  
 b. နီၤနီၤ *kelang* ‘chain’ → နီၤနီၤနီၤ *kelangan* ‘connection’  
 c. နီၤ *nanga* ‘house’ → နီၤနီၤ *nangān* ‘household’  
 d. မ်း *ten* ‘life’ → မ်းနီၤ *tenan* ‘soul’

There are also some apparent nominalizations in :ဆ်း -*am* and :ဆ်း -*ang*, although these are irregular and non-productive:

- (73) a. :ဆ်း *aja*- ‘play’ → :ဆ်း *ajam* ‘toy’  
 b. :ဆ်း *gin*- ‘drink’ → :ဆ်း *ginam* ‘glass’  
 c. :ဆ်း *mik*- ‘poison (v.)’ → :ဆ်း *mikam* ‘poison (n.), venom’  
 d. :ဆ်း *nuna*- ‘fly’ → :ဆ်း *nunam* ‘feather’
- (74) a. :ဆ်း *bayha*- ‘rule’ → :ဆ်း *bayhang* ‘government’  
 b. :ဆ်း *bapa* ‘remaining’ → :ဆ်း *bapang* ‘remainder’  
 c. :ဆ်း *kada*- ‘collect’ → :ဆ်း *kadang* ‘committee; alliance’  
 d. :ဆ်း *mima* ‘possible’ → :ဆ်း *mimang* ‘access’

Agentive nouns can be formed from regular nouns with the suffix :မာယာ -*maya*, compare the examples in (75). An epenthetic /a/ may be introduced to break up consonant clusters that would otherwise be either difficult to pronounce or violating phonotactics. When the stem of the word the agentive suffix is attached to ends in a consonant or /Ca/, it is also often found fused with the root, sometimes with the first /a/ of -*Caya* lengthened, see (76). Specifically feminine agentive nouns can be derived with the related suffix :မာယာ -*vaya*; two examples are given in (77).

- (75) a. :ဆ်း *anl*- ‘bring’ → :ဆ်း *anlamaya* ‘waiter’  
 b. :ဆ်း *hora* ‘sin’ → :ဆ်း *horamaya* ‘sinner’  
 c. :ဆ်း *nasy*- ‘follow’ → :ဆ်း *nasyamaya* ‘follower’  
 d. :ဆ်း *teba*- ‘bake’ → :ဆ်း *tebamaya* ‘baker’
- (76) a. :ဆ်း *asa*- ‘travel’ → :ဆ်း *asāya* ‘traveler’  
 b. :ဆ်း *ibut*- ‘trade’ → :ဆ်း *ibutaya* ‘trader, merchant’  
 c. :ဆ်း *lant*- ‘lead’ → :ဆ်း *lantaya* ‘leader; driver’  
 d. :ဆ်း *tang*- ‘listen’ → :ဆ်း *tangaya* ‘listener’
- (77) a. :ဆ်း *gan* ‘child’ → :ဆ်း *ganvaya* ‘governess’  
 b. :ဆ်း *lanya* ‘king’ → :ဆ်း *lanvaya* ‘queen’

Besides these, there is also a derivative suffix for makers of things, :ဆ်း -*ati* (contracting to /atʃ/ -*ac* before a vowel), though this is not too productive, and sometimes irregular, as :ဆ်း *sirtangati* ‘youth’ shows:

- (78) a. :ဆ်း *gindi* ‘poem’ → :ဆ်း *gindati* ‘poet’  
 b. :ဆ်း *sirtang* ‘young’ → :ဆ်း *sirtangati* ‘youth’  
 c. :ဆ်း *taban*- ‘write’ → :ဆ်း *tabanati* ‘scribe’  
 d. :ဆ်း *vehim* ‘piece of clothing’ → :ဆ်း *vehimati* ‘tailor’

A few instances also exist where a tool of sorts is derived with a suffix  $\text{ᠠᠨ}$  *-an*, which is related to the instrumental suffix  $\text{ᠡᠷᠢ}$  *-eri* in combination with the nominalizer  $\text{ᠠᠨ}$  *-an*:

- (79) a.  $\text{ᠭᠣᠷ}$  *gur-* ‘turn’ →  $\text{ᠭᠣᠷᠠᠨ}$  *guryan* ‘coil, cylinder’  
 b.  $\text{ᠮᠢᠰ}$  *mis-* ‘behave’ →  $\text{ᠮᠢᠰᠢᠷᠠᠨ}$  *miseryan* ‘method, strategy’  
 c.  $\text{ᠨᠠᠫ}$  *nap-* ‘burn’ →  $\text{ᠨᠠᠫᠢᠷᠠᠨ}$  *naperyan* ‘tinder’  
 d.  $\text{ᠫᠠᠷ}$  *pra-* ‘glitter, gleam’ →  $\text{ᠫᠠᠷᠢᠷᠠᠨ}$  *praryan* ‘spark’

While  $\text{ᠠᠨ}$  *-an* derives nouns from verbs to produce nouns that act as such in every way, it may sometimes be preferable to refer to the action as such by a noun, compare in English:

- (80) a. Manhattan is famous for its tall **buildings**.  
 b. **Building** a house is an expensive endeavor.

In (80a), *building* is simply a noun derived from the verb *build*. It acts as a noun in every way, for example, it can serve as a subject and object, it can be pluralized, it can take determiners, and can be modified by adjectives. The form of *building* in (80b), however, is a gerund, and as such underlies the restriction that it cannot be pluralized (Payne 1997: 35). As we have seen at the beginning of this section on nominalization, Ayeri can derive  $\text{ᠪᠡᠬᠠᠨ}$  *vehan* ‘building, construction’ from the verb  $\text{ᠪᠡᠬ}$  *veh-* ‘build’, which acts like every other common noun, much like in the English example in (80a):

- (81) a. *Lesāra*                      *sirimang*   *vehānreng*                      *tado*.  
 lesa-ara                      sirimang   vehān-reng                      tado  
 collapse-3SG.INAN   about.to   building-A.INAN   old  
 ‘The old building is about to collapse.’  
 b. *Le*                      *vacyang*                      *eda-vehān*.  
 le                      vac=yang                      eda=vehān-Ø  
 PT.INAN   like-1SG.A   this=building-TOP  
 ‘This building, I like it.’  
 c. *Ang latayo*                      *bayhang*                      *vehānyeley*                      *yona*.  
 ang   lata-yo                      bayhang-Ø                      vehān-ye-ley                      yona  
 AT   sell-3SG.N   government-TOP   building-PL-P.INAN   3SG.N.GEN  
 ‘The government is selling its buildings.’  
 d. *Le*                      *ming*   *kuysāran*                      *vehān-kay*                      *dirasyam*                      *ran*.  
 le                      ming   kuysa-aran                      vehān-Ø=kay                      diras-yam                      ran  
 PT.INAN   can   compare-3PL.INAN   building-TOP=few   splendor-DAT   3SG.INAN.GEN  
 ‘Few buildings can compare to its splendor.’

The above examples condense several properties into one for illustration. Thus, (81a) shows that ၵျံး *vehān* can serve as basically a subject of a clause, and that it can as well be modified by an adjective—the choice of adjectives is not subject to any distributional restrictions other than those imposed by the semantic frame of HOUSE. In the next example, (81b), ၵျံး *vehān* serves as the object of the clause and is being determined by the demonstrative prefix ၵိ *eda-* ‘this’. The third example, (81c), shows ၵျံး *vehān* both pluralized and modified by a possessive pronoun, ၵိ *yona* ‘of it’. And finally, in (81d) we see ၵျံး *vehān* quantified by the suffix ၵိ *-kay* ‘few’.

Similar to the English example in (80b), Ayeri can also derive nouns from the participle of a verb describing the action as such—a gerund. For an example, I will again draw on the Ayeri translation of Kafka’s short story “Eine kaiserliche Botschaft” (Becker 2012: 2, 14):

- (82) ... *nay ang pətangongva*                      *ankyu haruyamanas*                      *nanang megayena*  
 ... *nay ang pə-tang-ong=va.Ø*                      *ankyu haru-yam-an-as*                      *nanang mega-ye-na*  
 ... and AT NFUT-hear-IRR=2SG.TOP truly beat-PTCP-NMLZ-P great fist-PL-GEN  
*yana kunangya vana.*  
*yana kunang-ya vana*  
 3SG.M.GEN door-LOC 2SG.GEN

‘... and you would indeed hear his magnificent beating at your door very soon.’

The annotations to this translation contain a comment on the grammatical rules which operate in this passage, more specifically also on the gerund derivation ၵျံး *haruyaman* ‘beating’:

Furthermore, I wrote *haruyaman* ‘beating’ instead of *haruan* ‘beat(ing)’ because I wanted to emphasize the process of beating as an incomplete action. This is possible here because the word is not topicalized and neither is it marked as a dative, which would also require *haruyamanyam* ‘beat-PTCP-NMLZ-DAT’ to become *haru-anyam* ‘beat-NMLZ-DAT’ (the participle marker *-yam* is derived from the dative case ending *-yam*). (14–15)

We can read from this description that the participle marker in Ayeri has possibly been grammaticalized from the dative case marker, or that it is at least synchronically homonymous. In order for case marking to operate, this formation has to be nominalized, which is done the usual way by appending ၵိ *-an*, thus yielding the suffix cluster ၵျံး *-yaman* for the derivation of verbs as gerunds. If the gerund is marked for dative case, the suffix cluster ၵိ *\*-yamanyam* basically undergoes haplology to a simple nominalized form with the suffix cluster ၵိ *-anyam*:

- (83) *haru-*                      *haruyam*                      *haruyaman*                      *\*haruyamanyam*                      *haruanyam*  
       *haru-* → *haru-yam* → *haru-yam-an*                      → *haru-yam-an-yam* → *haru-an-yam*  
       beat                      beat-PTCP                      beat-PTCP-NMLZ                      beat-PTCP-NMLZ-DAT                      beat-NMLZ-DAT

The comment on the translation also makes a little note on the gerund being possible because the word is not topicalized. This is based on an old rule that gerunds cannot be topicalized unless nominalized first, however, usage has since changed so that earlier,  $\text{ᑭᐱᑦᑭᐱᑦ}$  *haruyam* would have constituted the gerund form, while even by the time of translating the short story, it had changed to  $\text{ᑭᐱᑦᑭᐱᑦ}$  *haruyaman*. Thus, it is no surprise to see the following example, from the partial translation of Saint-Exupéry's story "Le petit prince" (Becker 2015 [2013]: 3, 13):

- (84) *Sa koronyang palungyaman na Baysānterpeng nay na Bayokivo*  
 sa koron=yang palung-yam-an-Ø na Baysānterpeng nay na Bayokivo  
 PT knew=ISG.A distinguish-PTCP-NMLZ-TOP GEN Realm.Middle and GEN Spring.Little  
*menaneri nivānyena.*  
 menan-eri nivān-ye-na  
 first-INS glimpse-PL-GEN

'I knew how to distinguish between China and Arizona at first sight.'

A more literal translation of this sentence would be 'The distinguishing of China and Arizona, I knew it at first sight', so the whole passage  $\text{ᑭᐱᑦᑭᐱᑦ} - \text{ᑭᐱᑦᑭᐱᑦ}$  *palungyaman ... na Bayokivo* forms the topic of the sentence here, headed by the gerund  $\text{ᑭᐱᑦᑭᐱᑦ}$  *palungyaman* 'distinguishing'. According to the old rule obliquely quoted in the comment to the passage in (82), this should not be possible. As said before, though, usage has changed.

A rule we can gather from the above example from Saint-Exupéry is that gerunds are treated as animate nouns. Since they are impersonal, they trigger neuter agreement on verbs. They can also be the objects of sentences. The passage in (82) furthermore illustrates that gerunds can be modified by adjectives. The following example shows a gerund used as an agent—basically a subject—as well (Becker 2014):

- (85) *Dilayamanang kalamena babalanas ayonena ...*  
 dila-yam-an-ang kalam-ena bahalan-as ayon-ena ...  
 find.out-PTCP-NMLZ-A truth-GEN goal-P man-GEN ...

'(If) finding out the truth is the goal of the man ...'

What all the passages on gerunds quoted before show is that gerunds in Ayeri do not behave like transitive verbs as in English. Thus, what would be the object of the former verb appears in the genitive case in Ayeri. As in English, however, gerunds in Ayeri cannot be pluralized:

- (86) \**Noyo vebayamanjang nangayena.*  
 noyo veba-yam-an-ye-ang nanga-ye-na  
 expensive build-PTCP-NMLZ-PL-A house-PL-GEN

'\*The buildings of houses are expensive.'

It is possible, however, to quantify gerunds insofar as the quantifier does not imply countable quantities of the action. Moreover, it is possible for gerunds to be modified by possessors. The following two sentences exemplify this use:

- (87) *Ang lugayan                      delacamanas-ikan                      kayanya pang.*  
 ang luga=yan.Ø                      delak-yam-an-as=ikan                      kayan-ya pang  
 AT go.through=3PL.M.TOP suffer-PTCP-NMLZ-P=much war-LOC after

‘They went through a lot of suffering after the war.’

- (88) *Krico                      malyyamanang                      muya tan.*  
 krit-yo                      maly-yam-an-ang                      muya tan  
 annoy-3SG.N sing-PTCP-NMLZ-A wrong 3PL.M.GEN

‘Their off singing is annoying.’

## 4.2 Pronouns

Ayeri possesses different kinds of pronouns in the sense that there is a closed class of words which contains anaphora of various types—personal pronouns, demonstrative pronouns, interrogative pronouns, relative pronouns, as well as reflexive and reciprocal expressions. Each class of pronouns will be discussed in the following.

### 4.2.1 Personal pronouns

As Table 4.6 shows, Ayeri possesses quite a large number of personal pronouns with (maybe unnaturally) little syncretism between the different paradigm slots overall (the second person is a notable exception); there are also no gaps in the paradigm. Ayeri’s personal pronouns reflect the grammatical features also found in nouns, that is, number, gender, and case, and person is added to that. The individual forms range from completely fused to fully transparent even within the same case paradigm, for instance, *yā* ‘(to/for) me’ (1SG.DAT) on the one hand, and *yayam* ‘(to/for) him’ (transparently 3SG.M-DAT) on the other. Originally, all pronouns have been regular formations based on the respective unmarked pronominal element listed in the TOP column of Table 4.6 declined by adding a case suffix (see section 4.1.3). Use has caused many of these formations to contract and erode as grammaticalization progressed:

- (89) a. *ayang*                      → *yāng*  
           ay-ang                      yāng  
           1SG-A                      1SG.A
- b. *iyatena*                      → *tan*  
           iy-a-t-ena                      tan  
           3SG-M-PL-GEN                      3SG.M.GEN<sup>11</sup>



Table 4.6: Personal pronouns

Person	TOP	A	P	DAT	GEN	LOC	CAUS	INS
1SG	ay	yang	yas	yām	nā	yā	sā	rī
2SG	va	vāng	vās	vayam	vana	vaya	vasa	vari
3SG.M	ya	yāng	yās	yayam	yana	yāy	yasa	yari
3SG.F	ye	yeng	yes	yeyam	yena	yea	yesa	yeri
3SG.N	yo	yong	yos	yoyam	yonā	yoa	yosa	yorī
3SG.INAN	ra	reng	rey	rayam	ran	raya	rasa	rari
1PL	ayn	nang	nas	nyam	nana	nyā	nisa	ni
2PL	va	vāng	vās	vayam	vana	vaya	vasa	vari
3PL.M	yan	tang	tas	cam	tan	ca	tis	ti
3PL.F	yen	teng	tes	teyam	ten	teya	tēs	teri
3PL.N	yon	tong	tos	toyam	ton	toya	tōs	tori
3PL.INAN	ran	teng	tey	racam	ten	raca	ratas	ray

The plural series used to be derived by adding  $\text{𐌆}$  *-n* or, in the third person,  $\text{𐌆}$  *-t-* to the pronoun stem, which can still be easily observed in the unmarked pronouns as well as in the alternation between  $\text{𐌆}$  *y-* and  $\text{𐌆}$  *t-* in the third person pronouns. The same goes for the gender-marking thematic vowel in the animate third person pronouns, which has been retained as a distinctive feature even in the non-core pronouns despite sometimes heavy modifications. A further interesting property of Ayeri is that synchronically, singular and plural are distinguished, except for the second person, where the forms are the same, basically like in English. Lehmann (2015) explains, however, that this is not an unusual route for languages to take:

New pronouns, especially for the second person singular, are often obtained by shifting pronouns around in the paradigm, especially by substituting marked forms for unmarked ones. This explains, for instance, the use of [...] English *you* for the second person singular (42)

The second person singular subject pronoun of English used to be *thou*, cognate to German *du*, which can still be found in Shakespeare, for instance. Something along the lines of English *you* as a second person plural pronoun replacing second person singular *thou* by way of a deferential singular use of a plural pronoun (OED 2016: *you*, pron., adj., and n.) may have happened in Ayeri as well.

<sup>11</sup> Strictly speaking, this could as well be glossed as  $t<a>n$  (3SG.GEN<M>). I chose to gloss the pronoun in the above way, however, in order to not overly complicate things.

The personal pronouns are used in just the same way as their full-NP counterparts would be, also in the non-core cases:

- (90) a. *Sa barya ang Paradan tandās kaleri.*  
 sa har-ya ang Paradan tanda-as kal-eri  
 AT beat-3SG.M A Paradan fly-P rag-INS  
 ‘Paradan beats the fly with a rag.’
- b. *Sa baryāng tandās kaleri.*  
 sa har=yāng tanda-as kal-eri  
 AT beat=3SG.M.A fly-P rag-INS  
 ‘He beats the fly with a rag.’
- c. *Sa barya ang Paradan yos kaleri.*  
 sa har-ya ang Paradan yos kal-eri  
 AT beat-3SG.M A Paradan 3SG.N.P rag-INS  
 ‘Paradan beats it with a rag.’
- d. *Sa barya ang Paradan tandās rari.*  
 sa har-ya ang Paradan tanda-as rari  
 AT beat-3SG.M A Paradan fly-P 3SG.INAN.INS  
 ‘Paradan beats the fly with it.’

In the above set of examples, (90a) shows a sentence with full NPs, with the agent, *ang Paradan* replaced by the third person singular masculine agent pronoun *yāng* in (90b); in (90c) the patient, *tandās*, is replaced with the third person singular neuter patient pronoun *yos*; in (90d), lastly, the instrument, *kaleri* is replaced with the third person singular inanimate instrumental pronoun *rari*. Furthermore, complex NPs are in complementary distribution with pronouns as anaphora to NPs, that is, an NP which contains an adjective is wholly replaced by an NP containing a personal pronoun:

- (91) a. *Ang ninye vehimley veno.*  
 ang nin=ye.Ø vehim-ley veno  
 A wear=3SG.F.TOP dress-P.INAN beautiful  
 ‘She wears a beautiful dress.’
- b. \**Ang ninye adaley veno.*  
 ang nin=ye.Ø ada-ley veno  
 A wear=3SG.F.TOP that-P.INAN beautiful  
 ‘\*She wears a beautiful it.’
- c. *Ang ninye adaley.*  
 ang nin=ye.Ø ada-ley  
 A wear=3SG.F.TOP that-P.INAN  
 ‘She wears it.’

Comparing the example sentences in (90) with the TOP column in Table 4.6 an important property of personal pronouns becomes apparent. That is, the ‘un-marked’ (or rather, zero-marked) pronoun forms are also the ones showing as verb agreement. An important difference in this respect, however, is that the third person singular inanimate verb agreement marker is not  $\text{ra}$ , but  $\text{-ara}$ . The following two examples illustrate the parallel more clearly—observe the person marking on the verb in (92) and the corresponding object pronouns in (93):

- (92) a. *Sa manya ang Ajān Pila.*  
 sa man-ya ang Ajān Ø Pila  
 PT greet-3SG.M A Ajān TOP Pila  
 ‘Pila, Ajān greets her.’
- b. *Sa manye ang Pila Ajān.*  
 sa man-ye ang Pila Ø Ajān  
 PT greet-3SG.F A Pila TOP Ajān  
 ‘Ajān, she greets him.’
- (93) a. *Sa manye ang Pila ya.*  
 sa man-ye ang Pila ya.Ø  
 PT greet-3SG.F A Pila 3SG.M.TOP  
 ‘Pila greets him.’
- b. *Sa manya ang Ajān ye.*  
 sa man-ya ang Ajān ye.Ø  
 PT greet-3SG.M A Ajān 3SG.F.TOP  
 ‘Ajān greets her.’

Another important property of both pronouns and verbs is that agent pronouns (and patient pronouns under certain circumstances) replace person agreement by cliticizing to the verb stem. As person agreement morphology is a domain of verbs, it will be dealt with in more detail in the chapter on verbs proper. The following example illustrates the mainly relevant process, however:

- (94) a. *Sa manya ang Ajān Pila.*  
 sa man-ya ang Ajān Ø Pila  
 PT greet-3SG.M A Ajān TOP Pila  
 ‘Pila, Ajān greets her.’
- b. *Sa manyāng Pila.*  
 sa man=yāng Ø Pila  
 PT greet=3SG.M.A TOP Pila  
 ‘Pila, he greets her.’

Possessive pronouns are special compared to regular personal pronouns in that they act basically as possessive adjectives; pronominal uses of possessive pronouns need 𐌳𐌰: *da-* as a supporting particle. The main case for the pronouns listed above in the genitive column of Table 4.6 is that of possessive adjectives, which means that unlike personal pronouns, they are by themselves not in complementary distribution with nominal NPs, compare (91). Instead, they may be used as modifiers like, or alongside, adjectives, as (95) shows.

- (95) *nangaya ledō nā*  
       nanga-ya ledō nā  
       house-LOC blue ISG.GEN  
       ‘in my blue house’

Yet, however, possessives do not fully share the properties of adjectives, namely, they cannot be compared (\*𐌳𐌰:𐌳𐌰:𐌳𐌰: \**nā-eng* ‘\*myer’, \*𐌳𐌰:𐌳𐌰:𐌳𐌰: \**nā-vā* ‘\*myest’). Fronting them in actual pronominal fashion is possible, for instance, in predicative statements like the one in (96a). Alternatively, a phrasal construction with 𐌳𐌰:𐌳𐌰:𐌳𐌰: *vilyang-* ‘belong’ as indicated in (96b) may be used.

- (96) a. *Ada-nangāng da-nā.*  
       ada=nanga-ang da-nā  
       that=house-A one=ISG.GEN  
       ‘That house is mine.’  
       b. *Ang vilyangyo ada-nanga yas.*  
       ang vilyang-yo ada=nanga-Ø yas  
       AT belong-3SG.N that=house-TOP ISG.P  
       ‘That house belongs to me.’

#### 4.2.2 Demonstrative pronouns

Demonstrative pronouns in Ayeri are formed with the demonstrative prefixes: 𐌳𐌰: *eda-* ‘this’ (proximal), 𐌳𐌰: *ada-* ‘that’ (distal), and 𐌳𐌰: *da-* ‘such’ (indefinite). These are combined with a morpheme 𐌳𐌰: *nya*, which is related to the word for ‘person’, 𐌳𐌰:𐌳𐌰: *nyān*. Table 4.7 gives the declined forms for all of them. Those forms attested in the corpus gathered from dictionary entries and example texts also used for the syllable structure analyses in section 1.2 appear in upright type, those that should be grammatical as well otherwise are given in italic type. The corpus is very small, but the prevalence of some forms is possibly reflecting varying degrees of grammaticalization at least to some extent. Table 4.8 gives the token frequencies of the various attested forms.

Table 4.7: Demonstrative pronouns

Case	Proximal	Distal	Indefinite
TOP	edanya	adanya	danya
A	edanyāng	adanyāng	<i>danyāng</i>
A.INAN	edareng, <i>edanyareng</i>	adareng, adanyareng	<i>danyareng</i>
P	edanyās	adanyās	danyās
P.INAN	edaley	<i>adaley</i>	danyaley
DAT	<i>edayam</i>	adayam	<i>danyayam</i>
GEN	edanyana	adanyana	danyana
LOC	<i>edanyaya</i>	adanyaya	<i>danyaya</i>
CAUS	<i>edanyasa</i>	<i>adanyasa</i>	<i>danyasa</i>
INS	<i>edanyari</i>	<i>adanyari</i>	<i>danyari</i>

Of all the cases, the agent demonstratives have the highest token frequency at a combined 52.53%, especially the distal pronouns are very frequent in the sample. Moreover, the distal inanimate agent demonstrative occurs twice as often as its animate counterpart, the shortened form ᐃᐃᐃᐃ *adareng* ‘that (one)’ being far more current than the full form ᐃᐃᐃᐃᐃᐃ *adanyareng*. Interestingly, the shortened form ᐃᐃᐃᐃ *edareng* ‘this one’ is also the only one attested for the inanimate proximate agent; similarly, the only dative demonstrative attested once is shortened as well: ᐃᐃᐃᐃ *adayam* ‘(to/for) that’. For non-core cases, only ‘long’ demonstratives are attested, albeit sparingly so.

Regarding the variation between ‘long’ and ‘short’ forms, it is not surprising that those demonstratives with a high frequency of use are eroded in some way: it seems that Ayeri prefers them to stay trisyllabic, which is achieved by dropping the ᐃᐃ *nya* part.<sup>12</sup> A further reason for dropping the ᐃᐃ *nya* part especially in the inanimate demonstratives may be that it is perceived as a marker of animacy—it has been noted above already that it is related to the word ᐃᐃᐃᐃ *nyān* ‘person’. Both factors, high frequency and semantic mismatch, may thus promote contraction.

Still, the question for the reason for the high frequency especially of ᐃᐃᐃᐃ *adareng* remains open. It may be explained by looking at a few typical examples of this word in context, however.

<sup>12</sup> According to the so-called Zipf’s law, word length and token frequency correlate in that the most frequently used words in a language also tend to be the shortest (Zipf 1935: 25–27).

Table 4.8: Token frequencies of attested demonstrative pronouns

Pronoun	Gloss	Absolute	Relative
<i>edanya</i>	this.TOP	1	1.69%
<i>adanya</i>	that.TOP	9	15.25%
<i>danya</i>	such.TOP	1	1.69%
<i>edanyāng</i>	this.A	4	6.78%
<i>adanyāng</i>	that.A	8	13.56%
<i>edareng</i>	this.A.INAN	3	5.08%
<i>adareng</i>	that.A.INAN	15	25.42%
<i>adanyareng</i>	that.A.INAN	1	1.69%
<i>edanyās</i>	this.P	1	1.69%
<i>adanyās</i>	that.P	2	3.39%
<i>danyās</i>	such.P	2	3.39%
<i>edaley</i>	this.P.INAN	2	3.39%
<i>danyaley</i>	such.P.INAN	2	3.39%
<i>adayam</i>	that.DAT	3	5.08%
<i>edanyana</i>	this.GEN	1	1.69%
<i>adanyana</i>	that.GEN	2	3.39%
<i>danyana</i>	such.GEN	1	1.69%
<i>adanyaya</i>	that.LOC	1	1.69%
Total		59	100%

- (97) a. *Nay ang nelyo-ikan sungkorankibas, adareng tono.*  
 nay ang nel-yo=ikan sungkorankihas ada-reng tono  
 and AT help-3SG.N=much geography that-A.INAN certain  
 ‘And geography, that’s for sure, helped me a lot.’ (Becker 2015 [2013]: 13)
- b. *Adareng merambay-ikan, le sundalvāng sasān vana ...*  
 ada-reng merambay=ikan le sundal=vāng sasān-Ø vana ...  
 that-A.INAN useful=very PT.INAN lose=2SG.A way-TOP 2SG.GEN ...  
 ‘It’s very useful if you get lost [...]’ (14)
- c. *Adareng danyaley segasena boa tinka.*  
 ada-reng danya-ley segas-ena boa tinka  
 that-A.INAN such-P.INAN snake-GEN boa closed  
 ‘The one of the closed boa snake.’<sup>13</sup> (22)

In all of the example sentences in (97), *adareng* ‘that (one)’ serves as a dummy pronoun together with a predicative adjective or NP, which is the main reason why it occurs so frequently. This is to say, Ayeri prefers the demonstrative pronoun *adareng* as the dummy agent in predicative contexts over the personal pronoun *reng* ‘it’. Otherwise, however, demonstrative pronouns work regularly as deictic anaphora: ‘this’, ‘that’, and ‘such (a)’, except that as nominal elements they are declined for case—but not for number or animacy, which is a notable difference between demonstrative pronouns and personal pronouns:

- (98) a. *Ang vehya Ajān nangās.*  
 ang veh-ya Ø Ajān nanga-as  
 AT build-3SG.M TOP Ajān house-P  
 ‘Ajān builds a house.’
- b. *Nangās? Sa vehyāng may danya.*  
 nanga-as sa veh=yāng may danya-Ø  
 house-P PT build=3SG.M.A AFF such-TOP  
 ‘A house? He builds one indeed.’
- (99) a. *Sā hasuyeng eda-migorayye.*  
 sā hasu=yeng eda=migoray-ye-Ø  
 CAUT sneeze=3SG.F.A this=flower-PL-TOP  
 ‘These flowers make her sneeze.’
- b. *Ang tipinyon nivaye yena adanyari naynay.*  
 ang tipin-yon niva-ye-Ø yena adanya-ri naynay  
 AT itch-3PL.N eye-PL-TOP 3SG.F.GEN that-CAUS as.well  
 ‘Her eyes are itching due to that/them/those [the flowers] as well.’

As mentioned in the previous chapter (p. 67), the prefix *da-* ‘such, so’ can combine with a range of syntactic phrase types, but most notably NPs, to serve as an indefinite demonstrative:

- (100) *Adareng da-dipakanas.*  
 adareng da=dipakan-as  
 that-A.INAN such=pity-P  
 ‘That is such a pity.’

*da-* can be used to express English ‘one’ in the sense of a deictic anaphora as well. Thus, to express ‘the X one’, if X is an adjective, it is necessary to use the full

<sup>13</sup> More literal translations of this sentence are ‘That is the one of the closed boa snake’ or ‘That is one of a closed boa snake’.

demonstrative pronoun, 𐌲𐌺𐌰 *danya*, since adjectives do not decline, and Ayeri largely avoids undeclined NPs:<sup>14</sup>

- (101) a. *Silvyo ku-mino-ing danyāng kivo.*  
 silv-yo ku=mino=ing danya-ang kivo  
 look-3SG.N like=happy=so such-A little  
 ‘The little one looks so happy.’
- b. *Sa noyang danya tuvo.*  
 sa no=yang danya-Ø tuvo  
 PT want=1SG.A such-TOP red  
 ‘I want the red one.’

In language less formal than in (101b), it is also possible to abbreviate 𐌲𐌺𐌰 *danya* to the prefixed form 𐌲𐌺𐌰 *da-*, which may be complemented by adjectives and possessive pronouns alike. The adjective or pronoun basically forms a complex anaphora, then, which in most circumstances can be marked for case and topic like any other nominal element, as demonstrated in (102).

- (102) *Sa noyang da-tuvo.*  
 sa no=yang da=tuvo.Ø  
 PT want=1SG.A such=red.TOP  
 ‘I want the red one.’

If incorporated in this way, the adjective cannot take comparison morphology: (103a) is not possible since inflections cannot be appended to clitics (if we analyze 𐌲𐌺𐌰 *-eng* and 𐌲𐌺𐌰 *-vā* as such in this context); and the meaning of (103b) differs from what was intended, since the 𐌲𐌺𐌰 *-vā* clitic is appended not to the adjective but to the composite nominal as such.

- (103) a. \**da-tuvo-vāley*  
 da=tuvo=vā-ley  
 one=red=SUPL-P.INAN  
 Intended: ‘the reddest one’
- b. *!da-tuvoley-vā*  
 da=tuvo-ley=vā  
 one=red-P.INAN=most/\*SUPL  
 ‘most red ones’  
 Intended: ‘the reddest one’

<sup>14</sup> See section 4.1.3 above for examples of situations where nouns regularly do not exhibit case marking.



Table 4.9: Interrogative pronouns

Pronoun	Literal meaning	Idiomatic meaning
ṣṣṣ <i>sinya</i>	which one (ṣṣṣ <i>nyān</i> ‘person’)	‘who’, ‘what’, ‘which’
ṣṣṣ <i>sikan</i>	how much (ṣṣṣ <i>ikan</i> ‘much’)	‘how much’, ‘how many’
ṣṣṣ <i>sikay</i>	with what (ṣṣṣ <i>kayvo</i> ‘with’)	‘how’ (tool, circumstance)
ṣṣṣ <i>simin</i>	which way (ṣṣṣ <i>miran</i> ‘way’)	‘how’ (way, procedure)
ṣṣṣ <i>sitaday</i>	which time (ṣṣṣ <i>taday</i> ‘time’)	‘when’
ṣṣṣ <i>siyan</i>	which place (ṣṣṣ <i>yano</i> ‘place’)	‘where’

#### 4.2.3 Interrogative pronouns

The intererrogative pronouns are all formed with ṣṣ- *si-*, combined with a lexical element or a case marker; ṣṣ- *si-* is also related to the relativizer ṣṣ *si*. The interrogative pronouns are listed in Table 4.9.

A property which all interrogative pronouns share is that they are placed *in situ*. That is, they appear in the same position as the phrase they stand in for, so there will not be movement of the question word to the front as in English. Additionally, impersonal interrogative pronouns cannot be topicalized since they also do not inflect for case, which preempts the difference between zero-marked topicalized and overtly case-marked untropicalized forms.

- (104) a. *Sa petigavāṅg inun sikan?*  
 sa petiga=vāṅg inun-Ø sikan  
 PT catch=2SG.A fish-TOP how.much  
 ‘How much fish did you catch?’
- b. *Sa-sahavāṅg sitaday?*  
 sa~saha=vāṅg sitaday  
 ITER~come=2SG.A when  
 ‘When will you return?’

In the table on interrogative pronouns above, ṣṣṣ *sinya* ‘who, what, which’ is separated from the other pronouns because it behaves differently. Namely, it can be declined for all cases according to the syntactic or semantic role of the NP it replaces, and it can also be topicalized, since the element asked about is likely high in discourse salience.

Table 4.10: Declension paradigm for ṣṣṣṣ *sinya* ‘who, what’

Case	Pronoun	Translation
TOP	ṣṣṣṣ <i>sinya</i>	‘who’, ‘what’
A	ṣṣṣṣṣṣ <i>sinyāng</i>	‘who’, ‘what’
A.INAN	ṣṣṣṣṣṣ <i>sinyareng</i>	‘who’, ‘what’
P	ṣṣṣṣṣ <i>sinyās</i>	‘whom’, ‘what’
P.INAN	ṣṣṣṣṣṣ <i>sinyaley</i>	‘whom’, ‘what’
DAT	ṣṣṣṣṣṣ <i>sinyayam</i>	‘for/to whom’, ‘for/to what’
GEN	ṣṣṣṣ <i>sinyana</i>	‘whose’, ‘from whom’, ‘from what’
LOC	ṣṣṣṣṣ <i>sinyaya</i>	‘in/at/on whom’, ‘in/at/on what’
CAUS	ṣṣṣṣṣṣ <i>sinyisa</i>	‘due to/because of whom’, ‘due to/because of what’
INS	ṣṣṣṣṣṣ <i>sinyari</i>	‘by whose help’, ‘with what’

- (105) a. *Ang yomayo sinya adaya?*<sup>16</sup>  
ang yoma-yo sinya-Ø adaya  
AT exist-3SG.N who-TOP there  
‘Who is there?’
- b. *Sa narayeng sinyā?*  
sa nara=yeng sinyā-Ø  
PT say=3SG.F.A what-TOP  
‘What did she say?’

Ayeri does not strictly distinguish animate and inanimate entities in its interrogative pronouns, so there is no distinction between ‘who’ and ‘what’. ṣṣṣṣ *sinya* and/or the verb will simply inflect according to context and to the speaker’s expectations or knowledge (compare Table 4.10). Thus, there is also no dedicated question word for ‘why’, since in Ayeri one can simply ask ‘due to what/whom’ by inflecting ṣṣṣṣ *sinya*; ṣṣṣṣṣṣ *sinyisa* is essentially ṣṣṣṣ *sinya* marked for causative case by the suffix ṣṣṣṣ *-isa*.

- (106) a. *Le kayāng adanya sinyayam?*  
le ka=yāng adanya-Ø sinyā-yam  
PT.INAN throw.away=3SG.M.A that-TOP what-DAT  
‘Why (= what for) did he throw that away?’

<sup>16</sup> This may be shortened to just ṣṣṣṣṣṣ *sinyāng adaya?* ‘who (is) there?’ (who-A there).

- b. *Ang prantoyva sinyisa?*  
 ang prant-oy=va.Ø siny-a-isa  
 AT ask-NEG=2SG.TOP what-CAUS  
 ‘Why (= because of what) did you not ask?’

While there is no dedicated ‘why’, Ayeri distinguishes between two kinds of ‘how’:  $\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}$  *simin* asks about the way by which—or the circumstances under which—an action is carried out, whereas  $\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}$  *sikay* asks for the means or tools used to carry out an action:

- (107) a. *Le tiyavāng vadisān simin?*  
 le tiya=vāng vadisān-Ø simin  
 PT.INAN make=2SG.A bread-TOP how  
 ‘How do you make bread?’  
 b. *Le peralvāng sagan sikay?*  
 le peral=vāng sagan-Ø sikay  
 PT.INAN grind=2SG.A flour-TOP how  
 ‘How do you grind flour?’

The correct answer to the question in (107a) needs to treat the process of making bread, since  $\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}$  *simin* asks about the way; a correct answer to the question in (107b), on the other hand, will likely mention grinding utensils, like a mill or a pestle.

Comparing Tables 4.9 and 4.10, it may strike the reader’s eye that there are two possibilities to express ‘where’—lexical  $\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}$  *siyan* and synthetic  $\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}$  *sinyaya*. It is important to note, however, that these are not strictly interchangeable, even though some variation is to be expected. While  $\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}$  *siyan* refers to *places* in general, the  $\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}\text{ḥ}^{\text{h}}$  *sinya* series refers to *entities* both animate and inanimate more specifically:

- (108) a. *Saravāng siyan? — Ya Sikatay.*  
 sara=vāng siyan — Ya Sikatay  
 go=2SG.A where — LOC Sikatay  
 ‘“Where are you going?”—“To Sikatay.”’  
 b. *Ya divvāng sinyā? — Ya Haki.*  
 ya div=vāng sinyā-Ø — Ya Haki  
 LOCT stay=2SG.A who-TOP — LOC Haki  
 ‘“With whom are you staying?”—“At Haki’s”’

#### 4.2.4 Indefinite pronouns

Haspelmath (1997: 56) notes how descriptions of languages often do not document indefinite pronouns—whether they simply do not exist in this language or whether



Figure 4.1: The implicational map for indefinite pronoun functions (Haspelmath 1997: 4)

they escaped the author's attention remains unknown in these cases. It may thus be duly noted here that Ayeri does indeed possess indefinite pronouns.<sup>17</sup> In order to classify languages, Haspelmath (1997) generalizes the map displayed in Figure 4.1 based on a sample of 100 languages from all continents, although he notes that this sample has a European bias due to the availability of data (2). Languages typically form contiguous areas on the map, even though they may carve it up quite differently, and with overlaps between the different semantic groupings 1–9.

An interesting question that Haspelmath (1997) poses towards the end of his book is whether there are any correlations between word order typology and the preference for generic nouns ('person', 'thing', 'place', 'time', 'manner') or, for instance, interrogative-based systems (239–241). While from Haspelmath (1997)'s concluding statistics it looks as though there is a slight preference of languages with which Ayeri shares basic typological traits—such as verb-initial, verb-object, and noun-genitive word order, also having prepositions—for basing indefinite pronouns on generic nouns, Haspelmath (1997) concedes that these seeming correlations are skewed by areal effects, “because indefinite pronouns have a strongly areal distribution” (241).<sup>18</sup> He still presumes, however, that word-order typology may have an effect on the formation of indefinites insofar as it correlates with grammaticalization more generally (239).

<sup>17</sup> As it is a fictional language, the value of this assertion to linguistic typology remains doubtful, however.

<sup>18</sup> The map in *WALS*, Haspelmath (2013), suggests areal clusters at least for generic-noun based systems in Africa and Southeast Asia. *WALS* classifies 194 out of 326 languages (60%) as possessing interrogative-based indefinite pronouns to date, with evidence for this type quoted for all continents except Africa. The next smaller group, generic-noun based, falls behind at 85 data points (26%). The curious lack of evidence for the interrogative type in Africa despite its being the most frequent one in the set on all other continents may be due to the unavailability of data. Crossreferencing the *WALS* data for constituent-order with the map on indefinite-pronoun systems did not yield a result which obviously suggested a correlation.

Table 4.11: Indefinite pronouns

	every	some	none
PERSON	enya	arilinya	ranya
THING	enya	arilinya, arilya	ranya
PLACE	yanen	yāril	yanoy
TIME	tadayen	tajaril; metay	tadoy; jānyam
MANNER	arēn	miranaril	are moy
REASON	—	yāril	—

Haspelmath (1997) mentions generic nouns, and these can be combined with the quantifying expressions ‘every’, ‘any’, ‘some’, and ‘none’ into an array like the one presented in Table 4.11. Ayeri does not distinguish ‘every’ from ‘any’ as English does; there is also no distinction in polarity (affirmative versus negative) the way English has it:

- (109) a. \*I don’t know something about this.  
b. I don’t know anything about this.

Likewise, Ayeri does not distinguish between animate and inanimate indefinite referents—the same pronouns are used for either, although the shortening of *arilinya*, *arilya*, can only be used for inanimates, similar to the distinction in the demonstrative pronouns between *adanyāng* ‘that one’ (that.one-A) and *adareng* ‘that one’ (that.one-A.INAN; see section 4.2.2). Two further features stand out, however.

Firstly, most of the pronouns in the chart have a lexical part—Ayeri’s indefinite pronouns are based on generic nouns. Thus, the pronouns referring to people and things all have the *-nya* element in common, which we also find in the interrogative and demonstrative pronouns, and which also appears in the word *nyān* ‘person’. In the same way, the pronouns related to the notion of place have a *ya-* or *yan-* part, which we also find in *yano* ‘place’.<sup>19</sup> In a regular continuation of this pattern, the indefinite pronouns of time all have an element related to *taday* ‘time’ in common, which is obscured somewhat by palatalization in *tajaril*. The exception to this series, then, is *jānyam*, which is the multiplicative numeral formed from *ja* ‘zero’, thus means ‘zero times’ or ‘not once’ rather than

<sup>19</sup> *yano* itself is an old nominalization and very likely related as a morpheme to the locative suffix *-ya*.

‘never’, although it can also be used emphatically for the latter. The series of manner pronouns is an absolute exception in that it must be a residue from an older layer of grammaticalization since *are-* is not a recognizable morpheme in the modern language.<sup>20</sup> *miranaril* is a regular formation of *miran* ‘way, manner’ combined with the quantifier (!) for indefinite amounts, *-aril* ‘some’.

This observation leads to the second regular feature, that is, affixes as modifiers to generic nouns. The ‘every’ series regularly features the morpheme *en*, either prefixed or suffixed, which is related to the quantifier *-ben* ‘every, all, each’ and can presumably be found even on *arēn* in spite of its obscure lexical base. In the same manner, the series related to inspecific generic-noun referents is marked by the affix *aril* which, as we have just seen above, is otherwise used to refer to inspecific quantities, for instance, *vadisān-aril* ‘some bread’ (bread=some). In the case of *miranaril*, the suffix seems somewhat of an odd choice, since manner is not a quantifiable variable in the same way people, things, locations or moments are. Possibly, it is chosen rather in analogy with the other pronouns in this series than on semantic grounds. In any event, *metay* has the semantically more ‘proper’ *me-* prefix, relating it to absolute inspecificity.<sup>21</sup> This alternation is employed to distinguish between the meaning of ‘sometime’, that is, occurring once an unspecified point in time, and *tajaril* ‘sometimes’, which refers to repeated occurrence at inspecific times. The alternation between *miranaril* and regularly derived *mə-miran* can be leveraged to express a specificity difference as well. While the former suggests that an action is carried out or an event is happening by means of a specific, though unknown procedure, the latter suggests just any possible procedure. Lastly, the negative series is regularly marked by the negative suffix *-oy*, which also occurs with adjectives and verbs (see sections 4.3.2 and 4.5.4). An outlier in this series is the person/thing-related indefinite pronoun, *ranya*. The etymologic connections of the *ra* part are not presently known, perhaps the postposition *ran* ‘against’ is related.

The chart in Table 4.11 only tells half the truth by not giving any information on use contexts for the individual forms, so how do they fit in with the chart from Haspelmath (1997) quoted at the beginning of this section? Regarding the functions of indefinite pronouns annotated to the numbers on the map, Haspelmath (1997) gives the following examples, which, however, mostly only give one example for either the ‘person’ or ‘thing’ category at a time. It is up to the reader to generalize from this (2–3):<sup>22</sup>

<sup>20</sup> I probably made this up as I was going, many years ago, and without considering systematic implications, as I was unaware of them at the time.

<sup>21</sup> Compare German *irgendjemand* and French *n’importe qui* ‘no matter who’.

<sup>22</sup> These appear here reordered according to numerical order. The book lists them according to their logical order as tracing the map, the enumeration somewhat confusingly tied in with the



Figure 4.2: Map of indefinite pronoun functions in Ayeri

- (110)
1. specific, known to the speaker:  
*Somebody called while you were away: guess who!*
  2. specific, unknown to the speaker:  
*I heard something, but I couldn't tell what kind of sound it was.*
  3. non-specific, irrealis:  
*Please try somewhere else.*
  4. polar question:  
*Did anybody tell you anything about it?*
  5. conditional protasis:  
*If you see anything, tell me immediately.*
  6. indirect negation:  
*I don't think that anybody knows the answer.*
  7. direct negation:  
*Nobody knows the answer.*
  8. standard of comparison:  
*In Freiburg the weather is nicer than anywhere in Germany.*
  9. free choice:  
*Anybody can solve this simple problem.*

As we have seen in Table 4.11 above, Ayeri does not make a difference between 'every' and 'any', which is why the 'some' series can be applied to all of (1)–(5); it can also be used for indirect negation (6). The pronouns from the 'none' column, then, are used to express direct negation (7). Since double negation—that is, agreement in negation between verbs and indefinite pronouns for purposes of emphasis rather than double negation in the strictly logical sense—is possible, the 'none' series may also be employed for indirect negation (6). Moreover, Ayeri uses the 'every' series

running enumeration of examples.

for both standard of comparison (8) and free choice (9). Besides this, absolute-indefinite  $\text{me-}$  can be used for (3)–(6) in combination with a (generic) noun to attach to. It needs to be noted that only the indefinite pronouns with person or thing reference (those including  $\text{nya}$ ) decline; they can also be topicalized. The other indefinites, relating to place, time and manner, are indeclinable and also cannot be topics for this reason. For the ‘specific’ categories (1) and (2) it is furthermore possible to use the plain generic nouns,  $\text{nyān}$  ‘person’,  $\text{linya}$  ‘thing’,  $\text{yano}$  ‘place’,  $\text{taday}$  ‘time’,  $\text{miran}$  ‘way’, however. Figure 4.2 shows the groupings for Ayeri; (III) gives examples of all types.

(III) 1. specific, known to the speaker:

- a. *Ang sabaya arilinya, leku, sinyāng adaley!*  
 ang saha-ya arilinya-Ø lek-u sinyang ada-ley  
 AT come-3SG.M someone-TOP guess-IMP who-A that-P.INAN  
 ‘Someone came, guess who it is!’
- b. *Le ilta ningyang linya vayam.*  
 le ilta ning=yang linya-Ø vayam  
 PT.INAN need tell=1SG.A thing-TOP 2SG.DAT  
 ‘I need to tell you something.’

2. specific, unknown to the speaker:

- a. *Ang pegaya arilinya pangisley nā.*  
 ang pega-ya arilinya-Ø pangis-ley nā  
 AT steal-3SG.M someone-TOP money-P.INAN 1SG.GEN  
 ‘Someone stole my money.’
- b. *Ang sarayan yanoa agon.*  
 ang sara=yan yano-ya agon  
 AT go=3PL.M.TOP place-LOC foreign  
 ‘They are going somewhere foreign.’

3. non-specific, irrealis:

- a. *Pinyan, prantu yāril palung.*  
 pinyan prant-u yāril palung  
 please ask-IMP somewhere different  
 ‘Please ask somewhere else.’
- b. *Le ilta miranang adanya mā-miraneri palung.*  
 le ilta mira=nang adanya-Ø mā-miran-eri palung  
 PT.INAN need do=1SG.A that.one-TOP some-way-INS different  
 ‘We need to do that in some other way.’



## 4. polar question:

- a. *Ang koronva arilinyaley edanyana?*  
 ang koron=va.Ø arilinya-ley edanya-na  
 AT know=2SG.TOP something-P.INAN this.one-GEN

‘Do you know anything about this?’

- b. *Yomaya mə-nyānang si ang vaca mirongya edanyaley?*  
 yoma-ya mə-nyān-ang si ang vaca mira-ong=ya.Ø edanya-ley  
 exist-3SG.M some-person-A REL AT like do-IRR=3SG.M.TOP this-P.INAN

‘Is there *anyone* who would like to do this?’

## 5. conditional protasis:

- a. *Ang ming pengalayn sitanyās yāril, adareng pray-ven.*  
 ang ming pengal=ayn.Ø sitanya-as yāril ada-reng pray=ven  
 AT can meet-IPL.TOP each.other-P somewhere that-A.INAN great=pretty

‘If we can meet somewhere that would be pretty great.’

- b. *Sa na-naravāng mə-lentan, ang haray vās!*  
 sa na~nara=vāng mə-lentan-Ø ang har=ay.Ø vās  
 PT ITER~say=2SG.A some-sound-TOP AT punch-1SG.TOP 2SG.P

‘You make any more sound, I’m gonna punch you!’

## 6. indirect negation:

- a. *Paronoyyang, ang no tabaya arilinya adaley.*  
 paron-oy=yang ang no taha-ya arilinya-Ø ada-ley  
 believe-NEG=1SG.A AT want have-3SG.M anyone-TOP that-P.INAN

‘I don’t think anyone wants to have that.’

- b. *Paronoyyang, le ming sungvāng adanya yanoy.*  
 paron-oy=yang le ming sung=vāng adanya-Ø yanoy  
 believe-NEG=1SG.A PT.INAN can find=2SG.A that.one-TOP nowhere

‘I don’t think you can find that *anywhere*.’

## 7. direct negation:

- a. *Ang koronya ranya guratanley.*  
 ang koron-ya ranya-Ø guratan-ley  
 AT know-3SG.M nobody-TOP answer-P.INAN

‘Nobody knows the answer.’

- b. *Le ming sungvāng adanya yanoy.*  
 le ming sung=vāng adanya-Ø yanoy  
 PT.INAN can find=2SG.A that.one-TOP nowhere

‘You can’t find that anywhere.’

Table 4.12: Relative pronouns

Case	Pronoun	Pronoun with secondary inflection				
		DAT	GEN	LOC	CAUS	INS
Ø	si	siyām	sinā	siyā	sisā	sirī
A	sang	sangyam	sangena	sangya	sangisa	sangeri
A.INAN	sireng	sirengyam	sirengena	sirengya	sirengisa	sirengeri
P	sas	sasyam	sasena	sasya	sasisa	saseri
P.INAN	siley	sileyyam	sileyena	sileyya	sileyisa	sileyeri
DAT	siyam	siyamyam	siyamena	siyamyā	siyamisa	siyameri
GEN	sina/sena	sinayam	sinana	sinaya	sinaisa	sinari
LOC	siya	siyayam	siyana	siyaya	siyaisa	siyari
CAUS	sisā	sisayam	sisana	sisaya	sisaisa	sisari
INS	seri	seriyam	serina	seriya	serīsa	seriri

## 8. standard of comparison:

- a. *Sa engyeng larau enya palung.*  
 sa eng=yeng larau enya-Ø palung  
 PT be.more=3SG.F.A nice anyone different  
 ‘She is nicer than anyone else.’
- b. *Ang engyo ban eda-riman yanen palung.*  
 ang eng-yo ban eda=riman-Ø yanen palung  
 AT be.more=3SG.N good this=city-TOP anywhere different  
 ‘This city is better than anywhere else.’

## 9. free choice:

- a. *Ang ming guraca enya eda-prantanley.*  
 ang ming gurat-ya enya-Ø eda=prantan-ley  
 AT can answer=3SG.M anyone-TOP this=question-P.INAN  
 ‘Anyone can answer this question.’
- b. *Epayeng tadayen si sa pinyaya ye ang Tapan.*  
 epa=yeng tadayen si sa pinya-ya ye ang Tapan  
 refuse=3SG.F.A everytime REL PT ask-3SG.M 3SG.F.TOP A Tapan  
 ‘She refused everytime Tapan asked her.’

Table 4.13: Token frequencies of attested complex relative pronouns

Pronoun	Gloss	Absolute
<i>siyā</i>	REL.Ø.LOC	7
<i>sirī</i>	REL.Ø.INS	3
<i>sinā</i>	REL.Ø.GEN	1
<i>siyām</i>	REL.Ø.DAT	1
Total		12

#### 4.2.5 Relative pronouns

As has been described before, Ayeri connects relative clauses to main clauses with the relativizer ꞑ *si*. This relativizer can be declined for case in accordance with the relative clause's head in the matrix clause. The respective forms can be gathered from Table 4.12 (column 'Pronoun').

- (112) a. *Eryyo tarela natrangās si tado.*  
 ery-yo tarela natranga-as si tado  
 use-3SG.N still temple-P REL old  
 'The temple, which is old, is still being used.'
- b. *Edanyāng ayonas sirtang sas ang sihabaya mondoas nana.*  
 edanya-ang ayon-as sirtang si-as ang sihaba=ya mondo-as nana  
 this-A man-P young REL-P ang tend=3SG.M.TOP garden-P IPL.GEN  
 'This is the young man who tends our garden.'

As explained in section 3.3, if the relativizer is immediately following its lexical head, only the base form ꞑ *si* is used, which is illustrated in (112a). Here, the head of the relative clause is *natrangās* 'the temple', which is immediately followed by the relative clause. If word material is intervening, however, which is the case in (112b), the relative pronoun may be inflected to agree in case with its antecedent in more formal language for referential clarity: *sas* agrees in case with *ayonas* two words over to the left. Relative pronouns do not agree in number with their heads, though, and in gender only insofar as it is relevant to nominal case inflection, that is, agents and patients are distinguished for animacy.

A special property of the relative pronoun is that it can be declined for its role in the relative clause as well to express more complex relationships between the main clause and the relative clause. The respective forms can be found in the columns titled 'pronoun with secondary inflection' in Table 4.12. The token frequency of the actually occurring complex relative pronouns in the very small corpus gathered from example texts and dictionary entries (see section 1.2) is given in Table 4.13.

Compared to the unmarked relativizer  $\text{ḥ} si$ , which occurs 50 times in the sample (all relative pronouns from Table 4.12 occur 80 times in total), the complex relative pronouns have a very low frequency. This is not surprising, since ‘for whom’, ‘by which’, etc. are quite specialized expressions. It also seems that those forms unmarked for their antecedent are preferred, since those are the only ones attested—the sample is really much too small to make actually meaningful judgements here, however. Examples of complex relative pronouns are:

- (113) a. *Le vacyang koya yana sileyya ang layāy adanyana.*  
 le vac=yang koya-Ø yana si-le-yā ang laya=ay.Ø adanya-na  
 PT.INAN like=ISG.A book-TOP 3SG.M.GEN REL-P.INAN-LOC A read=ISG.TOP that-GEN  
 ‘I like his book in which I read about it.’
- b. *Yā saratang yano siyām sarasatang.*  
 ya sara=tang yano-Ø si-Ø-yām sara-asa=tang  
 LOCT go=3PL.M.A place-TOP REL-LOC-DAT go-HAB=3PL.M.A  
 ‘They went to the place to which they always went.’

It needs to be pointed out that a complex relative pronoun cannot form the topic of the relative clause even though it is marked for case according to the relative clause’s syntactic domain. Furthermore, the relative pronoun cannot receive inflection for an agent or a patient of the embedded clause. The following examples illustrate these points:

- (114) \**Mica edaya sobayāng si (‘sī) na ibayang koyaley.*  
 mit-ya edaya sobaya-ang si-Ø-Ø na iha=yang koya-ley  
 live-3SG.M here teacher-A REL-A-TOP GENT borrow=ISG.A book-P.INAN  
 ‘Here lives the teacher from whom I borrowed a book.’
- (115) \**Mica edaya sobayāng sāng le sobya payutān yām.*  
 mit-ya edaya sobaya-ang si-Ø-ang le sob-ya payutān-Ø yām  
 live-3SG.M here teacher-A REL-A-A PT.INAN teach-3SG.M math-TOP ISG.DAT  
 ‘Here lives the teacher who taught me math.’
- (116) \**Mica edaya sobayāng sās ya kradasayang kardang.*  
 mit-ya edaya sobaya-ang si-Ø-as ya krad-asa=yang kardang-Ø  
 live-3SG.M here teacher-A REL-A-P LOCT hate-HAB=ISG.A school-TOP  
 ‘Here lives the teacher whom I used to hate in school.’

Example (114) displays a sentence in which the relative pronoun, ungrammatically, forms the controller of topic agreement on the verb in the relative clause:  $\text{ḥ} na$  as a genitive topic is supposed to refer to  $\text{ḥ} sobayāng$  in the matrix clause by way of the relativizer  $\text{ḥ} si$  which would then necessarily carry a zero-morpheme

topic marker. There is no resumptive pronoun in the relative clause, so the relative pronoun itself forms the anaphora in the relative clause referring to the relativized argument of the matrix clause. This is not possible.

In example (115), the relative pronoun \*ṣāṅ carries no overt case agreement as it directly follows its antecedent (\*ṣāṅṅ otherwise) but the long vowel shows that it is declined as the agent of the relative clause; the verb agrees using *-ya* accordingly. There is no resumptive agent pronoun, so the relative pronoun would stand in for the agent NP that would be necessary if the relative clause were an independent sentence. The use of the relative pronoun as an agent-NP replacement in this sentence is equally ungrammatical, though, and so is the agreement between verb and declined relative pronoun.

Similarly, in (116), the relative pronoun carries case marking for the patient of the relative clause, since the agent of the matrix clause serves as the patient NP of the embedded clause. This is not grammatical either.

Altogether, it seems that in Ayeri, core arguments of intransitive and transitive clauses—agents and patients—cannot precede the embedded verb of a relative clause; the verb firmly forms the head of the embedded clause in this regard. The relative pronoun also cannot receive secondary marking for agents or patients, and neither can it stand in directly as the agent and patient NP of the relative clause, respectively. It is interesting in this regard that Ayeri *does* allow this for recipients, however, maybe since by their nature as goals they carry something of a locative connotation (compare (113b)) and are thus less tightly integrated with verbs, occupying a middle ground between core arguments and adverbials like the locative proper.<sup>23</sup>

#### 4.2.6 Reflexives and reciprocals

As mentioned previously, Ayeri forms its reflexives with the prefix *ṣitang-* in combination with a personal pronoun, compare (117). If the agent of the action is the same as the reflexive patient—that is, the agent acts on itself—the reflexive prefix can also migrate onto the verb instead, which is demonstrated in (118).

- (117) *Ang silvye                    sitang=yes    puluyya.*  
          ang silv=yes.Ø        sitang=yes    puluy-ya  
          AT    see=3SG.F.TOP   self=3SG.F.P   mirror-LOC

‘She sees herself in the mirror.’

<sup>23</sup> This would be interesting to explore in terms of grammaticalization, as it is very possible that this behavior reflects a stage of the language before *-yam* had been grammaticalized as the dative marker. In this respect, it would as well be necessary to explore whether the similarity between the dative marker *-yam* and the locative marker *-ya* is indeed etymological or merely incidental.

- (118) *Ang sitang-silvye puluyya.*  
 ang sitang=silv=ye.Ø puluy-ya  
 AT self=see=3SG.F.TOP mirror-LOC  
 ‘She sees herself in the mirror.’

Doing the same with a non-patient pronoun does not work, however, so the sentence in (118) with the reflexive *sitang* marked on the verb is not equivalent to the following one, in which *sitang* appears together with a personal pronoun in the locative case, even though here as well, the agent and the locative pronoun refer to the same entity:

- (119) *Ang silvye sitang-yea puluyya.*  
 ang silv=ye.Ø sitang=yea puluy-ya  
 AT look=3SG.F.TOP self=3SG.F.LOC mirror-LOC  
 ‘She looks at herself in the mirror.’

It may be noted furthermore that the genitive/possessive pronoun series conveys the meaning of ‘one’s own’, which is completely regular in meaning (‘of X-self’), however:

- (120) *Le no eryongyang pakay sitang-nā.*  
 le no ery-ong=yang pakay-Ø sitang=nā  
 PT.INAN want use-IRR=1SG.A umbrella-TOP self=1SG.GEN  
 ‘I’d like to use my own umbrella.’

*sitang* is also used to carry quantifiers pertaining to a pronominal suffix (121a), since appending a quantifier directly to the conjugated verb stem itself can be ambiguous (121b).

- (121) a. *Ang koronay sitang-nyama guratanley.*  
 ang koron=ay.Ø sitang=nyama guratan-ley  
 AT know=1SG.TOP self=even answer-P.INAN  
 ‘Even I know the answer.’  
 b. *!Ang koronay-nyama guratanley.*  
 ang koron=ay.Ø=nyama guratan-ley  
 AT know=1SG.TOP=even answer-P.INAN  
 ‘I even know the answer.’  
*Intended:* ‘Even I know the answer.’

It appears as though *sitang* does not act as the controller of the verbal topic marker; this is illustrated by the fact that *sitang* and a non-topic agent pronominal suffix can appear side by side. As described previously (section 3.2.5, p. 89), lexical NPs and pronominal suffixes on the verb are mutually exclusive.

Adjectives are one of the parts of speech in Ayeri which do not inflect for any of the grammatical properties of their heads, that is, there is no agreement relation between adjectives and nominal heads. They do inflect for comparison under certain circumstances, however, and can also take various affixes that modify the meaning of the adjective stem.

### 4.3.1 Comparison

In cases where a comparee is left unexpressed or the patient forms the standard of comparison, Ayeri uses clitic suffixes on adjectives. The suffixes involved are ၼ် -*eng* (COMP) and ၼ် -*vā* (SUPL):

- (125) a. *Yeng ganyena men si alingo-eng.*  
 yeng gan-ye-na men si alingo=eng  
 3SG.F.A child-PL-GEN one REL clever=COMP  
 ‘She is one of the more clever children.’
- b. *Ang tavya Diyan tingracas ban-eng na Maha.*  
 ang tav-ya Ø Diyan tingrati-as ban=eng na Maha  
 AT become-3SG.M TOP Diyan musician-P good=COMP GEN Maha  
 ‘Diyan became a better musician than Maha.’
- c. *Naratang, yāng pokamayās para-vā.*  
 nara=tang yāng pokamaya-as para-vā  
 say=3PL.M.A 3SG.M.A shooter-P fast-SUPL  
 ‘They said he is the fastest shooter.’

In (125a) the comparee is missing, while in (125b), the quality under comparison, ၼ် -*tingracas ban-eng* ‘a better musician’, is a patient NP; the standard, ၼ် -*Maha*, is expressed by an adverbial genitive NP. The example in (125c) similarly expresses an absolute without giving a group of entities to draw from. In all these cases, it is, however, also possible to use a more complex analytic construction using verbs.

### 4.3.2 Negation

Adjectives in Ayeri can be negated in two ways: categorially with ၼ် -*arya*, and pragmatically with ၼ် -*oy*. These correspond to English *un-*, and *in-*, *il-*, *ir-* etc. for categorial negation, and to *not* for pragmatic negation. ၼ် -*oy* absorbs the vowel of the root it is attached to if said root ends in a vowel.

- (126) *Telbaya miseryanang ku-ardārya.*  
 telba-ya miseryan-ang ku=arda-arya  
 show-3SG.M method-A like=suitable-NEG  
 ‘The method proved unsuitable.’
- (127) *Pakoy eda-yanoreng.*  
 paka-oy eda=yano-reng  
 safe-NEG this=place-A.INAN  
 ‘This place is not safe.’



Example (126) displays an adjective which carries the categorial negation marker ၼ် -*arya*; the adjective in (127) carries the simple, pragmatic negation marker ၼ် -*oy*. Which one to use is up to the speaker, since both negate the described property. The categorial marker puts an emphasis more on expressing a general opposite, while the pragmatic marker simply negates, so that it is not necessarily implied that the negative state persists. The place that is ကုၼ် *pakoy* ‘not safe’ now is not necessarily ကုၼ် *pakārya* ‘unsafe’ in general, but simply not safe in the context of the here and now of the utterance.

Besides *ad hoc* derivation of categorial negatives with ၼ် -*arya*, there are also a few lexicalized instances. These have an idiomatic meaning and the negator or the word itself may be irregularly reduced. Examples are, among others:

- (128) a. ခၼ် *ban* ‘good’ → ခၼ် *banaya* ‘ill, sick’  
 b. ကုၼ် *kovaro* ‘easy’ → ကုၼ် *kovarya* ‘awkward’  
 c. ခိၼ် *sirimang* ‘straight’ → ခိၼ် *sirimaya* ‘passive’

#### 4.3.3 Adjectivization

Adjectives in Ayeri are very commonly zero derivations, that is, there is rather free conversion between nouns and adjectives,<sup>24</sup> for instance:

- (129) a. ၼ် *Ayeri* ‘Ayeri’ ~ ၼ် *Ayeri* ‘Ayeri’  
 b. ၼ် *disa* ‘soap, lye’ ~ ၼ် *disa* ‘soapy, alkaline’  
 c. ၼ် *gino* ‘drink’ ~ ၼ် *gino* ‘drunk’  
 d. ကုၼ် *pahamay* ‘danger’ ~ ကုၼ် *pahamay* ‘dangerous’  
 e. ခိၼ် *sempay* ‘peace’ ~ ခိၼ် *sempay* ‘peaceful’

Adjectives can also be derived from verbs with the causative suffix ၼ် -*isa*, which often corresponds to adjectives derived from the past participle form—the meaning is often, but not necessarily, relating to an achieved state. The suffix may change the last vowel to ၼ် *u* or drop it; a specific pattern to these changes is not recognizable. The derivations may be idiomatic occasionally, as some derivations in the example below show.

- (130) a. ကုၼ် *kelang-* ‘connect’ → ကုၼ် *kelangisu* ‘connected, related’  
 b. ကုၼ် *palung-* ‘distinguish’ → ကုၼ် *palungisa* ‘various’  
 c. ခိၼ် *sundala-* ‘lose’ → ခိၼ် *sundalisu* ‘lost’  
 d. ၼ် *taban-* ‘write’ → ၼ် *tahanis* ‘literary’  
 e. ၼ် *vesa-* ‘give birth’ → ၼ် *vesisa* ‘native’

<sup>24</sup> Adjectives and split-off modifiers in noun–noun (compare section 4.1.5) compounds are thus similar at least superficially.

There are also at least two words where an *-isa* adjective is derived not from a verb, but a word of a different part of speech—in this case, a noun, and another adjective:

- (131) a. *ṣṣin* *apin* ‘luck’ → *ṣṣinisa* *apinisa* ‘lucky’  
 b. *ṣṣay* *iray* ‘high’ → *ṣṣayisa* *irayisu* ‘exalting’

#### 4.3.4 Other affixes

As with nouns, other affixes which can be attached to adjectives as clitic hosts, are the prefix *ku-*, expressing semblance, as well as quantifying and grading suffixes, of which the suffixes used to express comparative and superlative are, essentially, a grammaticalized variety, since *-eng* can also be used like ‘rather’.

- (132) *Ku-pikisu paray-parayang.*  
*ku=pikisu paray~paray-ang*  
*like=scared DIM~cat-A*  
 ‘The kitten is like scared.’

- (133) *Napay-eng eda-prikanreng.*  
*napay=eng eda=prikan-reng*  
*spicy=rather this=soup-A.INAN*  
 ‘This soup is rather spicy.’

## 4.4 Adpositions

Adpositions are another part of speech in Ayeri whose stem itself does not inflect. Ayeri’s most basic adpositions are derived from relational nouns, which is likely the reason why Ayeri mostly employs prepositions, with postpositions and ambipositions being less important placement patterns (Hagège 2010: 110–111; Lehmann 2015: 81 ff.). Adpositions in their most basic use trigger locative marking on the governed NP, the prepositional object; for allative and ablative meanings, the prepositional object may also appear in the dative and the genitive case, respectively, as described in section 4.1.3.<sup>25</sup> The cognitive metaphor ‘time equals space’ with the future conceptualized as lying ahead and the past behind also holds in Ayeri, so that some of the words describing locations also double to describe temporal relations.

	Preposition	Etymology (or related to)
ခံၤခါၣ် <i>agonan</i>	outside	ခံၤခါၣ် <i>agonan</i> ‘outside’
ခံၤၣ် <i>avan</i>	bottom, ground	ခံၤၣ် <i>avan</i> ‘ground, bottom; soil’
ခံၤခါၣ် <i>eyran</i>	under, below	ခံၤခါၣ် <i>eyran</i> ‘sole’
ခံၤခါၣ် <i>eyrarya</i>	over	ခံၤခါၣ် <i>eyran</i> ‘sole’ + ခံၤ <i>-arya</i> (NEG)
ခံၤခါၣ် <i>kayvo</i>	with, beside <sup>26</sup>	ခံၤခါၣ် <i>kayv-</i> ‘accompany’
ခံၤခါၣ် <i>kong</i>	inside, within	ခံၤခါၣ် <i>kong</i> ‘inside’
ခံၤခါၣ် <i>ling</i>	on	ခံၤခါၣ် <i>ling</i> ‘top’
ခံၤခါၣ် <i>luga</i>	among, between	ခံၤခါၣ် <i>luga-</i> ‘pass, penetrate’
ခံၤခါၣ် <i>mangasaba</i>	towards, in + <i>time</i>	ခံၤခါၣ် <i>manga saba-</i> ‘coming’
ခံၤခါၣ် <i>mangasara</i>	away	ခံၤခါၣ် <i>manga sara-</i> ‘going’
ခံၤခါၣ် <i>marin</i>	front, on (walls etc.)	ခံၤခါၣ် <i>marin</i> ‘face, surface’
ခံၤခါၣ် <i>miday</i>	around	ခံၤခါၣ် <i>miday-</i> ‘surround’
ခံၤခါၣ် <i>nasay</i>	near, close	ခံၤခါၣ် <i>nasay</i> ‘proximity’
ခံၤခါၣ် <i>nuveng</i>	left	ခံၤခါၣ် <i>nuho</i> ‘liver’
ခံၤခါၣ် <i>pang</i>	behind, ago	ခံၤခါၣ် <i>pang</i> ‘back’
ခံၤခါၣ် <i>patameng</i>	right	ခံၤခါၣ် <i>patam</i> ‘heart’

Table 4.14 gives all the words in Ayeri which may be used as prepositions. As mentioned above, most of these are derived transparently from nouns, so they have probably been grammaticalized relatively recently—their non-preposition meaning is still transparent, they are still phonologically rather complex, and some of them are even polysyllabic in spite of not being composed and covering rather basic meanings.<sup>27</sup> Since these nouns have ceased to function as common nouns in this context

<sup>26</sup> There is also a preposition  $\eta\beta\gamma$  *dayrin* ‘side’ listed in the dictionary, however, this has never seen much use. Instead,  $\eta\beta\gamma$  *kayvo* has come to cover ‘beside, to the side of’ as well.

Unsurprisingly, Hagege (2010: 129) references Zipf regarding speech economy and token frequency. According to Lehmann (2015: 134–141), the phonological integrity of morphemic units reduces as grammaticalization is progressing (with token frequency increasing due to increasing obligatoriness). Bybee and Hopper (2001) see the reason for phonological reduction of highly frequent phonological material “in the automatization of neuro-motor sequences [...]”. Such reductions are systematic across speakers; that is, they do not represent ‘sloppy’ or ‘lazy’ speech” (11). Hence, for example, English’s most basic prepositions are extremely short and

due to grammaticalization, however, it is not possible to inflect them in the way described in section 4.1. Thus, for example, while it is possible to say (134a), it is not really possible to say (134b):

- (134) a. *Le yomareng kanka lingya rivanena.*  
 le yoma=reng kanka-Ø ling-ya rivan-ena  
 PT.INAN exist=3SG.INAN.A snow-TOP top-LOC mountain-GEN  
 ‘There is snow on the top of the mountain.’<sup>28</sup>
- b. \**Ang nedraye lingya nedrānena.*  
 ang nedra=ye.Ø ling-ya nedrān-na  
 AT sit=3SG.F.TOP top-LOC chair-GEN  
 ‘<sup>2</sup>She sits on the top of a chair.’

Instead, the grammatical way to express (134b) is the following, using ማሽግ *ling* as a preposition with the object in the locative case:

- (135) *Ang nedraye ling nedrānya.*  
 ang nedra=ye.Ø ling nedrān-ya  
 AT sit=3SG.F.TOP top chair-LOC  
 ‘She sits on a chair.’

In this case, since *on* is the expected position of sitting with regards to chairs, the preposition can even be dropped:

- (136) *Ang nedraye nedrānya.*  
 ang nedra=ye.Ø nedrān-ya  
 AT sit=3SG.F.TOP chair-LOC  
 ‘She sits on a chair.’

With regards to (134a) it is also necessary to mention what Hagège (2010) calls the ‘Proof by Anachrony Principle’ (158–159). According to this principle, when an adposition is very grammaticalized, speakers can use both the adposition and its etymological ancestor side by side without taking offense in the double occurrence. This is notably not the case in Ayeri, where it is not possible to say things like

simple words, for instance, *of*, *at*, *in*, which derive from the slightly more complex PIE forms *\*h<sub>2</sub>ep-ó*, *\*h<sub>2</sub>ed*, *\*h<sub>1</sub>en(-i)*, respectively (Kroonen 2013: 1, 39, 269). Since adpositions frequently grammaticalize into case markers, it may be assumed that the phonologically much more simple case affixes of Ayeri constitute an older layer of basic adpositions. Their non-suffixed forms may be remnants of this use.

<sup>28</sup> The corresponding sentence with a preposition is ማሽግ ላይ ማሽግ ላይ ማሽግ ላይ *Le yomareng kanka ling rivanaya* ‘There is snow on top of the mountain’ (PT.INAN exist=3SG.INAN.A snow-TOP top mountain-LOC).

(137a), where 𐀓𐀕 *pang* is used in both its meanings so that the preposition 𐀓𐀕 *pang* ‘behind’ governs the original noun 𐀓𐀕 *pang* ‘back’.

- (137) a. \**Le ranice ang Maha adanya pang pangya yena.*  
 le ranit-ye ang Maha adanya-Ø pang pang-ya yena  
 PT.INAN hide-3SG.F A Maha that-TOP back back-LOC 3SG.F.GEN  
 ‘\*Maha hides it at the back of her back.’
- b. *Le ranice ang Maha adanya pangya yena.*  
 le ranit-ye ang Maha adanya-Ø pang-ya yena  
 PT.INAN hide-3SG.F A Maha that-TOP back-LOC 3SG.F.GEN  
 ‘Maha hides it at her back,’  
 or: ‘Maha hides it behind herself.’

Examples like (134b), on the other hand, show that there is nonetheless a tendency in Ayeri towards grammaticalization of nouns which used to be relational. Grammaticalization is visible in that formerly relational nouns have become restricted in the way they can be used syntactically (Lehmann 2015: 174). This specialization is also apparent in morphology from the fact that prepositions in Ayeri, in spite of their nominal origin, cannot be modified by adjectives and relative clauses like regular nouns. Thus, for instance, while 𐀓𐀕 *avan* as a noun can mean ‘soil’ or ‘ground’ and can be modified by semantically coherent adjectives like 𐀓𐀕 *kabu* ‘fertile’, the preposition 𐀓𐀕 *avan* cannot. Again, a grammatical way to express (138b) would have to use 𐀓𐀕 *avan* as a relational noun, that is, 𐀓𐀕𐀓𐀕𐀓𐀕𐀓𐀕 *avanya kabu similena* ‘at the fertile bottom of the country’ (bottom-LOC fertile country-GEN). The fact that topicalized heads lack case marking makes adpositions derived from nouns, like 𐀓𐀕 *avan* homophonous with the respective etymologically related preposition.

- (138) a. *Sa yomareng avan kabu ibangya yana.*  
 sa yoma=reng avan-Ø kabu ibang-ya yana  
 PT exist=3SG.INAN.A.INAN ground-TOP fertile field-LOC 3SG.M.GEN  
 ‘Fertile ground is on his field.’
- b. \**Ang mican avan kabu similya*  
 ang mit=yan.Ø avan kabu simil-ya  
 AT live=3PL.M.TOP bottom fertile country-LOC  
 ‘\*They live at the fertile bottom of the country.’

At the beginning of this section it was shown that prepositions in Ayeri cannot receive number and case marking, which are otherwise typical features of nouns. What is possible with regards to affixes, however, is adding degree suffixes to prepositions, since these suffixes are clitics rather than inflections (compare section 3.2.5, p. 94):

Table 4.15: Prepositions (directional)

Preposition	<i>manga</i> + PREP
ᐱᐱᐱᐱ <i>agonan</i> ‘outside’	out
ᐱᐱᐱ <i>avan</i> ‘at bottom’	to the bottom; <i>with DAT/GEN</i> : down to/from
ᐱᐱᐱᐱ <i>eyran</i> ‘under’	under
ᐱᐱᐱᐱᐱ <i>eyrarya</i> ‘over’	across, over
ᐱᐱᐱᐱ <i>kayvo</i> ‘with, beside’	along
ᐱᐱᐱᐱ <i>kong</i> ‘inside’	into
ᐱᐱᐱᐱ <i>ling</i> ‘on top’	onto, while; <i>with DAT/GEN</i> : up to/from
ᐱᐱᐱ <i>luga</i> ‘between’	through, during, for + <i>time</i>
ᐱᐱᐱᐱ <i>marin</i> ‘in front’	to the front
ᐱᐱᐱᐱ <i>miday</i> ‘around’	circling around
ᐱᐱᐱ <i>nasay</i> ‘near’	into the near
ᐱᐱᐱᐱ <i>nuveng</i> ‘left’	to the left
ᐱᐱᐱᐱ <i>pang</i> ‘behind’	behind, to the back
ᐱᐱᐱᐱᐱ <i>patameng</i> ‘right’	to the right

- (139) *Ang mitasaye pang-ikan mandayya tado.*  
 ang mit-asa=ye.Ø pang=ikan manday-ya tado  
 AT live-HAB=3SG.F.TOP back=much forum-LOC old

‘She used to live way behind the old forum.’

As demonstrated before, another quasi-inflection adpositions in Ayeri can carry is the directional marker ᐱᐱᐱ *manga* (see section 3.1). While most of the prepositions in Table 4.14 have a static meaning, ᐱᐱᐱ *manga* indicates a motion in the direction of the respective location, thus ᐱᐱᐱᐱ *kong* ‘inside’ becomes ᐱᐱᐱᐱᐱᐱ *manga kong* ‘into’, for instance. Table 4.15 repeats the table of prepositions above for the most part and gives the respective directional meanings. The prepositions ᐱᐱᐱᐱᐱᐱ *mangasaba* and ᐱᐱᐱᐱᐱᐱ *mangasara* are missing from this list and appear in the previous table instead, even though they express motion rather than position, because they are only used in this base form and cannot be prefixed by ᐱᐱᐱ *manga*, which they already contain. Note, however, that ᐱᐱᐱᐱᐱᐱ *mangasaba* and ᐱᐱᐱᐱᐱᐱ *mangasara* are not synonymous to an adjunct in the dative and the genitive case, respectively. Rather, the prepositions add a more deliberate or literal meaning:

- (140) a. *Ang nimpay kardangyam.*  
 ang nimp=ay.Ø kardang-yam  
 AT run=ISG.TOP school-DAT  
 ‘I’m running to (a/the) school.’  
 (e.g. for class, or just up to the building)
- b. *Ang nimpay mangasaba kardangya.*  
 ang nimp=ay.Ø mangasaba kardang-ya  
 AT run=ISG.TOP towards school-LOC  
 ‘I’m running towards (a/the) school.’  
 (up to the building)
- (141) a. *Ang lampay kardangena.*  
 ang walk=ay.Ø kardang-ena  
 AT walk=ISG.TOP school-GEN  
 ‘I’m walking from (a/the) school.’  
 (e.g. home, or somewhere else from there)
- b. *Ang lampay mangasara kardangya.*  
 ang lamp=ay.Ø mangasara kardang-ya  
 AT walk=ISG.TOP away school-LOC  
 ‘I’m walking away from (a/the) school.’  
 (away from the building)

Also note that while Germanic languages like English make frequent use of set expressions which combine a verb with an intransitive preposition, such as *run away*, *go by*, *raise up*, *track down*, sometimes with rather idiomatic meanings, this pattern does not occur as frequently in Ayeri. Some exceptions are:

- (142) a. ႁၢၦၢၦၢၦ *il-* mangasara ‘surrender’ (give away),  
 b. ႁၢၦၢၦၢၦ *lant-* mangasara ‘distract’ (lead away),  
 c. ႁၢၦၢၦၢၦ *nimp-* mangasara ‘escape’ (run away),  
 d. ႁၢၦၢၦၢၦ *tapy-* dayrin ‘save (valuable assets)’ (put aside),  
 e. ႁၢၦၢၦၢၦ *tapy-* miday ‘put on’ (put around),  
 f. ႁၢၦၢၦၢၦ *tura-* mangasaba ‘forward’ (send towards).

These verbs do not govern a prepositional object in the locative case in their idiomatic meaning, as displayed by the next example, in which ႁၢၦၢၦၢၦ *batangiman* and ႁၢၦၢၦၢၦ *sa Ajān* do neither serve as arguments of ႁၢၦၢၦ *lanco* or ႁၢၦၢၦ *mangasara*, but of the phrasal verb ႁၢၦၢၦၢၦ *lant-* mangasara:<sup>29</sup>

<sup>29</sup> Colloquially, ႁၢၦၢၦၢၦ *mangasaba* and ႁၢၦၢၦၢၦ *mangasara* may be shortened to just ႁၢၦၢၦ *saba* and ႁၢၦၢၦ *sara*, respectively.

- (143) *Ang lanco mangasara batangiman sa Ajān.*  
 ang lant-yo mangasara batangiman-Ø sa Ajān  
 AT lead-3SG.N away mosquito-TOP P Ajān  
 ‘The mosquito distracted Ajān.’

Very often, where the verb expression in English contains a preposition, there is a separate verb in Ayeri, or the same verb is used in Ayeri for both the plain English verb and the one extended by a preposition:

- (144) a. ႁႁႁ: *apand-* ‘descend, climb down’,  
 b. ႁႁႁ: *dila-* ‘figure out, find out’,  
 c. ႁႁႁ: *ling-* ‘ascend, mount, climb up’,  
 d. ႁႁႁ: *naga-* ‘watch after’,  
 e. ႁႁႁ: *pab-* ‘remove, take away’,  
 f. ႁႁႁ: *subr-* ‘cease, give up’.
- (145) a. ႁႁ: *ka-* ‘throw (away)’,  
 b. ႁႁ: *mat-* ‘warm (up)’,  
 c. ႁႁႁ: *sikl-* ‘rip (up)’.

In cases where the preposition does not have a prepositional object otherwise, its double nature as a noun comes to the fore in that the preposition word will be treated like a noun if it is denominal and carries the appropriate case marker itself:

- (146) a. *Ang sabayan manga pang nangaya.*  
 ang saha=yan.Ø manga pang nanga-ya  
 AT go=3PL.TOP DIR back house-LOC  
 ‘They go behind the house.’
- b. *Ang sabayan pangyam.*  
 ang saha=yan.Ø pangyam  
 AT go=3PL.TOP back-DAT  
 ‘They go behind (it),’  
 or: ‘They go to the back.’

#### 4.4.2 Postpositions

While Ayeri mainly uses prepositions—which is by far the most common order for VO languages (Dryer 2013b)—it also uses a number of postpositions, which are given in Table 4.16. As can be read from the table, postpositions do not usually have a nominal origin but are derived either from other prepositions, from adverbial



Table 4.16: Postpositions

	Postposition	Etymology (or related to)
𑌕𑌃𑌕𑌃𑌕𑌃	<i>da-nārya</i>	despite, in spite of
𑌕𑌃𑌕𑌃𑌕𑌃	<i>kayvay</i>	without
𑌕𑌃𑌕𑌃𑌕𑌃	<i>masabatay</i>	since
𑌕𑌃𑌕𑌃𑌕𑌃	<i>nasyam</i>	according to
𑌕𑌃𑌕𑌃𑌕𑌃	<i>pang</i>	beyond, after, past
𑌕𑌃𑌕𑌃𑌕𑌃	<i>pesan</i>	until
𑌕𑌃𑌕𑌃𑌕𑌃	<i>ran</i>	against
𑌕𑌃𑌕𑌃𑌕𑌃	<i>rayu</i>	diagonally across
𑌕𑌃𑌕𑌃𑌕𑌃	<i>yamva</i>	instead of

phrases, or even from an adjective in the case of 𑌕𑌃𑌕𑌃𑌕𑌃 *rayu*. The etymologies of 𑌕𑌃𑌕𑌃𑌕𑌃 *pesan* and 𑌕𑌃𑌕𑌃𑌕𑌃 *yamva* are unclear to date.

The postposition 𑌕𑌃𑌕𑌃𑌕𑌃 *pang* is special in that it also exist as a preposition meaning ‘behind, in the back of’, though as a postposition it acquires the related but slightly different meaning ‘beyond, after, past’. It might thus be better treated as a homonym to the preposition rather than as an ambiposition (Hagège 2010: 115). Example (147a) illustrates the use of 𑌕𑌃𑌕𑌃𑌕𑌃 *pang* as a preposition, (147b) the use of 𑌕𑌃𑌕𑌃𑌕𑌃 *pang* as a postposition. This is in contrast to typical ambipositions like German *wegen* ‘because of, due to’ in (148), which has the same meaning in either position and the position variant is just a matter of style.

- (147) a. *Sa lancāng pel manga pang penungya.*  
 sa lant=yāng pel-Ø manga pang penung-ya  
 PT lead=3SG.M.A horse-TOP DIR back barn-LOC  
 ‘The horse, he leads it behind the stable.’
- b. *Lesyo pelang si sã nimp̃yong penungya pang yan.*  
 les-yo pel-ang si sã nimp=yong penung-ya pang yan.Ø  
 fall-3SG.N horse-A REL CAUT run=3SG.N.A stable-LOC back 3PL.TOP  
 ‘The horse they raced past the barn fell.’

- (148) a. *wegen des schlechten Wetters* [German]  
 wegen des schlecht-en Wetter-s  
 because.of DEF.GEN.N.SG bad-GEN.N.SG weather-GEN  
 ‘because of the bad weather’

Table 4.17: Adpositions with temporal meaning

Adposition	Spatial meaning	Temporal meaning
<b>Prepositions</b>		
ቋጉ <i>kong</i>	inside	within
ሰገጉ <i>ling</i>	on top of	while
ጠገገ <i>marin</i>	in front of	before
ጠገገገ <i>manga luga</i>	through	during
ጠገገገገ <i>mangasaba</i>	towards	in + <i>time</i>
ገገ <i>pang</i>	behind	ago
<b>Postpositions</b>		
ጠገገገገገ <i>masahatay</i>	—	since
ሰገገ <i>pesan</i>	—	until
ገገ <i>pang</i>	beyond, after	after, past

- b. *des schlechten Wetters wegen*  
 des schlecht-en Wetter-s wegen  
 DEF.GEN.N.SG bad-GEN.N.SG weather-GEN because.of  
 (idem)

Besides the difference in placement, the morphological properties of postpositions are the same as those of prepositions. That is, where postpositions are derived from nouns at all, they do not receive case and number marking and cannot themselves be modified by adjectives or relative clauses. Generally, it is possible for them to be hosts of quantifier clitics where semantics permit it.

#### 4.4.3 Adpositions and time

It has been mentioned above that location also serves as the conceptual metaphor for expressing temporal relationships. Notably the prepositions ቋጉ *kong* ‘inside’, ሰገገ *ling* ‘on’, ጠገገ *marin* ‘in front of’, ጠገገገ *manga luga* ‘through’, ጠገገገገ *mangasaba* ‘towards’, and ገገ *pang* ‘behind’ come to mind as doubling for ‘within’, ‘while’, ‘before’, ‘during’, ‘in + *time*’, and ‘ago’, respectively (also see Table 4.17). Since postpositions are not primarily derived from nouns, there are dedicated forms for expressing temporal relationships, namely, ጠገገገገገ *masahatay* ‘since’, ሰገገ *pesan* ‘until’, and as the only form with a double function, ገገ *pang* ‘after, past’.

- (149) a. *Miranang kong bihanya sam.*  
 mira=nang kong bihan-ya sam  
 do=IPL.A inside week-LOC two  
 'We will do it within two weeks.'
- b. *Girenjang mangasaha pidimya-kay.*  
 girend=yang mangasaha pidim-ya=kay  
 arrive=3SG.M.A towards hour-LOC=few  
 'He will arrive in a few hours.'
- c. *Layaye-ikan ang Pila ling yeng pakur.*  
 laya-ye=ikan ang Pila ling yeng pakur  
 read-3SG.F=much A Pila on 3SG.F.A sick  
 'Pila read a lot while she was sick.'

Of the examples above, the use of ꨀꨣꨣ *kong* in (149a) is probably still closest to a local preposition in that the time span is conceptualized as a container, or the distance between two points. The use of ꨀꨣꨣꨣꨣ *mangasaha* in (149b), on the other hand, is more idiomatic. While the prepositions in these two examples each govern an NP, example (149c) shows that it is also possible for prepositions expressing a temporal relationship to govern a subclause. This ability is even more prominent with temporal postpositions in that all of the words listed above can govern either an NP or a clause, for instance, ꨀꨣꨣꨣꨣ *masahatay*:

- (150) a. *Ang manga hangya lakayperinya masahatay.*  
 ang manga hang=ya.Ø lakayperin-ya masahatay  
 AT PROG stay=3SG.M.TOP solstice-LOC since  
 He has been staying since the solstice.
- b. *Yeng giday sarayāng masahatay.*  
 yeng giday sara=yāng masahatay  
 3SG.F.A sad leave=3SG.M.A since  
 'She has been sad since he left.'

## 4.5 Verbs

Besides nouns, verbs constitute the other main part of speech in Ayeri which carries inflections. Verbs show person and number agreement, but may also inflect for tense, aspect, mood, and modality as grammatical categories of the verb itself. Personal pronouns may furthermore cliticize to the verb stem, and the verb phrase is also often marked with a clitic indicating the topic of the sentence and the topic NP's role in Ayeri's case system, which can be interpreted as a second agreement

Table 4.18: Conjugation paradigm for ᱥᱟᱱ: *sob*- ‘learn, teach’ (monoconsonantal root)

Person	Topicalized <sup>30</sup>	Clitic agent	Translation
1SG	<i>sobay</i>	<i>sobyang</i>	‘I learn’
2SG	<i>sobva</i>	<i>sobvāng</i>	‘you learn’
3SG.M	<i>sobya</i>	<i>sobyāng</i>	‘he learns’
3SG.F	<i>sobye</i>	<i>sobyeng</i>	‘she learns’
3SG.N	<i>sobyō</i>	<i>sobyong</i>	‘it learns’
3SG.INAN	<i>sobara</i>	<i>sobreng</i>	‘it learns’
IPL	<i>sobayn</i>	<i>sobnang</i>	‘we learn’
2PL	<i>sobva</i>	<i>sobvāng</i>	‘you learn’
3PL.M	<i>sobyan</i>	<i>sobtang</i>	‘they learn’
3PL.F	<i>sobyen</i>	<i>sobteng</i>	‘they learn’
3PL.N	<i>sobyon</i>	<i>sobtong</i>	‘they learn’
3PL.INAN	<i>sobaran</i>	<i>sobteng</i>	‘they learn’
IMP	<i>sobu!</i>	‘learn!’	
HORT	<i>sobu-sobu!</i>	‘let’s learn!’	
ITER	<i>so-sob-</i>	‘learn again, relearn’	
PTCP	<i>sobyam</i>	‘learning’	

relation. Further clitics may indicate reflexive actions, progressive aspect, likeness, logical connection, as well as degree and measure. Verbs are thus probably the most versatile part of speech on the one hand, but also the one with the heaviest workload on the other. The following sections will dissect the morphology of verbs category by category. Because verbs inhabit a central position in syntax and exhibit agreement morphology, it will be necessary in this section to merge syntax and morphology on occasion in order to describe morphosyntactic effects.

#### 4.5.1 Person marking

As described in section 3.3, Ayeri conjugates its main verbs, canonically in agreement with the agent NP, and verb conjugation as such is extremely pervasive. The basic conjugation paradigms are given in Tables 4.18–4.20.<sup>31</sup> Agreement causes verbs to reflect grammatical categories of nominal entities, thus, verbs show agreement in

<sup>30</sup> Third-person topicalized forms are homonymous with third-person agreement forms.

<sup>31</sup> Due to the agglutinating structure of Ayeri it makes little sense to list the whole paradigm of verb inflection for all possible affix combinations here, as the table would become unreasonably

Table 4.19: Conjugation paradigm for ၼ်: *anl-* ‘bring’ (biconsonantal root)

Person	Topicalized	Clitic agent	Translation
1SG	<i>anlay</i>	<i>anlyang</i>	‘I bring’
2SG	<i>anlava</i>	<i>anlavāng</i>	‘you bring’
3SG.M	<i>anlya</i>	<i>anlyāng</i>	‘he brings’
3SG.F	<i>anlye</i>	<i>anlyeng</i>	‘she brings’
3SG.N	<i>anlyo</i>	<i>anlyong</i>	‘it brings’
3SG.INAN	<i>anlara</i>	<i>anlareng</i>	‘it brings’
IPL	<i>anlayn</i>	<i>anlanang</i>	‘we bring’
2PL	<i>anlava</i>	<i>anlavāng</i>	‘you bring’
3PL.M	<i>anlyan</i>	<i>anlatang</i>	‘they bring’
3PL.F	<i>anlyen</i>	<i>anlateng</i>	‘they bring’
3PL.N	<i>anlyon</i>	<i>anlatong</i>	‘they bring’
3PL.INAN	<i>anlaran</i>	<i>anlateng</i>	‘they bring’
IMP	<i>anlu!</i>	‘bring!’	
HORT	<i>anlu-anlu!</i>	‘let’s bring!’	
ITER	<i>an-anl-</i>	‘bring again, bring back’	
PTCP	<i>anlyam</i>	‘bringing’	

person (1, 2, 3) and number (SG, PL); third persons are again differentiated by gender (M, F, N, INAN; compare section 4.1.1). Verbs only have agreement proper with third persons; their form, then, is the same as that of verbs with topicalized pronominal inflection (see section 4.2.1).

Regarding person–number inflection, verbs may be divided into three classes: monoconsonantal, biconsonantal, and vocalic stems. As discussed in section 1.2, Ayeri restricts the number of successive non-glides consonants to two, which has repercussions in the second person, since the conjugation suffix there is *-va*. Monoconsonantal roots are unaffected by this restriction, however, hence the conjugation suffixes can simply be appended as they are; this is illustrated with the verb ၼ်: *sob-* ‘teach, learn’ in Table 4.18. Verb stems ending in dental and velar plosives will naturally undergo palatalization in the third person animate, so for instance, the third person singular masculine of the verb ၼ်: *gurat-* ‘answer’ is ၼ်: *guraca* ‘(he) answers’, and the third person feminine plural of ၼ်: *abag-* ‘roam, wander’ is ၼ်: *abajen* ‘(they) roam, (they) wander’. Verbs whose stem ends in an affricate are

large. Instead, the various sections below will contain examples of use for all affixes.

Table 4.20: Conjugation paradigm for 𐌵: *no-* ‘want’ (vocalic root)

Person	Topicalized	Clitic agent	Translation
1SG	<i>noay</i>	<i>noyang</i>	‘I want’
2SG	<i>nova</i>	<i>novāng</i>	‘you want’
3SG.M	<i>noya</i>	<i>noyāng</i>	‘he wants’
3SG.F	<i>noye</i>	<i>noyeng</i>	‘she wants’
3SG.N	<i>noyo</i>	<i>noyong</i>	‘it wants’
3SG.INAN	<i>noara</i>	<i>noreng</i>	‘it wants’
IPL	<i>noayn</i>	<i>nonang</i>	‘we want’
2PL	<i>nova</i>	<i>novāng</i>	‘you want’
3PL.M	<i>noyan</i>	<i>notang</i>	‘they want’
3PL.F	<i>noyen</i>	<i>noteng</i>	‘they want’
3PL.N	<i>noyon</i>	<i>notong</i>	‘they want’
3PL.INAN	<i>noaran</i>	<i>noteng</i>	‘they want’
IMP	<i>nu!</i>	‘want!’	
HORT	<i>nu-nu!</i>	‘let’s want!’	
ITER	<i>no-no-</i>	‘want again’	
PTCP	<i>noyam</i>	‘wanting’	

treated as monoconsonantal roots as well, since the affricate occupies one consonant phoneme segment. Thus, the second person of 𐌵: *ic-* ‘glide, slide’ is not \*𐌵: *icava*, but 𐌵: *icva* ‘you glide, you slide’.

Since /v/ is neither a vowel nor a glide, as present in the non-second person suffixes, an epenthetic *-a-* is inserted between the stem and the second-person suffix for verbs whose stem ends in *-CC*.<sup>32</sup> This is illustrated in Table 4.19 for the verb 𐌵: *anl-* ‘bring’. The second person conjugation of this verb is not \*𐌵: *anlva*, since the cluster *-nlv-* is illegal, but 𐌵: *anlava*. Since Ayeri treats two successive instances of the same consonant as a single segment—there is no gemination—verbs like 𐌵: *silv-* ‘see’ conjugate like monoconsonantal roots with regards to consonant clusters. That is, the second person of 𐌵: *silv-* is not \*𐌵: *silvava*, as one might expect, but 𐌵: *silvva*. A further exception to this are verbs ending in *-Cs*, since *-Cs-C-* is commonly resyllabified as *-C-sC-* (see chapter 1, footnote 13). Thus, the

<sup>32</sup> A *root* is understood here as the uninflected verb morpheme, for instance, 𐌵: *anl-*, 𐌵: *ic-*, 𐌵: *no-*, or 𐌵: *sob-*. A *stem* may contain inflections and further inflectional affixes attach to it; it may also host clitics.

Table 4.21: Conjugation paradigm for ၼာ: *apa-* ‘laugh’ (vocalic root in -a)

Person	Topicalized	Clitic agent	Translation
1SG	<i>apāy</i>	<i>apayang</i>	‘I laugh’
2SG	<i>apava</i>	<i>apavāng</i>	‘you laugh’
3SG.M	<i>apaya</i>	<i>apayāng</i>	‘he laughs’
3SG.F	<i>apaye</i>	<i>apayeng</i>	‘she laughs’
3SG.N	<i>apayo</i>	<i>apayong</i>	‘it laughs’
3SG.INAN	<i>apāra</i>	<i>apareng</i>	‘it laughs’
IPL	<i>apāyn</i>	<i>apanang</i>	‘we laugh’
2PL	<i>apava</i>	<i>apavāng</i>	‘you laugh’
3PL.M	<i>apayan</i>	<i>apatang</i>	‘they laugh’
3PL.F	<i>apayen</i>	<i>apateng</i>	‘they laugh’
3PL.N	<i>apayon</i>	<i>apatong</i>	‘they laugh’
3PL.INAN	<i>apāran</i>	<i>apateng</i>	‘they laugh’
IMP	<i>apu!</i>	‘laugh!’	
HORT	<i>apu-apu!</i>	‘let’s laugh!’	
ITER	<i>ap-apa-</i>	‘laugh again’	
PTCP	<i>apayam</i>	‘laughing’	

second-person form of ၼာ: *kars-* ‘freeze’ is not \**ၼာ: \*karsava* as expected, but *ၼာ: karsva* ‘you freeze’.

Lastly, verb stems may end in a vowel, most commonly *-a*. In these cases as well, the conjugation suffixes may simply be appended to the stem. The conjugation of this class is illustrated in Table 4.20 with the verb ၼာ: *no* ‘want’. Verb stems ending in *-a* undergo the regular vowel lengthening process for the first person suffixes, hence, the topicalized first-person singular form of ၼာ: *apa-* ‘laugh’ is ၼာ: *apāy* ‘I laugh’ (compare Table 4.21). Verb stems ending in a diphthong in /ɪ/ are essentially treated as a hybrid of monoconsonantal and vocalic stems, since the diphthong’s final /ɪ/ is treated as /j/ before a vowel: ၼာ: *palayay* ‘I rejoice’, ၼာ: *palayva* ‘you rejoice’.

As mentioned above, the person marking on verbs is essentially the same as the topic-marked personal pronouns. This has further ramifications for person-marking on verbs, however, in that in canonical cases, even fully case-marked agent pronouns may act as person marking by means of cliticization. Thus, any person-marking on verbs except third-person agreement is, in fact, a topicalized pronoun clitic not only by diachronic origin. Unlike English, Ayeri does not use agent

pronouns in addition to person agreement on verbs. Consider these two example sentences in English:

- (151) a. *John greets Mary.* [English]  
           John greet-s Mary  
           John greet-3SG.PRS Mary
- b. *He greets Mary.*  
           he greet-s Mary  
           3SG.M greet-3SG.PRS Mary

In these examples, the verb has an agreement suffix *-s* which indicates third person singular, present tense, whether the subject of the sentence is a noun (*John*) or a pronoun (*he*), which acts as a free morpheme in English. Now consider the Ayeri equivalents of these two examples, on the other hand:<sup>33</sup>

- (152) a. *Ang manya Ajān sa Pila.*  
           ang man-ya Ø Ajān sa Pila  
           AT greet-3SG.M TOP Ajān[3SG.M] P Pila[3SG.F]  
           ‘Ajān greets Pila.’
- b. *Ang manya sa Pila.*  
           ang man=ya.Ø sa Pila  
           AT greet=3SG.M.TOP P Pila[3SG.F]  
           ‘He greets Pila.’

It is probably uncontroversial to analyze *ang -ya* in (152a) as person agreement: *ᑭᑭᑭ Ajān* is a male name in Ayeri while *ᑭᑭ Pila* is a feminine one; the verb inflects for a masculine third person, which tells us that it agrees with the one doing the greeting, Ajān. Ajān is also who this is about, which is shown on the verb by marking for an agent topic. In the second case, there is only anaphoric reference to Ajān; the full agent NP is not realized. Very broadly thus, the verb marking here seems to be like in Spanish, where you can drop the subject pronoun:

- (153) a. *Juan saluda a María.* [Spanish]  
           Juan salud-a a María  
           John greet-3SG ACC Mary  
           ‘John greets Mary.’

<sup>33</sup> Most of the following account is taken nearly verbatim from a previously published blog article, Becker (2016d). Some of the Ayeri examples used in the following come from a list of samples I provided for a bachelor’s thesis at the University of Kent in March 2016, in private conversation, on request.



- b. *Saluda a María.*  
 salud-a a María  
 greet-3SG ACC Mary  
 ‘He greets Mary.’

Example (152b) probably does not seem conspicuous if we assume that Ayeri is pro-drop, except that there is also topic marking for an agent there, the controller of which I have so far assumed to be the person inflection on the verb, in analogy with examples like the following:

- (154) *Lampyāng.*  
 lamp=yāng  
 walk=3SG.M.A  
 ‘He walks.’

This raises the question whether in Ayeri, there is dropping of an agent pronoun involved at all, which is why the person suffix in (152b) was glossed as =*ya.Ø* (=3SG.M.TOP) rather than just as *-ya* (3SG.M). In turn, this question leads us to consider another characteristic of Ayeri, namely that the topic morpheme on noun phrases is zero. That is, the absence of overt case marking on a nominal element indicates that it is a topic; the verb in turn marks the case of the topicalized NP with a (case) particle preceding it. Pronouns as well show up in their unmarked form when topicalized, which is why I am hesitant to analyze the pronoun in (155b) as a clitic on the VP rather than an independent morpheme:<sup>34</sup>

- (155) a. *Sa manya ang Ajān Pila.*  
 sa man-ya ang Ajān Ø Pila  
 PT greet-3SG.M A Ajān TOP Pila  
 ‘It’s Pila that Ajān greets.’  
 b. *Sa manyāng ye.*  
 sa man=yāng ye.Ø  
 PT greet=3SG.M.A 3SG.F.TOP  
 ‘It’s her that he greets.’

What is remarkable, then, is that *ye* (3SG.F.TOP) is the very same form that appears as an agreement morpheme on the verb, just like *-ya* (3SG.M) in various examples above (also compare the examples in section 4.2.1):

<sup>34</sup> Also, perhaps a little untypically, topic NPs in Ayeri are not usually pulled to the front of the phrase (at least not in the written language; see Lehmann 2015: 120–122), so topic-marked pronouns stay in-situ. Which NP constitutes the topic of the phrase is marked on the verb right at the head of the clause. How and whether this can be justified in terms of grammatical weight (see, for instance, Wasow 1997: 95–98) remains to be seen.

- (156) *Ang purivaye yāy.*  
 ang puriva=ye.Ø yāy  
 AT smile=3SG.F.TOP 3SG.M.LOC  
 ‘She smiles at him.’

This also holds for all other personal pronouns. Moreover, *yāy* as seen in examples (154) and (155b) may also be used as a free pronoun in equative statements with predicative nominals, as well as other such case-marked personal forms:

- (157) a. *Yeng mino.*  
 yeng mino  
 3SG.F.A happy  
 ‘She is happy.’  
 b. *Yāng naynay.*  
 yāng naynay  
 3SG.M.A too  
 ‘He is, too.’

As for case-marked person suffixes on verbs, the assumption so far has been that they are essentially clitics, especially since the following marking strategy is the grammatical one in absence of an agent NP (compare section 3.2.5, p. 89):

- (158) a. *Manye sa Pila.*  
 man-ye sa Pila  
 greet-3SG.F P Pila  
 ‘Pila is being greeted.’  
 b. *Manyes.*  
 man=yes  
 greet=3SG.F.P  
 ‘She is being greeted.’

The verb here agrees with the patient—or is it that person agreement suffixes on verbs are generally clitics in Ayeri, even where they do not involve case marking? There seems to be a gradient here between what looks like regular verb agreement with the agent on the one hand, and agent or patient pronouns just stacked onto the verb stem on the other hand. For an overview, compare Table 4.22. In this table, especially the middle, transitional category is interesting in that what looks like verb agreement superficially can still govern topicalization marking, which is indicated in column II by an index ‘I’. Note that this behavior only occurs in transitive contexts; there is no topic marking on the verb if the verb only has a single NP dependent. Also consider that for example (b) in the type III transitive cell the question is,

whether this should not better be analyzed as AT ...-3SG.M.TOP ...-TOP ...-P, with co-indexing of the topic on the person inflection of the verb, making it structurally closer to type II.

As for personal pronouns fused with the verb stem like in the first column, Corbett (2006) points out that

In terms of syntax, pronominal affixes are arguments of the verb; a verb with its pronominal affixes constitutes a full sentence, and additional noun phrases are optional. If pronominal affixes are the primary arguments, then they agree in the way that anaphoric pronouns agree [...] In terms of morphology, pronominal affixes are bound to the verb; typically they are obligatory [...]. (99–100)

This seems to be exactly what is going on for instance in (154) and (158b), where the verb forms a complete sentence. It needs to be pointed out that Corbett includes an example from Tuscarora, a native American polysynthetic language, in relation to the above quotation. Ayeri should not be considered polysynthetic, however, since its verbs generally do not exhibit relations with multiple NPs, at least as far as person and number agreement is involved (Comrie 1989: 45–46).<sup>35</sup>

Taking everything written above so far into account, it looks much as though Ayeri is in the process of grammaticalizing personal pronouns into person agreement (Lehmann 2015: 42–45; Gelderen 2011: 493–497). Corbett (2006: 76–77) illustrates an early stage of such a process:

- (159) a. *Ke móe ke=fue.* (\**Ke móe fue.*) [Skou]  
           3SG.M fish 3SG.M=see.3SG.M  
           ‘He saw a fish.’
- b. *Pe móe pe=fu.* (\**Pe móe fu.*)  
           3SG.F fish 3SG.F=see.3SG.F  
           ‘She saw a fish.’

What Gelderen (2011) calls the *subject cycle*, the “oft-noted cline expressing that pronouns can be reanalyzed as clitics and agreement markers” (493) applies here, and as well in Ayeri. However, while she continues to say that in “many languages, the agreement affix resembles the emphatic pronoun and derives from it” (494), Ayeri does at least in part the opposite and uses the case-unmarked form of personal pronouns for what resembles verb agreement most closely. This, however, should

<sup>35</sup> The topic NP marked on the verb may be a different from the one with which the verb agrees in person and number, so technically, Ayeri verbs *may* agree with more than one NP in a very limited way (compare section 3.3). Still, I would not analyze this as polypersonal agreement, since there is only canonical verb agreement with one constituent, that is, the agent NP. Topic marking should, in my opinion, be viewed as a separate agreement relation, as pointed out in the quoted section above.

Table 4.22: Verb inflection types in Ayeri

	Type I: Clitic pronouns	Type II: Transitional	Type III: Verb agreement
Inflectional categories	Person Number Case	Person Number Case/Topic	Person Number
Examples (intransitive)	...=yāng ...=3SG.M.A	—	...-ya <sub>1</sub> ...-ang <sub>1</sub> ...-3SG.M ...-A
Examples (transitive)	sa <sub>1</sub> ...=yāng ...-Ø <sub>1</sub> PT ...=3SG.M.A ...-TOP	ang <sub>1</sub> ...=ya.Ø <sub>1</sub> ...-as AT ...=3SG.M.TOP ...-P	a. ang <sub>1</sub> ...-ya <sub>1</sub> ...-Ø <sub>1</sub> ...-as AT ...-3SG.M ...-TOP ...-P  b. a <sub>1</sub> ...-ya <sub>2</sub> ...-ang <sub>2</sub> ...-Ø <sub>1</sub> PT ...-3SG.M ...-A ...-TOP

Syntax:	non-argument	argument	
Linguistic element:	'pure' agreement marker	pronominal affix	free pronoun
Morphology:	inflectional form		free form

Figure 4.3: The syntax and morphology of pronominal affixes (Corbett 2006: 101)

not be too controversial either, considering that, for instance, semantic bleaching and phonetic erosion go hand in hand with grammaticalization (Lehmann 2015: 136–137; Gelderen 2011: 497).

As pointed out above in (158), Ayeri usually exhibits verbs as agreeing with agents and occasionally patients, not topics as such. This may be a little counterintuitive since the relation between topics and subjects is close, but is possibly due to the fact that the unmarked word order is VAP. This means that agent NPs usually follow the verb, and does not seem too unnatural to have an agreement relation between the verb and the closest NP also when non-conjoined NPs are involved (Corbett 2006: 180). This conveniently explains why verbs can agree with patients as well if the agent NP is absent. Taking into account that the grammaticalization process is still ongoing so that there is still some relative freedom in how morphemes may be used if a paradigm has not yet fully settled (Lehmann 2015: 148–150) also makes this seem less strange. Verbs simply become agreement targets of the closest semantically plausible nominal constituent. Ayeri seems to be shifting from topics to subjects, and as a consequence the bond between agents and verbs is strengthened due to their usual adjacency; developing verb agreement with agents may be seen as symptomatic of this change.<sup>36</sup>

Signs so far point towards Ayeri's person agreement in fact being more likely enclitic pronominal affixes. The question is, then, how this might be corroborated. Corbett (2006) offers a typology here, see Figure 4.3. According to this typology, a pronominal affix is syntactically an argument of the verb but has the morphology of an inflectional form (compare section 3.2.5, p. 89). If we compare this to the

<sup>36</sup> When translating things in Ayeri, I find myself very often using agent topics, which may be because I am used to subjects proper. Supposing that this is also what Ayeri prefers in-universe, it would make sense to assume the usual grammaticalization path by which topics become subjects, thereby also leading to subject-verb agreement by means of resumptive pronouns referring back to left-dislocated topics (Lehmann 2015: 121–122; Gelderen 2011: 499–500). Lehmann (2015: 120) gives colloquial French *Jean, je l'ai vu hier* 'John, I saw him yesterday' as an example here: the object clitic *l'* (← *le* 'him') may well develop into an agreement affix (also see Gelderen (2011: 498) on a dialect of Spanish in which, she argues, this has happened). Ayeri in its present form does not use clitic doubling (Spencer and Luís 2012: 153–161), however.

gradient given in Table 4.22 above, it becomes evident that type I definitely fulfills these criteria, and type II does so as well, in fact, in that there is no agent NP that could serve as a controller if the verb inflection in type II were ‘merely’ an agreement target. The inflection in type III, on the other hand, appears to have all hallmarks of agreement in that there is a controller NP that triggers it, with the verb serving as an agreement target. Moreover, the person marking on the verb is not a syntactic argument of the verb in this case. As example (158a) shows, however, marking of type III permits the verb to mark more than one case role, which makes it slightly atypical, although verbs can only carry a single instance of person marking (Corbett 2006: 103). Regarding referentiality, the person suffixes on the verb in table 1, columns I and II are independent means of referring to discourse participants mentioned earlier, whereas the person suffix in III needs support from an NP in the same clause as a source of morphological features to share:

- (160) a. *Ajān ... Ang manya sa Pila.*  
           *Ajān ... Ang man=ya.Ø sa Pila*  
           *Ajān ... AT greet=3SG.M.TOP P Pila*  
           ‘*Ajān ... He greets Pila.*’
- b. *Ajān ... Sa manyāng Pila.*  
           *Ajān ... Sa man=yāng Ø Pila*  
           *Ajān ... PT greet=3SG.M.A TOP Pila*  
           ‘*Ajān ... It’s Pila that he greets.*’
- c. \**Ajān ... Manya sa Pila.*  
           *Ajān ... Man-ya sa Pila*  
           *Ajān ... greet-3SG.M P Pila*

Since person marking of the type I and II is *referential*, as shown in example (160a) and (160b), it can be counted as a cliticized pronoun (103). Pronouns in Ayeri can also refer to non-people—there are both a ‘neuter’ gender for non-people considered living (or being closely associated with living things), and an ‘inanimate’ gender for the whole rest of things (compare section 4.1.1). Since mere agreement as in type III needs support from an NP within the verb’s scope, though, it does not have *descriptive/lexical content* of its own. That is, it *only* serves a grammatical function (104), not strictly as an anaphora. As for Corbett (2006)’s *balance of information* criterion, Table 4.22 also highlights differences in what information is provided by the person marking. Nouns in Ayeri inherently bear information on person, number, and gender, and all three types of person inflection on verbs share these features. However, there are no extra grammatical features indicated by the first two inflection types that are not expressed by noun phrases, although under a very close understanding of Corbett (2006), the following example (161) may still qualify as person-marking on the verb realizing a grammatical feature shared with

an NP that is not openly expressed by the NP. Corbett (2006) writes that in the world's languages, this frequently is number (105). This, however, does not apply to Ayeri because the only time verbs display number not expressed overtly by inflection on a noun is in agreement like in type III (a):<sup>37</sup>

- (161) *Ang sabayan ayon kay kong nangginoya.*  
 ang saha-yan ayon-Ø kay kong nanggino-ya  
 AT come-3PL.M man-TOP three into tavern-LOC

‘Three men come into a pub.’

As shown above, verb marking of the types I and II is independent as a reference, so there is *unirepresentation* of the marked NP. In contrast, verb marking of type III requires a controlling NP in the same clause to share grammatical features with, so that there is *multirepresentation* typical of canonical agreement (106). A further property that hinges on types I and II being independent pronouns glued to verbs as clitics is that they are not coreferential with another NP of the same grammatical relation, but are in complementary distribution, as commonly assumed with pronominals (108). Hence, either of these two examples is ungrammatical:

- (162) a. \**Lampyāng ang Ajān.*  
 lamp=yāng ang Ajān  
 walk=3SG.M.A A Ajān  
 b. \**Ang lampyāng Ajān.*  
 ang lamp=yāng Ø Ajān  
 AT walk=3SG.M.A TOP Ajān

However, verb agreement with a free pronoun is also not possible even though it might be expected according to (109)—also compare example (151b) above. Instead, the agent pronoun replaces any possible person agreement on the verb (see section 3.2.5, p. 89 ff., for an attempt to explain this effect from a syntactic point of view):

- (163) a. *Lampyāng.*  
 lamp=yāng  
 walk=3SG.M  
 ‘He walks.’  
 b. \**Lampya yāng.*  
 lamp-ya yāng  
 walk-3SG.M 3SG.M.A

<sup>37</sup> From a Lexical-Functional Grammar point of view, the number feature of *kay* in (161) coalesces with the semantic features provided by *ayon* in the maximal projection; agreement is thus with the whole agent NP rather than just with *ayon* as the NP's categorial head.

In conclusion, we may assert that Ayeri appears to be in the process of grammaticalizing pronouns as verb inflections, however, how far this grammaticalization process has progressed is dependent on syntactic context. Ayeri displays a full gamut from personal pronouns (usually agents) glued to verbs as clitics to agreement with coreferential NPs that is transparently derived from these personal pronouns. With the latter, the complication arises that coreferential pronoun NPs are not allowed as agreement controllers as one might expect, but only properly nominal NPs. Slight oddities with regards to Austronesian alignment—Ayeri's actors bear more similarities to subjects than expected, but still without fully conflating the two notions—can possibly be explained by a strengthening of the verb-agent relationship pointed out as a grammaticalization process in this article as well. Information on agreement with committee nouns and coordinated NPs with incongruent agreement features can be found in the section on VPs.

#### 4.5.2 Tense

Tense in Ayeri is often not explicitly marked, but has to be inferred from context. However, where marked, Ayeri distinguishes past and future as referring to past and future events, respectively. Both past and future tenses come with three degrees each: near, recent/impending, and remote. Ayeri's distinguishing three degrees of both past and future time is a little unusual with regards to typology according to the survey conducted by Dahl (1985: 127). The decision for which subtier of the past and the future to use is up to pragmatics, that is, there are no definitive and clear-cut lines. The near-time markers are most commonly used for immediate scope, that is, things which have just happened or will happen in a moment. The recent/impending-time markers may then be used for anything else which does not qualify as remote, that is, a long time into the past or the future from the point of view of the speaker.

Dahl (1985: 117) further notes that among the languages in the surveyed sample, past tenses are mostly marked by suffixes, the marking of this category being extremely common in addition. Ayeri may thus be a little unusual crosslinguistically again by exclusively using prefixes for tense marking. This makes sense, however, if we assume that historically, the tense prefixes once were auxiliary verbs. Ayeri applies head-first word order to subordinating verbs, as we will see further below, so these prefixes may just have begun to *procliticize* instead of slipping into a position behind their head (that is, Wackernagel's position).

Of the triad tense–aspect–mood this section will only cover basic uses of the marked tense categories, followed by a discussion of complex tense combinations such as past-in-future. The subsequent section 4.5.3 will provide more insight into the morphological marking of aspectual categories; section 4.5.4 deals with the morphology of mood marking in Ayeri.



*Present tense*

Verbs in Ayeri are unmarked for present tense, as it is the normal mode of speaking. Besides being used to comment or report on current events, the present tense is also used to make statements of general truth:

- (164) *Sa arapyo tabanyamanang koyana nogalam-ikan.*  
 sa arap-yo tahanyaman-ang koya-na nogalam-Ø=ikan  
 PT require-3SG.N writing-A book-GEN patience-TOP=much

‘Writing a book requires much patience.’

Moreover, Ayeri does not strictly mark its verbs for past tense in narrative discourses—verbs may thus appear as though with a present-time reference in spite of recounting past events, whether historical or fictional. See the next subsection on the past tense.

*Past tense*

The past tense indicates actions in the past if not further modified. The three degrees of past tense are marked with  $\text{ḱ}$ : *kə-* (near/immediate),  $\text{ḱ}$ : *mə-* (recent), and  $\text{ḱ}$ : *və-* (remote), which attach right in front of a verb root. In spite of the customary spelling of the past tense prefixes with ⟨ə⟩, which reflects pronunciation, they have an underlying /a/ vowel in this place. This means that the vowel of the tense prefixes coalesces with a following /a/ to form a long vowel (see section 1.1.2), which is demonstrated in example (165b) below:

- (165) a. *Ang kəsilvay yes motonya.*  
 ang kə-silv=ay.Ø yes moton-ya  
 AT NPST-see=1SG.TOP 3SG.F.P store-LOC  
 ‘I’ve just seen her at the store.’  
 b. *Le mādruyāng ikan biratay.*  
 le mə-adru=yāng ikan biratay-Ø  
 PT.INAN PST-break=3SG.M.A wholly pot-TOP  
 ‘The pot, he completely broke it.’  
 c. *Vəmittang edaya.*  
 və-mit=tang edaya  
 RPST-live=3PL.M.A here  
 ‘They lived here (a long time ago)’

Note that the recent and the remote past tense are not generally marked if the past context is clear, for instance, when a past context has already been established in discourse. This may also happen explicitly by using a time adverbial such as  $\text{ḱ}$ : *ḱ*

*tamala* ‘yesterday’ or ၵၵ်းပီၵ်းပီၵ်း *pericanya menang pang* ‘a hundred years ago’. In the presence of an explicit time adverbial, redundant tense marking is also dropped subsequently:

- (166) *Ang kondayn kadanya terpasānley bihanya sarisa.*  
 ang kond=ayn.Ø kadanya terpasān-ley bihan-ya sarisa  
 AT eat=IPL.TOP together lunch-P.INAN week-LOC previous  
 ‘We had lunch together last week.’

The reference to a past time frame is explicitly given in this example by the adverbial phrase ၵၵ်းပီၵ်းပီၵ်း *bihanya sarisa* ‘last week’, hence the verb appears here simply as ၵၵ်း *kondayn*, rather than with redundant past-tense marking as ၵၵ်းမၵ်း *məkondayn*. Since past tense is often underspecified in Ayeri, the language also does not employ past forms in narrative contexts like English, among others, commonly does:

- (167) The sky above the port was the color of television, tuned to a dead channel.  
 (Gibson 1995 [1984]: 9)

This quote is, of course, the first sentence of Gibson (1995 [1984])’s novel *Neuromancer*, which never mentions any definite dates, but is clearly set in a future world, maybe somewhere in the latter half of the twenty-first century.<sup>38</sup> Yet, however, Gibson (1995 [1984]) recounts events which are logically happening in an imagined future as having already happened in the past: he uses the past tense as a convention of storytelling. What Ayeri, then, does in contrast to English is to basically treat stories as though happening in the present; adverbials referring to past time may, again, set up the correct time frame if required. Ayeri is in good company here, since according to Dahl (1985) “[m]ore common than marking narrative contexts [...] is not marking them—quite a considerable number of languages use unmarked verb forms in narrative contexts” (113). This, however, is yet different from a narrative present, that is, the use of present tense within a past context, which languages like English may use in narrative contexts to increase the feeling of immediacy and thus raise suspense. The following example from an Ayeri translation of the well-known Aesopian fable, ‘The North Wind and the Sun’ (compare International Phonetic Association 2007: 39), illustrates Ayeri’s non-marking of tense on verbs in narrative contexts:

<sup>38</sup> Christian (2017) reports that Gibson himself pictured his novel as set around 2035, though he has since realized that this cannot be right. One of the characters, the Finn, “makes an offhand reference to the ‘Act of ‘53’ as a law [which] deals with the citizenship status of artificial intelligences” (Christian 2017; also compare Gibson 1995 [1984]: 92)—this is very unlikely to refer to 1953.

- (168) *Ang manga ranyon adauyi Pintemis nay Perin, engyo*  
 ang manga ran-yon adauyi Ø Pintemis nay Ø Perin eng-yo  
 AT PROG argue-3PL.N then TOP North Wind and TOP Sun, be.more-3SG.N  
*mico sinyāng luga toya, lingya si lugaya asāyāng si*  
 mico sinyā-ang luga toya ling-ya si luga-ya asāya-ang si  
 strong who-A among 3PL.N.LOC, while-LOC REL pass-3SG.M traveler-A REL  
*sitang-naykonyāng kong tovaya mato.*  
 sitang-naykon=yāng kong tova-ya mato  
 self-wrap=3SG.M.A inside cloak-LOC warm.

‘The North Wind and the Sun were then arguing which among them is stronger, all the while a traveler passed by who had wrapped himself in a warm cloak.’

#### Future tense

Future tense marks explicit references to future time in Ayeri, that is, “someone’s plans, intentions or obligations” (Dahl 1985: 103), as well as predictions. The future prefixes behave analogously to the ones indicating past tense: *ṛ* *pə-* indicates immediate/near future (NFUT), *ṣ* *sə-* indicates impending future (FUT), and *ṣ̃* *ni-* indicates remote future (RFUT). Underlying the reduced vowels in *ṛ* *pə-* and *ṣ* *sə-* are /a/ and /e/, respectively, so that these prefixes cause adjacent vowels of the same type to lengthen as usual; the same, of course, applies to *ṣ̃* *ni-* regarding /i/. The following examples show the future tense markers in context:

- (169) a. *Pəsahayang!*  
 pə-saha=yang  
 NFUT-come=1SG.A  
 ‘I’m coming (in a moment)!’  
 b. *Ang səkarsayn kankaya.*  
 ang sə-kars=ayn.Ø kanka-ya  
 AT FUT-freeze=1SG.TOP snow-LOC  
 ‘We will freeze in the snow.’  
 c. *Paronatang, nisa-sabaya dibakayāng.*  
 parona=tang ni-sa~saha-ya dihakaya-ang  
 believe=3PL.M.A RFUT-ITER~come-3SG.M prophet-A  
 ‘They believe that the prophet will return (one day).’

Like the past tense, the future is often not explicitly marked if the time frame is clear from context or has been clarified with such adverbials as *ṣ̃ṣ̃ṣ̃* *tasela* ‘tomorrow’, *ṣ̃ṣ̃ṣ̃ṣ̃* *mangasaba pericanya* ‘in a year’, or *ṣ̃ṣ̃ṣ̃* *metay* ‘sometime’:

- (170) *Ang raypāy vāya bihanya mararya.*  
 ang raypa=ay.Ø vāya bihan-ya mararya  
 AT stop=ISG.TOP 2SG.LOC week-LOC next  
 ‘I’m stopping by you next week.’

It is possible here to explicitly mark the verb for future tense as well, for example, to make a promise, or to otherwise emphasize that the future condition will come to pass:

- (171) *Səsidejang tasela, diran.*  
 sə-sideg=yang tasela diran  
 FUT-repair=ISG.A tomorrow uncle  
 ‘I *will* repair it tomorrow, uncle.’

#### Past in past

So far, we have only dealt with tense marking from the point of view of the present. However, it is also possible to refer to an event which precedes another event in the past. Ayeri makes little use of auxiliary verbs, and thus the regular morphological and pragmatic means of tense marking have to cover this relation as well. In order to indicate pre-past events, it is customary to explicitly mark the verb for past time in Ayeri, in difference to the common lack of morphological marking for plain past tense. However, as it is possible for the *ə: mə-* prefix to be used to refer to ‘regular’ past events from a present point of view as well, context again has to provide the information that the frame of reference is past in this case, rather than the speaker’s present.

- (172) CONTEXT: Ajān’s past travels  
*Ya məsaraya iri maritay ang Ajān Tasankan*  
 ya ma-sara-ya iri maritay ang Ajān Ø Tasankan  
 LOCT PST-go-3SG.M already before A Ajān TOP Tasankan  
 ‘Tasankan, Ajān had already gone there before.’

The above example is essentially ambiguous as to the reference point. The explicit tense marking draws attention to the fact that the event definitely lies in the past and the adverbs underline this fact. Instead of reading the sentence as referring to a pre-past event, it is equally possible to read it from a present-time point of view as ‘Ajān has already gone to Tasankan before’, although under these circumstances, it would be more common to leave the *ə: mə-* out, as described in section 4.5.2:

(173) CONTEXT: Ajān's current traveling plans

*Ya saraya iri maritay ang Ajān Tasankan*  
 ya sara-ya iri maritay ang Ajān Ø Tasankan  
 LOCT go-3SG.M already before A Ajān TOP Tasankan

'Tasankan, Ajān already went there before.'

Likewise, it is possible to make plans in the past with the intention of them coming to fruition only later, possibly at a point before the current time or even further in the future. The English idiom to express this time relation is 'was going to'; in Ayeri, the relation cannot be expressed by morphological means, but only by lexical ones. Thus, 𐑦 *no-* 'want; plan to' must be used, together with explicit past marking. Since 𐑦 *no-* is a modal particle (see section 4.5.5), inflection is placed on the content verb.

(174) CONTEXT: Ajān's having gone to Tasankan

*Ang no m̐inca tosangyeley hiro yam Pila.*  
 ang no ma-int=ya tosang-ye-ley hiro yam Pila  
 AT want PST-buy=3SG.M.TOP earring-PL-P.INAN new DAT Pila

'He had planned to buy new earrings for Pila.'

The time relation expressed here is, thus, essentially that of a pre-past event again, since the planning of the action of buying took place before the time of going to Tasankan.

#### *Past in future*

Of course, it is also possible to refer to future actions or events which will already have happened before a point further in the future. From the point of view of the later event, the closer event will thus already lie in the past, forming its prerequisite. As with future-in-past, there is no way in Ayeri to mark this relation morphologically, but lexical means have to be used, that is, first and foremost the adverb 𐑦 *iri* 'already', which indicates that an action has been completed in the past. As with other future actions, the time frame must be inferred from context if it is not indicated explicitly by temporal adverbs or future-tense marking (compare section 4.5.2).

(175) CONTEXT: Ajān's traveling to Tasankan

*Ang girenja iri nilay sirutayya tamala pesan.*  
 ang girend=ya.Ø iri nilay sirutay-ya tamala pesan  
 AT arrive=3SG.M.TOP already probably evening-LOC tomorrow before

'He will probably already (have) arrive(d) before tomorrow evening.'

Strictly speaking, the above example does not make it explicit whether Ajān *will arrive* before evening or *will have arrived*. In order to indicate that the action is all but complete, the cessative adverb 𐎠𐎢𐎩 *mayisa* ‘be done; ready’ may be added:

- (176) *Girenjāng      mayisa    iri.*  
 girend=yāng    mayisa    iri  
 arrive=3SG.M.A   be.done   already

‘He already has arrived’, or: ‘He will already have arrived.’

### 4.5.3 Aspect

Aspectually unmarked verb forms indicate general statements, which may be completed or ongoing, depending on the meaning of the verb itself. Ayeri seems not to make strict formal distinctions with regards to either, perfectivity or lexical aspect. It needs to be noted, however, that at least to date, it is not entirely clear how Ayeri fares with regards to conceptualizing perfectivity, which Dahl (1985: 76) in reference to Comrie (1976: 16) characterizes as being based on the conceptualization of actions or events as bounded or otherwise limited wholes, versus a lack of closure. Dahl (1985) also notes that “it seems rather to be a typical situation that even in individual languages, we cannot choose one member of the opposition [perfective–imperfective] as being clearly unmarked” (69). He further argues that

The difficulty of deciding which member of the opposition is marked and which is unmarked is connected with the tendency for PFV:IPFV to be realized not by affixation or by periphrastic constructions but rather by less straightforward morphological processes. (73)

In other words: it *is* a difficult category to assess, in spite of being “often taken to be ‘the’ category of aspect” (69), mostly since languages often do not realize it by straightforward means. In Ayeri, the most tangible way of expressing completeness of an action is to use adverbs like 𐎠𐎢𐎩 *mayisa* ‘ready, done’, 𐎠𐎢𐎩 *iri* ‘already’, 𐎠𐎢𐎩 *ikan* ‘completely, wholly’ (also as an adjective); a quantifier like 𐎠𐎢𐎩 *-hen* ‘all’; verbs like 𐎠𐎢𐎩 *samir-* ‘finish’, 𐎠𐎢𐎩 *panga-* ‘end’, and 𐎠𐎢𐎩 *raypa-* ‘stop’; or an indefinite pronoun like 𐎠𐎢𐎩 *enya* ‘everything, everybody’:

- (177) *Le           kondjeng    enya.*  
 le           kond=yeng    enya-Ø  
 PT.INAN   eat=3SG.F.A   everything

‘She ate everything.’ or: ‘She ate it all up.’

Apart from the more general dilemma of determining how perfectivity is expressed in detail, Ayeri marks verbs openly by morphological means to indicate

progressive, habitual, and iterative actions—by their nature all conceptualizing actions as being composed of a series of two or more related actions of the same kind, though not necessarily implying a strong semantic connection to the past. The following sections will discuss each of these categories.

### Progressive

In order to indicate an ongoing action explicitly, Ayeri employs the marker *ang* *manga*, which we have already seen with directional prepositions above (section 4.4.1). This marker is a clitic within the verb phrase and precedes the verb word:

- (178) *Ang manga ilye karonas nakajyam.*  
 ang manga il=ye.Ø karon-as naka-ye-yam  
 AT PROG give=3SG.F.TOP water-P plant-PL-DAT  
 ‘She is giving water to the plants.’

Going by the data presented by Dahl (1985: 91), Ayeri is typologically unremarkable in marking progressive aspect with a periphrastic construction, although it is remarkable in possessing morphological progressive marking at all—it only occurs in 27% of the languages in Dahl (1985)’s sample. Typical of progressives, this form of the verb is not limited to present contexts in Ayeri as exemplified in (178) above. Instead, it is possible to also use the progressive in past (179a) and future (179b) contexts, the latter being probably less typical, though:

- (179) a. *Ang manga gumya Ajān tadayya si ya kongaye ang Pila*  
 ang manga gum-ya Ø Ajān taday-ya si ya konga-ye ang Pila  
 AT PROG work-3SG TOP Ajān time-LOC REL LOCT enter-3SG.F A Pila  
*gumanga tamala.*  
 gumanga-Ø tamala  
 workshop-TOP yesterday  
 ‘Ajān was working when Pila entered the workshop yesterday.’  
 b. *Ang manga nimpay rangya nā tadayya si cunyo bekalang*  
 ang manga nimp=ay.Ø rang-ya nā taday-ya si cun-yo bekal-ang  
 AT PROG run=1SG.TOP home-LOC 1SG.GEN time-LOC REL begin-3SG.N festival-A  
*tasela.*  
 tasela  
 tomorrow  
 ‘I will be running home when when the festival starts tomorrow.’

Ignoring the constructedness of the above examples, the time adverb is located in the relative clause in both sentences in this case. For illustrative purposes, let us assume that a narrative context with the respective time frames has already been

established in (179). As noted above, Ayeri prefers to not mark every verb for tense explicitly when the context is clear already, insofar the argument that progressive aspect works independent of *tense* needs corroboration; the question being whether constructions like  $\text{māṅ} \text{ m} \text{---}$  *māṅ m*... (PROG PST-...) are possible. Strictly speaking, there is nothing to prevent this construction, however, we have to wonder if it is actually *natural* to phrase things this way. What can be said at least is that progressive marking is possible within a context referring to past or future actions and events irrespective of their explicit marking on the verb. Furthermore, the examples in (179) illustrate a very typical use of the progressive as a structuring means, that is, an ongoing background action may be expressed using a progressive form, while an interrupting action receives no special marking (compare the past progressive in English).

#### Habitual

Unlike the few instances of habitual marking in Dahl (1985)'s survey (96), Ayeri possesses a suffix for marking habitual actions on the verb:  $\text{asa}$  *-asa*, where the first *-a* replaces the terminal vowel of a verb stem if present, compare example (180b) below. The habitual aspect in Ayeri stresses that an action is carried out as a habit, that is, not just a few times, but with regular frequency. Essentially, verbs marked with the habitual in Ayeri can be translated by adding the adverb *usually* in English (97). The habitual aspect is not restricted to present actions or absolute statements like the one in (180a), but can also be used in past contexts to express that something *used to* be done in the past as in (180b). While the contexts are probably very few, there are no restrictions about using the habitual also in contexts relating to future actions which are predicted to be carried out habitually. The following sentences illustrate typical contexts in which the habitual may be used:

- (180) a. *Le kondasayāṅ bemaṃe pruyya nay napayya kayvay.*  
 le kond-asa=yāṅ hema-ye-Ø pruy-ya nay napay-ya kayvay  
 PT.INAN eat-HAB=3SG.M.A egg-PL-TOP salt-LOC and pepper-LOC without  
 'He always eats his eggs without salt and pepper.'
- b. *Ang aṣayān ranisungas tadayya si yāṅ ganas.*  
 ang aja-asa=ayn.Ø ranisung-as taday-ya si yāṅ gan-as  
 AT play-HAB=1PL.TOP hide.and.seek-P time-LOC REL 1SG.A child-P  
 'We used to play hide-and-seek when I was a child.'

Importantly, the verb root with habitual marking forms a new verb stem to which affixes may be attached. This is relevant to mood suffixes, which follow aspectual marking.



In addition to a simple iterative meaning, a frequentative meaning like ‘walk around’, ‘cry all the time’, or ‘keep asking’ can be achieved by combining the iterative and progressive aspects, that is, the verb is both modified by *am-* *manga* for progressive aspect and partial initial reduplication for iterative aspect:

- (184) a. *Ang manga la-lampay saba-sara manga luga babisya-ben.*  
ang manga la~lamp=ay.Ø saha-sara manga luga bahis-ya=hen  
AT PROG ITER~walk=ISG.TOP back.and.forth DIR while day-LOC=all  
‘I was walking around back and forth all day long.’
- b. *Ang manga si-sipyé kimay sirutayya.*  
ang manga si~sip-ye kimay.Ø sirutay-ya  
AT PROG ITER~cry-3SG.F baby.TOP night-LOC  
‘The baby, she is crying all the time at night.’
- c. *Manga pa-prantu!*  
manga pa~prant-u  
PROG ITER~ask-IMP  
‘Keep asking!’

#### Lexically marked aspectual categories

Besides using morphological means, Ayeri expresses some aspectual categories by way of lexical items, that is, verbs and adverbs. The relevant words in this respect are the adverbs *sirimang* ‘about to’ (prospective) and *mayisa* ‘ready; be done’ (cessative), as well as the verb *cun-* ‘begin, start’ (inchoative):

- (185) *Saratang sirimang.*  
sara=tang sirimang  
leave=3PL.M.A about.to  
‘They are about to leave.’
- (186) *Konjang mayisa.*  
kond=yang mayisa  
eat=ISG.A be.done  
‘I am done eating.’
- (187) *Pəcunreng seyaryam.*  
pə-cun=reng seyar-yam  
NFUT-begin=3SG.INAN.A rain-PTCP  
‘It is going to start raining any moment.’

Prospective *sirimang* (185) and cessative *mayisa* (186) are expressed by adverbs which are regularly following verbs as their heads. They precede other adverbs due to a higher amount of semantic bondedness, by tendency, than other descriptive adverbs. For this reason, as well as for expressing a grammatical function rather than lexical meaning with the original meaning still transparent, they appear

to be on the verge of grammaticalization. In contrast, the verb inchoative  $\text{ᄃᆞᆫ}$  *cun-* (187) is part of a periphrastic verb construction, that is,  $\text{ᄃᆞᆫ}$  *cun-* requires a content-verb VP as a complement rather than an NP. The content/main verb appears in a non-finite form marked by  $\text{ᄃᆞᆫ}$  *-yam*, which will be described below.

#### 4.5.4 Mood

Besides various aspects, Ayeri also marks mood other than realis: irrealis, imperative, hortative, and negative. These are expressed by suffixes on the verb and typically follow aspectual marking where it is expressed by a suffix, that is, the habitative suffix  $\text{ᄃᆞᆫ}$  *-asa*. The following subsections will discuss each of the modal categories expressed by suffixes; modals proper will be discussed in section 4.5.5.

##### *Irrealis*

Irrealis marking in Ayeri is indicated by the suffix  $\text{ᄃᆞᆫ}$  *-ong* and marks that an action is thought of as hypothetical by the speaker, whether he or she expects it to be fulfilled or not:

- (188) *Sabongvāng edaya, ming silvongvāng sitang-vāri.*  
 saha-ong=vāng edaya ming silv-ong=vāng sitang=vāri  
 come-IRR=2SG.A here can see-IRR=2SG.A REFL=2SG.INS

‘If you came/had come here, you could see/have seen it yourself.’

As (188) shows, irrealis marking is especially prominent in conditional clauses which express a hypothetical cause and effect. Both condition/protasis and consequence/apodosis are marked with the irrealis suffix in this case. The example sentence also shows that, again, the initial vowel of the suffix replaces the last vowel of the verb stem, if there is one, so that  $\text{ᄃᆞᆫ}$  *saha-* becomes  $\text{ᄃᆞᆫ}$  *sabong-*, to which further mood suffixes may be added, and finally, person marking.

The same suffix,  $\text{ᄃᆞᆫ}$  *-ong*, is also used in other contexts expressing inactual events, for instance, in reported speech, or complement clauses expressing a wish about the actualization of a hypothetical event:

- (189) *Narayeng, ang menongye demās yena.*  
 nara=yeng ang menu-ong=ye.Ø dema-as yena  
 say=3SG.F.A AT visit-IRR=3SG.F.TOP aunt-P 3SG.F.GEN

‘She said she were visiting her aunt.’

- (190) *Hanuyang, koronongyang maritay.*  
 hanu=yang koron-ong=yang maritay  
 wish=1SG.A know-IRR=1SG.A before

‘I wish I had known this before.’

Irrealis marking does not, however, appear in contexts that express requirements on or wishes about a third person's actions, that is, typical subjunctive contexts; the verb in the complement clause rather appears in the indicative in these contexts. To add a sense of expectation of compliance about the action, the modal *g mya* 'be supposed to' may be added, see section 4.5.5.

- (191) a. \**Arapnang, sa garongyāng hatay.*  
           arap=nang sa gara-ong=yāng hatay-Ø  
           require=IPL.A PT call-IRR=3SG.M.A police-TOP
- b. *Arapnang, sa (mya) garayāng hatay.*  
           arap=nang sa (mya) gara=yāng hatay-Ø  
           require=IPL.A PT (be.supposed.to) call=3SG.M.A police-TOP
- 'We require that he call the police.'

### Negative

The negative mood is used to negate verbs, which is separate from irrealis marking: negation of verbs is marked by the suffix *-oy*, which has an allomorph *-u* before diphthongs in romanization and also in pronunciation. The Tahano Hikamu spelling is more conservative here and keeps the spelling *-oyay* for */-uay/* (-NEG=ISG.TOP). Like the irrealis suffix, the negative suffix deletes the last vowel of the verb stem if present, which is exemplified in (192b) besides this example showing the *-u* allomorph. Moreover, example (192c) shows that negative marking usually follows irrealis marking when suffixes are stacked: *-ong + -oy → -ongoy*.

- (192) a. *Ang silvoyyan nasiyamanas tan.*  
           ang silv-oy=yan.Ø nasi-yam-an-as tan  
           AT see-NEG=3PL.M.TOP approach-PTCP-NMLZ-P 3PL.M.GEN
- 'They did not see them approaching.'
- b. *Ang peguay kalam adaley!*  
           ang pega-oy=ay.Ø kalam ada-ley  
           AT steal-NEG=ISG.TOP honestly that-P.INAN
- 'I didn't steal it, honestly!'
- c. *Ang tendongoyva sarayam adaya.*  
           ang tend-ong-oy=va.Ø sara-yam adaya  
           AT dare-IRR-NEG=2SG.TOP go-PTCP there
- 'You would not dare to go there.'

If negated verbs appear together with negative indefinite pronouns (compare section 4.2.4), multiple negatives do not cancel each other out, but amplify the

negation instead. This is to say that Ayeri allows for multiple negation as a means to emphasize the impossibility of something.

- (193) *Le gamaroyya tadoy ranyāng adanya.*  
 le gamar-oy-ya tadoy ranyāng adanya-Ø  
 PT.INAN manage-NEG-3SG.M never nobody-A that-TOP

‘Nobody ever managed that’,  
 literally: ‘Nobody never didn’t manage that.’

### Imperative

The imperative mood is used to mark orders to an unspecified second person, that is, imperative verbs do not require an overt second person agent; if an addressee is included, it is oblique and unmarked for case, see section 4.1.3. Moreover, no distinction is made between singular and plural second-person addressees, so that the marker is 𐌸 -u in either case. Like the other mood suffixes, the vowel of the imperative suffix replaces the vowel of the verb stem if there is one.

- (194) a. *Giru māy!*  
 gira-u māy  
 hurry-IMP INT  
 ‘Hurry up!’
- b. *Tangu yām, Yan!*  
 tang-u yām Yan  
 listen-IMP 1SG.DAT Yan  
 ‘Listen to me, Yan!’
- c. *Tangu yām, ledanye nā!*  
 tang-u yām ledan-ye nā  
 listen-IMP 1SG.DAT friend 1SG.GEN  
 ‘Listen to me, my friends!’

It is important to note that imperative-marked verbs behave essentially as infinitive forms in that they do not exhibit any agreement in person, number, gender, and topic, and also cannot act as hosts for clitic personal pronouns. Imperative verbs may be marked for negative and hortative, however. Hence, for instance, (195) is grammatical, while the examples in (196) are not.

- (195) *Saroyu yas!*  
 sara-oy-u yas  
 leave-NEG-IMP 1SG.P  
 ‘Don’t leave me!’

- (196) a. \**Ya sa-sabu nanga!*  
           ya sa~saha-u nanga-Ø  
           LOCT ITER~go-IMP house-TOP  
           ‘Go back to the house!’
- b. \**Sa sutamuya kohanya tasela!*  
      sa sutam-u=ya.Ø kohan-ya tasela  
      PT hang-IMP=3SG.M.TOP sunrise-LOC tomorrow  
      ‘May he be hanged tomorrow at sunrise!’

Example (195) simply expresses a negative command, which is unproblematic in terms of logic, since commands may be issued to act in a certain way, or to forgo this action. Example (196a) shows the imperative verb as preceded by a locative topic marker, which is not logically impossible, but unacceptable by convention.<sup>39</sup> Example (196b) takes this one step further in displaying a cliticized object pronoun in the fashion of morphological passives (compare section 4.5.1, page 186).

#### *Hortative*

The hortative is a special kind of imperative which addresses a group including the speaker. Its implied referent is thus first-person plural. Again, it is not necessary to mark the verb for the addressee here. As the hortative is related in meaning to the imperative, the verb also uses the imperative inflection with *-u*, but it is fully reduplicated in addition to mark the difference. As regards agreement morphology, the same restrictions as those of imperatives apply.

- (197) a. *Sabu!*  
           saha-u  
           go-IMP  
           ‘Go!’
- b. *Sabu-sabu umangya!*  
      sahu~saha-u umang-ya  
      HORT~go-IMP beach-LOC  
      ‘Let’s go to the beach!’

#### 4.5.5 Modals

Modals in Ayeri express the notions of ability, desire, permission, requirement, obligation, and also of continuation, as indicated by Table 4.23. They can gener-

<sup>39</sup> The translation of ‘Ozymandias’ in section B.3 deviates from this rule in the line *sa silvu gumo nā* ‘my works, regard them’. This is poetic license, however.

Table 4.23: Modal verbs and particles

Category	Verb	Particle	Translation
ABILITY	မိာ်: <i>ming-</i>	မိာ် <i>ming</i>	‘be able to, can’
DESIRE, INTENTION	ၵံး: <i>vac-</i>	ၵံး <i>vaca</i>	‘like to’
	ၵံး: <i>no-</i>	ၵံး <i>no</i>	‘want to’
PERMISSION	ခိာ်: <i>kila-</i>	ခိာ် <i>kila</i>	‘be allowed to, may’
REQUIREMENT	ခိာ်: <i>ilta-</i>	ခိာ် <i>ilta</i>	‘need to’
OBLIGATION	မုး: <i>mya-</i>	မုး <i>mya</i>	‘be supposed to, shall’
	မုး: <i>rua-</i>	မုး <i>rua</i>	‘have to, must’
CONTINUATION	မုး: <i>div-</i>	မုး <i>diva</i>	‘stay, remain’

ally act as both fully inflectable intransitive verbs, as well as clitics which occur in combination with fully inflected content verbs:

- (198) a. *Rua babavāng babo, ang bihanoyya mirampaluy nas.*  
 rua baha=vāng baho ang bihan-oy=ya.Ø mirampaluy nas  
 must shout=2SG.A loudly AT understand-NEG=3SG.M.TOP otherwise IPL.P  
 ‘You have to shout loudly, otherwise he does not understand us.’
- b. *Ruasanang.*  
 rua-asa=nang  
 must-HAB=IPL.A  
 ‘We usually have to.’

As (198a) shows, the modal does not inflect in combination with another verb; as a clitic it rather acts similarly to a prefix, like the progressive marker မိာ် *manga*, which is also presumably deverbal (compare section 3.1, footnote 3). In difference to မိာ် *manga*, which as a preverbal element only serves a grammatical function, the semantic component of the modals is still prevalent, as is shown by (198b), where မုး *rua-* appears in its function as an intransitive verb with the same meaning of strong obligation as in (198a), though it carries regular person and aspect inflection here. Inflecting the modal in the context of cooccurrence with a content verb is, however, considered unacceptable:

- (199) \**Ruavāng babayam babo.*  
 rua=vāng baha-yam baho  
 must=2SG.AT shout-PTCP loudly  
 ‘You have to shout loudly.’

Regarding example (198b) and its ability to inflect, Ayeri also has a verb that generally means ‘do’, namely,  $\text{mir}$ : *mira-*. However, it is not common to use this as a dummy verb to carry the inflection instead of the modal verb either. While such a construction is not ungrammatical *per se*, it is simply not the preferred way to express intransitive modal verbs:

- (200)  $\text{?Rua mirasanang}$ .  
            $\text{rua mira-asa=nang}$   
           must do-HAB=IPL.A  
           ‘We usually have to.’

While most of the verbs listed in Table 4.23 should look reasonable to English speakers, Ayeri uses two verbs for modal particles which may seem odd:  $\text{vac}$  *vaca* ‘like to’, to express taking pleasure in doing something, and  $\text{div}$  *diva* ‘stay, remain’, to express that the action is being prolonged.<sup>40</sup> The latter verb thus also has an aspectual component to its meaning.

- (201) a.  $\text{Ang vacay betayley}$ .  
            $\text{ang vac=ay.}\emptyset$  betay-ley  
           AT like=ISG.TOP berry-P.INAN  
           ‘I like berries.’  
       b.  $\text{Ang vaca konday betayley}$ .  
            $\text{ang vaca kond=ay.}\emptyset$  betay-ley  
           AT like eat=ISG.TOP berry-P.INAN  
           ‘I like to eat berries.’
- (202) a.  $\text{Ang divay rangya nā tasela}$ .  
            $\text{ang div=ay.}\emptyset$  rang-ya nā tasela  
           AT stay=ISG.TOP home-LOC ISG.GEN  
           ‘I will stay home tomorrow.’  
       b.  $\text{Ang diva bengya ku-danyās kebay}$ .  
            $\text{ang diva beng=ya.}\emptyset$  ku=danya-as kebay  
           AT stay stand=3SG.M.TOP like=one-P alone  
           ‘He remained standing as the only one.’

The fact that modal particles in Ayeri retain their verbal semantics in spite of shedding verb morphology is probably even more obvious from the above examples (201) and (202), which show the alternation between full-verb use (a) and modal use

<sup>40</sup> The verb stems indeed end in a consonant while the modal particles need an epenthetic *-a* to form permissible words.



(b) for both ᳚᳚᳚: *vac-* and ᳚᳚᳚: *div-*. In comparison to the other modals in Table 4.23, these two verbs in particular also stand out by virtue of their roots ending in a consonant instead of a vowel like in the other cases. This suggests that they may have been grammaticalized as modals only relatively recently, and there appears to be variation at least for ᳚᳚᳚: *vac-*, for instance:

- (203) ... *yam vacongyang ilisayam eda-koyās gan* ...  
 ... *yam vac-ong-yang ilisa-yam eda=koya-as gan-Ø* ...  
 DATT like-IRR-1SG.A dedicate-PTCP this=book-P child-TOP

‘... I would like to dedicate this book to the child ...’ (Becker 2015 [2013]: 1, 8)

Moreover, as illustrated previously in (191b), ᳚᳚᳚ *mya* ‘be supposed to, shall’ can be used to express indirect commands where English may use the subjunctive mood; essentially the function of this modal is that of the jussive mood. For convenience, the above example will be repeated here:

- (204) *Arapnang, sa (mya) garayāng hatay.*  
*arap=nang sa (mya) gara=yāng hatay-Ø*  
 require=1PL.A PT (be.supposed.to) call=3SG.M.A police-TOP

‘We require that he call the police.’

In addition to this use, ᳚᳚᳚ *mya* is also used in commands to third persons, whether direct or indirect. English may use *shall* here as an equivalent.

- (205) a. *Ningu cam, mya saratang.*  
*ning-u cam mya sara=tang*  
 tell-IMP 3PL.M.DAT shall leave=3PL.M.A  
 ‘Tell them to leave.’  
 b. *Mya vehara nekanley.*  
*mya veh-ara nekan-ley*  
 shall build-3SG.INAN bridge-P.INAN  
 ‘A bridge shall be built.’  
 c. *Mya yomāra makangreng.*  
*mya yoma-ara makang-reng*  
 shall exist-3SG.INAN light-A.INAN  
 ‘Let there be light.’

#### 4.5.6 Participle

Besides the imperative—and, by extension, the hortative—Ayeri also possesses another infinite form called the participle. This form is marked by appending ᳚᳚᳚

-yam to the verb root. The participle is generally the form of verbal complements of intransitive subordinating verbs other than modal verbs (compare section 4.5.5). For instance,  $\text{𑜋𑜰𑜫}$  *cun-* ‘begin’ or  $\text{𑜇𑜡𑜫}$  *manang-* ‘avoid’ both allow complementation with another verb:

- (206) a. *Cunyo makayam perinang.*  
 cun-yo maka-yam perin-ang  
 begin-3SG.N shine-PTCP sun-A  
 ‘The sun began to shine.’
- b. *Manangyeng pengalyam badanas saba yena.*  
 manang=yeng pengal-yam badan-as saha yena  
 avoid=3SG.F.A meet-PTCP father-P in.law 3SG.F.GEN  
 ‘She avoids to meet her father-in-law.’

Since subordinated verbs may be transitive like in (206b), the problem of center-embedding arises when the agent NP of the subordinating verb is not simply a cliticized pronoun (see section 3.2.5, p. 89; 4.5.1), since arguments of the subordinating verb follow the embedded clause as in (206a):

- (207) a.  $\text{?Ang pinyaya [konjam inunas] Yan sa Pila.}$   
 ang pinya-ya kond-yam inun-as Ø Yan sa Pila  
 AT ask-3SG.M eat-PTCP fish-P TOP Yan P Pila  
 ‘Yan asks Pila to eat the fish.’
- b.  $\text{??Ang pinyaya [ilyam koyaley ledanyam yana] Yan sa Pila.}$   
 ang pinya-ya il-yam koya-ley ledan-yam yana Ø Yan sa Pila  
 AT ask-3SG.M give-PTCP book-P.INAN friend-DAT 3SG.M.GEN TOP Yan P Pila  
 ‘Yan asks Pila to give the book to his friend.’

In order to avoid too much complexity at the expense of ease of composition on the speaker’s side, and intelligibility on the listener’s, it is much preferable to express the embedded clause as a complement clause instead.<sup>41</sup> The particle  $\text{𑜇𑜡}$  *da-* may be

<sup>41</sup> The German linguist Otto Behaghel (1854–1936) coined a number of laws—albeit with German in focus—three of which are relevant to information flow: “Das oberste Gesetz ist dieses, daß das geistig eng Zusammengehörige auch eng zusammengestellt wird.” (Behaghel 1932: 4) [‘The supreme law is such that the mentally closely related is also arranged in close proximity.’]—“Ein zweites machtvolleres Gesetz verlangt, daß das Wichtigere später steht als das Unwichtige, dasjenige, was zuletzt noch im Ohr klingen soll.” (4) [‘A second powerful law demands that more important information appear at a later point than what is less important: the which is supposed lastly to resonate in the listener’s ear.’]—“Gesetz der wachsenden Glieder [...] es besagt, daß von zwei Gliedern, soweit möglich, das kürzere vorausgeht, das längere nachsteht.” (6) [‘Law of the growing constituents [...] it signifies that of two con-

added to the formerly subordinating verb in order to signal that a complement clause is following.

- (208) a. *Ang da-pinyaya Yan sa Pila, [le konjeng inun].*  
 ang da=pinya-ya Ø Yan sa Pila le kond=yeng inun-Ø  
 AT such=ask-3SG.M TOP Yan P Pila PT.INAN eat=3SG.F.A fish-TOP  
 ‘Yan asks Pila to eat the fish.’
- b. *Ang da-pinyaya Yan sa Pila, [le ilyeng koya ledanyam yana].*  
 ang da=pinya-ya Ø Yan sa Pila le il=yeng koya-Ø ledan-yam  
 AT such=ask-3SG.M TOP Yan P Pila PT.INAN give-3SG.F book-TOP friend-DAT  
 yana  
 3SG.M.GEN  
 ‘Yan asks Pila to give the book to his friend.’

#### 4.5.7 Other affixes

In the section on noun morphology we have already encountered a number of clitic prefixes that may attach to noun heads (see section 3.2.5, section 4.1.4), and some of these can also attach to verbs. Furthermore, verbs may also be modified by certain adverbial quantifier clitics. The latter are dealt with in more detail in the section on adverbs; only a few relevant examples will be given here.

##### Prefixes

We have already encountered the prefix 𐀄: *da-* ‘so, such’ in the previous section, as well as in the section on noun prefixes (see section 3.2.5, p. 81; 4.1.4; 4.5.6). With nouns, 𐀄: *da-* ‘such’ patterns as a demonstrative with the deictic prefixes 𐀄: *eda-* ‘this’ and 𐀄: *ada-* ‘that’. Distinguishing between near and far is not possible with verbs,<sup>42</sup> but pointing out that something is happening ‘in this way’, ‘so’ is still possible, hence 𐀄: *da-* is also applicable to verbs. 𐀄: *da-* can thus act essentially as a pro-verb. As a clitic, it leans on the verb, preceding all other inflectional prefixes, that is, any tense prefixes that may possibly precede the verb root.

stituents, if possible, the shorter one precedes, the longer one follows.’] Also compare Wasow (1997) on the cooperation between speaker and listener in the face of syntactically complex, ‘heavy’ constituents.

<sup>42</sup> Unless you distinguish between actions performed in the speaker’s proximity versus ones that are performed at a distance. Ayeri, however, does not make such a distinction.

- (209) a. *Da-mingya ang Diyan.*  
 da=ming-ya ang Diyan.  
 so=can-3SG.M A Diyan  
 ‘Diyan can (do it).’
- b. *Ang da-məpinyaya Yan sa Pila.*  
 ang da=mə-pinya-ya Ø Yan sa Pila  
 AT such=PST-ask-3SG.M TOP Yan P Pila  
 ‘Yan asked Pila to (do so).’

Another possible use of the prefix 𐄢: *da-* with verbs is related to the colloquial abbreviation of 𐄢𐄣 *danya* ‘such one’ as described in sections section 3.2.5 (p. 78) and section 4.2.2, where the demonstrative part, 𐄢: *da-* may be split off the pronoun and attached in front of the adjective directly to express ‘the ADJ one’. This practice has possibly been extended to verbs in analogy to the use just illustrated in (209). Example (102) from the mentioned section is repeated here for the reader’s convenience:

- (210) *Sa noyang da-tuvo.*  
 sa no=yang da=tuvo.Ø  
 PT want=1SG.A such=red.TOP  
 ‘I want the red one.’

When 𐄢: *da-* is used as an abbreviation for 𐄢𐄣𐄢 *danyās* (such.one-P) or 𐄢𐄣𐄢𐄣 *danyaley* (such.one-P.INAN), as in the following example, it may also appear prefixed to the verb:

- (211) *Mya da-vehoyyāng.*  
 mya da=veh-oy=yāng  
 supposed.to one=build-NEG=3SG.M  
 ‘He is not supposed to build one.’

As mentioned above, 𐄢: *da-* can also be used in an expletive way, to express ‘in this way’ or ‘like that’. It does not encode an anaphoric relation in this case, but merely serves as a discourse particle to highlight the action.

- (212) a. *Da-sabāra seyaneng.*  
 da=saha-ara seyan-eng  
 thus=come-3SG.INAN rain-A.INAN  
 ‘Here comes the rain.’
- b. *Le no da-subroyya ang Hasanjan tiga kaytan yana.*  
 le no da=subr-oy-ya ang Hasanjan tiga kaytan-Ø yana  
 PT want there=give.up-NEG-3SG.M A Hasanjan honorable right-TOP 3SG.M.GEN  
 ‘Mr. Hasanjan did not want to cease his right just there.’

Besides  $\text{da-}$ , verbs may also take the  $\text{ku-}$  ‘like’ prefix, which we have already seen with both nouns and adjectives (compare section 3.2.5, 4.1.4, 4.3.4). The English translation in context may rather be ‘as though’ than ‘like’ here, but the function is the same: expressing likeness and resemblance.

- (213) *Misyeng, ang ku-tangoyye yās.*  
 mis=yeng ang ku=tang-oy=ye.Ø yās  
 act=3SG.F.A AT like=hear-NEG=3SG.F.TOP 3SG.M.P

‘She acts as though she does not hear him.’

As previously described (compare section 3.2.5, p. 80, and 4.2.6),  $\text{sitang}$  ‘self’, the reflexive clitic, can appear as a prefix on verbs as well. This may be the case when the patient/undergoer of a transitive sentence signifies the same entity as the actor. Example (118) is repeated here for convenience:

- (214) *Ang sitang-silvye puluyya.*  
 ang sitang=silv=ye.Ø puluy-ya  
 AT self=see=3SG.F.TOP mirror-LOC

‘She sees herself in the mirror.’

The image of the agent in the mirror is that of the agent herself, so she is seeing her own reflection. Both agent and patient thus reference the same person, which means that instead of using the reflexive object pronoun  $\text{sitang-yes}$  ‘herself’ (self-3SG.F.P), it is possible to drop the pronoun and to place the reflexive prefix on the verb instead.

### Suffixes

Besides taking clitic prefixes, verbs may also take clitic suffixes, namely, adverbial suffixes denoting degree, such as  $\text{-ani}$  ‘not at all’,  $\text{-eng}$  ‘rather’,  $\text{-ikan}$  ‘much’,  $\text{-ikoy}$  ‘not much’,  $\text{-kay}$  ‘a little’,  $\text{-nama}$  ‘just, only, merely’,  $\text{-ngas}$  ‘almost’, and  $\text{-nyama}$  ‘even’ (see section 3.2.5, p. 94). Some of these overlap with quantifiers applicable to nouns, and all of them are also applicable to adjectives. As enclitics, these suffixes lean on the inflected verb:

- (215) a. *Ang rua apaya-kay Latun adanyaya.*  
 ang rua apa-ya=kay Ø Latun adanya-ya  
 AT must laugh-3SG.M=a.little TOP Latun that.one-LOC

‘Latun had to laugh a little at that.’

- b. *Ya no narayang-nama va.*  
 ya no nara=yang=nama va.Ø  
 LOCT want speak=1SG.A=just 2SG.TOP

‘It is you I just want to talk to.’

## 4.6 Adverbs

Adverbs in Ayeri are the counterparts of adjectives with regards to modification of verbs and phrases. Like adjectives, they do not display agreement, though attributive adverbs may take suffixes for comparison ('run *faster*', 'climb *better*'). Adverbs may equally be modified by the usual degree suffixes. Generally, there is no rigid distinction between adverbs and adjectives, so the latter may easily be used as the former. The following subsections will discuss the different kinds of adverbs and their possible uses as modifiers.

### 4.6.1 Attributive adverbs

Attributive adverbs are words expressing the manner in which an action is carried out, or the circumstances of an event. Like adjectives, adverbs follow their heads, that is, verbs. If near-grammaticalized adverbs are involved, namely, adverbs whose function predominates over their semantic content, attributive adverbs follow these. This case is illustrated in (216a), where the attributive adjective 𑀧𑀺 *ban* 'good' follows the more functional adverb 𑀧𑀺 *iri* 'already'. In (216b), on the other hand, the descriptive adjective 𑀧𑀺 *cabo* 'late' can directly follow the verb. Further adverbs may follow in decreasing order of semantic relation to their head. With regards to grammaticalization, Lehmann (2015: 157 ff.) speaks of *bondedness* or *fugungsenge* ('closeness of construction'): the closer the bond between two juxtaposed terms is, the higher is its degree of grammaticalization. This explains why 𑀧𑀺 *iri* must follow the verb in (216a) while descriptive adverbs less central to the verb's meaning typically follow with increasing optionality.

- (216) a. *Ri rija iri ban ang Tapan palān yena.*  
           ri rig-ya iri ban ang Tapan palān-Ø yena  
           INST draw-3SG.M already well A Tapan age-TOP 3SG.F.GEN  
           'For her age, Tapan already draws well.'
- b. *Sabasaya cabo ang Niyas.*  
           saha-asa-ya cabo ang Niyas  
           come-HAB-3SG.M late A Niyas  
           'Niyas is usually late.'

Adverbs do not show agreement, however, attributive adverbs can be negated. This makes them very similar to adjectives, except that they do not modify nouns. The negative suffix for attributive adverbs is 𑀧𑀺 *-oy*, which is demonstrated in (217).

- (217) *Ersasayan napayoy ang Temisi.*  
 ers-asa-yan napay-oy Ø Temisi  
 cook-HAB-3PL.M spicy-NEG A Northerner  
 ‘The Northerners cook in an uns spicy way.’

The adjective  $\text{ᑭᑭ}$  *napay* ‘spicy’ has been seamlessly converted into an adjective here and negated to  $\text{ᑭᑭᑭ}$  *napayoy* ‘uns spicy(ly)’. The semantic difference from the same sentence with the verb negated instead of the adverb is marginal and up to the choice of the speaker:

- (218) *Ersasoyyan napay ang Temisi.*  
 ers-asa-oy-yan napay Ø Temisi  
 cook-HAB-NEG-3PL.M spicy A Northerner  
 ‘The Northerners don’t cook in a spicy way.’

#### Comparison of adverbs

Since actions are usually gradable in the way they are carried out, it is possible to compare adverbs in the same way as adjectives. Here, however, only the particle-based strategy described in section 4.3.1 can be used.

- (219) a. *Ang rije ban-eng Sipra na Tapan.*  
 ang rig-ye ban=eng Ø Sipra na Tapan  
 AT draw-3SG.F good=COMP TOP Sipra GEN Tapan  
 ‘Sipra draws better than Tapan.’  
 b. *Rije ban-vā ang Nava.*  
 rig-ye ban=vā ang Nava  
 draw-3SG.F good=SUPL A Nava  
 ‘Nava draws best.’

In order to form the comparative (219a), the suffix  $\text{ᑭᑭ}$  *-eng* is appended to the adverb; the superlative (219b) carries the suffix  $\text{ᑭᑭᑭ}$  *-vā* as a marker.

#### *Māy and voy*

The discourse particles  $\text{ᑭᑭ}$  *māy* ‘yes’ and  $\text{ᑭᑭᑭ}$  *voy* ‘no’ can also appear as adverbs, though since they act mainly as functional morphemes here, it is not possible for them to undergo comparison in spite of their attributive use. While  $\text{ᑭᑭ}$  *māy* ‘yes’ and  $\text{ᑭᑭᑭ}$  *voy* ‘no’ normally express affirmative and negative responses as answers to closed questions,  $\text{ᑭᑭ}$  *māy*, for one, can be used adverbially as an intensifier (220). In a similar way,  $\text{ᑭᑭᑭ}$  *voy* can be used for negative intensification (221). The negative

intensifier replaces negation on the verb in this case, though the verb may still be negated as well for very forceful negation.

- (220) a. *Nay le konja māy epang ang Kaji nernan barina sebu!*  
 nay le kond-ya māy epang ang Kaji nernan-Ø bari-na sebu  
 and PT.INAN eat-3SG.M INT then A Kaji piece-TOP meat-GEN rotten

‘And then Kaji totally ate the piece of rotten meat!’

- b. *Yāng māy karomayās nārya.*  
 yāng māy karomaya-as nārya  
 3SG.M.A INT doctor-P though

‘He *is* a doctor, though.’

- (221) a. *Le vacyo voy veneyang kondan.*  
 le vac-yo voy veney-ang kondan-Ø  
 PT.INAN like-3SG.N INT.NEG dog-A food-TOP

‘The food, the dog did not like it at all.’

- b. *Adareng voy babisley niru.*  
 ada-reng voy bahis-ley niru  
 that-A.INAN INT.NEG day-P.INAN bad

‘That is not a bad day at all.’

Besides this use, both ʔᵐ māy and ʔᵐ voy can also be used in tag questions, to reflect the expectation of a person asking with regards to the answer:

- (222) a. *Sa konjon māy patasjang keynam?*  
 sa kond-yon māy patas-ye-ang keynam-Ø  
 PT eat-3PL.N AFF bear-PL-A people-TOP

‘People, bears eat them, don’t they?’

- b. *Sa ginyon voy patasjang nimpur?*  
 sa gin-yon voy patas-ye-ang nimpur-Ø  
 PT drink-3PL.N NEG bear-PL-A wine-TOP

‘Wine, bears don’t drink it, do they?’

Example (222a) poses the question with the expectation of an affirmative answer. This is indicated by using the affirmative particle ʔᵐ māy after the verb. Example (222b), on the other hand, indicates that the asker has doubts about the issue in question and expects their opposite to decline. The negative particle ʔᵐ voy is placed in adverb position after the verb accordingly.



## 4.6.2 Degree and quantity

While attributive adverbs follow their heads as independent words, the most common adverbs expressing degree and quantity (quantifiers) do not only follow verbs, nouns, adpositions, adjectives, or other adverbs, but they cliticize on them, that is, they are dependent morphemes (compare section 3.2.5, p. 94). The word stem—a lexical head which is usually inflected except in the case of adjectives—serves as the host for the clitic in all these cases. Examples of degree and quantifier suffixes and how they interact with different parts of speech were already given in all the relevant sections; an example from each section is repeated here for convenience. As we will see below, there are common grading and quantifying adverbs which behave like regular adverbs as well. It is possible to combine both the suffixed and the free kinds with other adverbs as long as those adverbs permit modification with regards to degree and/or quantity. Purely functional adverbs like မရှိ *iroy* ‘not yet’ or မိမိ *sirimang* ‘about to’ may allow degree adverbs to modify them at least.

- (223) a. With a noun (7b):

*Ajayon ganang-ikan kivo.*  
 aja-yon gan-ang=ikan kivo.  
 play-3SG.N child-A=many small

‘Many small children are playing.’

- b. With an adjective (133):

*Eda-prikanreng napay-eng*  
 eda=prikan-reng napay eng  
 this=soup-A.INAN spicy rather

‘This soup is rather spicy.’

- c. With an adposition (139):

*Ang mitasaye pang-ikan mandayya tado.*  
 ang mit-asa=ye.Ø pang=ikan manday-ya tado  
 AT live-HAB=3SG.F.TOP back=much forum-LOC old

‘She used to live way behind the old forum.’

- d. With a verb (215a):

*Ang rua apaya-kay Latun adanyaya.*  
 ang rua apa-ya=kay Ø Latun adanya-ya  
 AT must laugh=3SG.M=a.little TOP Latun that.one-LOC

‘Latun had to laugh a little at that.’

A number of quantifier adverbs can be used to express both quantity and degree, especially prominent in this regard is မြန် *-ikan*, which comprises all of ‘many’, ‘much’ and ‘very’, as displayed in examples (223a) and (223c), where in the former

Table 4.24: Adverbial degree and quantifier suffixes

Suffix	Degree	Quantity
ᐱᐱᐱ - <i>ani</i>	not at all	none at all
ᐱᐱᐱᐱ - <i>aril</i>		some
ᐱᐱᐱ - <i>eng</i>	rather, more	more
ᐱᐱᐱ - <i>hen</i>	completely	all, every, each
ᐱᐱᐱᐱ - <i>ikan</i>	much, very	many, much
ᐱᐱᐱᐱ - <i>ikoy</i>	not very, less	not many, not much
ᐱᐱᐱ - <i>ing</i>	so	
ᐱᐱᐱ - <i>kay</i>	a bit, little	few
ᐱᐱ - <i>ma</i>	enough	enough
ᐱᐱᐱ - <i>mas</i>	some kind of	
ᐱᐱᐱ - <i>nama</i>	just, merely	just, only
ᐱᐱᐱ - <i>ngas</i>	almost	
ᐱᐱᐱᐱ - <i>nyama</i>	even	
ᐱᐱᐱ - <i>vā</i>	most	most
ᐱᐱᐱ - <i>ven</i>	pretty, quite	

case it appears as a quantifier of a countable entity (ᐱᐱᐱᐱᐱᐱ *ganang-ikan* ‘many children’) and in the latter case as a degree adverb (ᐱᐱᐱᐱᐱᐱ *pang-ikan* ‘way behind’). The complete set of degree and quantifier suffixes is listed in Table 4.24. Note that not all suffixes have both a grading and a quantifying meaning.

Grading and quantifying expressions which deviate from the pattern of cliticization and instead are used as independent words are, most notably: ᐱᐱᐱᐱ *ankyū* ‘really’, ᐱᐱᐱᐱ *diring* ‘several’, ᐱᐱᐱᐱ *ekeng* ‘over-, overly, too’, ᐱᐱᐱᐱ *bengas* ‘almost all’, ᐱᐱᐱᐱᐱᐱᐱ *ikan-ikan* ‘altogether, totally’, ᐱᐱᐱᐱᐱᐱᐱ *ikanvānya* ‘at most, by and large’, ᐱᐱᐱᐱ *kagan* ‘excessively, far too’, ᐱᐱᐱᐱᐱᐱ *menikaneng* ‘another (one more)’, ᐱᐱᐱᐱ *miday* ‘approximately’, ᐱᐱᐱᐱ *palung* ‘another (a different kind)’, ᐱᐱᐱᐱᐱᐱ *regandey* ‘bit by bit, gradually’, ᐱᐱᐱᐱ *sano* ‘both’, ᐱᐱᐱᐱᐱᐱ *varyānya* ‘at least’. Besides, adjectives denoting a degree, like ᐱᐱᐱ *ipan* ‘drastic, extreme, radical’ can of course also be used as adverbial modifiers. Since adverbs do not inflect, the conversion happens without the need to explicitly mark it. ᐱᐱᐱ *ipan* can thus also be used to mean ‘extremely’:

- (224) *Yang valuy ipan, sa silvyang va.*  
 yang valuy ipan sa silv=yang va.Ø  
 1SG.A glad extremely PT see=1SG.A 2SG.TOP

‘I’m extremely glad to see you.’

## 4.6.3 Sentence adverbs

Ayeri allows adverbs to modify sentences, for instance, to express the stance of the speaker, to concede an argument, or simply to structure an argumentative chain.

*Stance adverbs*

Adverbs indicating the stance of the speaker towards an assertion or a statement are, for instance: *ᐱᑦᑦᑦ ankyu* ‘really’, *ᑦᑦᑦᑦ cuyam* ‘actually, indeed, in fact’, *ᐱᑦᑦᑦ kalam* ‘honestly’, *ᐱᑦᑦᑦ kuban* ‘fortunately’, *ᐱᑦᑦᑦ kuniru* ‘unfortunately’, *ᑦᑦᑦᑦ nilay* ‘probably’, *ᐱᑦᑦᑦ yoming* ‘maybe, perhaps’. These adverbs are usually placed after the verb like any other attributive adverb, even though their scope is over the whole clause. It is also possible to place them towards the end of the clause they are used in, however. Example (225) gives an example of either position.

- (225) a. *Ang ming bengya kuban Tipal vabamya bavesangena nā.*  
 ang ming beng-ya kuban Ø Tipal vaham-ya bavesang-ena nā  
 AT can attend-3SG.M fortunately TOP Tipal party-LOC birthday-GEN 1SG.GEN  
 ‘Fortunately, Tipal can attend my birthday party.’
- b. *Sahayāng cabo-kay nilay nārya.*  
 saha=yāng cabo=kay nilay nārya  
 come=3SG.M late=a.little probably though  
 ‘He will probably come a little late, though.’

*Discourse-structuring adverbs*

Ayeri does not have a great number of concessive adverbs, that is, *ᐱᑦᑦᑦ arēn* ‘however, anyway’ and *ᑦᑦᑦᑦ nārya* ‘although, though; nevertheless’ do most, if not all the work. Like adverbs expressing stance, they may follow the verb or be placed at the end of the clause. Example (225b) above already shows an example of *ᑦᑦᑦᑦ nārya* being used as a sentence adverb. With regards to this word, it is important to note that *ᑦᑦᑦᑦ nārya* may also be used as a general contrastive conjunction which can mostly be translated as ‘but’. In this sense, its placement in a clause creates a slight difference in meaning, as illustrated by example (226) below.<sup>43</sup>

<sup>43</sup> Possibly it is easier not to distinguish between conjunction and adverb at all in this case and instead to treat *ᑦᑦᑦᑦ nārya* as an adverb with a general contrastive meaning which can exceptionally be found at the beginning of a clause as well.

- (226) a. *Garayang, nārya guraca ranyāng.*  
 gara=yang nārya gurat-ya ranya-ang  
 call=1SG.A but answer-3SG.M nobody-A  
 ‘I called, but nobody answered.’
- b. *Garayang, guraca nārya ranyāng.*  
 gara=yang gurat-ya nārya ranya-ang  
 call=1SG.A answer-3SG.M although nobody-A  
 ‘I called, although nobody answered.’

Besides the two adverbs mentioned above, there is also *da-nārya* ‘even though, in spite of, despite’ as a postposition with a contrastive meaning (see section 4.4.2). As an adposition it accepts either a noun phrase or a complement phrase (CP) as a complement. In the latter case, which is shown in (227b), there is no locative case agreement of the whole CP with the postposition, since there is no fitting agreement target to attach it to.

- (227) a. *Ya precang nanga yena* [PP [NP *sarānya yena*] *da-nārya*].  
 ya pret=yang nanga-Ø yena sarān-ya yena da-nārya  
 LOCT knock=1SG.A house-TOP 3SG.F.GEN absence-LOC 3SG.F.GEN in.spite  
 ‘I knocked at her house in spite of her absence.’
- b. *Precang* [PP [CP *ang yomoyye rangya yena*] *da-nārya*].  
 pret=yang ang yoma-oy=ye.Ø rang-ya yena da-nārya  
 knock=1SG.A AT exist-NEG=3SG.F.TOP home-LOC 3SG.F.GEN even.though  
 ‘I knocked, even though she wasn’t at home.’

Further adverbs which are commonly used as adverbial expressions and which may appear in the presentation of arguments include: *deramyam* ‘after all’, *kaybunay* ‘by the way’, *ku-nasya* ‘as follows’, *menanya* ‘on the one hand’, *mirampaluy* ‘otherwise’, *nāreng* ‘rather’, *naynay* ‘(and) also, moreover, furthermore’, *palunganya* ‘on the other hand’, *panca* ‘finally, eventually, in the end’, *pinyan* ‘please’, *subing* ‘naturally, of course’. It should be apparent by the complexity and relative length of some of these words that they are fossilized expressions, for instance, *deramyam* ‘after all’ transparently derives from *deram* ‘matter of fact’ declined for dative case (*yam*, see section 4.1.3); *ku-nasya* is derived from a phrase literally meaning ‘as (it) follows’; and *palunganya* ‘on the other hand’ literally means ‘in difference’, from *palungan* ‘difference, distinction’. Of the list given above, it may be noted that *pinyan* ‘please’ (from *pinya-* ‘ask’) is often found at the beginning of polite requests:

- (228) *Pinyan, sahu kongya!*  
 pinyan saha-u kong-ya  
 please come-IMP inside-LOC  
 ‘Please come inside!’

### Conjunctive adverbs

The term ‘conjunctive adverb’ here refers to sentence adverbs which have the distribution of a conjunction. Whereas sentence adverbs are normally placed either after the verb or at the end of a clause, these words are usually found as introducing clauses since they connect two otherwise independent statements to show their relation to each other. Their meaning extends that of the ‘pure’, logical conjunctions *ꠞꠞ* *nay* ‘and’ and *ꠞꠞꠞ* *soyang* ‘or’, however.<sup>44</sup> Part of this small class of words are the expressions *ꠞꠞꠞ* *bata* ‘if, whether’,<sup>45</sup> *ꠞꠞꠞ* *kada* ‘then, thus’, *ꠞꠞꠞꠞꠞ* *kada-kada* ‘so that ... again’, *ꠞꠞꠞꠞ* *kadāre* ‘so that’, *ꠞꠞꠞꠞ* *naynay* ‘moreover, furthermore, and also’, *ꠞꠞꠞꠞ* *nāreng* ‘(but) rather’, *ꠞꠞꠞꠞ* *nāroy* ‘but not’, *ꠞꠞꠞꠞ* *nārya* ‘but, except that, though, yet’, *ꠞꠞꠞꠞ* *sining* ‘that is’, and *ꠞꠞꠞꠞ* *yanoyam* ‘because, for, since’.

- (229) a. *Le rimasayang kunang sirutayya, kadāre ming toryang*  
 le rima-asa=yang kunang-Ø sirutay-ya kadāre ming tor=yang  
 PT.INAN shut-HAB=ISG.A door-TOP night-LOC so.that can sleep=ISG.A  
*ban-eng.*  
 ban=eng  
 good=COMP  
 ‘I usually close the door at night so that I can sleep better.’
- b. *Ilta toryeng, nārya da-kilisoyyon nilanjang yena.*  
 ilta tor=yeng nārya da=kilis-oy-yon nilan-ye-ang yena  
 need sleep=3SG.F.A but so=allow-NEG-3PL.N thought-PL-A 3SG.F.GEN  
 ‘She needed to sleep, but her thoughts did not allow her to.’
- c. *Ang ming bangoyya Yan padangas, yanoyam yāng pisu.*  
 ang ming hang-oy-ya Ø Yan padang-as yanoyam yāng pisu  
 AT can keep-NEG-3SG.M TOP Yan mind-P because 3SG.M.A tired  
 ‘Yan cannot concentrate because he is tired.’

Regarding (229b), it needs to be pointed out that *ꠞꠞꠞꠞ* *nārya* can also be used as a regular adverb. In those cases it is considered to have less contrastive force,

<sup>44</sup> Logical ‘not’ is usually expressed by a negative suffix on the adjective or the verb, compare sections 4.3.2 and 4.5.4, respectively. For conjunctions proper, see section 4.8.

<sup>45</sup> Conditional protasis and apodosis are usually unmarked in Ayeri, however, it may still be desirable occasionally to use a particle to indicate them explicitly.

however: postposed  $\text{၁၃၃၈}$  *nārya* is best translated as ‘though, although’ (compare section 4.6.3).

Since verbs can be negated and reduplicated for grammatical purposes, the adverbs  $\text{၁၃၃၈}$  *kada-kada* ‘so that ... again’ and  $\text{၁၃၃၈}$  *nāroy* ‘but not’ are mostly used with predicative adjectives, since equative statements lack a verb to apply verb morphology to. These two conjunctive adverbs thus can convey the most important distinctions otherwise expressed by the verb as a substitute. This ability, however, is not a productive grammatical process, but specific to  $\text{၁၃၃၈}$  *kada-kada* and  $\text{၁၃၃၈}$  *nāroy*, respectively.

- (230) a. *Rua nibaya ang Pulan, kada-kada yāng sapin tadayya kivo.*  
            $\text{rua}$   $\text{niba-ya}$   $\text{ang}$   $\text{Pulan}$   $\text{kada}\sim\text{kada}$   $\text{yāng}$   $\text{sapin}$   $\text{taday-ya}$   $\text{kivo}$   
           must rest-3SG.M A Pulan ITER~so.that 3SG.M.A healthy time-LOC little  
           ‘Pulan must rest so that he will be healthy again very soon.’
- b. *Yang temisena cuyam, nāroy yang petau.*  
            $\text{yang}$   $\text{temis-ena}$   $\text{cuyam}$   $\text{nāroy}$   $\text{yang}$   $\text{petau}$   
           1SG.A north-GEN indeed but.not 1SG.A stupid  
           ‘I may be from the north, but I am not stupid.’

As described above (compare section 4.5.3), partial reduplication of the verb expresses iterative aspect, which in Ayeri is used to mean ‘VERB again’ and ‘VERB back’, depending on context. The reduplicated form  $\text{၁၃၃၈}$  *kada-kada* as displayed in (230a) is irregular in its formation if we assume that it is formed from  $\text{၁၃၃၈}$  *kadāre* ‘so that’; the regular outcome with iterative reduplication applied would be  $\text{၁၃၃၈}$  *\*ka-kadāre*. As a conjunction, however, it is relatively frequent, so it does not seem odd that it has assumed a phonologically more simple, yet distinct form (compare, for instance, Bybee and Hopper 2001: 11–12). The conjunctive adverb in (230b) exhibits likewise a slightly irregular formation if we consider that it is essentially the negated form of  $\text{၁၃၃၈}$  *nārya* ‘but’; the regular outcome would have been  $\text{၁၃၃၈}$  *\*nāryoy*, which underwent simplification to  $\text{၁၃၃၈}$  *nāroy*, presumably as well due to its relatively high token frequency.

#### 4.6.4 Demonstrative adverbs

Besides demonstrative pronouns like  $\text{၁၃၃၈}$  *adanya* ‘that (one)’ (see section 4.2.2), and indefinite pronouns like  $\text{၁၃၃၈}$  *yāril* ‘for some reason; somewhere’ (see section 4.2.4), Ayeri also possesses demonstrative pronouns for the adverbial categories place, time, manner, and reason. The full paradigm is given in Table 4.25. Compared to the paradigm for demonstrative pronouns relating to persons or things, the paradigm of adverbial demonstratives is incomplete in that forms with  $\text{၁၃၃၈}$  *da-* ‘such’ are unattested. Thus, instead of the hypothetical form with  $\text{၁၃၃၈}$  *da-*, a full-NP adverbial

Table 4.25: Demonstratives relating to adverbial categories

Category	Proximal		Distal	
PLACE	<i>edaya</i>	‘here’	<i>adaya</i>	‘there’
TIME	<i>edaui</i>	‘now’	<i>adaui</i>	‘then’
MANNER	<i>edāre</i>	‘hereby’	<i>adāre</i>	‘thereby’
REASON	<i>edayam</i>	‘herefore’	<i>adayam</i>	‘therefore’

with a generic noun has to be used: \* $\text{ᠰᠡᠳᠠ}$  \**daya*  $\rightarrow$   $\text{ᠰᠡᠳᠠᠶ᠋ᠠᠨᠣᠶ᠋ᠠ}$  *da-yanoya* ‘in such a place’ (such-place-LOC). Adverbial demonstratives are, like pronouns, in complementary distribution with full NPs, since they are pro-forms. Thus, using them as modifiers to NPs as in (231a) is not possible, while using simple demonstrative  $\text{ᠰᠡᠳᠠ}$  *eda* ‘this’ together with a noun as in (231b) or using  $\text{ᠰᠡᠳᠠᠶ᠋ᠠ}$  *edaya* ‘here’ as a pro-form fully replacing the NP  $\text{ᠰᠡᠳᠠᠨᠠᠭᠠᠶ᠋ᠠ}$  *eda-nangaya* ‘in this house’ as in (231c) is generally unproblematic.

- (231) a. \**Ang mice Pada nangaya edaya.*  
           ang mit-ye Ø Pada nanga-ya edaya  
           AT live-3SG.F TOP Pada house-LOC here
- b. *Ang mice Pada eda-nangaya.*  
      ang mit-ye Ø Pada eda=nanga-ya  
      AT live-3SG.F TOP Pada this=house-LOC  
      ‘Pada lives in this house.’
- c. *Mice ang Pada edaya.*  
      mit-ye ang Pada edaya  
      live-3SG.F A Pada here  
      ‘Pada lives here.’

## 4.7 Numerals

The vast majority of the 196 sampled languages in Comrie (2013) either counts in tens or employs a mixed vigesimal-decimal system, while only five languages in the sample use a different base than 10. Ayeri uses a duodecimal system and is thus very untypical compared to real-world languages in using a number base other than 10—none of the languages in Comrie (2013)’s sample are listed as duodecimal.<sup>46</sup>

<sup>46</sup> I chose to use 12 as a numerical base because I simply wanted to toy with it. Also, I originally conceived of Ayeri speakers as humanoid but not necessarily human, which meant that they

Ayeri's number words are mostly semantic primes, that is, their meanings cannot be readily recognized as derived from body parts (Dixon 2012: 74) or from internal arithmetic like 9 as 'ten lacking one', for instance. The numerals ꠘꠗ *kay* 'three', ꠘꠗ *iri* 'five', and ꠘꠗ *ben* 'eight' may be an exception: as a quantifying adverb, ꠘꠗ *kay* means 'a little, few'; ꠘꠗ *iri* means 'already', which might refer to the fact that a full hand has been counted off; and ꠘꠗ *ben* also means 'all'. Ayeri moreover appears extremely sophisticated as far as the upper limit of counting systems is concerned in possessing a way of forming large numerals by a theoretically open-ended, recursive process.

#### 4.7.1 Cardinal numerals

Since people and concrete things are usually present in a countable manner, I want to comment first on how countable entities are handled with regards to numerals. After this, a discussion of how to express fractional amounts will follow.

##### *Integers*

Cardinal numerals work much like adjectives in that they modify nouns. As modifiers, they are placed after nouns. The full table of cardinal numerals from  $0 \times 12^\circ$  (o) to  $11 \times 12^\circ$  (B) is given in Table 4.26.<sup>47</sup> An example of simple modification by a numeral is given in (232):

- (232) *Ang tenyaya pang bihanya yo soyang miye.*  
 ang tenya=ya pang bihan-ya yo soyang miye  
 AT die=3SG.M.TOP ago week-LOC four or six

'He died four or six weeks ago.'

would not need to have evolved to have five fingers on each hand—for an earlier fictional language of mine, Daléian, I used an octal system reasoning that speakers would only have four digits. In any case, a duodecimal system could work reasonably well with human hands if you counted not only the fingers, but also the hands themselves. Finger-counting in Ayeri's duodecimal system would probably be similar to counting in the senary system of Nén described in Evans (2009) (as quoted in Dixon 2012): "In counting, Nén speakers 'first count off the five fingers with a finger of their other hand, and then on the sixth they place their counting finger on the inside of the wrist'" (73–74). Even though duodecimal numeral systems only occur rarely in natural languages, they are not entirely unheard of. Thus, for instance, Cain and Gair (2000) report that in Maldivian, the "decade plus numeral system is currently in fashion, but with some remnants of an older system as well. The numeral *fas dolas* '60' (lit., 'five twelves') comes from a duodecimal system that has all but disappeared in the Maldives. This number system was used for special purposes such as counting coconuts" (21).

<sup>47</sup> For the sake of typographic simplicity, A and B will henceforth be used to mean  $10 \times 12^\circ$  and  $11 \times 12^\circ$ , respectively. An index '10' after a figure indicates base 10, while an index '12' indicates base 12.



Table 4.26: Basic cardinal numerals

Numeral	Word	Numeral	Word
0	ညံ့ <i>ja</i>	6	မိပ် <i>miye</i>
1	မံင် <i>men</i>	7	မိပ် <i>ito</i>
2	မံ <i>sam</i>	8	မံင် <i>hen</i>
3	မံင် <i>kay</i>	9	မံင် <i>veya</i>
4	ယို <i>yo</i>	A	မံင် <i>mal</i>
5	မိပ် <i>iri</i>	B	မံင် <i>tam</i>

Table 4.27: Numerals for factors of 12

Numeral	Word	Numeral	Word
		60	မိပ်မံင် <i>miyelan</i>
10	မံင်မံင် <i>menlan</i>	70	မိပ်မံင် <i>itolan</i>
20	မံင်မံင် <i>samlan</i>	80	မံင်မံင် <i>henlan</i>
30	မံင်မံင် <i>kaylan</i>	90	မံင်မံင် <i>veyalan</i>
40	ယိုမံင် <i>yolan</i>	A0	မံင်မံင် <i>mallan</i>
50	မိပ်မံင် <i>irilan</i>	B0	မံင်မံင် <i>tamlan</i>

In this example, the numeral ယို *yo* ‘four’ modifies the noun မိပ် *bihan* ‘week’. Notably, however, plural marking is missing on the noun, since the notion of plurality is provided by the numeral itself; the numeral is thus normally sufficient to mark the whole NP as plural.

Multiples of 12<sup>1</sup> between 10 and 80 are formed by appending the suffix *-lan* to the numbers from 0 to B, which are given in Table 4.27. These numerals themselves act as heads for forming compounds with lower numerals to fill in the 12° numerals 11, 12, 13, ..., 21, 22, 23, etc. Thus, one counts on from မံင်မံင် *menlan* ‘dozen’ in the following way:

- (233) a. မံင်မံင်မံင် *menlan-men* (11),  
 မံင်မံင်မံင် *menlan-sam* (12),  
 မံင်မံင်မံင် *menlan-kay* (13),  
 etc.
- b. မံင်မံင်မံင် *samlan-men* (21),  
 မံင်မံင်မံင် *samlan-sam* (22),  
 မံင်မံင်မံင် *samlan-kay* (23),  
 etc.

- c.  $\text{ṭamlan-men}$  (B1),  
 $\text{ṭamlan-sam}$  (B2),  
 $\text{ṭamlan-kay}$  (B3),  
 etc.

In order to form yet higher numbers, the suffix  $\text{-nang}$  is appended to numerals:  $\text{ṭenang}$  ( $\leftarrow \text{ṭen}$  '1' +  $\text{-nang}$ ),  $\text{samang}$  '2' ( $\leftarrow \text{sam}$  +  $\text{-nang}$ ),  $\text{kaynang}$  ( $\leftarrow \text{kay}$  '3' +  $\text{-nang}$ ), etc. While  $\text{ṭenang}$  is used for 100, higher forms in the *nang* series each multiply the numeral from which they are derived by the factor of a duodecimal myriad ( $= 20736_{10}$ ). Thus, we get the following series:

- (234)  $\text{samang}$   $12^{(2-1) \times 4} = 12^4$  myriad  
 $\text{kaynang}$   $12^{(3-1) \times 4} = 12^8$  myriad myriads  
 $\text{yonang}$   $12^{(4-1) \times 4} = 12^{12}$  myriad myriad myriads  
 $\text{irinang}$   $12^{(5-1) \times 4} = 12^{16}$  myriad myriad myriad myriads  
 etc.

The numeral which the *nang* series word is based on essentially indicates the number of myriad groups, thus, 1-*nang* maximally contains BBBB; 2-*nang* maximally contains BBBB BBBB; 3-*nang* maximally contains BBBB BBBB BBBB, etc. Furthermore, the *nang* series words serve as unit words, and thus can be modified by numerals again, for instance:

- (235) a.  $\text{ṭenang sam veyalan-kay}$   
           100   2   90   3  
 $293_{12} = 399_{10}$   
 b.  $\text{samang benlan-miye ṭenang sam veyalan-kay}$   
    10000   80   6   100   2   90   3  
 $860293_{12} = 2115471_{10}$

In (235a),  $\text{sam}$  modifies  $\text{ṭenang}$  to indicate that there are two sets of  $100_{12}$ . Likewise, in (235b),  $\text{samang}$  is modified by  $\text{benlan-miye}$  to mean  $86_{12}$  times  $10000_{12}$ . Unit words like  $\text{ṭenang}$ ,  $\text{samang}$ , etc. may also be used as (inanimate) nouns, so it is possible to speak of  $\text{ṭenangye}$  'hundreds'. To express 'hundreds of people', however, the head of the genitive NP is pluralized exceptionally, even if it is a *plurale tantum*:

- (236) a.  $\text{Ang bengyon keynam ṭenang kanānya desay iray.}$   
           ang beng-yon keynam-Ø ṭenang kanān-ya desay iray  
           AT attend-3PL.N people-TOP hundred wedding-LOC royal  
           'A hundred people attended the royal wedding.'

- b. *Ang bengyon keynamye menang kanānya desay iray.*  
 ang beng-yon keynam-ye-Ø menang kanān-ya desay iray  
 AT attend-3PL.N people-PL-TOP hundred wedding-LOC royal  
 ‘Hundreds of people attended the royal wedding.’

In (236), *keynam* is morphologically a singular form referring semantically to a multitude. It is usually treated as a **plurale tantum** in that it triggers plural agreement in spite of being morphologically singular, which is illustrated in (236a). In (236b), the word still receives otherwise redundant plural marking to express the difference in meaning from (236a).

In order to indicate that myriad groups have been skipped, the conjunction *ṛṇay* ‘and’ is used to avoid confusion, as shown in (237), or simply to avoid having two single-digit numerals following each other, as illustrated by (238).

- (237) a. *samang menang men benlan-miye*  
 1 0000 100 1 80 6  
 $186\ 0000_{12} = 5\ 101\ 056_{10}$
- b. *samang menang men ṇay benlan-miye*  
 1 0000 100 1 and 80 6  
 $100\ 0086_{12} = 2\ 986\ 086_{10}$

- (238) a. *ṇmenang mal ito*  
 100 A 7
- b. *menang mal ṇay ito*  
 100 A and 7  
 $A07_{12} = 1\ 447_{10}$

### Fractions

So far, we have explored only integers, that is, whole numbers. Since things can often be divided up into smaller sections as well, this section is going to deal with how to express fractional amounts. The main way to express common fractions like  $\frac{1}{2}$ ,  $\frac{1}{3}$ ,  $\frac{1}{4}$ , etc. is to prepend *men* ‘one’ to the denominator; the full paradigm for  $\frac{1}{2}$  to  $\frac{1}{B}$  is given in Table 4.28. Note that a number of these fractions have slightly irregular forms due to assimilation in consonant clusters. In order to introduce a numerator, the fraction numeral is used as a unit word which is modified by a regular cardinal numeral, as (239) shows.

Table 4.28: Simple fractions from  $\frac{1}{2}$  to  $\frac{1}{6}$ 

Numeral	Word	Numeral	Word
$\frac{1}{2}$	မေဗေ <i>mesam</i>	$\frac{1}{7}$	မေဗို <i>menito</i>
$\frac{1}{3}$	မေဗ် <i>menkay</i>	$\frac{1}{8}$	မေဗို <i>menyen</i>
$\frac{1}{4}$	မေဗို <i>menyo</i>	$\frac{1}{9}$	မေဗို <i>menveya</i>
$\frac{1}{5}$	မေဗို <i>meniri</i>	$\frac{1}{A}$	မေဗို <i>memal</i>
$\frac{1}{6}$	မေဗို <i>memiye</i>	$\frac{1}{B}$	မေဗို <i>mentam</i>

- (239) a. *Ang ilca Yan vadiśānley mesam.*  
ang ilt-ya Ø Yan vadiśān-ley mesam  
AT buy-3SG.M TOP Yan bread-P.INAN half  
‘Yan bought half a loaf of bread.’
- b. *Ang ilce Mali sikanley menyo kay kipunena.*  
ang ilt-ye Ø Mali sikan-ley menyo kay kipunena  
AT buy-3SG.M TOP Mali pound-P.INAN fourth three cheese-GEN  
‘Mali bought a three-quarter pound of cheese.’

In order to express compound numerals, မေဗို *men-* may still be added to the denominator head word, for instance

- (240) *memallan-ben*  
men-mallan-hen  
 $1/10 \times 12^1 + 8$   
 $\frac{1}{A8_{12}} = \frac{1}{128_{10}}$

However, this may become confusing if numerators are used, so

- (241) *memenang ito menlan-yo kay*  
men-menang ito menlan-yo kay  
 $1/12^2$  7  $1 \times 12^1 + 4$  3  
 $\frac{3}{714_{12}} = \frac{3}{1024_{10}}$

would be expressed less ambiguously as

- (242) *menangan ito menlan-yo nernanyena kay*  
menang-an ito menlan-yo nernan-ye-na kay  
 $12^2$ -NMLZ 7  $1 \times 12^1 + 4$  part-PL-GEN 3  
‘three of the 1024th part’

Table 4.29: Basic ordinal numerals

Numeral	Word	Numeral	Word
0th	གཤམ་ <i>jān</i>	6th	མྱུ་ <i>miyan</i>
1st	མེན་ <i>menan</i>	7th	ཇིན་ <i>itan</i>
2nd	མམ་ <i>saman</i>	8th	མེན་ <i>henan</i>
3rd	མུ་ <i>kayan</i>	9th	མུ་ <i>veyān</i>
4th	ཡ་ <i>yan</i>	10th	མུ་ <i>malan</i>
5th	མུ་ <i>iran</i>	11th	མུ་ <i>taman</i>

using the ordinal form of the denominator.

#### 4.7.2 Ordinal numerals

Ordinal numerals are formed by nominalization from cardinal numerals. This may be another slightly odd strategy, however, it is in fact attested in Classical Tibetan, according to Chung et al. (2014), in reference to Beyer (1992):

The suffix *-pa* forms a noun from another noun, meaning ‘associated with N’ (e.g. *rta* ‘horse,’ *rta-pa* ‘horseman,’ *yi-ge* ‘letter,’ *yi-ge-pa* ‘one who holds a letter of office,’ cf. Beyer 1992: 117). When suffixed to ordinal numbers this suffix forms ordinals (e.g. *gsum* ‘three,’ *gsum-pa* ‘third’; *bcu* ‘ten,’ *bcu-pa* ‘tenth’). (Chung et al. 2014: 626)

Unfortunately, neither Chung et al. (2014) nor Beyer (1992) say whether Classical Tibetan treats these derived forms as nouns or as numerals, or whether it makes that distinction at all. In Ayeri, ordinals are firmly treated as noun-like nominal elements due to the derivational suffix *-an* (compare section 4.1.7). Since nominals are the heads of NPs, this also means that the ordinal numeral forms the head of the NP it occurs in, instead of modifying the entity being counted like an ordinal numeral does. This is illustrated in (243) below. The paradigm for the ordinal numerals from 0 to 8 can be found in Table 4.29.

- (243) a. *Ang Mahān menanas.*  
 ang Mahān menan-as  
 A Mahān first-P  
 ‘Mahan is the first.’
- b. *Ang Mahān menanas si girenjāng.*  
 ang Mahān menan-as si girend=yāng  
 A Mahān first-P REL arrive=3SG.M.A  
 ‘Mahān is the first to arrive.’

- c. *Ang girenja Mahān babalanya ku-menan diyan*  
 ang girend-ya Ø Mahān bahalan-ya ku=menan diyan  
 A arrive-3SG.M TOP Mahān finish-LOC like=first worthy  
 ‘Mahān arrives at the finish as a worthy first.’
- d. *Ang tavya Mahān menanas ganyena yana.*  
 ang tav-ya Ø Mahān menan-as gan-ye-na yana  
 AT get-3SG.M TOP Mahān first-P child-PL-GEN 3SG.M.GEN  
 ‘Mahān gets his first child’,  
 literally: ‘Mahān gets the first of his children.’

As (243a) shows, the ordinal may serve as an anaphora meaning ‘the *n*th (one)’. In these cases, animacy is determined by the word the ordinal references, as far as case marking and person agreement are concerned. Since ordinals are treated as nominals, they can also be modified by both a relative clause, as (243b) shows, and an adjective, as shown in (243c). In order to include an entity whose rank in a series is given, the counted entity appears as a genitive attribute, which is illustrated by (243d).

So far, only single-digit ordinals have been described. In order to form higher ordinals, the head unit word receives the nominalizer with the rest of the term trailing as a modifier, otherwise the number word as such is nominalized. Essentially, an ordinal in the ‘teens’<sup>48</sup> behaves like a ‘tight’ noun compound, while ordinals involving unit words for powers of 12 higher than 1 behave as ‘loose’ compounds (compare section 4.1.5, p. 130).

- (244) a. *Adareng kaylan-miyanley bahisyena pericanena.*  
 ada-reng kaylan-miye-an-ley bahis-ye-na perican-ena  
 that-A.INAN  $3 \times 12^1 + 6$ -NMLZ-P.INAN day-PL-GEN year-GEN  
 ‘It is the 36th (= 42nd) day of the year.’
- b. *Adareng menanganley kaylan-miye bahisyena pericanena.*  
 ada-reng menang-an-ley kaylan-miye bahis-ye-na perican-ena  
 that-A.INAN  $12^2$ -NMLZ-P.INAN  $3 \times 12^1 + 6$  day-PL-GEN year-GEN  
 ‘It is the 136th (= 186th) day of the year.’

In (244a), the whole numeral  $\text{𑜀𑜢𑜤𑜰𑜫 𑜏𑜢𑜤𑜰𑜫}$  *kaylan-miye* is nominalized and inflected for case, yielding  $\text{𑜀𑜢𑜤𑜰𑜫 𑜏𑜢𑜤𑜰𑜫}$  *kaylan-miyanley*. This is analogous to such nouns as  $\text{𑜀𑜢𑜤𑜰𑜫 𑜏𑜢𑜤𑜰𑜫}$  *betaynimpur* ‘grape’ (literally ‘wine-berry’), which inflects as a single unit—a ‘tight’ compound. In (244b), on the other hand, only the first unit word,  $\text{𑜀𑜢𑜤𑜰𑜫}$  *menang* is nominalized and inflected, yielding  $\text{𑜀𑜢𑜤𑜰𑜫 𑜏𑜢𑜤𑜰𑜫}$  *menanganley* with  $\text{𑜀𑜢𑜤𑜰𑜫}$

<sup>48</sup> More specifically,  $a \times 12^1 + b$  with  $\{a, b \in \mathbb{Z} \mid 0 < a, b < 12^1\}$ .

- (246) a. *kaylan-tamanyam*  
           *kay-lan-tam-an-yam*  
            $3 \times 12^1 + 11\text{-NMLZ-DAT}$   
           ‘ $_{3B}$  (=  $47_{10}$ ) times’
- b. *menanganyam*    *men samlan-kay*  
       *menang-an-yam*    *men sam-lan-kay*  
        $12^2\text{-NMLZ-DAT}$      $1$      $2 \times 12^1 + 3$   
       ‘ $_{123}$  (=  $171_{10}$ ) times’

#### 4.7.4 Distributive numerals

Distributive numerals are formed similar to multiplicative numerals in that they are based on a derivation of the respective ordinal numeral, which itself has the form of a nominalized cardinal numeral (compare Table 4.29). The derivative affix in this case is the instrumental marker,  $\text{ᄒᆞᆫ}$  *-eri* (compare section 4.1.3). Distributive numerals refer to groups of *n*, as example (247) shows.

- (247) *Ang sarayon burangjang kong besonya samaneri.*  
 ang sara-yon burang-ye-yang kong beson-ya sam-an-eri  
 AT go-3PL.N animal-PL-A inside ship-LOC two-NMLZ-INS

‘The animals went inside the ship two by two.’

The formation of composite numerals mirrors that of multiplicative numerals, in that composite numerals below  $12^1$  are treated as single units whereas composite numerals of orders of magnitude larger than  $12^1$  mark only the head word while the remainder of the phrase follows as an uninflected modifier.

- (248) a. *benlan-yaneri*  
 hen-lan-yo-an-eri  
 $8 \times 12^1 + 4$ -NMLZ-INS  
 ‘84 by 84 (=  $100_{10}$ )’  
 b. *menanganeri miye tamlan-yo*  
 menang-an-eri miye tam-lan-yo  
 $12^2$ -NMLZ-INS 6  $11 \times 12^1 + 4$   
 ‘6B4 by 6B4 (=  $1000_{10}$ )’

#### 4.7.5 Number ranges

Ranges of cardinal numbers may be viewed as conceptually similar to stretches of way, hence they are predestined to be expressed by prepositional phrases, or in Ayeri, by any of the cases which can be used for locative purposes (dative, genitive, locative). However, Ayeri treats cardinal numerals more like adjectives than nouns, so using means of case marking is not possible. On the other hand, adpositions take both NPs and CPs as complements, so that an adjective should be able to act as a complement to an adposition as well. Since the numeral in the PP is treated like an adjective, it is not marked for locative case, since adjectives do not inflect for nominal categories (compare section 4.3). Ranges of cardinal numbers may hence be expressed using the postposition  $\text{ᄒᆞᆫ}$  *pesan* ‘(up) until’. When counting starts at  $\text{ᄒᆞᆫ}$  *men* ‘one’, this numeral may be dropped, like in English ‘count to ten’ instead of ‘count from one to ten’.



- (249) *Kurye ang Pila (men) tam pesan.*  
 kur-ye ang Pila (men) tam pesan  
 count-3SG.F A Pila (1) B until  
 ‘Pila counts from 1 to B (= 1 ... 11<sub>10</sub>).’

Since ordinal numerals are treated as nouns, they may receive case marking. This means that, in contrast to cardinal numerals, it is possible to express a range using a combination of the genitive and the dative case, or again *pesan* with its prepositional object in the locative case. Context is needed to disambiguate whether the dative form of the numeral is a multiplicative derivation or an actual ordinal numeral in the dative case. Examples for this are given in (250).

- (250) a. *Ang gumasaya samanena pidimyena da-malanyam.*  
 ang gum-asa=ya saman-ena bahis-ye-na da=malan-yam  
 AT work-HAB=3SG.M.TOP second-GEN hour-PL-GEN such=tenth-DAT  
 ‘He usually works from the second hour to the tenth.’
- b. *Ang yomaya Magay diyan edaya benanena babisyena*  
 ang yoma-ya Ø Magay diyan edaya henan-ena bahis-ye-na  
 AT exist-3SG.M TOP Magay worthy here eighth-GEN day-PL-GEN  
*da-menlananya pesan.*  
 da-menlanan-ya pesan  
 such-dozen-th-LOC until  
 ‘Mr. Magay is here from the eighth to the dozen-th day.’

*samanena* ‘from the first’ in (250a) and *benanena* ‘from the eighth’ in (250b) use the genitive case marker *-ena* (compare section 4.1.3) to indicate the starting point. *da-malanyam* ‘to the tenth one’ and *da-menlananya pesan* ‘up until the dozen-th one’ indicate the end points. Since *menlan* in (250b) is embedded in a PP headed by the postposition *pesan*, it appears in the locative case instead of the dative case like *malan* in (250a).

## 4.8 Conjunctions

section 4.6.3 already dealt with conjunctive adverbs as sentence adverbs and their conjunction-like behavior. The present section is about the ‘purely logical’ conjunctions *ꠘꠗ* *nay* ‘and’ and *ꠘꠗꠗꠗ* *soyang* ‘or’, as well as their combination with *ꠘꠗꠗꠗ* *kamo* ‘equal(ly)’ to form correlative conjunctions.

### 4.8.1 Simple conjunction and disjunction

Coordination is commonly achieved by the conjunction  $\text{ꠘ}$  *nay* ‘and’. It is placed in between the conjuncts, and works on all syntactic levels. Namely, it may coordinate lexical heads, as well as phrases, and whole clauses.

- (251) a.  $[\text{AP}[\text{A } \text{Tarán}] \text{ } \text{nay} \text{ } [\text{A } \text{saco}]] \text{ } \text{nāngāng}$ .  
           Tarán    nay    saco    nāngāng  
           quiet   and   cool   house-A  
           ‘The house is quiet and cool.’
- b.  $\text{Ajayan} \text{ } [\text{NP } \text{yanang}] \text{ } \text{nay} \text{ } [\text{NP } \text{layang}]$ .  
      aja-yan       yan-ang   nay   lay-ang  
      play-3PL.M   boy-A     and   girl-A  
      ‘The boy and the girl are playing.’
- c.  $[\text{S } \text{nāng } \text{pisu}] \text{ } \text{nay} \text{ } [\text{S } \text{tapannang}]$ .  
      nāng   pisu   nay   tapan-nang  
      IPL.A   tired   and   be.thirsty-IPL.A  
      ‘We are tired and are thirsty.’

The example sentences in (251) are ordered by increasing level of coordination: (251a) combines two adjective-phrase (AP) heads,  $\text{ꠘ}$  *tarán* ‘quiet’ and  $\text{ꠘ}$  *saco* ‘cool’, which together make up the predicative AP that is equated to  $\text{ꠘ}$  *nāngāng* ‘a/the house’. In (251b), then, two agent NPs,  $\text{ꠘ}$  *yanang* ‘a/the boy’ and  $\text{ꠘ}$  *layang* ‘a/the girl’, together form the subject of the verb  $\text{ꠘ}$  *ajayan* ‘play’. Lastly, (251c) shows two main clauses coordinated, that is,  $\text{ꠘ}$  *nāng pisu* ‘we are tired’ on the one hand, and  $\text{ꠘ}$  *tapannang* ‘we are thirsty’ on the other.

Just as  $\text{ꠘ}$  *nay* expresses *conjunction*,  $\text{ꠘ}$  *soyang* ‘or’ expresses *disjunction*. It is equally placed between two disjuncts and equally works at all levels—lexical heads, phrases, and clauses. Inclusive and exclusive ‘or’ are not formally distinguished in Ayeri by the disjunction  $\text{ꠘ}$  *soyang* alone, so context is necessary to contrast between them. Alternatively, a construction akin to English ‘either ... or’ may be used to make the distinction explicit (see section 4.8.2).

- (252) a.  $\text{Pasyyang, } \text{yāng} \text{ } [\text{AP}[\text{A } \text{mino}] \text{ } \text{soyang} \text{ } [\text{A } \text{giday}]]$ .  
      pasy=yang   yāng       mino   soyang   giday  
      wonder=1SG.A   3SG.M.A   happy   or   sad  
      ‘I wonder whether he is happy or sad.’
- b.  $\text{Le } \text{no } \text{ginvāng} \text{ } [\text{NP } \text{karon}] \text{ } \text{soyang} \text{ } [\text{NP } \text{gali}] \text{ } ?$   
      le   no   gin=vāng       karon-Ø   soyang   gali-Ø  
      PT.INAN   want   drink=2SG.A   water-TOP   or   juice-TOP  
      ‘Do you want to drink water or juice?’

- c. [<sup>s</sup>Beratu edauyi] soyang [<sup>s</sup>sa-sahu rangya ]!  
 berata-u edauyi soyang sa~saha-u rang-ya  
 decide-IMP now or return-IMP home-LOC  
 ‘Decide now or go home!’

As above, (252) shows different syntactic contexts for *soyang*. In (252a), two adjectives, *mino* ‘happy’ and *giday* ‘sad’ are put in opposition as phrasal heads making up a predicative AP. Then, in (252b), the choice is between two NPs, *le ... karon* ‘water’ and *le ... gali* ‘juice’, which jointly form the object of *ginvāng* ‘you drink’. Lastly, in (252c), two main clauses are in opposition—either disjunct forms a complete sentence on its own.

#### 4.8.2 Complex conjunction and disjunction

English has a number of correlative conjunctions, that is, conjunctions made up of multiple parts which work together as one expression. Among these are, notably, *as ... as*, *both ... and*, *either ... or*, *neither ... nor*, *rather ... than*, and *the ... the*. Ayeri uses the adverb *kamo* ‘equally, same’ together with a conjunction for many of these.

*kamo ... nay* ‘equally ... and’ is equivalent to ‘both ... and’: the correlative construction emphasizes that two options are equal to each other. Syntactically, resulting sentences are equal to those presented in (251). *sano* ‘both’ may be used as a synonym to *kamo* as well here.

- (253) *Ang vacay kamo piyuley nay obanley.*  
 ang vac=ay.Ø kamo piyu-ley nay oban-ley  
 AT like=ISG.TOP equally grain-P.INAN and bean-P.INAN  
 ‘I like both grains and beans.’

Alternatively, it is possible to use a construction with *naynay* ‘(and) also’:

- (254) *Ang vacay piyuley, obanley naynay.*  
 ang vac=ay.Ø piyu-ley oban-ley naynay  
 AT like=ISG.TOP grain-P.INAN bean-P.INAN also  
 ‘I like grains and also beans.’

The example in (253) may be translated more literally as ‘I like grains and beans equally’, with two NPs in alternation, both being objects of a transitive verb, *vac-* ‘like’. With predicative adjectives, the verb *kama-* ‘(be) equal’ may be used:

- (255) *Ang kamayan mabo nay giday.*  
 ang kama=yan.Ø mabo nay giday  
 AT be.equal=3PL.M.TOP hungry and thirsty  
 ‘They are both hungry and thirsty.’

ꨀꨣ: *kama-* is one of Ayeri’s copular verbs used to express equality between two properties of its subject. The literal meaning of (255) is thus, roughly, ‘They are as hungry as they are thirsty’. The construction slightly differs from that used to do comparison of NPs, however, in that the conjunction ꨀꨣ *nay* is placed between both predicative terms here. In order to express literal ‘be ... as ... as’, thus, the conjunction is dropped:

- (256) *Kamareng matikan belanas agonanya.*  
 kama=reng matikan helan-as agonan-ya  
 be.equal=3SG.INAN.A hot oven-P outside-LOC  
 ‘It’s as hot as an oven outside.’

ꨀꨣ—ꨀꨣ *kamo ... nay* is used to express ‘the ... the’, that is, a proportional or antiproportional relationship between two amounts, sizes, or properties; using ꨀꨣ *sano* ‘both’ here is judged less fitting. In order to express a relationship of equal increase/decrease in this way, conjuncts are additionally marked with the comparative suffix ꨀꨣ *-eng* ‘more, rather’ or its opposite, ꨀꨣ *-ikoy* ‘less’:

- (257) *Ang tavyan kamo nakēng nay konjāng-eng.*  
 ang tav=yan.Ø equal naked-eng nay kond=yāng=eng  
 AT become=3SG.M.TOP equally tall-COMP and eat=3SG.M.A=more  
 ‘The taller they get, the more they eat.’

The type of correlative conjunction which selects one of two alternatives but not both—that is, exclusive ‘or’ (XOR)—is expressed by the construction ꨀꨣ—ꨀꨣ *kamo ... soyang* ‘equally ... or’, as illustrated by (258).<sup>49</sup> For its negative opposite, ‘neither ... nor’, negation must be used, which is displayed in (259).

<sup>49</sup> Interestingly, it looks as though I slightly plagiarized English here in idea, albeit unwittingly: the etymology of *either* is given as being from Old English *ēghwæðer*, *ēgðer*, from Germanic *\*aiwon* ‘always’ + *\*gihwafaro-z* ‘each of two’, cf. OED (2016: *either*, adj. (and pron.) and adv. (and conj.)). On the other hand, the collective wisdom of the internet’s fictional-language community holds that one cannot truly innovate grammatical structures; there will always be a natural language which evolved a given construction before, and possibly with more complications involved. This situation is referred to as ‘ANADEW’, an acronym for ‘a nat[ural] lang[uage] already dunnit except worse’ (Teoh 2003).

- (258) a. *Ang miraya kamo Ajān adaley eda-konkyanya soyang da-mararya.*  
 ang mira-ya kamo Ajān adaley eda=konkyan-ya soynag da=mararya  
 AT do-3SG.M equally Ajān that-P.INAN this=month-LOC or such-next  
 ‘Ajān does it either this month or next.’
- b. *Kamayong mabo soyang krito mirampaluy.*  
 kama=yong mabo soyang krito mirampaluy  
 be.equal=3SG.N.A hungry or angry otherwise  
 ‘They are either hungry or otherwise angry.’
- (259) a. *Ang taboyye kamo Sipra netuas soyang kinās.*  
 ang taha-oy-ye kamo Ø Sipra netu-as soyang kinās  
 AT have-NEG-3SG.F equally TOP Sipra brother-P or sister-P  
 ‘Sipra has neither a brother nor a sister.’
- b. *Ang kamuay layas soyang veno.*  
 ang kama-oy=ay.Ø lay-as soyang veno  
 AT be.equal-NEG=1SG.A girl-P or beautiful  
 ‘I am neither a girl nor beautiful.’



## 5 Phrase structures

---

While the previous chapter dealt largely with the various parts of speech and their various distributive and inflectional properties, the present chapter will elaborate on how these words combine into syntactic phrases.<sup>1</sup> Since Ayeri is a verb-initial language, it is probably most comfortably analyzed in terms of Lexical-Functional Grammar (Bresnan 1982 ff.; more recently Bresnan et al. 2016; Dalrymple 2001), since LFG does not require complicated derivations behind the surface structure of sentences.<sup>2</sup> It will be assumed here that, even though Ayeri is basically VSO with predicate and predication not adjacent to each other, it is configurational in that there is a VP which c-commands a number of other constituents as complements in transitive sentences.

In principle, LFG assumes that grammar operates on different structural levels: mainly, these are a(rgument) structure, c(onstituent) structure, and f(unctional) structure; other layers have been proposed by different researchers for different purposes (Butt and King 2015: 862–865). Bresnan et al. (2016) define three core design principles for LFG:

**Variability:** “The principle of variability states that *external structures vary across languages*. The formal model of external structure in LFG is the *c-structure*, which stands for ‘constituent structure’ or ‘categorical structure’” (41). C-structures are commonly represented by context-free phrase-structure rules; constituency trees are based on an extended version of X-bar theory (42).<sup>3</sup>

<sup>1</sup> It should be noted that, since trees are large, a number of regularly numbered examples in this chapter will appear not in place but floated, like tables and figures in the previous chapters.

<sup>2</sup> Passivization, for instance, is assumed to be a lexically motivated alternation in predicate structure (SUBJ is blocked, so the nominative is assigned to OBJ, and the original SUBJ is expressed by an ADJ), rather than an internal derivational process (Bresnan et al. 2016: 23 ff.).

<sup>3</sup> The basic recursive rules of X-bar theory are observed:

1.  $XP \rightarrow YP, X'$  (specifier rule)
2.  $X' \rightarrow X', YP$  (adjunct rule)

**Universality:** “The principle of universality states that *internal structures are largely invariant across languages*. The formal model of internal structure in LFG is the *f-structure*, which stands for ‘functional structure’” (Bresnan et al. 2016: 42). The f-structure is depicted as an argument-value matrix (AVM) which maps the relations between ‘subject’ (SUBJ), ‘object’ (OBJ), ‘predicator’ (PRED), etc. as functional abstractions of NP, VP, V, etc. (42). Verbs are also presented with their *a-structure* spelled out. That is, which arguments a verb has relations to is formally stated (15). The f-structure collates semantic features associated with heads of grammatical functions (GFs), such as case (CASE), person (PERS), number (NUM), which are abstract features and as such need not have morphological realization (43).

**Monotonicity:** “Constituent structure form is simply not the same in all languages [...] In LFG the correspondence mapping between internal and external structures does not preserve sameness of form. Instead, *it is designed to preserve inclusion relations between the information expressed by the external structure and the content of the internal structure*” (43). Due to the principle of monorepresentation, information distributed over different morphemes which logically belongs to a single grammatical function is presented in the f-structure as unified.

To illustrate the different parallel structures in operation, Bresnan et al. (2016: 15) give the schema in Figure 5.1 to demonstrate which part of the a- and c-structure respectively corresponds (‘links’, ‘maps’) to which part of the f-structure. Regarding the different functions distinguished, LFG assumes the following hierarchies (97, 100):

- (1) a. Grammatical functions (GFs):  

$$\overbrace{\text{SUBJ} > \text{OBJ} > \text{OBJ}_\theta}^{\text{core}} > \overbrace{\text{OBL}_\theta > \text{XCOMP}, \text{COMP} > \text{ADJ}}^{\text{noncore}}$$
- b. (Non)argument functions (AFS/AFs):  

$$\overbrace{\text{TOP FOC}}^{\text{non-a-fns}} \overbrace{\text{SUBJ OBJ OBJ}_\theta \text{ OBL}_\theta \text{ XCOMP COMP}}^{\text{a-fns}} \overbrace{\text{ADJ}}^{\text{non-a-fns}}$$
- c. Discourse functions (DFS):  

$$\overbrace{\text{TOP FOC SUBJ}}^{\text{d-fns}} \overbrace{\text{OBJ OBJ}_\theta \text{ OBL}_\theta \text{ XCOMP COMP ADJ}}^{\text{non-d-fns}}$$

3.  $X' \rightarrow X^0, YP$  (complement rule)

The principle of economy of expression furthermore dictates that essentially, trees be pruned of empty terminal nodes and non-branching preterminal nodes, since these do not provide structurally or semantically relevant information (Bresnan et al. 2016: 119–128).





Figure 5.1: F-structure mappings (Bresnan et al. 2016: 15)

The elements listed in (1) will also appear in phrase-structure rules and c-structure trees together with arrows. These arrows symbolize inheritance of feature information from the current level ( $\downarrow$ ) of the tree to the next ( $\uparrow$ ), so for instance,  $(\uparrow \text{SUBJ}) = \downarrow$  means that the information subsumed by the current node ('down') is passed on as the subject function of the next higher node ('up') in the tree. Concise information on notational formalisms of LFG can be found, for instance, in Butt and King (2015).

## 5.1 Noun and determiner phrases

Noun phrases (NPs), and determiner phrases (DPs) as their functional counterpart, fulfill the functions of subject (SUBJ), object (OBJ), secondary object ( $\text{OBJ}_\theta$ ), as well as various oblique constituents ( $\text{OBL}_\theta$ ; see (2) below), and they can also form adjuncts (ADJ). DPs and NPs can also constitute topics (TOP). Which DP or NP receives which function is selected by the a-structure of the verb—this also has repercussions on case- and topic marking.

Since the constituent containing the various non-verbal elements in Ayeri is likely exocentric and constituents may move around within it in a restricted way, we have to assume that not constituent structure, but case marking identifies the grammatical functions of the various arguments of verbs. Thus, the following lexico-centric conditions operate on both DP and NP as exponents of case:

- (2) a.  $(\downarrow \text{ CASE}) = \text{A} \implies (\uparrow \text{ SUBJ}) = \downarrow$   
 b.  $(\downarrow \text{ CASE}) = \text{P} \implies (\uparrow \text{ OBJ}) = \downarrow$   
 c.  $(\downarrow \text{ CASE}) = \text{DAT} \implies (\uparrow \text{ OBJ}_\theta) = \downarrow \text{ OR } (\uparrow \text{ PCASE}) = \text{OBL}_{\text{goal}}$   
 d.  $(\downarrow \text{ CASE}) = \text{GEN} \implies (\uparrow \text{ POSS}) = \downarrow \text{ OR } (\uparrow \text{ PCASE}) = \text{OBL}_{\text{src}}$   
 e.  $(\downarrow \text{ CASE}) = \text{LOC} \implies (\uparrow \text{ OBL}_{\text{loc}}) = \downarrow \text{ OR } (\uparrow \text{ PCASE}) = \text{OBL}_{\text{loc}}$   
 f.  $(\downarrow \text{ CASE}) = \text{CAUS} \implies (\uparrow \text{ OBL}_{\text{caus}}) = \downarrow$   
 g.  $(\downarrow \text{ CASE}) = \text{INS} \implies (\uparrow \text{ OBL}_{\text{ins}}) = \downarrow$

The rules in (2) determine the typical mappings between case marking and grammatical functions, which are not always unambiguous. As explained above (compare section 4.1.3), the dative case does not only indicate that something is done to this referent or to their benefit, but it may also indicate motion towards this referent. Likewise, the genitive case does not only indicate possession, but also origin, and motion from this referent. Nominal adjuncts to nouns which specify what the noun consists also appear in the instrumental case, besides the instrumental being used to indicate the means or the circumstance by which an action comes about. Moreover, DPs or NPs may also lack case marking, which indicates that the respective phrase is a part of the topic of the verb, which is what (3) describes:

- (3)  $\neg(\downarrow \text{ CASE}) \implies \downarrow \in (\uparrow \text{ TOP})$

Instead of case marking on the DP or NP, there is a marker in front of the verb which provides information on the case and, if AT or PT, also about the animacy of the topicalized phrase. This means that grammatical information about the topic of a phrase is spread over two discontinuous locations. This issue does not pose a problem to an LFG-based analysis, however, since both locations unify their information content in the f-structure feature TOP. Since information located in multiple places is jointly feeding this feature, I am using the annotation ' $\downarrow \in (\uparrow \text{ TOP})$ ' for each location rather than simple ' $(\uparrow \text{ TOP}) = \downarrow$ '. Note that only one NP among the arguments of a verb may be the topic of the phrase, and a topic can only be marked if the verb is finite and the number of arguments to the verb is greater than one.

### 5.1.1 Noun phrases

#### *Constituent order within noun phrases*

Nouns are one of the main parts of speech of Ayeri, and nouns can be modified by a number of other free elements, as we have seen previously—adjectives, possessive adjectives, as well as relative clauses and nominal adjuncts. These typically follow nouns. It was also described before that Ayeri's nouns may host a number of clitics, among which are deictic prefixes and quantifiers, as well as enclitic case markers in the case of proper nouns. These clitics, however, will not be treated as targets

of syntactic operations, since LFG follows the approach of lexical integrity. Thus, bound elements like affixes and clitics are assumed not to be reflected or affected by syntax itself. The phrase structure of NPs should thus look like depicted in (4). The asterisk in (4d) is the Kleene star and indicates zero or more instances of the element. Elements which may occur zero or once are in brackets.

- (4) a.  $NP \rightarrow \begin{matrix} N' \\ \uparrow = \downarrow \end{matrix}$
- b.  $N' \rightarrow \begin{matrix} N' \\ \uparrow = \downarrow \end{matrix} \left( \begin{matrix} CP \\ (\uparrow \text{ COMP}) = \downarrow \end{matrix} \right)$
- c.  $N' \rightarrow \begin{matrix} N' \\ \uparrow = \downarrow \end{matrix} \left( \begin{matrix} \{NP \mid DP\} \\ (\uparrow \text{ POSS}) = \downarrow \end{matrix} \right)$
- d.  $N' \rightarrow \begin{matrix} N' & AP^* \\ \uparrow = \downarrow & \downarrow \in (\uparrow \text{ ADJ}) \end{matrix}$
- e.  $N' \rightarrow \begin{matrix} N' \\ \uparrow = \downarrow \end{matrix} \left( \begin{matrix} NP \\ (\uparrow \text{ XCOMP}) = \downarrow \end{matrix} \right)$

This rule defines that NPs have a lexical head which is on the left side, followed optionally by modifiers which may have various relationships to the noun: complement (instrumental nominal, NP/DP), adjunct (adjective phrase, AP), possessor (genitive NP/DP), relative clause (forming a complementizer phrase, CP). This can be represented as a constituent-structure tree in the way described in (5).



Here as well, we can see a nominal head on the left which may be modified by phrases of various types. The maximal projection of  $N^0$  (that is, NP) is annotated very generally for the function of the NP—basically, an NP can act as either a discourse function (DF) or a grammatical function (GF). (6) gives an example of each kind of modifier. Since there is no grammatical context given, NP is unmarked for function in these examples.

- (6) a. noun + adjective:

*ningan hiro*  
 ningan hiro  
 story new  
 ‘new story’



- b. noun + possessor:

*kegan ayonena*  
 kegan ayon-ena  
 hat man-GEN  
 ‘the man’s hat’



- c. noun + instrumental complement:

*kasu bariri*  
 kasu bari-ri  
 basket meat-INS  
 ‘basket of meat’



- d. noun + relative clause:

*nanga si incāng*  
 nanga si int=yāng  
 house REL buy=3SG.M.A  
 ‘the house he bought’



As described before (compare section 3.2.5), nouns can be modified by a number of clitics which are not represented through syntax. Since it is not possible for these clitic elements to be divided from their phonological hosts, they should be treated as being an integral part of the word they attach to. Hence, N<sup>0</sup> is given in (7) as split into ‘Cl’ and N<sup>0</sup>.

- (7) a. noun + deictic prefix:

*eda- nanga*  
 eda= nanga  
 this=house  
 ‘this house’



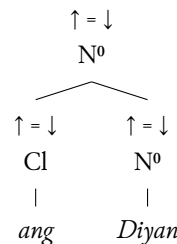
- b. noun + quantifier:

*nangās -ikan*  
 nanga-as=ikan  
 house-P =many  
 ‘many houses’



- c. proper noun + case:

*ang Diyan*  
 ang=Diyan  
 A= Diyan  
 ‘Diyan’



Of course, it is also possible to combine these nominal modifiers. In this case, there is a certain hierarchy, presumably based on Behaghel’s first law, “Das oberste Gesetz ist dieses, daß das geistig eng Zusammengehörige auch eng zusammengestellt wird” (Behaghel 1932: 4; ‘The supreme law is such that the mentally

closely related is also arranged in close proximity.’), and also grammatical weight (Wasow 1997):

1. instrumental NP indicating what the head consists of,
2. APs (also cardinals) and other NPs describing attributes,
3. possessive genitive NPs,
4. relative clauses.

Wasow (1997) writes that “[i]t is very hard to distinguish among various structural weight measures as predictors of weight effects. Counting words, nodes, or phrasal nodes all work well” (102), which means that no single metric can be used to describe the order of constituents in a phrase. However, for instance, relative clauses trail whenever possible presumably since they tend to contain whole sub-clauses and thus a lot of information. It seems advisable not to put an element with much less information content after them, especially when it refers to a different head than all the things inside the relative clause. The following example (8) illustrates the unmarked order of modifiers.

- (8) a. *diranang caban nā si ang mica ya Kārvisam*  
 diran-ang caban nā si ang mit=ya.Ø ya= Kārvisam  
 uncle-A favorite ISG.GEN REL AT live=3SG.M.TOP LOC=Kārvisam  
 ‘my favorite uncle who lives in Kārvisam’



As the c-structure tree in (8) shows, Ayeri prefers head–dependent word order with exceeding consistency. As illustrated by previous examples, both adjuncts and complements are, for the most part, consistently appended to the right of their

heads, which means that Ayeri may be classified as a rather consistently right-branching language. However, a certain number of postpositions form an exception to this classification (section 4.4.2). In the light of word order typology, we can formulate the following generalizations:

- (9) a. Order of noun and adjective: N Adj
- b. Order of noun and genitive: N Gen
- c. Order of noun and relative clause: N Rel

More important to LFG than c-structure trees, however, are function-structure matrices which gather all information in a given utterance and gather potentially disparate information into semantically coherent units within the matrix.<sup>4</sup> In the following, I will thus give a list of morpholexic specifications which give an overview of the different semantic and morphological features nouns basically provide (also compare section 4.1). These also form the basis for f-structure matrices of the kind already shown in (64), section 3.2.5, p. 91.

- (10) ...           N           (↑ PRED)   =   ‘...’  
                                   (↑ ANIM)   =   ±  
                                   (↑ CASE)   =   {A, P, DAT, GEN, LOC, INS, CAUS}  
                                   (↑ GEND)   =   {M, F, N, INAN}  
                                   (↑ NUM)   =   {SG, PL}  
                                   (↑ PERS)   =   3

Nouns generally imply a third-person reference; they distinguish number, gender and animacy, as well as case. Clitics, however, may also add information about deixis (11); likeness and quantity might be interpreted conveniently as adding to the list of a noun’s ADJ feature (12).

- (11)                           (↑ DEIX)   =   {*this, that, such*}

- (12) a.   *ganang-hen   mino*  
           *gan-ang=hen   mino*  
           *child-A=all    happy*  
           ‘all happy children’

<sup>4</sup> Essentially, c-structure is similar to the tree hierarchy of paragraphs, images, tables etc. in an HTML file, while f-structure describes semantic properties of elements in the tree similar to how CSS defines the layout properties of these elements.

$$b. \left[ \begin{array}{c} \text{SUBJ} \\ \left[ \begin{array}{c} \text{PRED} \text{ 'child'} \\ \text{ANIM} + \\ \text{CASE} A \\ \text{ADJ} \left\{ \begin{array}{c} \left[ \text{PRED} \text{ 'all'} \right], \\ \left[ \text{PRED} \text{ 'happy'} \right] \end{array} \right\} \end{array} \right] \end{array} \right]$$

### Morphosyntactic operations within the noun phrase

It has been pointed out above that nouns encode animacy. This has repercussions in the choice of case markers of the agent and patient cases, which need to agree with the lexical head they attach to. An example of this is given in (13).



Example (13a) shows a well-formed construction: the noun,  $\text{gan}$  'child', is animate, hence the case particle also needs to be animate—the case particle must thus be  $-\text{ang}$  to be coherent. In contrast to this, example (13b) is not well-formed in that the noun is animate but the case particle,  $-\text{reng}$ , signals that it is inanimate: the ANIM values of the noun stem and its suffix clash and cannot be conclusively unified for  $N^0$  itself. The same principle of coherence is, of course, also true for proper nouns, which receive a case-marking particle:





Furthermore, example (12) already showed that nouns may be modified by quantifiers, whether these are clitic suffixes (section 4.6.2) or numerals (section 4.7). In these cases, plural marking on the noun is suppressed by the presence of the modifier which supplies the information by itself so that further morphological plural marking by the suffix  $\text{ḥ}$  *-ye* on the noun stem itself would be redundant. As shown in section 4.7 (p. 226), however, there are very limited occasions where a noun may be marked for plural in spite of the presence of a numeral, for instance:

- (15) *Ang bengyon keynamye menang kanānya desay iray.*  
 ang beng-yon keynam-ye-Ø menang kanān-ya desay iray  
 AT attend-3PL.N people-PL-TOP hundred wedding-LOC royal

‘Hundreds of people attended the royal wedding.’

Here, the noun  $\text{ḥ}$  *keynam* ‘people’ is marked additionally for plural by the nominal plural suffix  $\text{ḥ}$  *-ye* in spite of being a *plurale tantum* and in spite of the presence of the numeral  $\text{ḥ}$  *menang* hundred. Without plural morphology, the meaning of  $\text{ḥ}$  *keynam menang* would be ‘a hundred people’, not generic ‘hundreds’.

### 5.1.2 Determiner phrases

Determiner phrases (DPs) are the functional equivalent of NPs; determiners as their heads ( $D^0$ ) are a closed class of function words (Bresnan et al. 2016: 102). In English, for instance, articles and pronouns are counted among them (Carnie 2013: 208–211). Ayeri, however, probably does not possess articles as such. The preposed case markers of proper nouns bear a superficial similarity to cased articles like in German (16) and the suffixed case markers look superficially similar to suffixed articles in Romanian (17). The presence or absence of case markers in Ayeri is moreover morphosyntactically controlled by topicalization and thus also interacts with definiteness. However, as we will see below, the distribution of these case particles differs from that of articles in languages like English or German.

- (16) a. A *ang Sān* ‘Sān’  
 P *sa Sān* ‘Sān’  
 DAT *yam Sān* ‘to Sān’  
 GEN *na Sān* ‘Sān’s’  
 LOC *ya Sān* ‘at Sān’  
 CAUS *sā Sān* ‘due to Sān’  
 INS *ri Sān* ‘with/by Sān’
- b. German:  
 NOM.SG *der Mann* ‘the man’  
 ACC.SG *den Mann* ‘the man’  
 DAT.SG *dem Mann* ‘to the man’  
 GEN.SG *des Mannes* ‘of the man’

- (17) a. A *ganang* 'a/the child'  
 P *ganas* 'a/the child'  
 DAT *ganyam* 'to a/the child'  
 GEN *ganena* 'of a/the child'  
 LOC *ganya* 'at a/the child'  
 CAUS *ganisa* 'due to a/the child'  
 INS *ganeri* 'with/by a/the child'
- b. Romanian (adapted from Lyons 1999: 75):
- |        |               |             |
|--------|---------------|-------------|
| PRI.SG | <i>cartea</i> | 'the book'  |
| OBL.SG | <i>cărți</i>  | 'the book'  |
| PRI.PL | <i>cărți</i>  | 'the books' |
| OBL.PL | <i>cărți</i>  | 'the books' |

While in German an article and a demonstrative pronoun, or also a possessive pronoun, cannot co-occur, this appears not to be a problem in Ayeri. As argued in section 3.2.5, both case markers and deictic/demonstrative prefixes in Ayeri are clitics; the similarity between possessive pronouns and adjectives has also been noted in section 4.2.1 (p. 147). Furthermore, the preposed case markers of nouns are an exception compared to the much more frequent occurrence of case-marking suffixes on generic nouns. It thus does not seem straightforward to analyze the case markers as heads of DPs.

- (18) a. *das* *Haus* [German]  
 das Haus  
 DEF.NOM.SG.N house  
 'the house'
- b. *dieses* *Haus*  
 dies-es Haus  
 this-NOM.SG.N.ST Haus  
 'this house'
- c. *mein* *Haus*  
 mein-Ø Haus  
 1SG.GEN-NOM.SG.N.ST house  
 'my house'
- d. \**das* *diese* *Haus*  
 das dies-e Haus  
 DEF.NOM.SG.N this-NOM.SG.N.WK house  
 'the this house'
- e. \**das* *meine* *Haus*  
 das mein-e Haus  
 DEF.NOM.SG.N 1SG.GEN-NOM.SG.N.WK house  
 'the my house'

- f. \**dieses*                      *meine*                      *Haus*  
      dies-es                      mein-e                      Haus  
      this-NOM.SG.N.ST    ISG.GEN-NOM.SG.N.WK    house  
      ‘this my house’
- g. #*dieses*                      *mein*                      *Haus*  
      dies-es                      mein-Ø                      Haus  
      this-NOM.SG.N.ST    ISG.GEN-NOM.SG.N.ST    house  
      ‘this house of mine’

The examples in (18) show that determining elements such as a definite article (*der* ‘the’), a demonstrative pronoun (*dieser* ‘this’) and a possessive pronoun (*mein* ‘my’) are in complementary distribution for some combinations. The only exception to this is the combination of demonstrative and possessive in (18g), which is grammatically marked, however.<sup>5</sup> On this phenomenon of complementary distribution of determiners, which also holds true for English, Carnie (2013: 208) writes, “One thing to note about determiners is that they are typically heads. Normally, there can only be one of them in an NP,” at least in English (and German). Demske (2001: 9–22) elaborates on this point for German as well. Regarding the examples in (17b), Dindelegan (2013) states about Romanian that

Prenominal demonstrative [*sic*] take a determinerless (articleless) head-noun complement [...] while postnominal demonstratives obligatorily occur in DPs with article-bearing noun heads [...] The postnominal construction is thus a polydefinite structure, since definiteness is realized twice [...], by the article and by the demonstrative.  
 (297)

Dindelegan (2013: 297) gives the following examples for these two placement variants (glosses extended based on further information in the grammar):<sup>6</sup>

- (19) a.    *acest*                      *om*                      [Romanian]  
          acest-Ø                      om-Ø  
          this-NOM+ACC.SG.M    man-SG  
          ‘this man’

<sup>5</sup> Example (18f) differs from (18g) in the declension paradigm of the adjective: (18f) uses the ‘weak’ (WK) declension regularly, since a determiner with strong (ST) declension precedes. (18g) appears to be an exception in permitting two determiners of the strong declension. Demske (2001: 160–161, 203–205) notes that, according to Plank (1992), possessive pronouns may apparently still act as modifiers, not only determiners, under certain circumstances. In modern Standard German, this construction is strongly marked, however. It is probably a remnant of earlier stages of German where there was no such restriction on the co-occurrence of demonstrative and possessive pronouns yet (Demske 2001: 173).

<sup>6</sup> In declension charts, Dindelegan (2013) indicates the cases as ‘NOM ≡ ACC’ and ‘DAT ≡ GEN’ where Lyons (1999) uses PRI and OBL.

- b. *omul*                                      *acesta*  
 om-ul                                      acest-a  
 man-DEF.NOM+ACC.SG.M    this-NOM+ACC.SG.M  
 ‘this man’

Ayeri, however, behaves differently than either German or Romanian in treating case markers and demonstrative elements as clitics. The case marker is always present for untopicalized NPs, whether there is modification by a demonstrative clitic or not. The demonstrative clitic merges with the head noun to the point where it is not certain whether it is still a clitic or already an inflectional prefix (section 3.2.5), that is, they do not have phrasal status like the postnominal determiners of Romanian, but they are not heads of DP like the prenominal determiners of Romanian either (Dindelegan 2013: 299), due to their status as clitics.

- (20) a. *ang Săn*  
 ang=Săn  
 A= Săn  
 ‘Săn’
- b. *ang eda- Săn*  
 ang=eda= Săn  
 A= this=Săn  
 ‘this Săn’
- c. *ang Săn nă*  
 ang=Săn nă  
 A= Săn 1SG.GEN  
 ‘my Săn’
- d. *ʔang eda- Săn nă*  
 ang=eda= Săn nă  
 A= this=Săn nă  
 ‘this Săn of mine’

In all cases listed in (20), the case marker is present and marks the NP simply for agent case, irrespective of other elements. Characteristically, neither the demonstrative prefixes, nor the possessive pronoun/adjective in Ayeri mark case, while they do in German. The case marker thus cannot be simply left out, because the information it provides is not redundant, strictly speaking. Where it is left out, it marks the NP as topicalized and it is required, then, that the verb mark the topicalized NP’s case. The same is also true of generic nouns:

- (21) a. *veneyang*  
veney-ang  
dog-A  
'a/the dog'
- b. *eda-veneyang*  
eda=veney-ang  
this=dog-A  
'this dog'
- c. *veneyang nā*  
veney-ang nā  
dog-A ISG.GEN  
'my dog'
- d. *eda-veneyang nā*  
eda=veney-ang nā  
this=dog-A ISG.GEN  
'this dog of mine'

While it has been argued that Ayeri does not possess articles, it does possess a large variety of pronouns. These, as pro-forms, appear in complementary distribution with NPs. Since they encode morphosyntactic functions rather than semantic content, they are ideal candidates for heads of DP. DPs can be modified by those phrases which can modify nouns as well: NPs, APs and CPs, collectively referred to as XP below. The phrase structure of DPs should thus look as illustrated in (22) and (23). The XP constituent is in brackets in both cases since it is optional, that is, it is an adjunct rather than a complement. The functional annotation identifies it as such.

- (22) a.  $DP \rightarrow \begin{matrix} D' \\ \uparrow = \downarrow \end{matrix}$
- b.  $D' \rightarrow \begin{matrix} D^0 \\ \uparrow = \downarrow \end{matrix} \left( \begin{matrix} XP \\ \downarrow \in (\uparrow_{ADJ}) \end{matrix} \right)$

- (23)  $\{(\uparrow_{DP}) = \downarrow \mid (\uparrow_{GF}) = \downarrow\}$
- DP
- |
- $\uparrow = \downarrow$
- D'
- / \
- $\uparrow = \downarrow$       $\left( \begin{matrix} \downarrow \in (\uparrow_{ADJ}) \\ D^0 \quad \quad \quad XP \end{matrix} \right)$

*Personal pronouns*

The morpholexic specifications for personal pronouns are given in (24). Personal pronouns, as a functional category, are a closed class of words; the chart of personal pronouns in Ayeri is given in section 4.2.1. Since personal pronouns are pro-forms, they do not have lexical content for a predicator, but only ‘*pro*’. Pronouns distinguish all grammatical categories of nouns—number, gender and animacy, and case; they agree with their antecedents in number, gender, and animacy. In addition to these person features, pronouns of course also encode person as a deictic category. The reflexive clitic  $\text{ṣṭang}$ : *sitang*- also redefines the personal pronoun as reflexive.

(24)	...	N	(↑ PRED)	=	‘ <i>pro</i> ’
			(↑ ANIM)	=	±
			(↑ CASE)	=	{A, P, DAT, GEN, LOC, INS, CAUS}
			(↑ GEND)	=	{M, F, N, INAN}
			(↑ NUM)	=	{SG, PL}
			(↑ PERS)	=	{1, 2, 3}
			((↑ PRONTYPE)	=	REFL)

Personal pronouns, as an exception to the phrase-structure definition in (22), cannot be modified by adjectives and nominal adjuncts, however, they may still be modified by relative clauses as well as clitic quantifiers. Modification by a quantifier clitic is only possible for personal pronouns when they are free morphemes, though; modification of pronominal suffixes by quantifiers is not possible, as described in section 4.2.6—a reflexive particle is needed here to attach the quantifier to. The examples in (25) show what is not possible as compared to nouns.

- (25) a. \*pronoun + adjective:

*yāng hiro*  
*yāng hiro*  
 3SG.M.A new  
 ‘new he’

- b. \*pronoun + possessor:

*reng ayonena*  
*reng ayon-ena*  
 3SG.INAN.A man-GEN  
 ‘the man’s it’

- c. \*pronoun + instrumental complement:

*nang bariri*  
*nang bari-ri*  
 IPL.A meat-INS  
 ‘we of meat’

As mentioned above, it is possible for pronouns to be modified by relative clauses, which is illustrated by (26). In this example, the pronoun ၵၼ် *yeng* ‘she’ is modified by the relative clause ၵိမိၼ် *si mino* ‘who is happy’.

(26) pronoun + relative clause:

*yeng si mino*  
*yeng si mino*  
 3SG.F.A REL happy  
 ‘she who is happy’

### Demonstrative pronouns

The morphology of demonstrative pronouns was described in section 4.2.2. In contrast to personal pronouns, demonstrative pronouns do not mark person; a third person reference is implied like with nouns, however. Instead, they mark deixis more generally as location in space. Notably, demonstrative pronouns also lack a number distinction. As previously discussed, Ayeri distinguishes proximal (ၵၼ် *eda-*) and distal (ၵၼ် *ada-*) as well as an indefinite ‘such’ (ၵ် *da-*), which is why the feature definitions in (27) list a DEIX feature encoding *this*, *that*, and *such*, rather than a binary PROX or DIST feature.

(27) ... N (↑ PRED) = ‘*pro*’  
 (↑ DEIX) = {*this*, *that*, *such*}  
 (↑ PERS) = 3  
 (↑ ANIM) = ±  
 (↑ CASE) = {A, P, DAT, GEN, LOC, INS, CAUS}

Regarding the ability of demonstrative pronouns to be modified, it is necessary to distinguish the proximal ၵၼ် *eda-* and distal ၵၼ် *ada-* series from the indefinite ၵ် *da-* series.<sup>7</sup> The issue at hand here is that the proximal and distal demonstrative pronouns proper are not usually modified, while the indefinite one can be, as demonstrated in section 4.2.2; example (101b) from this section is repeated here as (28c) for convenience:

<sup>7</sup> Based on the absence of evidence for languages which merge *that* and *such*, Lyons (1999: 152) concludes that demonstratives are inherently definite, that is, he apparently refutes the idea that *such* is an ‘indefinite demonstrative’. However, he does not make any suggestions for a better term, which is why I will keep ‘indefinite’ here.

- (28) a. <sup>?</sup> *Sa noyang edanya tuvo.*  
           sa no=yang edanya-Ø tuvo  
           PT want=ISG.A this.one-TOP red  
           ‘I want this red one.’
- b. <sup>?</sup> *Sa noyang adanya tuvo.*  
           sa no=yang adanya-Ø tuvo  
           PT want=ISG.A that.one-TOP red  
           ‘I want that red one.’
- c. *Sa noyang danya tuvo.*  
           sa no=yang danya-Ø tuvo  
           PT want=ISG.A such-TOP red  
           ‘I want the red one.’

Related to the ability of indefinite demonstrative pronouns to be modified by adjuncts is also the formation of complex demonstratives which incorporate an adjective, both generic and possessive. To repeat (102) from section 4.2.2 for illustration:

- (29) *Sa noyang da-tuvo.*  
           sa no=yang da=tuvo.Ø  
           PT want=ISG.A such=red.TOP  
           ‘I want the red one.’

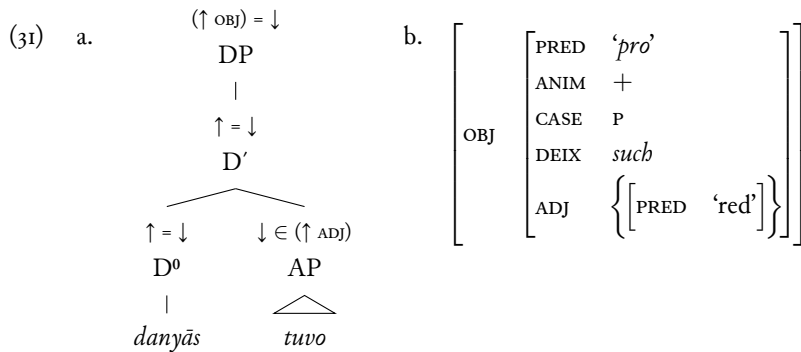
It has been argued in section 3.2.5 that  $\downarrow$  *da-* in this case is a simple clitic, as it appears in the same position as the full form  $\downarrow_{22}$  *danya*; the result is complex demonstrative form which is inflectable for case. How to represent this in terms of a feature matrix, though? For one, the adjective loses its ability to carry comparison morphology (whether it is interpreted as inflectional or clitic) when being incorporated into a demonstrative form, so it is not possible to say (30a) and the effective meaning of (30b) differs from what is intended, since  $\rightarrow$  *-vā* is interpreted in its regular, non-grammaticalized meaning here. Thus, in these cases, the demonstrative must be used in the full form (30c).

- (30) a. \**da-tuvo-vāas*  
           da=tuvo=vā-as  
           one=red=SUPL-P  
           Intended: ‘the reddest one’
- b. <sup>!</sup> *da-tuvoas-vā*  
           da=tuvo-as=vā  
           one=red-P=most/\*SUPL  
           ‘most red ones’  
           Intended: ‘the reddest one’



- c. *danyās*      *tuvo-vā*  
 danya-as      tuvo=vā  
 that.one-P    red=SUPL  
 ‘the reddest one’

The form with the incorporated adjective is basically the same as that of a noun modified by  $\downarrow$  *da-*, so it was assumed previously that the proclitic essentially acts as a nominalizer for the adjective. This could also explain why the adjective forfeits its ability to undergo comparison: comparison is not a morphological operation available to nouns, and the comparison morphemes are still loose enough to not jointly undergo derivation to a noun together with the root adjective in the way it is possible, for example, in German to form the deadjectival nouns *das Große* ‘what is big/great’ (for instance, *im Großen* ‘on a large scale’) and *das Größte* ‘the greatest thing’ from the adjective *groß* ‘big, large, great’ and its superlative form *größt-* ‘biggest, greatest’; the superlative suffix being *-(e)st*. Both c- and f-structure should thus look different for the unincorporated and the incorporated adjective, respectively. Example (31) illustrates what the c- and f-structure for the unincorporated adjective looks like.



The difference between  $\downarrow$  *da-* combined with a noun and the same combined with an adjective is, however, that with a noun, the meaning is ‘such a NOUN’, while with an adjective the meaning is not ‘such an ADJECTIVE one’, but ‘the ADJECTIVE one’. Thus, the deictic/anaphoric meaning remains, strictly speaking, which is manifest in the fact that the gender of the compound depends on that of the referred-to entity (32). For cases where the noun and the adjective are homophones—like  $\text{tuvo}$  ‘red’—the correct interpretation is dependent on context.

- (32) a. CONTEXT: an apple ( $\text{ᄡᆞᆫ ᄃᆞᆫ ᄃᆞᆫ}$  *seygo*, AN):

... *da-tuvoas*  
 da=tuvo-as  
 one=red-P.AN

‘... the red one’

- b. CONTEXT: a box ( $\text{ᄡᆞᆫ ᄃᆞᆫ ᄃᆞᆫ}$  *bin*, INAN):

... *da-tuwoley*  
 da=tuvo-ley  
 one=red-P.INAN

‘... the red one’

If we stay with the analysis that  $\text{ᄡᆞᆫ}$  *da-* is a simple clitic and thus equivalent to the full form  $\text{ᄡᆞᆫ}$  *danya* (albeit restricted in its use), the assumption stands to reason that the function embodied by  $\text{ᄡᆞᆫ}$  *danya* still forms the head of the phrase, so we still have a DP. Since  $\text{ᄡᆞᆫ}$  *da-* is no independent word, it cannot be the head of the phrase, so it must not be  $D^0$ . Since there is no other word material for nominal case marking to attach to, the adjective stem is inflected instead of the pronoun. A hypothesis about the structure of  $\text{ᄡᆞᆫ}$  *da-tuwoley* is given in (33). This is not unproblematic, though, as we will shortly see.

- (33) a. ?  $(\uparrow \text{OBJ}) = \downarrow$
- $$\begin{array}{c}
 \text{DP} \\
 | \\
 \uparrow = \downarrow \\
 \text{D}' \\
 \swarrow \quad \searrow \\
 \uparrow = \downarrow \quad \downarrow \in (\uparrow \text{ADJ}) \\
 \text{Cl} \quad \text{AP} \\
 | \quad | \\
 \text{da-} \quad \uparrow = \downarrow \\
 \quad \text{A}^0 \\
 \quad \swarrow \quad \searrow \\
 \quad \uparrow = \downarrow \quad \uparrow = \downarrow \\
 \quad \text{A}_{\text{stem}} \quad \text{A}_{\text{infl}} \\
 \quad | \quad | \\
 \quad \text{tuvo} \quad \text{-ley}
 \end{array}$$
- b. ?  $\left[ \begin{array}{c} \text{PRED} \quad \text{'pro'} \\ \text{DEIX} \quad \text{such} \\ \text{OBJ} \quad \left[ \begin{array}{c} \text{ADJ} \quad \left\{ \begin{array}{c} \text{PRED} \quad \text{'red'} \\ \text{ANIM} \quad - \\ \text{CASE} \quad \text{P} \end{array} \right\} \end{array} \right] \end{array} \right]$

One question is, whether the AVM in this case should look like the one in (31), since now it is the adjective which carries the nominal inflections. Strictly speaking, the functional annotation for the case marker should appear along with the

adjective's PRED under the ADJ relation (33b). Since LFG postulates that all semantic information is inherited upwards with an intersection operation applied at each node in a syntactic tree, DP will still contain all the relevant information in a consistent manner.

Another question is where in the tree  $\downarrow$  *da-* should appear as a clitic. A lexicalist approach like LFG favors would treat  $\downarrow$  *da-tuvoley* as a single unit in terms of syntax. The problem with this is, however, that  $\downarrow$  *da-* encodes ( $\uparrow$  PRED) = '*pro*' and the adjective encodes ( $\uparrow$  PRED) = '*red*'. Now, if they were both dominated by the same node ( $A^0$  in (33a)), that node would have clashing values for PRED. This in turn motivates the assumption that  $\downarrow$  *da-* should be represented in the c-structure at a higher level than AP, that is, adjoined to D'. Now it potentially violates lexical integrity, though. One has to wonder if it would be feasible instead to conceptualize  $\downarrow$  *da-tuvoley* as a portmanteau item with two PRED values, like (34).

$$(34) \quad * \left[ \begin{array}{c} \text{OBJ} \\ \left[ \begin{array}{ll} \text{PRED} & \left\{ \begin{array}{l} \text{'pro'}, \\ \text{'red'} \end{array} \right\} \\ \text{ANIM} & \text{—} \\ \text{CASE} & \text{P} \\ \text{DEIX} & \text{such} \end{array} \right] \end{array} \right]$$

The AVM in (34) violates the uniqueness condition, though: “Every attribute has a unique value” (Bresnan et al. 2016: 45), that is, no attribute may have more than one value assigned to it, but we just did this for PRED by assigning both '*pro*' and '*red*' to it. Thus, in order not to violate this principle, the question is whether a list of PRED attributes (like ADJ lists, or whenever coordination is involved) would be acceptable:

$$(35) \quad ? \left[ \begin{array}{c} \text{OBJ} \\ \left[ \begin{array}{l} \left[ \begin{array}{ll} \text{PRED} & \text{'pro'} \\ \text{DEIX} & \text{such} \end{array} \right] \\ \left[ \begin{array}{ll} \text{PRED} & \text{'red'} \end{array} \right] \end{array} \right] \\ \text{ANIM} & \text{—} \\ \text{CASE} & \text{P} \end{array} \right]$$

Another way to look at this problem is to analyze  $\downarrow$  *da-tuvoley* by means of inside-out functional uncertainty.<sup>8</sup> Dalrymple (2001: 144) gives an example from Warlpiri (from Nordlinger 1998: 136, attributed to Simpson 1991) which contains a noun with double case marking. This example may serve as a template for a solution

<sup>8</sup> To come back to our analogy with HTML et al., inside-out functional uncertainty is similar to traversing the DOM tree upwards in jQuery with `$(this).parent(selector).attr(key)`.

(36) *Japanangka-rlu luwa-rnu marlu pirli-ngka-rlu* [Warlpiri]  
 Japanangka-ERG shoot-PST kangaroo rock-LOC-ERG  
 'Japanangka shot the kangaroo on the rock.'

(37) a.

$$\left[ \text{SUBJ } f: \begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{OBL}_{loc} & g: \begin{bmatrix} \text{PRED} & \text{'rock'} \\ \text{CASE} & \text{LOC} \end{bmatrix} \end{bmatrix} \right]$$

b.

<i>pirli-ngka-rlu</i>	N	(↑ PRED)	= 'rock'
		(↑ CASE)	= LOC
		((OBL <sub>loc</sub> ↑) CASE)	= ERG
		(SUBJ OBL <sub>loc</sub> ↑)	

So how can we transfer this to our composite demonstrative  $\text{ၵ၄ၵ၄ၵ၄} da\text{-}tuvo\text{ley}$  ‘a/the red one’? We know that there is a lexical base which is an adjective  $\text{ၵ၄} tuvo$  which has ‘red’ for PRED. This is embedded as the ADJ to a functional head with deictic features (formerly  $\text{ၵ၄ၵ၄} danyale\text{y}$ , now reduced to a simple clitic  $\text{ၵ၄} da\text{-}$ ). The whole compound is marked for P case, which identifies it as an OBJ. Thus, I propose the f-structure and annotations in (38):

(38) a.  $\left[ \text{OBJ } f: \left[ \begin{array}{ll} \text{PRED} & \text{'pro'} \\ \text{ANIM} & \text{—} \\ \text{CASE} & \text{P} \\ \text{DEIX} & \text{such} \\ \text{ADJ} & g: \left\{ \left[ \text{PRED} \quad \text{'red'} \right] \right\} \end{array} \right] \right]$

b.	ᄒᄒᄒᄒᄒ <i>da-tuwoley</i>	N	(↑ PRED)	=	'red'
			((OBJ ↑) PRED)	=	'pro'
			((OBJ ↑) DEIX)	=	<i>such</i>
			((OBJ ↑) CASE)	=	P
			(OBJ ADJ ↑)		

Essentially, we can generate the same f-structure as in (31) above this way, since ᄒᄒᄒᄒᄒ *da-tuwoley* and ᄒᄒᄒᄒᄒᄒ *danyaley tuvo* should be functionally equivalent in spite of their different morphology. The approach with functional uncertainty permits us, for one, not to violate lexical integrity as in (33a), so the c-structure should more correctly look like in (39). Secondly, we do not have to assume that the adjective itself untypically inflects for case, at least not at a functional level. The case marker just happens to be stuck on it because it is the next best lexical base to attach it to, which is represented in the accompanying c-structure (non-branching bar levels and the empty  $D^0$  have been pruned).



### Interrogative pronouns

Like the other kinds of pronouns, interrogative pronouns are as well a closed class of words in Ayeri. The whole list of them is given in section 4.2.3. Interrogative pronouns, like demonstrative pronouns, only inflect for case, and they distinguish animacy in agreement with their referent. Since interrogative pronouns in Ayeri do not appear in clause-initial position, but *in situ*, they probably should not be analyzed as heading a CP like in English (Carnie 2013: 359–369; Dalrymple 2001: 405–408), but as heads of DP. The question word ᄒᄒᄒ *sinya* ‘who, what, which (one)’ may also serve as an adjective in cases like (40), however.

- (40) *Ang pretva kunangya sinyā?*  
 ang pret=va.Ø kunang-ya sinyā  
 AT knock=2.TOP door-LOC which  
 ‘Which door did you knock at?’

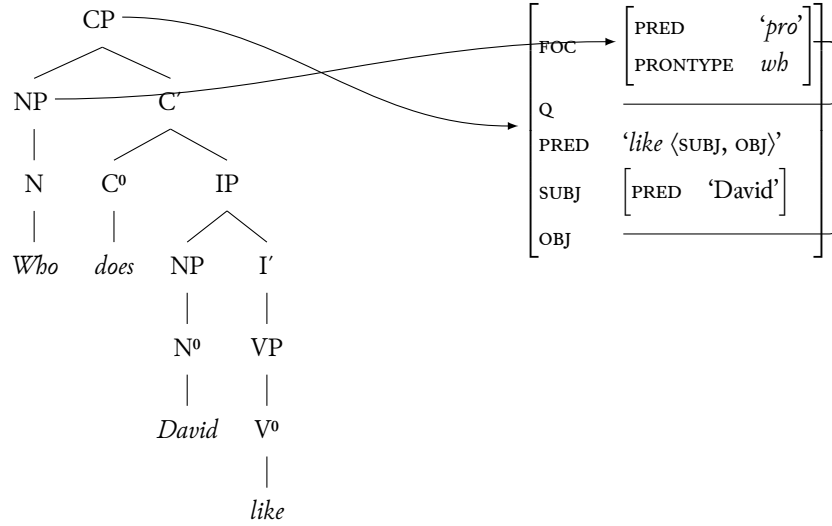
This case warrants special discussion later, as it differs from the way the majority of interrogative pronouns work. A more canonical example of ꠘꠘꠘ *sinyā* in which it acts as a pronoun proper is given in (41).

- (41) *Le tinkaya sinyāng kunang?*  
 le tinka-ya sinyā-ang kunang.Ø  
 PT.INAN open-3SG.M who-A door.TOP  
 ‘Who opened the door?’

Note also that neither ꠘꠘꠘ *sikay* ‘how (means, circumstance)’ nor ꠘꠘꠘ *simin* ‘how (way, procedure)’ can be combined with an adjective to ask about the extent of an attributive property. ꠘꠘꠘ *sikan* ‘how much, how many’ is only used for asking about quantity and cannot combine with adjectives either. Instead, questions like ‘how large’ or ‘how common’ must be phrased using (generic) nouns (42).

- (42) a. *Nahungreng mavayena sinyaley?*  
 nahung-reng mavay-ena sinyā-ley  
 size-A.INAN world-GEN what-P.INAN  
 ‘What is the size of the world?’  
 or: ‘How big is the world?’
- b. *Adareng vihay apānya sinyā?*  
 ada-reng vihay apān-ya sinyā  
 that-P.INAN common extent-LOC which  
 ‘To what extent is it common?’  
 or: ‘How common is it?’

Logically, what the question word queries for is both the implied object of the question and also new information. That is, what it asks for is the sentence’s focus. For English, Dalrymple (2001: 406) gives the following example:

(43) a. *Who does David like?*

$$b. \quad CP \rightarrow \left( \begin{array}{c} XP \\ (\uparrow \text{ FOC}) = \downarrow \\ (\uparrow \text{ FOC}) = (\uparrow \text{ QFOCUSPATH}) \\ (\uparrow \text{ Q}) = (\uparrow \text{ FOC WHPATH}) \\ (\uparrow \text{ Q PRONTYPE}) =_c wh \end{array} \right) \left( \begin{array}{c} C' \\ \uparrow = \downarrow \end{array} \right)$$

The AVM in (43a) indicates that information contained in FOC (as a discourse function) is shared with both the question particle Q and the OBJ of the clause. That is, Q is replacing the OBJ as a pronoun, and as such, embodies the FOC function with the respective properties. The phrase structure in (43b), then, tries to give a formula as general as possible for all question words in English, hence we see XP instead of NP like in example (43a), which is an example of a specific sentence containing the interrogative pronoun *who*. XP corresponds to any phrase type which can contain a question word, that is, NP, PP, AdvP, and AP (407). The rather intricate annotation for XP is due to English's fronting of the question word, which necessitates definitions to retrieve the correct corresponding information further down the tree. The annotation basically says that XP is the focus of the clause and specifies that the corresponding information must be found in a location accessible to the *wh*-word (that is, the *wh*-word must c-command it), and there is a requirement that a *wh*-word exist.

Since Ayeri does not front interrogative pronouns like English does, the functional annotations of the phrase structure rule should look a whole lot easier. Yet, however, we still need to account for Q, FOC and its associated GF sharing information. In order to give an example, let us have a look again at (41):





macy for agents and patients. Other pronouns like  $\text{ṣṣuṣ}$  *siyan* ‘where’ or  $\text{ṣṣṣṣ}$  *sikan* ‘how many’ are invariant, as described in section 4.2.3. Which morphological surface form is picked for Q if case is not specified basically depends on the GF it is linked to.

- (45) ... N      (↑ PRED)      = ‘*pro*’  
                                  ((↑ ANIM)      =  $\pm$ )  
                                  ((↑ CASE)      = {A, P, DAT, GEN, LOC, INS, CAUS})  
                                  (↑ PRONTYPE)      = *wh*

Since question pronouns stay *in situ*, it is not necessary to devise an outside-in functional uncertainty rule; the interrogative pronoun is already in the place the argument it stands in for would normally occupy. The interrogative pronoun as such wholly replaces an NP, so it is not possible to join modifiers to D’, like with personal pronouns:

- (46) a. \**Ang pengalye      sinyās      denisa?*  
           ang pengal=ye.Ø      sinyā-as      denisa  
           AT    meet=3SG.F.TOP    who-P    famous  
           ‘Whom famous did she meet?’  
       b. \**Ang sarava      siyan      veno?*  
           ang sara=va.Ø      siyan      veno  
           A    go=2.TOP    where    beautiful  
           ‘Where beautiful did you go?’

$\text{ṣṣuṣ}$  *sinya*’s other function as an adjective may come into play here together with a generic noun, however. In these cases,  $\text{ṣṣuṣ}$  *sinya* occupies the final position in the NP, that is, the first adjunct to N’:

- (47) a. *Ang pengalye      nyānas      denisa      sinyā?*  
           ang pengal=ye.Ø      nyān-as      denisa      sinyā  
           AT    meet=3SG.F.TOP    person-P    famous    which  
           ‘Which famous person did she meet?’



- b. *Ang sarava yanoya veno sinyā?*  
 ang sara=va.Ø yano-ya veno sinyā  
 AT go=2.TOP place-LOC beautiful which  
 ‘Which beautiful place did you go to?’



The interrogative pronoun asking for quantity, མཇུག་ *sikan* ‘how much, how many’, likewise acts like a nominal modifier, and only ever does—that is, unlike མཇུག་ *sinyā* it does not have a double role. It may combine with any countable or quantifiable noun:

- (48) *Sabayan keynamang sikan?*  
 saha-yan keynam-ang sikan  
 come-3PL.M people-A how.many  
 ‘How many people came?’

#### Indefinite pronouns

Indefinite pronouns (Table 4.11) cover a range of both syntactic phrases and semantic roles: they may substitute NPs, but also PPs; and they may also form various core ar-

guments (SUBJ, OBJ, OBJ<sub>recip</sub>) as well as various oblique arguments (POSS, OBL<sub>loc</sub>, OBL<sub>ins</sub>, OBL<sub>caus</sub>). Those indefinite pronouns substituting PERSON and THING decline while those encoding PLACE, TIME, MANNER, and REASON are invariant. Thus, ၵၵၵ *enya* ‘everyone/ -thing’, ၵၵၵၵၵ *arilinya* ‘somebody/-thing’, and ၵၵၵ *ranya* ‘nobody/-thing’ can be declined for all cases, while the groups around ၵၵၵ *yanen* ‘everywhere’ and ၵၵၵ *tadayen* ‘everytime’ always imply location (= LOC), the group around ၵၵၵ *arēn* ‘in every way’ implies a manner (= INS), and ၵၵၵ *yāril* ‘for some reason’ implies a reason (= CAUS).

Regarding the functional definition of indefinite pronouns, at least Dalrymple (2001) treats them as lexical items proper, that is, she does not transcribe them as [PRED ‘*pro*’] with additional functional annotations from which the surface form arises, but simply like [PRED ‘*somebody*’]. Since indefinite pronouns are not composed in a systematic way, I will treat them as lexical items proper here as well. For a discussion of a few regularities that nonetheless exist, see section 4.2.4 (p. 157). Example (49) attempts to model the c- and f-structure of a sentence containing an indefinite pronoun indicating place, ၵၵၵ *yanen* ‘everywhere’.

- (49) *Balangya yanen ang Amān.*  
 balang-ya yanen ang Amān  
 search-3SG.M everywhere A Amān  
 ‘Amān searched everywhere.’



The annotation in example (49) assumes that by its lexical meaning alone, ၵၵၵ *yanen* ‘everywhere’ is identified as the locative adverbial stated in the verb’s a-structure (a possible OBJ has been dropped from the example). ၵၵၵ *ang Amān* itself does not mark 3SG.M, but the SUBJ relation gets this information through uni-

son with the respective marking on the verb with which the subject NP proper coheres. The relevant information is implicitly contained in the NP's INDEX feature (Bresnan et al. 2016: 186–192), which is not shown in the example above. Example (50) gives a generalized list of features which indefinite pronouns may encode.

- (50) ... N (↑ PRED) = ‘...’  
 ((↑ ANIM) = ±)  
 ((↑ CASE) = {A, P, DAT, GEN, LOC, INS, CAUS})

As noted above, only those indefinite pronouns referring to persons or things decline for case, and thus also for animacy for agents and patients; for other indefinite pronouns, functional information is provided by the lexicon, as described for  $\text{ᵛᵛ} \text{yanen}$  in (49). Regarding their distribution, indefinite pronouns behave like personal pronouns in that it is not possible for them to be modified by nominal or adjectival adjuncts, at least formally:

- (51) a. \**Ang vacye enyaley leno*  
 ang vac=ye.Ø enya-ley leno  
 AT like=3SG.F.TOP everything-P.INAN blue  
 ‘She likes everything blue.’  
 b. \**Ang sarayan yāril agon.*  
 ang sara=yan.Ø yāril agon  
 AT go=3PL.M.TOP somewhere foreign  
 ‘They are going somewhere foreign.’

Instead, it is necessary to modify the indefinite pronoun with a relative clause as illustrated by (52). Alternatively, it is possible to use a generic noun instead of the indefinite pronoun, like  $\text{ᵛᵛ} \text{linya}$  ‘thing’ instead of  $\text{ᵛᵛ} \text{nya}$  ‘everyone’ and  $\text{ᵛᵛ} \text{yano}$  ‘place’ instead of  $\text{ᵛᵛ} \text{yāril}$  ‘somewhere’ in (53).

- (52) a. *Ang vacye enyaley si leno*  
 ang vac=ye.Ø enya-ley si leno  
 AT like=3SG.F.TOP everything-P.INAN REL blue  
 ‘She likes everything that is blue.’  
 b. *Ang sarayan yāril si agon.*  
 ang sara=yan.Ø yāril si agon  
 AT go=3PL.M.TOP somewhere REL foreign  
 ‘They are going somewhere that is foreign.’

It may happen occasionally that a pronominal reference is to third persons of mixed genders. English has no problem here since it does not distinguish gender in plural, so both 'John' and 'Mary' in (56a) can simply be referred to by the pronoun *them*, which is indifferent to gender (number resolution occurs). Since Ayeri's personal pronouns distinguish gender in the plural as well as in the singular, however, there needs to be a way to deal with groups whose person features cannot easily be unified,

- (55) *Ang silvyan Kan nay Maha sitanyās.*  
 ang silv-yan Kan nay Maha sitanya-as  
 AT see-3PL.M TOP Kan and Maha each.other-P

‘Kan and Maha see each other.’



which an example is given of in (56b). In addition to this, Ayeri has a two-tier system where three genders—masculine, feminine, neuter—are grouped together as animate, which is in opposition to inanimate gender, compare (57).

- (56) a. CONTEXT: John and Mary  
*I give **them** the keys.*  $\{M, SG\} \cup \{F, SG\} \implies \{\emptyset, PL\}$   
 b. CONTEXT: Ajān and Pila  
*Le ilyang tinkayye **cam.***  
 le il=yang tinkay-ye-Ø cam  
 PT.INAN give=ISG.A key-PL-TOP 3PL.M.DAT  
 ‘I give them the keys.’  $\{M, SG\} \cup \{F, SG\} \implies \{M, PL\}$



If the group referred to has already served as the controller of verb agreement, it is to be expected that the person features of whatever the verb agreement indicates is simply carried through the conversation, if the resolution is justifiable. Otherwise, a speaker will have to decide on which object or oblique pronoun to use. In either case, an animacy hierarchy operates in that animate referents outweigh inanimate ones. Mixed animate groups often default to masculine, though not in all cases. The rules which operate are the following:

1. If the conjuncts are of the same grammatical gender, use that. No gender resolution is necessary, since the person features of the conjuncts coincide.

$$\begin{array}{lcl} M & \cup & M \implies M \\ F & \cup & F \implies F \\ N & \cup & N \implies N \\ \text{INAN} & \cup & \text{INAN} \implies \text{INAN} \end{array}$$

2. If a conjunct referring to an animate referent is present, use the masculine form as the default for sexed entities; use neuter otherwise. Resolution to the animate conjunct's gender more generally is also possible.

$$\begin{array}{lcl} M \cup \text{INAN} & \implies & M \\ F \cup \text{INAN} & \implies & M \vee F \\ N \cup \text{INAN} & \implies & N \end{array}$$

3. For resolution of only animate referents, again, use the masculine form as the default. Resolution to the sexed conjunct's gender (M and F versus N) is possible here as well.

$$\begin{array}{lcl} M \cup F & \implies & M \\ M \cup N & \implies & M \\ F \cup N & \implies & M \vee F \end{array}$$

It is possible that the preference for the masculine form as a default is in part motivated by phonology, since /a/ is by far the most common vowel in Ayeri in all positions (section 1.2). Thus, it is also the least marked one, so that pronominal forms with /a/ are preferred over those with more marked vowels. Corbett (1983) makes an observation on Slovene, Polish, and, to a lesser extent, French that in these languages the first strategy in gender resolution is to resolve towards a semantically justified gender. If this fails, the second strategy is to resolve towards the form which least ambiguously marks plural. Both strategies are described as a general tendency in gender resolution (205). Since there is almost no syncretism in Ayeri's

third-person plural pronouns which would render certain forms ambiguous, the second strategy is moot.<sup>9</sup>

While coordinated NPs may be a very obvious case in which resolution occurs, there are even cases where simplex NPs trigger it. This is the case for lexical hybrids, for one. Hybrid nouns are nouns whose outward, morphological form encodes a gender different from the empirically observable gender of the denoted entity—there is a mismatch between form and meaning, grammatical and apparent gender.

Corbett (2006) quotes the Russian word *врач* *vrač* ‘(woman) doctor’, which is masculine at the level of syntax, but can be used to refer to female doctors alike at the level of semantics, hence triggering feminine agreement in more than half of examined relevant cases (158). Another example is German *Mädchen* ‘girl’, which is syntactically neuter, but semantically feminine.<sup>10</sup> Moreover, French *sentinelle* (quoted by Wechsler 2009 as an example) is a feminine noun, but often refers to a male person. English common nouns do not morphologically distinguish gender, but for instance, *child* may refer to both boys and girls, potentially triggering masculine *he* or feminine *she* rather than the syntactically required neutral *it* in anaphoric recourse.

Ayeri’s nouns are mainly distinguished by animacy; masculine and feminine only plays a role for those nouns which refer to beings, such as humans, animals, and gods (compare section 4.1). For GEND there is no cross-categorization as such, however, there is for the ANIM feature. Thus, for instance, *ḥēḥ* *subey* ‘slave’ is listed in the dictionary as inanimate even though it denotes a person. Conversely, *mimān* ‘opportunity’, *myaltan* ‘debt’, *natranga* ‘temple’, *napakaron* ‘acid’, and *piyu* ‘grain’ are all listed as animate in spite of not referring to beings, but to things, tangible or abstract. It may also be noted that it is a lot easier to find ‘miscategorized’ semantic inanimates than semantic animates. However, as noted before, Ayeri is rather generous about what it allows to be categorized as animate. Thus, not only beings count as animate, but also things closely associated with living things, such as events, concepts, or activities executed or connected to them; likewise things giving some semblance of life by growing or moving are often animate.

Of nouns referring to living beings of both sexes, there are *bayhi* ‘ruler’, *dapal* ‘boss, chief, superior’, *gan* ‘child’, *lajāy* ‘student’, *ledan* ‘friend’, and

<sup>9</sup> *teng* is used for both 3PL.F.A and 3PL.INAN.A; *tan* is used for both 3PL.F.GEN and 3PL.INAN.GEN. They all unambiguously indicate plural.

<sup>10</sup> The same goes for the German neuter *Weib* (compare Fleischer 2012: 165–166), which had been the long-time generic word for ‘woman’ (Old High German *wīb*, Middle High German *wîp*; cognate to English *wife*) before the feminine *Frau*, formerly denoting a woman of high social standing (OHG *frouwa*, MHG *vrouwe*), began to dominate in the first half of the 20th century. *Weib* is now derogatory.



ꠘꠞꠞ *sobaya* ‘teacher’, for instance. All of these should be treated as animate neuters, factually, however, agreement is usually semantic in these cases, if justified:

- (58) *Yam il-ilya badanang gan<sub>i</sub> ajamley yena<sub>i</sub>.*  
 yam il~il-ya badan-ang gan-Ø ajam-ley yena  
 DATT ITER~give-3SG.M father-A child-TOP toy-P.INAN 3SG.F.GEN

‘The child, the father gives her her toy back.’

The English translation of (58) is slightly odd in that ‘child’ is resumed by ‘her’, but in Ayeri ꠘꠞ *gan* ‘child’ is permissible as the referent of ꠘꠞ *yena* ‘her’. Since ꠘꠞꠞ *badan* ‘father’ denotes a male person, it would be odd for the feminine pronoun to be covariant with ꠘꠞꠞ *badan*—the only other permissible reading is that ꠘꠞ *yena* refers to a non-present feminine third person, the owner of the toy and the recipient not being the same person. However, as the indices show, this was not intended here; identity of the recipient and the owner was. There is no obligation for semantic agreement, so it is possible just as well to use ꠘꠞ *yona* ‘its’ here. It is also a choice of the speaker to use ꠘꠞ *gan* ‘child’ rather than ꠘꠞꠞ *lay* ‘girl’ more explicitly.

Besides hybrid nouns, there is also a class of nouns which Corbett (2006) refers to as ‘committee nouns’ after the observation that the word *committee* may trigger both singular and plural forms in agreement, with American English preferring singular agreement (*the committee has decided*), British English preferring plural agreement (*the committee have decided*), and Australian English allowing both with some preference for singular forms (212–213). Other typical committee nouns in English are *government*, *team*, and *family*. What they all have in common is that syntactically they are singular forms, but semantically they refer to a group of people. Examples of committee nouns in Ayeri are ꠘꠞꠞꠞꠞ *bayhang* ‘government’, ꠘꠞꠞꠞ *batay* ‘police’, ꠘꠞꠞꠞ *kadang* ‘committee, coalition’, and ꠘꠞꠞꠞ *pandaba* ‘family’. For committee nouns in Ayeri, canonical agreement is singular, since the body denoted by the word as such is taken as the unit of reference. This is illustrated by (59). The canonical gender of the words is thus also animate neuter.

- (59) *Ang menuyo pandaba pandāpanas yona.*  
 ang menu-yo pandaha-Ø pandāpan-as yona  
 AT visit-3SG.N family-TOP relatives-P 3SG.N.GEN

‘The family is visiting its relatives.’

Here we have animate neuter singular agreement on the verb, as well as on the possessive pronoun: both agree syntactically with ꠘꠞꠞꠞ *pandaba* ‘family’ as expected. However, Corbett (2006) points out that cross-linguistically, there is a likelihood of semantic agreement creeping in, and that different constituents are differently affected by this, so that he arrives with the gradient described in (60)—the agreement hierarchy.

(60) Agreement hierarchy (Corbett 2006: 206 ff.):

attribute > predicate > relative pronoun > personal pronoun

Corbett (2006) notes on the above chart that for “any controller that permits alternative agreements, as we move rightwards along the Agreement Hierarchy, the likelihood of agreement with greater semantic justification will increase monotonically (that is, with no intervening decrease)” (207). That is, agreement of a target according to the syntactic features of its agreement controller are becoming increasingly likely as we go from right to left, whereas agreement according to a controller’s semantic features become increasingly likely as we go from left to right. Fleischer (2012) also finds that at least with regards to agreement triggered by the German lexical hybrid *Weib* ‘woman’ as surveyed under a diachronic perspective that the distance between a controller and its target also plays some role in semantic agreement being triggered. This effect is also to be expected in Ayeri, and, according to Corbett (2006)’s agreement hierarchy, especially so for personal pronouns, which we have discussed in this subsection. Thus, (59) might be followed up with:

- (61) *Tang mino tadayya si girendtang panca.*  
 tang mino taday-ya si girend=tang panca  
 3PL.M.A happy time-LOC REL arrive=3PL.M.A finally  
 ‘They are happy when they finally arrive.’

In (61), suddenly, anaphora relating to ጥብሩ *pandaba* ‘family’ before—the subject pronoun ጥንጥ *tang* ‘they’ as well as the pronominal suffix of the verb ጥንጥጥ *girendtang* ‘they arrive’—switch to animate masculine plural; semantic agreement has been triggered, and gender and number resolution thus occur. The pronoun of choice is masculine, since this is the default gender for mixed-gender groups of living beings, and plural, since ጥብሩ *pandaba* denotes a group of people.

However, no statistics for Ayeri have been compiled on this issue to date, and least of all across a variety of literary genres and across different simulated historical stages of the language. Another difficulty in this regard is that all existing texts in Ayeri are translations from English or German, and conceived as more or less carefully crafted written texts. Any statistics on the effect of distance on semantic agreement (if it can at all be found) will thus likely be skewed as compared to the output of native speakers, both written and spoken.

### 5.1.3 Nominal clitics

For nominal clitics, we have already seen how preposed case markers work; some nominal clitics are also described in section 3.2.5 in a way to establish why they are clitics rather than affixes or words. In the following subsections, we will have

a closer look at the morphosyntax of each group of clitics which can interact with nouns, with regards to nouns specifically for those clitics which can interact with more than this one part of speech.

#### Demonstrative prefixes

For one, there is the series of demonstrative prefixes—or rather, deictic proclitics: ມີ *eda-* ‘this’, ພີ *ada-* ‘that’, and ດີ *da-* ‘such’. In section 3.2.1 (p. 67), it was reasoned that in Ayeri, to capture all three clitics, a feature DEIX with values *this*, *that*, and *such* should be assumed in the place of PROX, since [PROX ±] does not apply to ດີ *da-*:

- (62) a. *eda-nangās*  
            $\text{eda}=\text{nanga-as}$   
            $\text{this}=\text{house-P}$   
           ‘this house’
- $$\left[ \begin{array}{c} \text{OBJ} \\ \left[ \begin{array}{cc} \text{PRED} & \text{'house'} \\ \text{ANIM} & + \\ \text{CASE} & \text{P} \\ \text{DEIX} & \text{this} \end{array} \right] \end{array} \right]$$
- b. *ada-veneyya*  
        $\text{ada}=\text{veney-ya}$   
        $\text{that}=\text{dog-LOC}$   
       ‘at that dog’
- $$\left[ \begin{array}{c} \text{OBL}_{\text{loc}} \\ \left[ \begin{array}{cc} \text{PRED} & \text{'dog'} \\ \text{CASE} & \text{LOC} \\ \text{DEIX} & \text{that} \end{array} \right] \end{array} \right]$$
- c. *da-mebirena*  
        $\text{da}=\text{mehir-ena}$   
        $\text{such}=\text{tree-GEN}$   
       ‘of such a tree’
- $$\left[ \begin{array}{c} \text{POSS} \\ \left[ \begin{array}{cc} \text{PRED} & \text{'tree'} \\ \text{CASE} & \text{GEN} \\ \text{DEIX} & \text{such} \end{array} \right] \end{array} \right]$$

The functional annotations of the deictic proclitics are thus, very straightforwardly, given in (63). Demonstrative clitics cannot be combined with pronouns of any kind, only with nouns; the combination of ດີ *da-* with a possessive pronoun does not result in the meaning of *\*such my*, *\*such your*, etc., but it derives independent possessive pronouns such as *mine*, *yours*, etc.

- (63) ...            Cl            (↑ DEIX) = {*this*, *that*, *such*}

#### Likeness prefix *ku-*

There is furthermore a clitic expressing likeness and resemblance, ກຸ *ku-* ‘like (a)’, which precedes common nouns, and which precedes proper nouns if no overt case marker is present and follows them if there is. In all three cases, this might be conveniently interpreted as the noun having an adverbial modifier meaning ‘like’.

- (64) a. *ku-ayonang*  
 ku=ayon-ang  
 like=man-A  
 'like a man'
- b. *ku-Kaman.Ø*  
 ku=Kaman.Ø  
 like=Kaman.TOP  
 'like Kaman'
- c. *ang Apitu-ku*  
 ang=Apitu=ku  
 A= Apitu=like  
 'like Apitu'
- $$\left[ \begin{array}{c} \text{SUBJ} \left[ \begin{array}{c} \text{PRED} \text{ 'man'} \\ \text{ANIM} + \\ \text{CASE} A \\ \text{ADJ} \left\{ \left[ \text{PRED} \text{ 'like'} \right] \right\} \end{array} \right] \end{array} \right]$$
- $$\left[ \begin{array}{c} \text{TOP} \left[ \begin{array}{c} \text{PRED} \text{ 'Kaman'} \\ \text{ADJ} \left\{ \left[ \text{PRED} \text{ 'like'} \right] \right\} \end{array} \right] \end{array} \right]$$
- $$\left[ \begin{array}{c} \text{SUBJ} \left[ \begin{array}{c} \text{PRED} \text{ 'Apitu'} \\ \text{ANIM} + \\ \text{CASE} A \\ \text{ADJ} \left\{ \left[ \text{PRED} \text{ 'like'} \right] \right\} \end{array} \right] \end{array} \right]$$

In cases like (64c), if there are adjectives following the noun, the clitic actually follows the whole NP like the 's possessive clitic in English, at least from a logical point of view. Phonetically, however, the enclitic leans on the last word in the phrase. Regarding constituency, it is possible to replace all of *ang Apitu puti* with *ၵၵၵၵၵၵ danyāng* 'such one' and to still modify that by *ၵၵၵ ku-*, as illustrated by (65).

- (65) a. *ang Apitu puti-ku*  
 ang=Apitu puti=ku  
 A= Apitu diligent=like  
 'like diligent Apitu'
- b. *ku-danyāng puti*  
 ku=danya-ang puti  
 like=such.one-A diligent  
 'like the diligent one'
- c. *ku-danyāng*  
 ku=danya-ang  
 like=such.one-A  
 'like such one'



Regarding the examples in (65), we cannot simply state that the clitic should be adjoined to a bar level higher than the adjective in (66) even though the replacement tests point into this direction, since  $\text{ku-}$  cannot stand on its own as an independent word and LFG requires lexical integrity. Thus, we need to invoke the inside-out functional uncertainty principle again to associate the information provided by  $\text{-ku}$  with the right spot in our AVM: postposed  $\text{-ku}$  needs to feed not into the AP's ADJ list, but into that of the superordinate GF (or DF, for that matter). Example (67) furthermore demonstrates how the positioning variants of  $\text{ku-}$  and  $\text{-ku}$  differ in meaning and grammaticality.

- (67) a. *[ang Apitu puti ]-ku*  
           ang=Apitu puti =ku  
           A= Apitu diligent =like  
           'like diligent Apitu'
- b. *ang Apitu ku-puti*  
           ang=Apitu ku=puti  
           A= Apitu like=diligent  
           'diligent-like Apitu'
- c. \**ang Apitu [puti-ku]*  
           ang=Apitu puti=ku  
           A= Apitu diligent=like  
           Expected: 'like diligent Apitu'

In other words,  $\text{puti-ku}$  and  $\text{ku-puti}$  are also distinguished by the lack of postposed  $\text{-ku}$ 's possibility to modify the adjective, since adjectives only take  $\text{ku-}$  as a proclitic, but never as an enclitic. Thus, the functional difference is also reflected in a difference in the morphological distribution of the clitic. Regarding functional annotations—as already suggested by section 66—for the two kinds of  $\text{ku}$  we can posit:

- (68) a.  $\text{ku-}$  Cl ( $\uparrow$  ADJ PRED) = 'like'  
 b.  $\text{-ku}$  Cl (( $x \uparrow$ ) ADJ PRED) = 'like'  
 ( $x$  ADJ  $\uparrow$ )

Proclitic  $\text{ku-}$  in (68a) has a straightforward definition: the current phrase's ADJ is filled with an instance of [PRED 'like']. As described above, enclitic  $\text{-ku}$  in (68b), on the other hand, requires a slightly more complicated definition, since it does not feed into the ADJ list of the lexical head it is adjacent to, but into that of the phrase superordinate to the modifier it leans on phonetically. There thus is to be some phrase  $x$  which has an ADJ attribute which we can fill in with [PRED 'like'].

Regarding the ability of nominal clitics to combine with pronouns,  $\text{ku-}$  is probably most versatile in that it can combine with personal pronouns, as well as with demonstrative and indefinite pronouns, and the reciprocal pronoun  $\text{sitanya}$ —but not with all interrogative pronouns and the various forms of the relativizer  $\text{si}$ .

- (69) a. *Eng tangara ku-yes.*  
 eng tang=ara.Ø ku=yes  
 AT.INAN hear=3SG.INAN.TOP like=3SG.F.P  
 'It sounds like her.'
- b. *Le surpeng tadoy ku-adanya.*  
 le surp=reng tadoy ku=adanya.Ø  
 P.INAN seem=3SG.INAN.A never like=that.one-TOP  
 'Like that it has never seemed.'
- c. *Miratang ku-enyāng.*  
 mira=tang ku=enya-ang  
 do=3PL.M.A like=everyone-A  
 'They do it like everyone.'
- d. *Ang silvyan ku-sitanyās.*  
 ang silv=yan.Ø ku-sitanya-as  
 AT see=3PL.M.TOP like=each.other-P  
 'They look like each other.'

- e. *Ang silvya ku-sinyās?*  
 ang silv=ya.Ø ku=sinya-as  
 AT see=3SG.M.TOP like=who-P  
 ‘Who does he look like?’
- f. \**Girenjāng ku-sitaday?*  
 girend=yāng ku=sitaday  
 arrive=3SG.M.A like=when  
 ‘Like when does he arrive?’
- g. \**Adareng ayonas ku-si silvyāng.*  
 ada-reng ayon-as ku=si silv=yāng  
 that-A.INAN man-P like=REL see=3SG.M.A  
 Intended: ‘That is the man who he looks like.’

#### Inspecificity prefix *mə-*

While demonstrative elements make NPs more specific in selecting a particular specimen out of a group of entities, the proclitic 𑀓: *mə-* does the opposite: it expresses that the speaker refers to any representative of a group entities, not a specific one, as in (70). The functional definition for 𑀓: *mə-* should thus look as given in (71): the NP it is added to is neither definite nor specific reference, the speaker means just any representative of the kind the nominal head specifies.



- (71) 𑀓: *mə-* Cl (↑ DEF) = —  
 (↑ SPECIFIC) = —

Since according to Lyons (1999: 152), demonstratives are inherently definite—and their reference is specific, too—the deictic proclitics 𑀓: *eda-* ‘this’, 𑀓: *ada-*

‘that’, and 𐌆 *da-* ‘such’ are in complementary distribution with the inspecificity marker, as demonstrated by (72): neither combination of 𐌆 *mə-* and 𐌆 *eda-* in (72cd) results in an ungrammatical sentence because 𐌆 *eda-* encodes [DEF +, SPECIFIC +] while 𐌆 *mə-* encodes [DEF –, SPECIFIC –]. Attempting to assign opposing values to the same feature (DEF and SPECIFIC, respectively) must fail, since these cannot be united at the functional level.

- (72) a. *Ang nakasyon nibanye eda-mebirya.*  
 ang nakas-yon nihan-ye-Ø eda=mehir-ya  
 AT grow-3PL.N fruit-PL-TOP this=tree-LOC  
 ‘Fruits are growing on this tree.’
- b. *Ang nakasyon nibanye mə-mebirya.*  
 ang nakas-yon nihan-ye-Ø mə=mehir-ya  
 AT grow-3PL.N fruit-PL-TOP whichever=tree-LOC  
 ‘Fruits are growing on whichever tree.’
- c. \**Ang nakasyon nibanye mə-eda-mebirya.*  
 ang nakas-yon nihan-ye-Ø mə=eda=mehir-ya  
 AT grow-3PL.N fruit-PL-TOP whichever=this=tree-LOC
- d. \**Ang nakasyon nibanye eda-mə-mebirya.*  
 ang nakas-yon nihan-ye-Ø eda=mə=mehir-ya  
 AT grow-3PL.N fruit-PL-TOP this=whichever=tree-LOC

The inspecificity proclitic 𐌆 *mə-* cannot commonly be combined with pronouns, since personal pronouns as well as demonstrative pronouns have a definite reference; combining the clitic with an indefinite pronoun would be redundant. With interrogative pronouns it is feasible to use 𐌆 *mə-* as an intensifier, though without the vulgar tone of the English translation given in the following example:

- (73) *Amangreng mə-simin?*  
 amang=reng mə=simin  
 happen=3SG.INAN.A whichever=how  
 ‘How the fuck did that happen?’

#### Quantifying clitics

As described in section 3.2.5 (p. 94 ff.) and section 4.6.2, Ayeri has a number of suffixed—that is, enclitic—quantifiers which can attach to inflected nouns. These, along with their non-clitic counterparts, are modifiers and are thus part of an NP’s list of adjuncts, ADJ, along with adjectives and nominal adjuncts. An example of this is given in (74), using the clitic quantifier 𐌆 *-ikan* ‘much, many; very’. Since



the quantifier already lexically expresses a multitude in this case, the noun does not carry redundant plural marking.

- (74) a. *Sa tahayeng bisuay-ikan ban*  
 sa taha=yeng bisuay.Ø=ikan ban  
 PT have=3SG.F.A idea.TOP=many good

‘She has many good ideas.’

- b. 
$$\left[ \begin{array}{c} \text{TOP} \\ \text{OBJ} \end{array} \left[ \begin{array}{l} \left[ \begin{array}{l} \text{PRED} \text{ ‘idea’} \\ \text{ANIM} + \\ \text{CASE} \text{ P} \end{array} \right] \\ \text{ADJ} \left\{ \begin{array}{l} \left[ \begin{array}{l} \text{PRED} \text{ ‘many’} \end{array} \right] \\ \left[ \begin{array}{l} \text{PRED} \text{ ‘good’} \end{array} \right] \end{array} \right\} \end{array} \right] \right]$$

Clitic quantifiers as well can combine at least with independent personal pronouns especially in the plural, and demonstrative pronouns. A clitic quantifier can also still be recognized in the interrogative pronoun ꤁꤁꤁ *sikan* ‘how many’, although here it is incorporated into the pronoun itself, that is, there is no productive combination of interrogative pronouns and quantifiers, with the exception of ꤁꤁꤁ *sinya* ‘who’. With indefinite pronouns, quantifiers are somewhat redundant in some combinations, though it is feasible in this regard to use them for emphasis. The relativizer ꤁ *si* and its declined forms cannot be quantified; the simple and most common form of the relativizer, ꤁ *si*, is also unstressed, which makes it a bad host for a clitic. Combinations of pronouns and quantifying clitics are illustrated in (75).

- (75) a. *Ang koronay tas-ikan.*  
 ang koron=ay.Ø tas=ikan  
 AT know=1SG.TOP 3PL.M.P=many  
 ‘I know many of them.’
- b. *Kalam adanyana-ikoy*  
 kalam adanya-na=ikoy  
 true that.one-GEN=not.much  
 ‘Little about that is true.’
- c. *Yomāra sinyareng-ma?*  
 yoma-ara sinyar-eng=ma  
 exist-3SG.INAN what-A.INAN=enough  
 ‘What is there enough of?’

- d. *Ang il-ye Ø= Pada enyaley-kay cam.*  
 ang il-ye Ø= Pada enya-ley=kay cam  
 A give-3SG.F TOP=Pada everything-P.INAN=a.little 3PL.M.DAT  
 ‘Pada gave them a little of everything.’
- e. *Enyāng-ben siyan?*  
 enya-ang=hen siyan  
 everyone-A=all where  
*Literally:* ‘Where is all of everyone?’
- f. \**Le inttang piyu si-ma ya yomareng bukuno.*  
 le int=tang piyu-Ø si=ma ya yoma-reng bukuno-Ø  
 P.INAN buy=3PL.M.A grain-TOP REL=enough LOCT exist=3SG.INAN.A storage-TOP  
*Intended:* ‘They are buying the grain enough of which is in the storage’

## 5.2 Adjective and adverb phrases

Adjectives and adverbs in Ayeri are largely similar in that they can both be modified by adverbs like *very*, and they both modify heads: adjectives modify nouns, adverbs modify everything else. Carnie (2013: 51) urges his readers to think about whether it is sensible to distinguish between the two categories, so let us focus in this section on structural similarities and dissimilarities between the two, as well as their distribution as morphemes.

### 5.2.1 Adjective phrases

As described in the previous section, APs are usually found as adjuncts of NPs or DPs, where they describe properties of these nominal elements. Adjectives are likewise commonly found as open complements (XCOMP) in equative statements, which will be dealt with in section 5.4.4. Possessive pronouns can be used as adjective-like modifiers as well, though they are probably still better classified as DP heads, since they are functional morphemes. Possessive adjectives thus also do not share all the morphological properties of adjectives, for instance, they cannot be compared (see section 4.2.1, p. 147), also they can only be instanced once. In other words, it is not possible to modify the same noun with multiple possessive morphemes, which means that they are not part of ADJ. The phrase-structure rule in (76) and the c-structure tree in (77) show how an AP is constructed.

- (76) a.  $AP \rightarrow \begin{matrix} A' \\ \uparrow = \downarrow \end{matrix}$
- b.  $A' \rightarrow \begin{matrix} A^0 \\ \uparrow = \downarrow \end{matrix} \left( \begin{matrix} XP \\ (\uparrow \text{ GF}) = \downarrow \end{matrix} \right)$



Adjective phrases have an adjective as their lexical head. This head may be extended by modifiers adjoined to A'; A' repeats for the adjunction of multiple modifiers. Since modifiers follow their heads here as well, APs are also a right-branching constituent. Modifiers of adjectives are subsumed under the label XP here, which here stands for AdvP, NP, DP, and CP. An example of each phrase type modifying an adjective is given in (78).

- (78) a. adjective + AdvP adjunct:

*Adareng bisuayas sadayo kalam.*  
 ada-reng bisuay-as sadayo kalam  
 that-A.INAN idea-P crazy truly

'That is a truly crazy idea.'

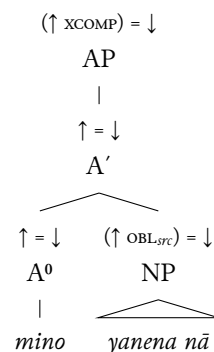


- b. adjective + NP complement:

*Yang mino yanena nā.*  
 yang mino yan-ena nā  
 ISG.A happy son-GEN ISG.GEN

‘I am happy about my son.’

XCOMP	[	PRED	'happy <OBL <sub>src</sub> >'		[	PRED	'son'		[	CASE	GEN		[	PRED	'pro'		[	CASE	GEN		[	NUM	SG		[	PERS	I		]	]	]	]	]	]

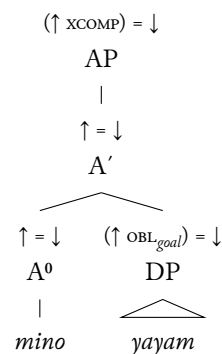


- c. adjective + DP complement:

*Yang mino yayam.*  
 yang mino yayam  
 ISG.A happy 3SG.M.DAT

‘I am happy for him.’

XCOMP	[	PRED	'happy <((↑ OBL <sub>goal</sub> ))>'		[	PRED	'pro'		[	CASE	DAT		[	NUM	SG		[	PERS	2		]	]	]	]	]	]	]	]	]	]	]	]

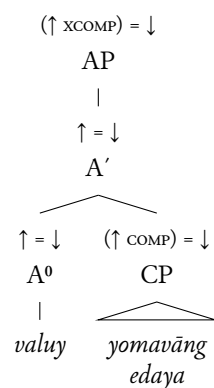


- d. adjective + CP complement:

*Yang valuy yomavāng edaya.*  
 yang valuy yoma=vāng edaya  
 ISG.A glad exist=2.A here

‘I am glad you are here.’

XCOMP	[	PRED	'glad <((↑ COMP))>'		[	PRED	'exist <...>'		[	SUBJ	[...]		[	OBL <sub>loc</sub>	[...]		]	]	]	]	]	]	]	]	]	]	]	]	]	]	]	]	]



Example (78a) gives the *f*- and *c*-structures for the adjective phrase to show that ADJ may be recursive: an adjective which serves as one of many modifiers to a noun can itself be modified by adverbs. Likewise, an adjective–adverb combination can be complemented by an NP, as shown in (78c). Especially in (78b) and (78c) we can see Ayeri’s propensity for using cases with complements where English would use prepositions. Thus, in (78b), ‘about’ is expressed by putting the nominal complement in the genitive case: the NP complement expresses the source by which the experiencing subject becomes happy. This, however, should not be conflated with a possessor, POSS, but should be labeled separately as an oblique complement, *OBL<sub>src</sub>*. Similarly, the recipient of the subject’s happiness appears as an NP complement in the (ethical) dative in (78c). Instrumental and causative NP complements instead of PPs may be found as well.

Since in LFG, empty *X'* nodes are collapsed, LFG’s version of X-bar theory does not strictly distinguish between complements and adjuncts (Bresnan et al. 2016: 127, footnote 52); the functional annotation provides information about whether a sister node to *X<sup>0</sup>* is a complement or an adjunct instead. How do we know that the extensions to the adjective in (78) are complements rather than adjuncts? X-bar theory posits that complements and adjuncts are on different *X'* branches—a complement is a sister to the head whereas an adjunct is, so to speak, its cousin:



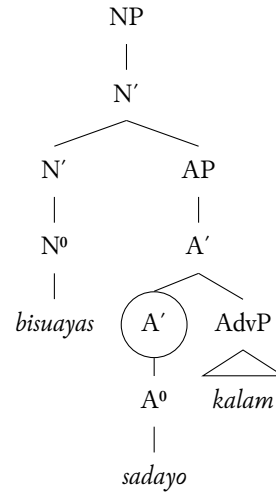
How does this hypothesis fare when applying constituency tests, though? With regards to complements and adjuncts to NP, Carnie (2013) explains, “Since complements are sisters to *X* and not *X'*, they cannot stand next to the word *one*. Adjuncts, by definition, can” (182). To target *A'* nodes, we need something corresponding to English *so*, which we find in Ayeri: *da-* ‘thus, so, such’. Since adverbs are adjuncts, it is possible to replace the adjective *sadayo* ‘crazy’ in (80a) with *da-* in (80b); the *A'* in (80a) being targeted is the sister to AdvP.

Continuing this replacement test for the other example sentences from (78), we can see that the outcomes are ungrammatical; the NPs are dependent on their lexical head, which cannot be omitted or replaced by a pro-form. The idiomatic English translations in (81) somewhat conceal this.<sup>11</sup> The literal translations try to

<sup>11</sup> At least to me, as a non-native speaker of English, these translations do not sound too odd if

- (80) a. *bisuayas sadayo kalam*  
 bisuay-as sadayo kalam  
 idea-P crazy truly  
 ‘a truly crazy idea’

- b. *Da-kalam bisuayang.*  
 da=kalam bisuay-ang  
 so=truly idea-A  
 ‘The idea is truly so.’



show that the complements are more tightly integrated into the sentence in Ayeri than in the idiomatic translations in an attempt to convey how they must seem non-sensical to an Ayeri speaker.

- (81) a. \**Yang da-yanena nā.*  
 yang da=yan-ena nā  
 ISG.A SO=SON-GEN ISG.GEN  
 ‘I am so about my son.’  
*Literally:* ‘I am my son’s so.’
- b. \**Yang da-yayam.*  
 yang da=yayam  
 ISG.A SO=3SG.M.DAT  
 ‘I am so for him.’  
*Literally:* ‘I am him so.’
- c. \**Yang da-yomavāng edaya.*  
 Yang da=yoma=vāng edaya  
 ISG.A SO=exist=2.A here  
 ‘I am so that you are here.’

context is taken into consideration. Carnie (2013: 181) gives an example \**the one of poems with a red cover*, referring to a book. His point is that *of poems* is a complement of *the book*, so *the book* cannot be replaced by a pro-form like *one*. However, he notes that at least some native English speakers find this example acceptable. Personally, I do not take offense either, but I am not a native speaker.

Further complications arise with regards to complementation in that non-clitic quantifiers like *an̄kyu* ‘really’ usurp the complement position so that an adjective’s complement XP appears as an adjunct. This is possibly due to syntactic weight (Wechsler 2009). Example (82) attempts to illustrate the c-structure for a displaced CP; the arrow shows the apparent movement from complement to adjunct.

(82) *Yang valuy an̄kyu, [yomavāng edaya].*

yang valuy an̄kyu yoma=vāng  
1SG.A glad really exist=2.A

‘I am really glad you are here.’



In LFG, however, there are no transformations by design; instead, c-structure is taken to be base-generated while acknowledging that different languages may have different phrase-structure plans, and the functional structure is what creates a sense of cohesion. Thus, even though the CP in (82) appears dislocated, the information is still found in the ‘right’ place in f-structure independent of the diverging c-structure, compare the AVM in (83), where the adjective is given as having an oblique nominal complement, and what surfaces as the CP adjunct is annotated accordingly as COMP, not as ADJ, even though in c-structure it appears in the position of an adjunct.

As pointed out in section 4.3, Ayer’s adjectives inflect very little, since there is no agreement morphology. However, it is possible for adjectives to be compared and to be negated by means of morphology. This is reflected in the functional

$$(83) \left[ \begin{array}{c} \text{XCOMP} \\ \left[ \begin{array}{c} \text{PRED} \quad \text{'glad' } \langle (\uparrow \text{COMP}) \rangle \\ \text{ADJ} \quad \left[ \begin{array}{c} \text{PRED} \quad \text{'really'} \\ \text{COMP} \quad \left[ \begin{array}{c} \text{PRED} \quad \text{'exist' } \langle (\uparrow \text{SUBJ}), (\uparrow \text{OBL}_{loc}) \rangle \\ \text{SUBJ} \quad \left[ \begin{array}{c} \text{PRED} \quad \text{'pro'} \\ \text{CASE} \quad \text{A} \\ \text{PERS} \quad 2 \end{array} \right] \\ \text{OBL}_{loc} \quad \left[ \begin{array}{c} \text{PRED} \quad \text{'here'} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]$$

annotations given in (84). The features COMPAR and NEG appear in brackets here, since they do not apply to every adjective: adjectives normally appear in the positive in both regards, comparison and polarity, and are morphologically unmarked in these cases.

$$(84) \quad \dots \quad \text{A} \quad \begin{array}{l} (\uparrow \text{PRED}) \quad = \quad \text{'...'} \\ ((\uparrow \text{COMPAR})) \quad = \quad \{\text{COMP}, \text{SUPL}\} \\ ((\uparrow \text{NEG})) \quad = \quad + \end{array}$$

Examples of the different ways an adjective may be morphologically marked and their respective representation as an AVM are given in (85).

$$(85) \quad \begin{array}{ll} \text{a.} & \begin{array}{l} \text{nake-vā} \\ \text{nake=vā} \\ \text{tall=sUPL} \\ \text{'tallest'} \end{array} & \left[ \begin{array}{c} \text{PRED} \quad \text{'tall'} \\ \text{COMPAR} \quad \text{SUPL} \end{array} \right] \\ \\ \text{b.} & \begin{array}{l} \text{mingoy} \\ \text{ming-oy} \\ \text{capable-NEG} \\ \text{'incapable'} \end{array} & \left[ \begin{array}{c} \text{PRED} \quad \text{'capable'} \\ \text{NEG} \quad + \end{array} \right] \\ \\ \text{c.} & \begin{array}{l} \text{pasīsoy-eng} \\ \text{pasīsa-oy=eng} \\ \text{interesting-NEG=COMP} \\ \text{'more uninteresting'} \end{array} & \left[ \begin{array}{c} \text{PRED} \quad \text{'interesting'} \\ \text{COMPAR} \quad \text{COMP} \\ \text{NEG} \quad + \end{array} \right] \end{array}$$

As described before (section 4.3.1), the morphemes used for synthetic comparison of adjectives are grammaticalized clitics literally meaning 'rather, more' (ཤར་ -eng) and 'most' (ཤར་ -vā) as lexical quantifiers. With adjectives, there is no clear-cut line between their functional and their lexical use. I have analyzed them here



as functional, since they may be interpreted as such, depending on context. As we have observed before (section 4.6.2), quantifier clitics may modify adjectives like any other adverbial modifiers, except that their surface form is clitic rather than free. Quantifying adverbs, both enclitic and free, thus also find themselves in ADJ.

- (86) a. *luyu-mas*  
           *luyu=mas*  
           strange=kind.of  
           ‘kind of strange’
- $$\left[ \begin{array}{l} \text{PRED} \quad \text{'strange'} \\ \text{ADJ} \quad \left\{ \left[ \text{PRED} \quad \text{'kind of'} \right] \right\} \end{array} \right]$$
- b. *valuy ipan.*  
      *valuy ipan*  
      glad extremely  
      ‘extremely glad’
- $$\left[ \begin{array}{l} \text{PRED} \quad \text{'glad'} \\ \text{ADJ} \quad \left\{ \left[ \text{PRED} \quad \text{'extremely'} \right] \right\} \end{array} \right]$$

### 5.2.2 Adverb phrases

Adverbs, as (86b) shows, can easily be converted from adjectives. Thus, *ṣṣṣ ipan* ‘extreme’, which is normally an adjective, is used there in an adverbial way, meaning ‘extremely’. The word stays the same, however: *ṣṣṣ ipan*, without a derivative affix akin to English *-ly* or French *-ment*. Since adverbs and adjectives are largely similar in that they provide additional information about the nature or circumstance of a noun or another part of speech, Carnie (2013) poses the question whether adjectives and adverbs should be better analyzed as being part of the same category. He reasons that

Both Adj and Adv can be modified by the word *very*, and they both have the same basic function in the grammar—to attribute properties to the items they modify. In fact, the only major distinction between them is syntactic: Adjectives appear inside NPs, while adverbs appear elsewhere. (51)

Adjectives and adverbs are in complementary distribution, which, he writes, would normally be taken as evidence that these two things are of the same category. In fact, the only reason Carnie (2013) adduces for keeping the two categories apart is “because they are familiar to most people”, and he prompts the reader to consider that uniting them in a single supercategory “might provide a better analysis and might be better motivated scientifically” (51). Bresnan et al. (2016: 126) also classify both adjectives and adverbs as heads of AP with reference to Emonds (1988). As described in section 4.6, the only morphology attributive adverbs take in Ayerí is comparison morphology and negation. This is the same as with adjectives indeed, hence the functional specifications appear equal:

- (87) ...            A            ( $\uparrow$  PRED)        =    ‘...’  
   (( $\uparrow$  COMPAR)    =    {COMP, SUPL})  
   (( $\uparrow$  NEG)        =    +)

If we look at the phrase structure (88) and the c-structure (89) for adverbs, however, there is a slight difference in that adverbs cannot serve as xCOMP in equative statements; they also can only be modified by other adverbs, but not by NP, DP, or CP.<sup>12</sup> On the other hand, adjectives are restricted to nominal contexts whereas adverbs may modify any other lexical category: verbs, adjectives, prepositions, as well as other adverbs.

- (88) a.    AdvP  $\rightarrow$  Adv'  
    $\uparrow = \downarrow$   
       b.    Adv'  $\rightarrow$  Adv<sup>0</sup>     $\left( \begin{array}{c} \text{AdvP} \\ \downarrow \in (\uparrow \text{ADJ}) \end{array} \right)$   
    $\uparrow = \downarrow$



Example (90) gives examples of modifiers of adverbs. These modifiers are often quantifying adverbs, whether clitic ones or free ones. Since adjectives and adverbs are not distinguished by morphology, the heads of both phrases, रक्षक *vakṣa* ‘careful’ and साधारण *bita* ‘ordinary, normal’, may be interpreted as well as adjectives, depending on context.

<sup>12</sup> For adjectives, compare (76), (77), and (84) in section 5.2.1.

- (90) a. *bita ikan-ikan*  
           bita       ikan.ikan  
           ordinarily completely  
           ‘completely ordinarily’

$$\left[ \text{ADJ} \left\{ \left[ \begin{array}{ll} \text{PRED} & \text{'ordinarily'} \\ \text{ADJ} & \left\{ \left[ \text{PRED} \text{'completely'} \right] \right\} \end{array} \right] \right\} \right]$$



- b. *vakisoy -ma*  
      vakisa-oy   =ma  
      carefully-NEG=enough  
      ‘not carefully enough’

$$\left[ \text{ADJ} \left\{ \left[ \begin{array}{ll} \text{PRED} & \text{'carefully'} \\ \text{NEG} & + \\ \text{ADJ} & \left\{ \left[ \text{PRED} \text{'enough'} \right] \right\} \end{array} \right] \right\} \right]$$



For the clitic  $\text{ku}$  ‘like’ described above (68b), it was necessary to make a separate rule for the prefixed and suffixed version. This, however, is not necessary for clitic quantifiers, since they attach immediately to the head they modify rather than to the last word of the AdvP. Besides their diverging morphological distribution, it may thus be assumed that the functional scheme given in (87) also holds for enclitic quantifiers. Quantifiers appear to be able to be modified in turn, as illustrated by (76) from section 3.2.5 (compare p. 98), which is repeated again here, abbreviated;  $\text{ikan}$  ‘much, many, very’ is modified in this case by  $\text{kagan}$  ‘far too’ to convey the meaning ‘far too many’.

- (91) *keynam -ikan kagan*  
      keynam-Ø =ikan   kagan  
      people-TOP=many far.too  
      ‘far too many people’

Besides the parenthetical insertion tests on  $\text{kagan}$  in section 3.2.5, the following examples show that  $\text{kagan}$  also cannot swap places with other nominal modifiers without a change in meaning (92a), that  $\text{kagan}$  does not work when

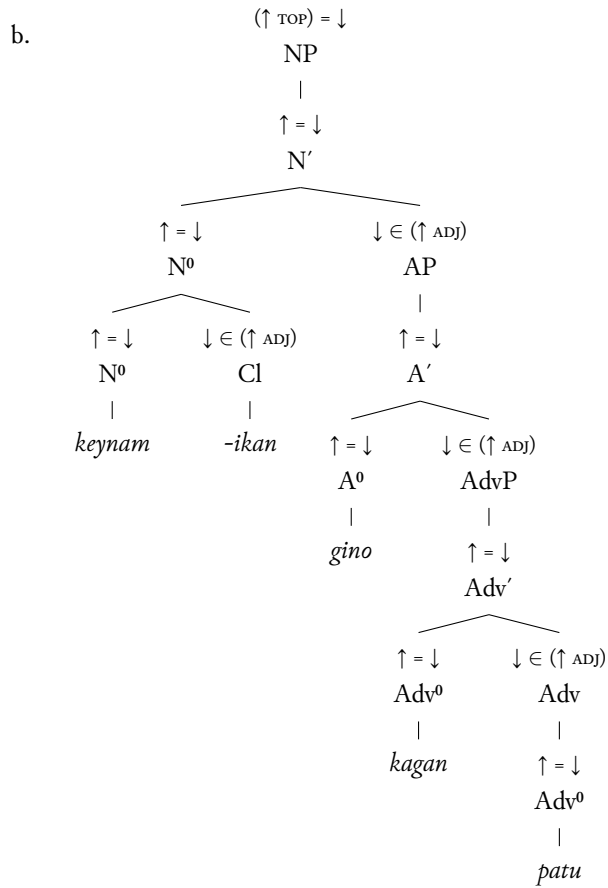
the rest of the NP is replaced by an anaphora (92b), and that 𑜁𑜪𑜫𑜰𑜫 -*ikan* cannot be replaced by an anaphora either (92c). It should be clear that 𑜁𑜪𑜫 *kagan* is dependent on 𑜁𑜪𑜫 *ikan*, and that 𑜁𑜪𑜫 -*ikan* has head-like qualities with regards to modification by 𑜁𑜪𑜫 *kagan*.

- (92) a. <sup>1</sup>*keynam-ikan gino kagan*  
           keynam=ikan gino kagan  
           people=many drunk far.too  
           ‘many far too drunk people’  
           *Intended:* ‘far too many drunk people’
- b. \**tas kagan*  
       tas kagan  
       3PL.M.P far.too  
       ‘far too of them’
- c. \**keynam da-kagan*  
       keynam da=kagan  
       people so=far.too  
       ‘far too so people’

As (92a) shows, 𑜁𑜪𑜫 *kagan* and 𑜁𑜪𑜫𑜰𑜫 *ekeng* also work with regular adjectives, and it is possible as well to add further adjuncts to the AdvP, for instance, 𑜁𑜪𑜫𑜰𑜫 *patu* ‘surprisingly’ in (93). 𑜁𑜪𑜫𑜰𑜫 *kagan patu* ‘surprisingly too’, thus, has to form a phrase which acts as a modifier to 𑜁𑜪𑜫 *gino* ‘drunk’. Likewise, it is not ungrammatical to say 𑜁𑜪𑜫𑜰𑜫𑜁𑜪𑜫𑜰𑜫𑜁𑜪𑜫𑜰𑜫 *keynam-ikan kagan patu* ‘surprisingly too many people’. Since 𑜁𑜪𑜫 *kagan* apparently projects an AdvP (as it can be modified), it cannot be a clitic, because clitics do not project phrases by their very nature. Thus, the c-structure in (94c) is possibly actually more correct, in that it is constructed in analogy to the one for an adjective modified by 𑜁𑜪𑜫 *kagan* in (93b).

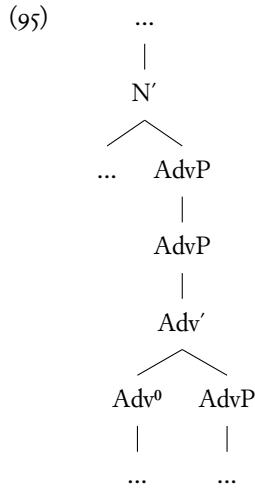
The c-structure in (94c) assumes that the head of the NP has an AdvP sister which has no head, but only an adjunct which is the AdvP headed by 𑜁𑜪𑜫 *kagan*. The head position left empty here is where one would find the adjective in (93b). It is also the position where one would expect a free quantifier, except that the quantifier in this example is clitic and found merged with the noun. In the functional representation, a clitic quantifier also logically places its information in the same spot a free quantifier would, by the distributed exponence principle: the ADJ of the maximal projection, that is, the TOP NP’s ADJ in (94). Since word order matters with regards to quantification, however, we know that 𑜁𑜪𑜫𑜰𑜫 *kagan patu* modifies its antecedent adjective or adverb, which is 𑜁𑜪𑜫 -*ikan* rather than 𑜁𑜪𑜫 *gino*. Furthermore, since empty terminals and non-branching bar levels are supposed to be pruned, the tree in (93b) should have the structure depicted in (95).

(93) a. 
$$\left[ \text{ADJ} \left\{ \begin{array}{l} \left[ \text{PRED 'many'}, \right. \\ \left. \left[ \text{PRED 'drunk'} \right. \right. \\ \left. \left. \left[ \text{ADJ} \left\{ \begin{array}{l} \left[ \text{PRED 'far too'} \right. \right. \\ \left. \left. \left[ \text{ADJ} \left\{ \left[ \text{PRED 'surprisingly'} \right] \right\} \right] \right\} \right] \right\} \right] \right\} \right] \right\} \right]$$



- (94) a. *keynam-ikan kagan patu gino*  
 keynam-Ø=ikan kagan patu gino  
 people-TOP=many far.too surprsingly drunk  
 ‘surprisingly far too many drunk people’





Building the tree this way, it is assumed that there is a hidden AdvP, which is a little unintuitive. However, if the AdvP headed by  $\text{ႁ႗႗႗} kagan$  ‘far too’ were attached directly to N’, the constituency would be wrong:  $\text{ႁ႗႗႗} kagan$  would modify the noun,  $\text{ႁ႗႗႗႗႗႗} keynam-ikan$  ‘many people’ as a unit, compare (96).  $\text{႗႗႗႗} -ikan$  ‘many’ would also wrongly appear at the same level as  $\text{ႁ႗႗႗} kagan$  instead of subordinated to it, unless one were to give  $\text{ႁ႗႗႗} kagan$ ’s maximal projection an annotation like  $\downarrow \in (\uparrow \text{ADJ ADJ})$ , again assuming that  $\text{ႁ႗႗႗} kagan$  is identified as modifying  $\text{႗႗႗႗} -ikan$  on the grounds of word order.



Based on the observation above that  $\text{႗႗႗႗} -ikan$  acts like a head, it might also be the case that both  $\text{႗႗႗႗} -ikan$  and its opposite  $\text{႗႗} -kay$  ‘few, little’ are acting as free morphemes in such contexts, thus eradicating the need for an empty slot in the c-structure in (94c). That is, these two forms would be simple clitics which appear in their full form in this context, the full form happening to be homonymous with the clitic form at this stage, as illustrated by (97).



On the other hand, when using other combinations, like  $\text{ᑭᑦᑎᑦᑎᑦᑎᑦ} -ma$  *nilay* ‘probably enough’, the dilemma starts anew, since  $\text{ᑭᑦᑎᑦᑎᑦᑎᑦ} -ma$  is even more likely a clitic than  $\text{ᑭᑦᑎᑦᑎᑦᑎᑦ} -ikan$  and  $\text{ᑭᑦᑎᑦᑎᑦᑎᑦ} -kay$ , which—untypically of clitics—carry stress, whereas  $\text{ᑭᑦᑎᑦᑎᑦᑎᑦ} -ma$  does not.

- (98) a. *keynam-ma nilay*  
           *keynam-Ø=ma nilay*  
           *people-TOP=enough probably*  
           ‘probably enough people’
- b. *Adareng-ma nilay.*  
       *ada-reng=ma nilay*  
       *that-A.INAN=enough probably*  
       ‘That is probably enough.’

Ayeri is different from English, but since English is often used as a model for generative grammar, let us also try to see what the ‘canonical’ case for modification of APs is for the sake of exploring possibilities. Sobin (2011: 110) likens intensifiers of adjectives like *very* (which I have subsumed under the label ‘quantifier’ and classified as adverbs) to prepositions. He analyzes both, intensifiers and prepositions, as specifiers:





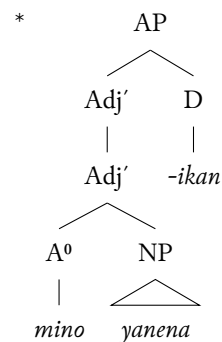


This, however, does not seem to be how Ayeri works, if we assume that there are no transformations and that Ayeri is consistently right-branching. An approach like Sobin (2011) delineates for English, thus, generates the wrong word order for free intensifying adverbs like ʔə́ anyku ‘really’, whether one assumes that an adverb in Spec is the first element in the phrase or the last. In fact, it consistently follows its head, as we have seen before.

- (100) a. \*Yāng **anyku** mingoy sibunana  
           yāng anyku mingoy sibunana  
           3SG.M.A really incapable syntax-GEN
- b. \*Yāng mingoy sibunana **anyku**  
           yāng mingoy sibunana anyku  
           3SG.M.A incapable syntax-GEN really
- c. Yāng mingoy **anyku** sibunana  
           yāng mingoy anyku sibunana  
           3SG.M.A incapable really syntax-GEN
- ‘He is really incapable of syntax.’

It is also incompatible with our analysis of the most common quantifying and degree adverbs as clitics, since clitics, at least in LFG, cannot exist disembodied from their hosts due to lexical integrity:

- (101) <sup>!</sup>mino yanena-ikan  
           mino yan-ena=ikan  
           happy boy-GEN=very/many
- ‘happy about the many boys’  
 Intended: ‘very happy about the boy’



Since Spec also cannot recursively branch (Carnie 2013: 184), examples like (91) would not be grammatical either if ʔə́ -ikan were in the position of a specifier.

As a clitic, ၵိၵ်ႉ *-ikan* cannot be a head either as a logical consequence of lexical integrity.



In conclusion, Ayeri probably does not follow the way of English with regards to quantifiers and intensifiers of adjectives in that assuming these elements are in Spec results in wrong predictions about word order. Since LFG does not assume movement, the hypothesis that Ayeri's quantifiers are in Spec can be discarded, since without movement it results in wrong predictions about word order. It is probably more likely that quantifiers are treated as adjuncts with a constraint that quantifiers cannot be iterated—if a quantifier is present, no other quantifier may modify the same head unless it is coordinated with the first, if this is semantically plausible at all. There also needs to be a constraint which says that quantifiers always need to be adjacent to their modification target. Quantifiers, however, are optional elements, that is, they are not required to complete their head's argument structure in the way a preposition or a verb require an NP complement. It is tempting to treat clitic quantifiers the same way as free ones, just with an empty head position; the quantifying adverb instead finds itself cliticized to the preceding head, compare (94c) and (95).

### 5.3 Adpositional phrases

As described in the section on the morphology of adpositions (section 4.4), Ayeri employs both prepositions and postpositions, though the former are a lot more common and basic than the latter. Thus, PPs are an exceptional domain in that

very limited left-branching is possible in that the complement of a postpositions precedes its head, the postposition. The complement is again right-branching, however. Since the label ‘AP’ has already been used for ‘adjective phrase’, I will use the common label ‘PP’ to refer to both prepositional and postpositional phrases, with their respective heads referred to as ‘P<sup>0</sup>’. As described earlier, there is no morphological difference between prepositions and postpositions; head placement is a syntactic issue and the preference of placement is rooted in the lexical entry for each adposition, albeit ParGram (2009–2016) does not list a feature to distinguish between preposition and postposition, probably because it does not have any relevance to semantics.

LFG categorizes the object of an adposition as  $OBL_{\theta}$ , where  $\theta$  stands for the thematic role of the embedded phrase (Dalrymple 2001: 9–10). As we have seen in the previous chapter, adpositions in Ayeri usually allow adpositional objects to mark one of three cases:

**Locative:** Standard case for prepositional objects, indicates a location (LOCATION). It usually corresponds to English ‘at’, ‘in’.

**Dative:** Indicates motion towards the direction the adposition indicates (GOAL, DIRECTION) in cases where just locative marking would be ambiguous, for instance, ‘up to’ instead of ‘onto’, since  $\text{ᄇᆞᆫᆫᆞᆫ}$  *manga ling* can mean both, by itself. It usually corresponds to English ‘to’, ‘for’.

**Genitive:** Indicates motion from the direction the adposition indicates (SOURCE, ORIGIN) in cases where just locative marking would be ambiguous, for instance, ‘down from’ instead of ‘to the bottom’, since  $\text{ᄇᆞᆫᆫᆞᆫ}$  *manga avan* can mean both, by itself. It usually corresponds to English ‘of’, ‘from’.

Ayeri does not make use of verbs which take a transitive PP as a complement, so there are no cases where an adpositional object is a required part of a verb’s argument structure as in English *talk to someone*; the semantic role of the PP complement is encoded by the PCASE attribute. The complement is also presented as not governing a prepositional object, but the preposition is purely functional (151–153). The case for English is illustrated by (103).

Ayeri, in contrast to English, often uses an NP complement marked with one of the cases in the list above, compare (104). In the case of  $\text{ᄇᆞᆫᆫᆞᆫ}$  *nara-* ‘speak, talk’, the complement thus appears in the locative case, but as an NP, not as a PP. Thus, there is no PCASE attribute necessary here, since there is no preposition to indicate the relation of the complement to the verb because case marking accomplishes this function. Which oblique case the complement appears in is determined by the a-structure of the verb, not by the semantics of the complement. Since the dative is also used for the BENEFICIARY role, the argument expressing the direction of the

(103) John talks to Mary.



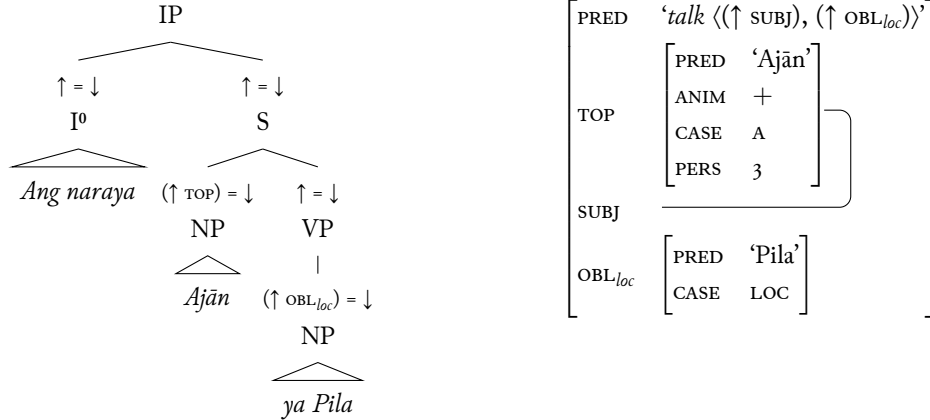
talking appears exceptionally in the locative here. In (104) it may seem as though  $\cup ya$  is a preposition—which is likely accurate historically—however, it has been established previously that preposed case markers are proclitics (see section 3.2.5, p. 75).  $\cup ya$  is a case marker, not a preposition; ‘at/to Pila’ is thus  $\cup \check{\text{r}} \check{\text{r}} \check{\text{r}} ya$  Pila, not  $*\cup \check{\text{r}} \check{\text{r}} \check{\text{r}} \cup ya$  Pila with additional locative marking on the noun.

In Ayeri, PPs proper may either be locative complements of the verb or locative adverbials ungoverned by the verb, though the number of verbs taking PP complements is smaller than in English due to case marking, as described above. As mentioned initially, Ayeri possesses prepositions as well as postpositions. With regards to word order typology, we can thus note:

(105) Order of noun and adposition: Adp N, N Adp

The fact that Ayeri has both prepositions and postpositions is also reflected in the phrase structure rules given in (106). Here, the P<sup>0</sup>s in brackets are supposed to mean that such an element can appear in either position. Since there are no circumpositions in Ayeri, only ever one site is occupied. XP is used again as a catch-all term for various phrase types which form complements of semantic adpositions in the form of NPs and DPs, but especially the postposition  $\check{\text{r}} \check{\text{r}} \check{\text{r}} \check{\text{r}} pesan$  ‘until’ may

- (104) *Ang naraya Ajān ya Pila.*  
 ang nara-ya Ø Ajān ya Pila  
 AT talk-3SG.M TOP Ajān LOC Pila  
 ‘Ajān talks to Pila.’



also have a CP complement, that is, a whole complement clause. The nominal complements are objects governed by the adposition.

- (106) a.  $PP \rightarrow P'$   
 $\uparrow = \downarrow$
- b.  $P' \rightarrow \begin{pmatrix} P' \\ \uparrow = \downarrow \end{pmatrix} \left( \begin{array}{c} \text{AdvP} \\ \downarrow \in (\uparrow_{\text{ADJ}}) \end{array} \right)$
- c.  $P' \rightarrow \begin{pmatrix} P^0 \\ \uparrow = \downarrow \end{pmatrix} \{ (\uparrow_{\text{OBJ}}) = \downarrow \mid (\uparrow_{\text{COMP}}) = \downarrow \} \begin{pmatrix} XP \\ \uparrow = \downarrow \end{pmatrix}$

The same as above is spelled out again in c-tree form in (107). Prepositional phrases, if not subcategorized by the verb, are optional information and thus not governed by the verb; the respective PP is thus part of the set of adjuncts. Some verbs, like  $\text{ᐅᑦᑭᑦ}$  *tapy-* ‘put’, may take PP complements headed by an adposition, for instance, in cases like *Mary puts the book* [<sub>PP</sub> [<sub>P</sub> *on*] [<sub>NP</sub> *the pile*]]. The PP *on the pile* is not an adjunct here because *She does so on the pile* makes no sense; the PP must be an argument of the verb in this case, not additional, optional information. Ayeri behaves the same way in this regard. These governed PPs are categorized as OBL<sub>θ</sub> with  $\theta$  replaced by the proper semantic macrorole (*loc*, *src* or *goal*).

With regards to functional structure, the difference between preposition and postposition does not matter, so the AVM in (108) does not make a formal distinction between prepositions and postpositions. As described in section 3.2.5 (p. 77)

(107) a. Prepositions:



b. Postpositions:



and section 4.4, there is a particle  $\text{མངཱ}$  *manga*, which has been analyzed as a clitic. This particle indicates that an adposition has a directional reading (specifically, into the direction of the preposition). This is encoded by the feature PSEM, which is represented morphologically by said particle and hence not generally present.

(108) a. ...	P	(↑ PRED)	=	'... <(↑ OBJ)>'
		(↑ PCASE)	=	{OBL <sub>loc</sub> , OBL <sub>goal</sub> , OBL <sub>src</sub> }
		(↑ OBJ CASE)	= <sub>c</sub>	LOC <sup>13</sup>
b. $\text{མངཱ}$ <i>manga</i>	Cl	(↑ PSEM)	=	<i>dir</i>
		(↑ PCASE)	=	OBL <sub>goal</sub>

German, for one, encodes the difference between locational and directional of a group of prepositions uses as well, but by an alternation in the prepositional object's case between DAT (locational) and ACC (directional). Thus, *auf dem Tisch* (on DEF.DAT.M.SG table) means 'on the table', whereas *auf den Tisch* (on DEF.ACC.M.SG table) means 'onto the table'. Butt (2005) defines this alternation (albeit for *in* 'in' and *an* 'at') as:

(109)	PSEM <i>dir</i> $\implies$	(↑ OBJ CASE) = <sub>c</sub> ACC	[German]
	PSEM <i>loc</i> $\implies$	(↑ OBJ CASE) = <sub>c</sub> DAT	

The PSEM attribute thus governs the case of the prepositional object by a requirement that its case be either ACC or DAT, depending on its value. Since nouns in German always carry case, the PSEM attribute needs to be present in the lexical

<sup>13</sup> This lexical rule goes for all adpositions except  $\text{ལྷོ}$  *ling* and  $\text{མཇུག}$  *avan*, which require special treatment with regards to case marking; see below.

rules for all prepositions which show the ACC/DAT alternation (*an, auf, hinter, in, neben, über, unter, vor, zwischen*). In Ayeri, however, the particle  $\text{ᄡᆞᆯ}$  *manga* alternates with nothing, so the absence of marking *dir* indicates a locational reading. Since the PSEM attribute finds no overt realization in the absence of the morphological marker, however, it may as well be omitted. Different than in German, the presence or absence of directionality marking also does not influence the case of the adpositional object.

Example (110) illustrates the alternation in annotation between a bare adposition and one modified by  $\text{ᄡᆞᆯ}$  *manga*:  $\text{ᄡᆞᆯ}$  *manga* adds the feature–value pair [PSEM *dir*] to the ADJ (or OBL<sub>θ</sub> if the PP is governed by a verb) of the partial f-structure in (110b), which is how the difference between locational ‘in’ and directional ‘into’ is represented. The case of the OBJ is LOC in both cases.

- (110) a. *kong nangaya*  
           *kong nanga-ya*  
           *inside house-LOC*  
           ‘in the house’
- $$\left[ \text{ADJ} \left\{ \begin{array}{l} \left[ \text{PRED} \text{ ‘inside } \langle (\uparrow \text{OBJ}) \rangle \text{’} \right] \\ \left[ \text{OBJ} \left[ \begin{array}{l} \text{PRED} \text{ ‘house’} \\ \text{CASE} \text{ LOC} \end{array} \right] \right] \end{array} \right\} \right]$$
- b. *manga kong nangaya*  
      *manga kong nanga-ya*  
      *DIR inside house-LOC*  
      ‘into the house’
- $$\left[ \text{ADJ} \left\{ \begin{array}{l} \left[ \text{PRED} \text{ ‘inside } \langle (\uparrow \text{OBJ}) \rangle \text{’} \right] \\ \left[ \text{PCASE} \text{ OBL}_{\text{goal}} \right] \\ \left[ \text{PSEM} \text{ dir} \right] \\ \left[ \text{OBJ} \left[ \begin{array}{l} \text{PRED} \text{ ‘house’} \\ \text{CASE} \text{ LOC} \end{array} \right] \right] \end{array} \right\} \right]$$

With the prepositions  $\text{ᄡᆞᆯ}$  *ling* ‘on top of’ and  $\text{ᄡᆞᆯ}$  *avan* ‘at the bottom of’ there is another alternation, based on case, for verbs which do not encode direction. This is rooted in the etymology of these words:  $\text{ᄡᆞᆯ}$  *ling* as a noun means ‘top’,  $\text{ᄡᆞᆯ}$  *avan* means ‘ground, bottom’. The directional variants of these prepositions mean ‘to the top, onto’ and ‘to the bottom’. These are telic in that they express arriving at the destination specified by the adpositional object. However, Ayeri does not possess separate adpositions to express ‘up’ and ‘down’ as atelic concepts, that is, without referring to arriving at a destination.<sup>14</sup> Instead,  $\text{ᄡᆞᆯ}$  *ling* and  $\text{ᄡᆞᆯ}$  *avan* double for ‘up’ and ‘down’ with dative complements, respectively, to mark the difference, compare Table 5.1. Note, however, that Ayeri does not possess intransitive adpositions, with the exception of a few verbs where the adposition is a lexicalized part of the expression, for instance,  $\text{ᄡᆞᆯ}$   $\text{ᄡᆞᆯ}$  *tapy-dayrin* ‘save (assets)’, with the fossilized, now

<sup>14</sup> I concede that the chosen terminology is slightly problematic since, strictly speaking, ‘up’ and ‘down’ are also directions rather than locations. Moreover, whether ‘up to’ and ‘down from’ qualify as atelic is probably debatable, since they as well suggest arriving at a destination eventually.

Table 5.1: Case alternations of လှံ *ling* and အံ့ *avan*

	+ LOC	+ DAT
<i>ling</i> ‘top’	on top of	up
<i>manga ling</i> ‘to top’	to the top of	up to
<i>avan</i> ‘bottom’	at the bottom of	down
<i>manga avan</i> ‘to bottom’	to the bottom of	down to

defunct preposition လှံ *dayrin* ‘beside, next to’ (modern လှံ *kayvo*). ‘Up’ and ‘down’ thus refer to their transitive uses as in *up the stairs* and *down the hill*, respectively. (III) provides example sentences for all configurations.

- (III) a. *Ang lampyo paray ling nayingya.*  
ang lamp-yo paray-Ø ling naying-ya  
AT walk-3SG.N cat-TOP top.of roof-LOC  
‘The cat is walking on the roof’
- b. *Ang puco paray manga ling mebirya.*  
ang puk-yo paray-Ø manga ling mehir-ya  
AT jump-3SG.N cat-TOP DIR top.of tree-LOC  
‘The cat jumps onto the tree.’
- c. *Ang nimpyan ganye ling turayyam.*  
ang nimp-yan gan-ye-Ø ling turay-yam  
AT run-3PL.M child-PL top.of hill-DAT  
‘The children are running up the hill.’
- d. *Ang saraya jarmaya manga ling pelangyam.*  
ang sara-ya jarmaya-Ø manga ling pelang-yam  
AT go-LOC pilgrim-TOP DIR top.of castle-DAT  
‘The pilgrim goes up to the castle.’
- [PRED ‘top of <((↑ OBJ))’]  
PCASE OBL<sub>loc</sub>  
PSEM loc  
TEL +  
OBJ [PRED ‘roof’]  
CASE LOC]
- [PRED ‘top of <((↑ OBJ))’]  
PCASE OBL<sub>goal</sub>  
PSEM dir  
TEL +  
OBJ [PRED ‘tree’]  
CASE LOC]
- [PRED ‘top of <((↑ OBJ))’]  
PCASE OBL<sub>loc</sub>  
PSEM loc  
TEL –  
OBJ [PRED ‘hill’]  
CASE DAT]
- [PRED ‘top of <((↑ OBJ))’]  
PCASE OBL<sub>goal</sub>  
PSEM dir  
TEL –  
OBJ [PRED ‘castle’]  
CASE DAT]

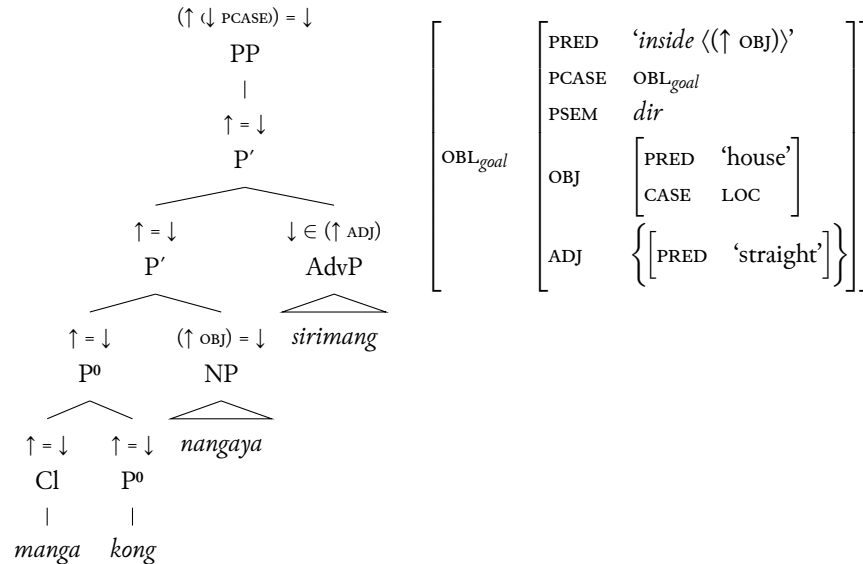


<sup>15</sup> A grammaticalization process similar to that of *inside of the X* to *inside the X* may be a logical next step from here. What would be possible moreover as a result is the functionalization of

Like adjectives, adpositions may be modified by adverbs, for instance, for intensification. This means that items from the class of adverbs categorized as quantifiers (section 4.6.2) are likely to occur. As we have seen previously, the most common such expressions are enclitic, that is, they merge with  $P^0$  rather than be adjuncts to  $P'$ . An analysis of the c- and f-structure of adpositional phrases with adverbial modifiers is shown in (114).

- (114) a. *Ang nimpya Tapan manga kong nangaya sirimang.*  
 ang nimp-ya Ø Tapan manga kong nangaya sirimang  
 AT run-3SG.M TOP Tapan DIR inside house-LOC straight

‘Tapan is running straight into the house.’



- b. *Ang galamyan panganya pesan-ben.*  
 ang galam=yan.Ø pangan-ya pesan=hen  
 AT wait=3PL.M.TOP end-LOC until=all

‘They waited until the very end.’

the triad LOC–DAT–GEN to indicate *loc*, *goal*, and *src* also with adpositions in general, either with *manga* grammaticalizing further to then solely mark direction, necessitating an obligatory DAT/GEN complement to indicate which way around, or with *manga* withering to zero, since the cases are already enough to mark direction. The way to express the difference between *top/up* and *bottom/down* would have to change for obvious reasons, though *ring* from *ring*: *ring*– ‘rise, lift’ and *lesa* from *lesa*: *lesa*– ‘fall’, or *rota* from *rota*: *rota*– ‘heave (up)’ and *kosa* from *kosa*: *kosa*– ‘drop (down)’ would be good candidates from which to generate new adpositions. Alternatively, *saba* could become an equivalent of *manga* to indicate direction *from* the indicated place.



Since with prepositions the adverb comes after the nominal complement as in (114a), there may be potential ambiguity as to constituency, since adjectives and adverbs do not differ in form. One such case is illustrated in (115), where the A-type modifier  $\text{အဝေး}$  *baras* ‘rough(ly)’ may be interpreted either as an adverb modifying the adpositional phrase or as an adjective modifying the prepositional object. The individual words are copied again to the very bottom of the c-structure trees in (115) to highlight that different syntactic structures may lead to the same outcome on the surface.

Ambiguity can be resolved in these cases by subordinating the adjective to the noun explicitly with the relativizer  $\text{အဘယ်}$  *si*, as shown in (116). The relative clause, then, essentially means ‘which is ADJECTIVE’. The fact that it is somewhat hard to come up with an example is probably telling of the likelihood of such ambiguity. In either case, wrapping an adjective into a relative clause CP to explicitly subordinate it is always a permissible strategy of clarification. This also means in turn that adpositions cannot be modified by relative clauses, however, this should rarely be necessary, if it makes any sense at all.

- (115) *marin altanya baras*  
 marin altan-ya baras  
 in.front rock-LOC rough(ly)

‘roughly in front of the rock’  
 or: ‘in front of the rough rock’



or



- (116) *marin altanya si baras*  
 marin altan-ya si baras  
 in.front rock-LOC REL rough

‘in front of the rough rock’  
 literally: ‘in front of the rock which is rough’



It was mentioned initially that certain adpositions are also able to take clausal complements. This is especially the case for when adpositions are used to describe points in or stretches of time. A list of adpositions which can be used for this purpose is given in Table 4.17. Essentially, a subset of spatial prepositions can be used metaphorically to refer to time, just as in English.

- (117) *Gamaryang marin sa-sahaye ang Sipra.*  
 gamar=yang marin sa~saha-ye ang Sipra  
 manage=1SG.A before ITER~come-3SG.F A Sipra  
 ‘I’ll get it done before Sipra returns.’

Both *gamaryang* ‘I manage (it), I get (it) done’ and *sa-sahaye ang Sipra* ‘Sipra returns’ are complete sentences; the preposition *marin* ‘in front of, before’ ties them together and indicates the second part’s relationship to the first: the embedded clause expresses a future state which serves as the background for the action expressed by the matrix clause. The corresponding constituent structure of the PP is shown in (118).



In cases like (118), an adverbial modifier in the fashion of (114) would attach in the adjunct position as well, following the clausal complement. For the example above, this is not very awkward, because the embedded clause is very short and does not branch too deeply. For cases where the length or depth of the CP does produce an awkward result, however, it is possible to extrapose it with *da-* on the phrase’s head representing the missing complement, altogether structurally similar to (82). An example of this strategy is given in (119).

- (119) *Gamaryāng*    *da-marin*    *sirimang*  
 gamar=yāng    da=marin    sirimang  
 manage=1SG.A    such=before    straight  
  
*sa-sabayē*            *ang Sipra.*  
 sa~saha-ye            ang Sipra  
 ITER~come-3SG.F A    Sipra  
  
 ‘I will do it right before Sipra returns.’



## 5.4 Inflectional and verb phrases

In this section I will finally describe the way verbs operate with regards to morphosyntax. Verbs are the part of speech which governs the relations of the various phrase types to each other and they are thus central to the formation of clauses. It comes as no surprise that LFG always lists a verb with its argument structure first in representations of a clause's functional structure.

Just from looking at the numerous examples in the previous section, it should have become clear that Ayeri's preferred word order is verb-first, which opens up a few typological questions, first and foremost, whether Ayeri actually has a verb phrase, or in terms of generative grammar: whether it is configurational in this regard. As we have seen, Ayeri definitely has a constituent structure as far as NPs, APs, PPs, etc. are concerned. However, due to VSO word order, it is not obvious whether verb and object actually form a constituent, since V and O are not adjacent to each other. Since Ayeri marks topics in terms of morphology, it will also be necessary to discuss how this mechanism works and how it relates to the notion of the subject.

A discussion of subject, topic, and configurationality is interesting also insofar as Ayeri's syntactic alignment was originally *inspired* by the Austronesian or Philippine alignment system—Tagalog, an Austronesian language of the Malayo-Polynesian branch, spoken mainly in the Philippines (Hammarström et al. 2017; Schachter and Otanes 1972), usually serves as the academic poster child in descriptions of this system. Ayeri departs from Tagalog's system in a number of ways, though, towards

the more conventional from the point of view of Standard Average European; the Austronesian alignment is thus not necessarily the best role model to liken Ayeri's syntax to. It will nonetheless be informative to compare both systems based on the work of Kroeger (1991, 1993), who provides an analysis of Tagalog's verb alignment at least roughly in terms of the LFG framework as well and uses some heuristics which may help in establishing what is actually going on in Ayeri.<sup>16</sup>

#### 5.4.1 Typological considerations

As mentioned above, Ayeri's unmarked word order has the verb come first, and then, in decreasing order of bondedness to the verb, the phrases which make up the verb's arguments: logical subject (agent), direct object (patient), indirect object (dative), followed by adverbials in the genitive, locative, instrumental, and causative case. Ayeri's basic word order is thus VSO, a trait it has in common with about 7% of the world's natural languages according to Dryer (2013a) (compare Table 5.2).<sup>17</sup> Following the format of previous statements on word order typology, we can declare the generalization in (120), which is consistent also with previous observations on word order typology, where the head preceded the modifier. The head is here represented by the verb, the modifier by the object—like English, Ayeri is a VO language, thus. In addition to this, however, Ayeri regularly puts the verb as the head of the clause itself first.

- (120) a. Order of subject, object, and verb: VSO  
b. Order of verb and object: VO

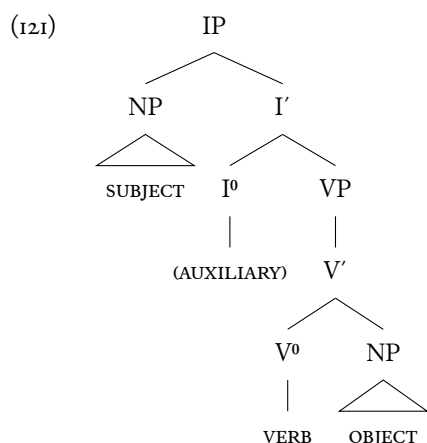
It is commonly assumed that languages have a logical subject which occupies a certain position in the constituent structure—the predicate—and which commands a constituent jointly formed by the verb and its dependents—the predication. An SVO sentence thus very generally looks like this (compare the examples in Bresnan et al. 2016: 101–111):

<sup>16</sup> As mentioned in the introductory chapter, I started Ayeri in late 2003—then still in high school and not knowing much about linguistics. Of course, I had to go and pick the syntactic alignment which has long been “a notorious problem for both descriptive grammarians and theoretical syntacticians” to the point where it “sometimes seems as if Austronesian specialists can talk (and write) of nothing else” (Kroeger 2007: 41).

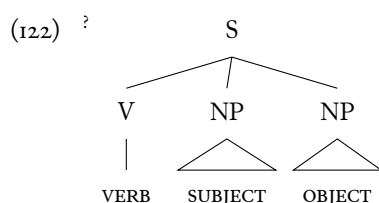
<sup>17</sup> It may be noted here that the ‘other’ category is actually ‘no dominant order’ in the original, and that this is maybe not quite accurate for at least a few languages in the sample, for instance, German. The unmarked word order of declarative statements in German main clauses is SVO with the finite verb mandatorily occupying the second position (V2). This position is commonly analyzed as C<sup>0</sup> in German, however, since the infinite content verb V<sup>0</sup> appears at the end of the clause when the C<sup>0</sup> position is occupied by an auxiliary or a modal (compare, for instance, Bresnan et al. 2016: 375–379, 447–450, and Fortmann 2006); subclauses are verb-last. It is not uncommon to analyze German as an OV language with V2 tendencies.

Table 5.2: Frequency of word order types  
(n = 1377; adapted from Dryer 2013a)

Order	Frequency	Percentage
SOV	565	41.0%
SVO	488	35.4%
VSO	95	6.9%
VOS	25	1.9%
OVS	11	0.8%
OSV	4	0.3%
Other	189	13.7%



However, Ayeri is an VSO language, so the question arises how the basic constituent structure should be diagrammed in tree form, since V and O are not adjacent. In an initial hypothesis one might assume that they cannot form a unit together, since S somehow stands in between the constituents it is supposed to command. A very first stab at diagramming would probably be to come up with a flat, non-configurational structure:





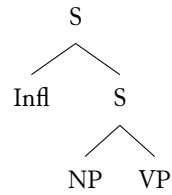
Such a structure, though, does not do Ayeri justice in that, for instance, right-node-raising of a subject and object NP together is possible, so there is evidence that they form a constituent subordinate to the verb. NP–XP constructions where XP is not a maximal projection of a verb also exist in isolation, so NP and XP are probably contained in a small-clause constituent S. The verb in the initial position furthermore shows inflection, so one might rather construe it as an  $I^0$ , projecting an IP, which frees up VP for other purposes while we can use IP to govern both  $I'$  and S. In fact, such a structure is basically the conclusion Chung and McCloskey (1987) come to for Irish, which is also a VSO language (123a). Bresnan et al. (2016) give the similar chart in (123b) for Welsh, equally a VSO language (also compare Dalrymple 2001: 66, sourcing Sadler 1997). Kroeger (1991) suggests the two structures depicted in (123c) for Tagalog, based on the suggested constituent structure for Celtic languages.

What all of these c-structures have in common is that the inflected verb appears in  $I'$ , which is a sister of S. S, in turn, is a small clause containing the arguments of the verb. In the cases of Irish and Welsh, however, there is a VP sister of the subject NP which itself does not have a head, but contains the object NP as a modifier. In the case of Tagalog, S is non-configurational, that is, while XP may contain an infinite verb, the subject and object NPs are on equal footing.

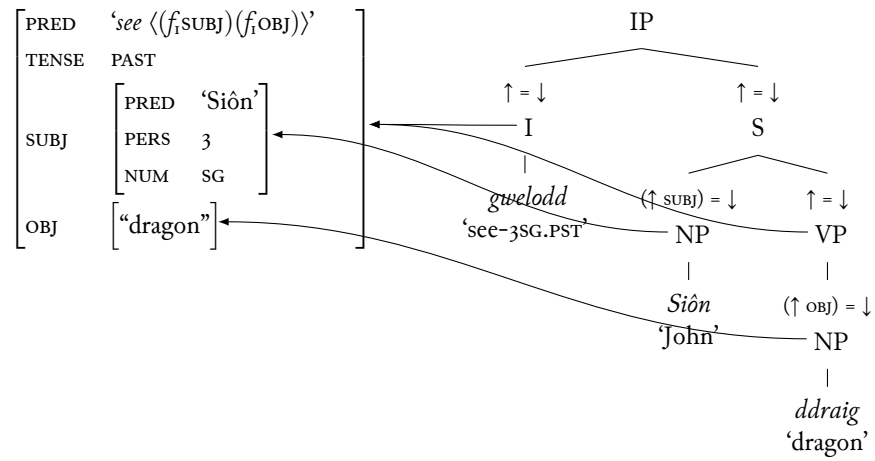
Bresnan et al. (2016: 129–138) informs that the phenomenon of the verb ending up in a different head position ( $V^0$  apparently moves to  $I^0$ ) in (123b) is commonly known as ‘head movement’, except that LFG is built specifically without any movement. Since LFG is based on the assumption that all nodes in a syntactic structure are base-generated, that is, that there are no transformational rules generating the surface structure from a deeper layer of representation underneath it, there cannot be a trace of V left behind in VP. LFG avoids empty categories, as there is no information contained in an empty node. The functional information provided by the verb is not lost, however, it is merely now provided by the verb in  $I^0$ . Essentially, the Welsh example does not violate endocentricity, since the finite verb in  $I^0$  still forms the verbal head in the functional structure representation of the clause. With regards to constituent structure,  $V^0$ , if present, c-commands its NP sister; both  $V^0$  and NP are dominated by VP:

- (124) a. Exhaustive domination (Carnie 2013: 121):  
 “Node A exhaustively dominates a set of terminal nodes {B, C, ..., D}, provided it dominates all the members of the set so that there is no member of the set that is not dominated by A and there is no terminal node G dominated by A that is not a member of the set.”
- b. C-command (127):  
 “Node A c-commands node B if every node dominating A also dominates B, and neither A nor B dominates the other.”

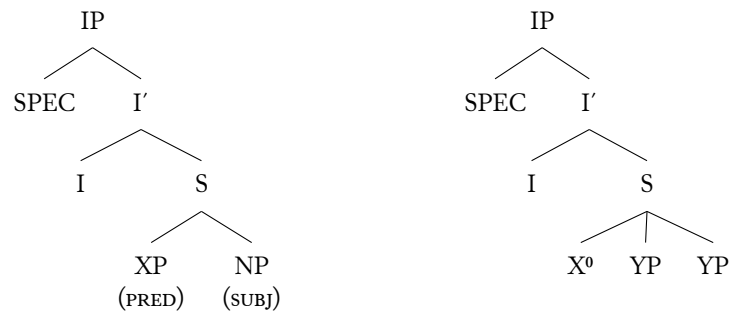
- (123) a. Irish (Chung and McCloskey 1987: 235):



- b. Welsh (adapted from Bresnan et al. 2016: 134):



- c. Tagalog (Kroeger 1991: 131):



The AVM in (123b) shows that the contents normally found in  $V^0$  are provided by the head of its equivalent functional category,  $I^0$ ;  $I^0$  and VP are said to map into the same f-structure (Bresnan et al. 2016: 136). Endocentricity still holds in that IP dominates all nodes below it, thus also  $I^0$  and the object NP. In addition,  $I^0$  c-commands its sister node and all of its children, hence also the object NP. As Bresnan et al. (2016) put it: “X is an extended head of Y if X is the X’ categorial head of Y [...], or if Y lacks a categorial head but X is the closest element higher up in the tree that functions like the f-structure head of Y” (136). For our example, replace X with  $I^0$  and Y with VP in the second half of the quote:  $I^0$  is the closest element higher up in the tree that functions like the f-structure head of VP, which itself lacks a categorial head.

The analysis of the sentence structure of Celtic languages shows that VSO languages do not automatically need to be considered ‘non-configurational’ and lacking a VP if the notion of extended heads is accepted. In any case, tests need to be performed to see whether one of the analyses presented in (123) holds true for Ayer as well.

#### *‘Trigger languages’ and Austronesian alignment*

The term ‘trigger language’ comes up in discussions on Conlang-L as early as 1995, where it may well have originated as an established term for what will be described below in brief. That is, I have not been able to find any earlier mentions of the term ‘trigger’ as referring to an alignment system in the archives; other mainstays of the fictional-language community, such as the ZBB, were established only about a decade later. In a message dated December 16, 1995, John Cowan writes that he wants “to propose a reform of Radilu, to make it use the Tagalog concept of a ‘trigger’” (Cowan 1995). By his definition, this entails that

each clause contains one noun phrase which is not marked for case, but rather has a distinct marking called the “trigger marker”. [...] The verb carries a marking (which of course looks nothing like the noun case markers) that tells the true case of the trigger. For example, the English sentence “I gave the dog the bone” can come out:

- 1) I-TRIGGER gave-A.T. dog-DATIVE bone-PATIENT
- 2) I-ACTOR gave-P.T. dog-DATIVE bone-TRIGGER
- 3) I-ACTOR gave-D.T. dog-TRIGGER bone-PATIENT

This involves changing the name of “nominative” and “accusative” to “actor” and “patent” [*sic*], since there is no longer a “subject” or “object” as such. Of course, word order is free, so:

- 4) dog-DATIVE gave-P.T. I-ACTOR bone-TRIGGER

(Cowan 1995)

He also notes that “Usually the trigger is definite (Tagalog doesn’t have articles)” (Cowan 1995); essentially, it seems that the motivation for Cowan’s system is that the ‘trigger’ indicates that a certain NP is definite. As we will see below, this

is very similar to how Tagalog marks one of its relations on the verb, with that relation being definite—things are slightly more complicated in reality, though, of course. The term ‘trigger’ seems to have currency in that, for instance, Schachter (2015) chooses it explicitly to refer to the “non-case-marked argument” (1659). In a parenthetical remark he adds that

Some previous treatments have referred to the argument in question as the *topic* and some as the *subject*. However, as will become clear below, each of these labels appears to carry some inappropriate connotation, making a neutral term like *Trigger* seem preferable [...] There also seems to be good reason to reject the term *focus*. (1659)

It may be noted that term ‘focus’ is used in Schachter and Otanes (1972), the main reference grammar of Tagalog. However, what is interesting in comparing Schachter (2015)’s and Kroeger (1991)’s respective analyses of Tagalog’s syntactic alignment is that both come to the same conclusion: Tagalog is ambiguous as to whether the subject notion is vested in the NP whose role is marked on the verb or the actor, since certain syntactic constructions typically associated with subjects apply to either or both. While this ambiguity leads Schachter (1976, 2015) to ultimately conclude that Tagalog lacks a single unified relation which can be analyzed as a syntactic subject,<sup>18</sup> Kroeger (1991) reaches the opposite conclusion by performing further tests and taking a functionalist rather than purely structuralist perspective:

- “Tagalog has a well-defined grammatical subject”. What Schachter (1976) lists as evidence against are special cases which can be explained by the high semantic and pragmatic prominence of actors more generally (Kroeger 1991: 225);
- “grammatical relations are defined independently of phrase structure” (225);
- “patients can become subjects even when the agent is expressed as a direct (non-oblique) argument of the verb” (225).
- “Subject selection in Tagalog does not work by demotion or suppression of thematically more prominent arguments. Rather, all arguments seem to be equally eligible for mapping onto the subject relation” (226).

He also provides evidence based on statistics and examples that the marked-for relation, which he classifies as being in the nominative case according to his hypothesis that it is the syntactic subject, is neither especially salient in terms of pragmatic topichood, nor does it show signs of carrying pragmatic focus specifically. He finds that rather, nominative marking works independent of discourse functions (56 ff.).

<sup>18</sup> Cowan (1995)’s sketch may be based on Schachter (1976). Curiously, Schachter (2015) does not acknowledge Kroeger (1991) at all, nor does he refer to any other research more recent than 1985. The reason may be that Schachter presumably retired in the early 1990s, as the UCLA linguistics department’s Department history (2017) suggests.

The tests for typical properties associated with grammatical subjects which Kroeger (1991) performs partially extend those presented in Schachter (1976). Moreover, his conclusions build on a more modern, functionally oriented approach than Schachter's. For this reason, I will follow Kroeger rather than Schachter. Either way, we will have to test quantifier float, relativization, number agreement, reflexive binding, raising and control verbs, control of secondary predicates, and the termhood of the agent NP in transitive sentences with what was called 'patient topic' so far. First of all, it will be helpful, however, to define some terms which will be used in the discussion below.

#### 5.4.2 Definition of terms

The terms 'subject', 'topic', and 'focus' were already used a number of times above, but it seems advisable to sketch out working definitions in order to preclude confusion before continuing to look at how Ayerí fares with regards to these notions. As we will see, all of subject, topic, and focus relate to different ways in which the relative prominence of certain NPs is raised; subject and topic are also closely related to each other. It ought to be noted that while LFG treats topic and focus as grammaticalized discourse functions outside of the argument-structure frame of a verb, it treats the subject as both a discourse function and an argument function; topic and focus, on the other hand, must be identified with a corresponding argument function (Bresnan et al. 2016: 99–100; also compare (1) above, p. 240).

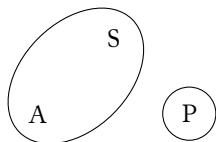
#### *Subject*

First things first, the **subject** can be defined in a variety of ways, and maybe especially because the notion of a subject is so basic, Comrie (1989) notes that if

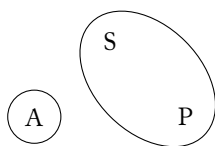
linguists were invariably in agreement in stating which noun phrase, in each construction in each language, is the subject, then we could, perhaps, accept this inter-subjective agreement, and devote correspondingly less energy to trying to find an explicit definition of subject. However, it turns out that, in a wide range of cases, this inter-subjective agreement is lacking. (104)

Dixon (2010) defines a subject as "the entity about which something is affirmed or denied" (76). He goes on to explain that, ignoring copula clauses like 'We are tired and thirsty', every language has two varieties of clauses, intransitive ones, where the verb has just one core argument, and transitive ones, where the verb has two core arguments. A basic definition based on this is given by the chart in (125). It shows the definition of the notion of subject for both nominative–accusative languages and ergative–absolutive languages. Languages of the world differ based on how they prefer to treat the two nominal relations of a transitive verb in relation to intransitive verbs: they may have a strong preference to either treat the agent

- (125) a. nominative–accusative alignment (S/A—P):



- b. ergative–absolutive alignment (S/P—A):



(A)—the entity that prototypically acts in some way—or the patient/undergoer (P)—the entity which is prototypically affected by the action in some way, that is, the object—the same as S. In the former case, the language is said to have NOM–ACC alignment (125a), whereas in the latter case, the language is said to have ABS–ERG alignment (125b). Comrie (1989) illustrates this difference with an example from Chukchi, which we will here contrast with English:<sup>19</sup>

- (126) a. *I came*  
 ISG.NOM come.PST  
 S
- b. *I saw her*  
 ISG.NOM see.PST 3SG.F.OBL  
 S/A P

- (127) Chukchi (adapted from Comrie 1989: 104):

- a. *ɣəm tə-yet-ɣʔek*  
 ISG.ABS came-1SG  
 S  
 ‘I came.’
- b. *ɣəm-nan ɣət tə-lʔu-ɣət*  
 ISG.ERG 2SG.ABS saw-1SG-2SG  
 A S/P  
 ‘I saw thee.’

<sup>19</sup> In English, *you* is the same for both singular and plural as well as subjective and objective case, which is why I replaced it with the unambiguous *her* in (126).

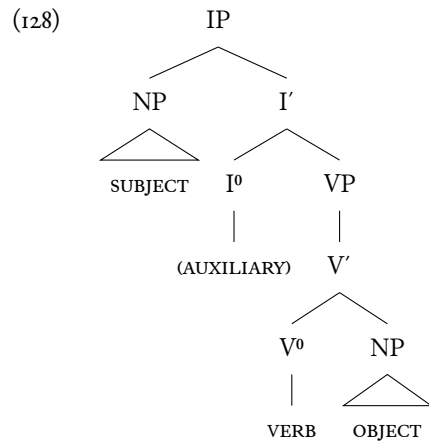
While English treats the subject of the intransitive sentence (126a) the same as that of the transitive one (126b)—both sentences use *I* in the nominative—Chukchi appears to use a different pronoun for the subject of the intransitive sentence (127a) than the subject of the transitive one—absolutive *yəm* versus ergative *yəmnən*, respectively. At least in Standard English, it would be ungrammatical to use the pronoun *me* in place of *I* in (126b), since *me* can only be used for first-person objects of the verb, but not for subjects of transitive clauses.

However, Comrie (1989) also urges to consider that grammatical relations and their representation in morphology are not always this clear cut as in the example above. While he characterizes the prototypical subject as the intersection of agent and topic as far as cross-linguistic evidence is concerned (107), he also points out that subjects do not necessarily have to unite all the properties typically associated with them (110). This is the case, for instance, with Tagalog, as observed by both Schachter (1976) and Kroeger (1991), and may considerably complicate making a definitive statement.

Moreover, he points out that statistically, languages of the world show a strong preference for NOM–ACC alignment, possibly due to the fact that human perception values agents as more relevant to discourse than patients, which is why agents are far more likely also to be pragmatic topics (Comrie 1989: 120). Yet, though, dominantly NOM–ACC-aligned languages may show a bias towards an ABS–ERG treatment, for instance, of resultative constructions. On the other hand, dominantly ABS–ERG languages show a bias towards a NOM–ACC treatment, for instance, of addressees of imperatives (116–119).

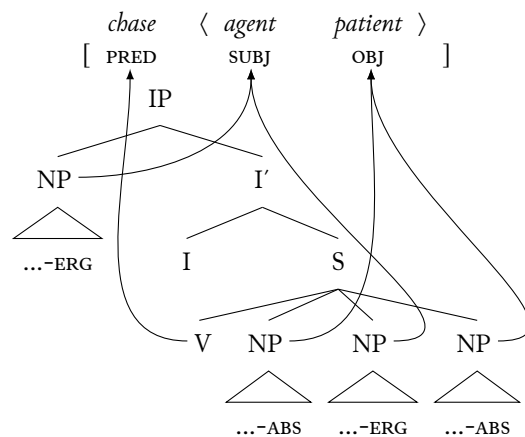
According to (Carnie 2013), from the point of view of constituent structure (which is key in Generative Grammar), a subject is conventionally understood to be a “DP that has the property indicated by the predicate phrase. What the sentence is about. In most sentences, this surfaces in the specifier of [the tense phrase]” (221). However, as we have learned above, this notion is challenged by languages such as Tagalog (Kroeger 1991: 225). What Carnie (2013) refers to in terms of constituent structure was basically already indicated by (121), except with different labels; the example is repeated here for convenience as (128). For systemic reasons, Carnie (2013) refers to a DP subject which serves as the specifier of a TP. This corresponds to the subject NP and the IP here. Unlike GG, LFG treats tense as a semantic feature, not as a functional head with a fixed position in constituent structure, hence the difference in labeling.

LFG defines both a syntactic-subject function, SUBJ, and a logical subject,  $\hat{\theta}$ , which operate in lockstep. Which semantic component the subject is mapped onto is understood to be based on the relative prominence of the subject argument along some dimension compared to other arguments. For instance, NOM–ACC languages prefer the semantically most prominent available role of a verb’s argument structure,



ABS-ERG languages instead pick the argument most affected by the agent's action, and active languages focus on the argument in control of the action (Bresnan et al. 2016: 95–96). The mapping between grammatical functions like SUBJ and the lexical components that make it up also does not need to be a one-to-one correspondence, since LFG allows for the distributed exponence of grammatical features like in the example of Warlpiri in (129). The only condition is that grammatical functions be uniquely defined within their minimal f-structure (45). As (129) shows, multiple NPs in different positions in the constituent structure may feed semantic information to a single function defined by the argument structure of the verb.

(129) Warlpiri (Bresnan et al. 2016: 325):





The logical subject  $\hat{\theta}$  is defined as “the most prominent semantic role of a predicator” (330). Furthermore, Bresnan et al. (2016) devise two a-structure features,  $[\pm o]$  (objective) and  $[\pm r]$  (restrictive). According to this classification, SUBJ is assigned the features  $[-r, -o]$ , since the subject is not restricted to a certain semantic role, nor needs to have a semantic role. Also, subjects do not complement transitive predicators like objects do, so they are not ‘objective’. Bresnan et al. (2016)’s lexical mapping theory assumes that all languages have subjects, which goes counter to Schachter (1976, 2015)’s claim that subjects are possibly not a universal (Bresnan et al. 2016: 330–331).

### Topic

The notion of **topic** refers essentially to who or what a longer stretch of conversation is about. Givón (1983) defines the topic of a ‘thematic paragraph’—as he calls a coherent unit of discourse above the level of a single sentence—to be “the continuity marker, the *leitmotif*” (8). The topic is thus

the participant *most crucially involved* in the action sequence running through the paragraph; it is the participant most closely associated with the higher-level “theme” of the paragraph; and finally, it is the participant most likely to be coded as the *primary topic*—or grammatical subject—of the vast majority of sequentially-ordered clauses/sentences comprising the thematic paragraph. (8)

This indicates that topic and subject are closely related concepts, as already mentioned above in reference to Comrie (1989). Languages employ various means to indicate topics; right- and left-dislocation, as known from English, are only one among many possibilities (Dixon 2010: 174).

Topicality also interfaces with definiteness in that chain-initial topics may be definite (already introduced into discourse) or indefinite (newly introduced into discourse), while chain-medial topics and chain-final topics are always expected to be definite (Givón 1983: 10). Dixon (2010: 171) adds that topic NPs are coreferential with arguments of clauses immediately preceding or following the current clause. Moreover, the strategy of passivization (in NOM–ACC languages) or of antipassivization (in ABS–ERG languages) exists in order to keep a certain discourse item persistent in the highly topical subject position even if it would otherwise be the object of the clause. This is related in turn to the notion of syntactic pivot in clause coordination (172).

### Focus

Regarding the definition of **focus**, Dixon (2010: 174) only mentions contrastive focus, which basically raises the prominence of a certain NP within a single clause. It is not necessary for the focussed NP to be coordinated with another NP by ‘or’.

Dixon (2010) also warns that focus is often confused with topic. Perhaps this is in part also, as Bresnan et al. (2016) mention, due to the fact that English may use the topic position for either topic or focus under certain circumstances:

(130) English (Bresnan 2016: 98):

Q: What did you name your cat?

A: ROSIE I named her. (*Rosie* = FOC)

The answer to a *wh*-question is considered focused, so *Rosie* in (130) is the focus in ‘I named her ROSIE’. However, in the example above, *Rosie* is fronted, which following Givón (1983), constitutes a disruptive action used to establish a new topic of conversation: left-dislocation in languages with rigid SVO word order such as English is typically associated with low topic continuity, and left-dislocated NPs can be found most often as initiating a topic chain (32).

### 5.4.3 Tests on subjecthood

As initially mentioned, Ayeri was originally conceived under an impression of what was described in the quotation of Cowan (1995) above in terms of ‘trigger language’ (also compare Schachter 2015). That is, in simple declarative statements, the semantic macrorole of a definite NP is marked on the verb. This is itself a very basic account of what can be observed in Tagalog and other Philippine languages, compare (131) (emphasis mine).

(131) Tagalog (Kroeger 1991: 14, adapted from Foley and Van Valin 1984: 135):

- a. *B-um-ili ang=lalake ng=isda sa=tindahan.*  
 PFV.AV-buy NOM=man GEN=fish DAT=store  
 ‘The man bought fish at the store.’
- b. *B-in-ili-Ø ng=lalake ang=isda sa=tindahan.*  
 PFV-buy-OV GEN=man NOM=fish DAT=store  
 ‘The man bought *the fish* at the store.’
- c. *B-in-ilh-an ng=lalake ng=isda ang=tindahan.*  
 PFV-buy-DV GEN=man GEN=fish NOM=store  
 ‘The man bought fish *at the store*.’
- d. *Ip-in-am-bili ng=lalake ng=isda ang=pera.*  
 IV-PFV-buy GEN=man GEN=fish NOM=money  
 ‘The man bought fish *with the money*.’

The examples in (131) show variations on the same sentence, differing in the distribution of the definite NP which Kroeger (1991) classifies as being the subject

of the respective sentence on syntactic grounds. The subject NPs are marked with the clitic *ang*, and their role in the clause is reflected by the voice marking on the verb (the root is *bili* ‘buy’): in (131a) the subject is the actor, in (131b) it is the object, in (131c) it is a location, and in (131d) it is an instrument. What is remarkable is that this voice marking goes beyond mere passivization, so even the oblique arguments of (131cd) can become subjects of their respective clauses. Ayeri is at least superficially similar, compare (132).

- (132) a. *ang=int-ya ayon-Ø inun-ley moton-ya*  
 AT=buy-3SG.M man-TOP fish-P.INAN store-LOC  
 ‘The man, he bought fish at the store.’
- b. *le=int-ya ayon-ang inun-Ø moton-ya*  
 PT.INAN=buy-3SG.M man-A fish-TOP store-LOC  
 ‘The fish, the man bought it at the store.’
- c. *ya=int-ya ayon-ang inun-ley moton-Ø*  
 LOCT=buy-3SG.M man-A fish-P.INAN store-TOP  
 ‘The store, the man bought fish there.’
- d. *ri=int-ya ayon-ang inun-ley pangis-Ø*  
 INST=buy-3SG.M man-A fish-P.INAN money-TOP  
 ‘The money, the man bought fish with it.’

Like Tagalog, Ayeri marks a privileged NP on the verb, however, in Ayeri, this is the topic, not the subject (this will be subject to further scrutiny below). Unlike Tagalog, the marked NP is not marked by a particle, but by the very absence of case marking on the NP itself. The case marker appears as a clitic marker in its prefix form at the left-most edge of the clause, before the verb (compare sections 4.1.3 and 4.5). While the marker on the verb is related to nominal case markers in Ayeri, Tagalog uses a number of prefixes, suffixes, and infixes for voice marking which are not obviously related to case markers on nouns. For instance, non-subject actors are marked by the genitive clitic *ng*, while actor voice is marked by *mag-* or *-um-* (Schachter and Otanes 1972: 74, 78). In Ayeri, on the other hand, non-topic animate agents are marked by *ᐃᐃᐃ* *-ang* or *ᐃᐃᐃ* *ang*, and animate agent-topics are marked by *ᐃᐃᐃ* *ang* as well.

#### Syntactic pivot

First of all, it may be interesting to find out whether basic declarative statements in Ayeri are following a NOM-ACC or ABS-ERG pattern. A simple experiment which Comrie (1989: 111–114) performs in this regard is to test coreference in coordinated clauses. Commonly, in coordinated clauses, the subject of the second conjunct

may drop out. Thus, in English, which behaves very much in terms of NOM-ACC alignment in this regard, we get the following result:

(133) English:

- a.  $\underbrace{\text{The cat}}_A \text{ hunts } \underbrace{\text{the mouse}}_P$ .
- b.  $\underbrace{\text{The cat}}_S \text{ comes here.}$
- c.  $\underbrace{\text{The mouse}}_S \text{ comes here.}$
- d.  $\underbrace{\text{The cat}}_A \text{ hunts } \underbrace{\text{the mouse}}_P \text{ and } \underbrace{\phantom{\text{the mouse}}}_S \text{ comes here.}$
- 

In the English example in (133), *the cat* constitutes the coreferential subject in (133d). This NP is the intransitive subject S of (133b) and the agent A of (133a). English thus typically has NOM-ACC alignment, since it treats S and A alike. In an ABS-ERG language, then, we would expect the opposite case: S and P should be treated alike. In Dyirbal, we find the following situation:

(134) Dyirbal (adapted from Comrie 1989: 112):

- a.  $\underbrace{\text{balan } d\acute{r}ugumbil}_{\text{DET woman-ABS } P} \underbrace{\text{bangul } ya\acute{r}angu}_{\text{DET man-ERG } A} \text{ balgan}$   
hit  
'The man hit the woman.'
- b.  $\underbrace{\text{baya } ya\acute{r}a}_{\text{DET man-ABS } S} \text{ banin}^{\text{ru}}$   
came.here  
'The man came here.'
- c.  $\underbrace{\text{balan } d\acute{r}ugumbil}_{\text{DET woman-ABS } S} \text{ banin}^{\text{ru}}$   
came.here  
'The woman came here.'
- d.  $\underbrace{\text{balan } d\acute{r}ugumbil}_{\text{DET woman-ABS } P} \underbrace{\text{bangul } ya\acute{r}angu}_{\text{DET man-ERG } A} \text{ balgan, } \underbrace{\phantom{\text{the woman}}}_{\text{S}} \text{ banin}^{\text{ru}}$   
hit came.here  
'The man hit the woman, and [the woman] came here.'
- 

In (134), we find that *balan d'rugumbil* 'the woman' is coreferential in (134d). This is the S of (134c), and the P of (134a). Dyirbal, thus, treats S and P alike, as predicted for an ABS-ERG language—at least in this case, since Comrie (1989:

(13) also explains that 1SG and 2SG pronouns in Dyirbal behave in terms of NOM-ACC.

For Tagalog, “the deletion is not obligatory but null nominative arguments are always interpreted as referring to the nominative argument of the main clause” (Kroeger 1991: 30). Due to the way Tagalog treats subjects, however, the nominative argument can be formed by either NP in (135) with the voice marked accordingly on the verb.<sup>20</sup>

(135) Tagalog (adapted from Kroeger 1991: 31, from Ramos and Cena 1990: 151–152):

- a. *tinanong ni=Derek si=Marvin, bago umalis*  
 PFV-ask-OV GEN=Derek NOM=Marvin before PFV.AV-leave S  
 A P  
 ‘Derek asked Marvin before [Marvin] left.’
- b. *nagtanog si=Derek kay=Marvin, bago umalis*  
 PFV.AV-ask NOM=Derek DAT=Marvin before PFV.AV-leave S  
 A P  
 ‘Derek asked Marvin before [Derek] left.’

...

#### Verb agreement

...

#### Quantifier float

...

#### Relativization

...

#### Reflexive binding

...

<sup>20</sup> Thus, compare the English passive sentence *Marvin<sub>i</sub> was asked by Derek before he<sub>i</sub> left* with (135a). In English, the reference of *he* is ambiguous between the syntactic subject *Marvin* and the agent *Derek*, however. As we have seen above, though, Tagalog would also be able to make a subject of an oblique argument, not just the patient.

*Raising*

...

*Control*

...

*Control of secondary predicates*

...

*Termhood of the non-topic agent NP*

...

- Schachter (2015: 1674–1676):
  - The relation the verb marks should be called ‘trigger’, as it is inappropriate to identify it as the syntactic subject (1674).
  - Trigger system is not a voice system; calling it that disregards differences between familiar voice systems and Tagalog’s strategy (1674).
  - Properties commonly associated with the syntactic subject are divided between actor and trigger: actor is the addressee of imperatives; actor controls reference of reflexive pronoun; only triggers and “certain constituents of triggers” may be relativized; only triggers can launch floated quantifiers; only actor *or* trigger may be deleted in so-called Equi-NP constructions (i.e. control constructions, where the matrix verb passes either its S [= subject control] or its O [= object control] down as the S of the subordinate verb) (1675).
- Kroeger (1991: 19–52):
  - Floated quantifiers always modify NOM argument as the highest term (in reference to Schachter 1976: 501; presumably also referring to LFG’s functional hierarchy, cf. Bresnan et al. 2016: 230–231?) (Kroeger 1991: 23).
  - Only NOM can be relativized, gap in relative clause always corresponds to the argument role signalled by the verb; relativized NP may be of any role (23–24).
  - What number agreement on the verb there is, is with the NOM argument, independent of whether it is the actor or not (24–25).
  - Raising verbs promote the NOM argument of the complement clause, which is not necessarily the actor (exs. 13, 14) which is especially clear

with clitic pronouns (26). Non-NOM actors may be raised only when a resumptive pronoun appears in the complement clause (28). No NP in the complement clause may be marked for NOM and the raised argument in the main clause must be marked NOM 29.

- Post-verbal secondary predicates must refer to NOM if occurring in post-verbal position; the NOM argument is not necessarily the actor (29–30).
- Dropped NOM arguments in subclauses are always interpreted as referring to the coreferential NOM argument in the main clause (30–31). → relevance?
- Only genitive attributes of NOM NPs can be topicalized (31). → relevance?
- Only NOM arguments can be deleted/omitted in ‘conjunction reduction’, and never non-NOM actors or non-NOM “direct objects”; this is different from both pro-drop and right node raising (32–34).
- Control verbs typically target the subject of their complement clause, and both NOM and actor arguments are eligible for this purpose under certain circumstances, i.e. the controllee subject of a transitive complement clause is not necessarily the actor (36–38).
- Non-NOM actors are not demoted to oblique status (‘chômeurs’) like in passives, i.e. they keep their status as terms and are thus treated as objects (38 ff.). → relevance?
- Existential, meteorological, and exclamatory sentences, and experiencer verbs with indefinite objects lack NOM arguments are grammatical nonetheless. Existential and weather verbs are cross-linguistically very common contexts for subjectlessness (46–49). Applicability of an inversion construction only applies to subjects, but can also apply to the single nominal non-NOM agent argument of exclamatory sentences involving predicative adjectives 50.
- Statistically, patients appear to be preferred subjects if definite, unless a syntactic operation is applied to another argument which requires that that relation be the subject; object-voice marked verbs are also less marked than active-voice ones (53).
- Tagalog is thus basically an ergative language, but the termhood of actors and their semantic and pragmatic prominence creates the ambiguity in subjecthood mentioned above (54).
- Statistics reveal that the NOM argument is not the topic and not the focus either; those NPs marked for NOM are not especially salient for topicality (56 ff.).
- NOM NP can also bear pragmatic focus (e.g. in *wh*-questions), which would be impossible if it were a grammaticalized topic (60 ff.). *ay*-inversion patterns much better with notions of pragmatic topichood

(Kroeger 1991: 63).

#### 5.4.4 Establishing the constituent structure

- How do we know S is a constituent?
- IP above S → verb in this position can't be infinite, so inflection is an obligatory condition in this position
- Do we know whether there's a "hidden" VP in S? → test for constituency
- Tests on cline between logical SUBJ NPs vs. rest → compare with Kroeger (1991: 112) (who cites Speas 1990: 137 for criteria, cf. Speas (1990: 123–201) for whole chapter)

*Inflectional phrases*

...

*Verb phrases*

...

*Equative statements*

...

*Existential statements*

...



## A Names

---

### A.1 Given names

#### A.1.1 Masculine names

အဲဗျဲ *Ajan*  
အဲဗျဲ *Ajān*  
အဲဗဲ *Akan*  
အဲဗျဲ *Amān*  
အဲဗဲ *Apan*  
အဲဗဲ *Apican*  
အဲဗဲ *Banan*  
အဲဗဲ *Bayhan*  
အဲဗဲ *Baykan*  
အဲဗဲ *Bibān*  
အဲဗဲ *Canya*  
အဲဗဲ *Denan*  
အဲဗဲ *Diyan*  
အဲဗဲ *Gabān*  
အဲဗဲ *Hanuan*  
အဲဗဲ *Hanvan*  
အဲဗဲ *Hinvo*  
အဲဗဲ *Hiro*  
အဲဗဲ *Ijān*  
အဲဗဲ *Ikan*  
အဲဗဲ *Kadijan*  
အဲဗဲ *Kagan*  
အဲဗဲ *Kaman*  
အဲဗဲ *Kan*  
အဲဗဲ *Kolun*  
အဲဗဲ *Krui*

အဲဗဲ *Kruyan*  
အဲဗဲ *Lantān*  
အဲဗဲ *Lanyan*  
အဲဗဲ *Latun*  
အဲဗဲ *Ledo*  
အဲဗဲ *Linko*  
အဲဗဲ *Lita*  
အဲဗဲ *Mahān*  
အဲဗဲ *Makang*  
အဲဗဲ *Mangan*  
အဲဗဲ *Mangān*  
အဲဗဲ *Mangyan*  
အဲဗဲ *Maran*  
အဲဗဲ *Mican*  
အဲဗဲ *Mico*  
အဲဗဲ *Nabang*  
အဲဗဲ *Nibān*  
အဲဗဲ *Niyas*  
အဲဗဲ *Pangal*  
အဲဗဲ *Peran*  
အဲဗဲ *Pinyān*  
အဲဗဲ *Pralan*  
အဲဗဲ *Prano*  
အဲဗဲ *Pulan*  
အဲဗဲ *Saylan*  
အဲဗဲ *Sān*

အဲဗဲ *Sedan*  
အဲဗဲ *Sirtang*  
အဲဗဲ *Sopan*  
အဲဗဲ *Suhing*  
အဲဗဲ *Taboy*  
အဲဗဲ *Tang*  
အဲဗဲ *Tapan*  
အဲဗဲ *Taryan*  
အဲဗဲ *Telbān*  
အဲဗဲ *Tenan*  
အဲဗဲ *Tendan*  
အဲဗဲ *Tenyam*  
အဲဗဲ *Tikim*  
အဲဗဲ *Tipal*  
အဲဗဲ *Togas*  
အဲဗဲ *Toryan*  
အဲဗဲ *Tukong*  
အဲဗဲ *Ulang*  
အဲဗဲ *Ven*  
အဲဗဲ *Vey*  
အဲဗဲ *Veykan*  
အဲဗဲ *Vipin*  
အဲဗဲ *Virang*  
အဲဗဲ *Yan*  
အဲဗဲ *Yonang*

**A.1.2 Feminine names**

၁၀၀၀ *Agivay*  
 ၁၀၀၁ *Anang*  
 ၁၀၀၂ *Apitu*  
 ၁၀၀၃ *Apituay*  
 ၁၀၀၄ *Apitvay*  
 ၁၀၀၅ *Avan*  
 ၁၀၀၆ *Babay*  
 ၁၀၀၇ *Bamis*  
 ၁၀၀၈ *Biling*  
 ၁၀၀၉ *Binis*  
 ၁၀၁၀ *Briha*  
 ၁၀၁၁ *Caysu*  
 ၁၀၁၂ *Dembay*  
 ၁၀၁၃ *Diras*  
 ၁၀၁၄ *Dita*  
 ၁၀၁၅ *Diya*  
 ၁၀၁၆ *Gada*  
 ၁၀၁၇ *Gindi*  
 ၁၀၁၈ *Gumkay*  
 ၁၀၁၉ *Kadisui*

၁၀၂၀ *Karon*  
 ၁၀၂၁ *Kemis*  
 ၁၀၂၂ *Kumang*  
 ၁၀၂၃ *Linglay*  
 ၁၀၂၄ *Lito*  
 ၁၀၂၅ *Magaya*  
 ၁၀၂၆ *Maha*  
 ၁၀၂၇ *Mali*  
 ၁၀၂၈ *Malivay*  
 ၁၀၂၉ *Mangavay*  
 ၁၀၃၀ *Migray*  
 ၁၀၃၁ *Misan*  
 ၁၀၃၂ *Nilan*  
 ၁၀၃၃ *Ninlay*  
 ၁၀၃၄ *Niva*  
 ၁၀၃၅ *Pada*  
 ၁၀၃၆ *Pakay*  
 ၁၀၃၇ *Palay*  
 ၁၀၃၈ *Panglay*  
 ၁၀၃၉ *Paso*

၁၀၄၀ *Pila*  
 ၁၀၄၁ *Pin*  
 ၁၀၄၂ *Pituay*  
 ၁၀၄၃ *Sempay*  
 ၁၀၄၄ *Sikavay*  
 ၁၀၄၅ *Silva*  
 ၁၀၄၆ *Silvan*  
 ၁၀၄၇ *Sinlay*  
 ၁၀၄၈ *Sinvay*  
 ၁၀၄၉ *Sipra*  
 ၁၀၅၀ *Tamay*  
 ၁၀၅၁ *Taniva*  
 ၁၀၅၂ *Tavisay*  
 ၁၀၅၃ *Teping*  
 ၁၀၅၄ *Trānay*  
 ၁၀၅၅ *Tunan*  
 ၁၀၅၆ *Tuvo*  
 ၁၀၅၇ *Vala*  
 ၁၀၅၈ *Vapa*  
 ၁၀၅၉ *Vomay*

**A.1.3 Gender-neutral names**

၁၀၆၀ *Anang*  
 ၁၀၆၁ *Banvā*  
 ၁၀၆၂ *Cān*  
 ၁၀၆၃ *Cisu*  
 ၁၀၆၄ *Dikun*  
 ၁၀၆၅ *Leno*  
 ၁၀၆၆ *Maka*  
 ၁၀၆၇ *Manting*  
 ၁၀၆၈ *Ming*

၁၀၆၉ *Natran*  
 ၁၀၇၀ *Nava*  
 ၁၀၇၁ *Nilyam*  
 ၁၀၇၂ *Pangra*  
 ၁၀၇၃ *Para*  
 ၁၀၇၄ *Parān*  
 ၁၀၇၅ *Perin*  
 ၁၀၇၆ *Piba*  
 ၁၀၇၇ *Pralan*

၁၀၇၈ *Pray*  
 ၁၀၇၉ *Sutay*  
 ၁၀၈၀ *Tabi*  
 ၁၀၈၁ *Taran*  
 ၁၀၈၂ *Toran*  
 ၁၀၈၃ *Tupoy*  
 ၁၀၈၄ *Veno*  
 ၁၀၈၅ *Veta*

## B Example Texts

## B.1 The North Wind and the Sun

(From Becker 2016c)

The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shone out warmly, and immediately the traveller took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two. (After Aesop; International Phonetic Association 2007: 39)

[illegible]

*Ang manga ranyon adauyi Pintemis nay Perin, engyo mico sinyāng luga toya, lingya si lugaya asāyāng si sitang-naykonyāng kong tovaya mato. Sakantong, engongyo mico danyās palung menanang sirī ang pahongya asāya tovaley yana. Ang gihayo Pintemis minganeri-hen yona. Nay gihayong mico nay mico-eng, nay ang da-naykonya rado nay rado-eng asāya tovaley yana. Subryo deramyam ang Pintemis. Cunyo makayam mato epang ang Perin, nay ang pahya edauyikan asāya tovaley yana. Kada rua bengyo ang Pintemis, ang engyo mico cuyam Perin luga toya sam.*

- (1) *Ang manga ranyon adaui Pintemis nay Perin, engyo*  
 ang manga ran-yon adaui Ø Pintemis nay Perin eng-yo  
 AT PROG argue-3PL.N then TOP North.Wind and Sun be.more-3SG.N  
*mico sinyāng luga toya, lingya si lugaya asāyāng si*  
 mico sinyā-ang luga toya ling-ya si luga-ya asāya-ang si  
 strong who-A among 3PL.N.LOC while-LOC REL pass-3SG.M traveler-A REL  
*sitang=naykonyāng kong tova ya mato.*  
 sitang=naykon-yāng kong tova-ya mato  
 self=wrap-3SG.M.A inside cloak-LOC warm

‘The North Wind and the Sun were then arguing which among them is stronger, all the while a traveler passed by who had wrapped himself in a warm cloak.’

- (2) *Sakantong, engongyo mico danyās palung menanang sirī ang*  
 sakan=tong eng-ong-yo mico danya-as palung menan-ang si-ri<i> ang  
 agree=3PL.N be.more-IRR-3SG.N strong one-P other first-A REL<-A>-CAUS AT  
*pahongya asāya tova ley yana.*  
 pah-ong-ya asāya-Ø tova-ley yana  
 remove-IRR-3SG.M traveler-TOP cloak-P.INAN 3SG.M.GEN

‘They agreed that the first one due to whom the traveler would take off his cloak would be stronger than the other.’

- (3) a. *Ang gibayo Pintemis minganeri-ben yona.*  
 ang giha-yo Ø Pintemis mingan-eri=hen yona  
 AT blow-3SG.N TOP North.Wind ability-INS=all 3SG.N.GEN

‘The North Wind blew with all of his might.’

- b. *Nay gibayong mico nay mico-eng, nay ang da=naykonya rado nay*  
 nay giha=yong mico nay mico=eng nay ang da=naykon-ya rado nay  
 and blow=3SG.N.A strong and strong=COMP and AT so=wrap-3SG.M tight and  
*rado-eng asāya tova ley yana.*  
 rado=eng asāya-Ø tova-ley yana  
 tight=COMP traveler-TOP cloak-P.INAN 3SG.M.GEN.

‘And it blew harder and harder, and the traveler so wrapped his cloak tighter and tighter.’

- c. *Subryo deramyam ang Pintemis.*  
 Subr-yo deramyam ang Pintemis  
 give.up-3SG.N after.all A North.Wind

‘The North Wind gave up after all.’

- (4) *Cunyo makayam mato epang ang Perin, nay ang pahya edauyikan*  
 cun-yo maka-yam mato epang ang Perin nay ang pah-ya edauyikan  
 begin-3SG.N shine-PTCP warm next A Sun and AT remove-3SG.M immediately  
*asāya tova ley yana.*  
 asāya-Ø tova-ley yana  
 traveler-TOP cloak-P.INAN 3SG.M.GEN

‘Next, the Sun began to shine warmly, and the traveler immediately took off his cloak.’

- (5) *Kada rua bengyo ang Pintemis, ang engyo mico cuyam Perin*  
*kada rua beng-yo ang Pintemis ang eng-yo mico cuyam Ø Perin*  
 thus must admit-3SG A North.Wind AT be.more-3SG.N strong indeed TOP Sun  
*luga toya sam.*  
*luga toya sam*  
 among 3PL.N.LOC two

‘Thus the North Wind had to admit that the Sun was indeed the stronger among both of them.’

## B.2 The Fox and the Rooster

(Adapted from Becker 2016b)

Once upon a time, a hungry fox came to a village. He said to the rooster: “Let me hear your beautiful voice!” The proud rooster closed his eyes and crowed loudly. There the fox grabbed him and carried him into the forest. When the farmers noticed this, they ran after the fox and cried “The fox is carrying away our rooster!” There the rooster said to the fox: “Tell them, ‘I am carrying my rooster and not yours!’” The fox released the rooster from his mouth and called “I am carrying my rooster and not yours!” There, however, the rooster quickly flew onto a tree. The fox called himself a fool and trotted off. (After Aesop)

မိုးလင်း၍ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။ ချောင်းဝှမ်း၌ အလှူအတန်းများစွာရှိသည်။

*Mə-bahisya, ang sahaya runay mabo minkayya. Ang naraya aguyanya: “Garu, sa ming tangyang kadāre sekay veno vana!” Ang rimaya aguyan viyu nivajas yana nay garayāng baho. Sa da-kacisaya runayang ya nay sa ninyāng ya manga kong vinimya. Tadayya si ang kengyan bedangye adaley, ang nimpyan manga pang runayya nay babatang: “Ang manga pahya runay aguyanas nana!” Nay ang naraya aguyan runayya: “Ningu cam: ‘Sa ninyang aguyan nā; ninoyyang da-vana.’” Ang bomya runay aguyanas bantana yana nay garayāng: “Sa ninyang aguyan nā; ninoyyang da-vana.” Ang nunaya para nārya aguyan manga ling mehirya. Sitang-gasiya runayang, yāng depangas, nay lampyāng mangasara.*

- (i) a. *Mə-bahisya, ang sabaya runay mabo minkayya.*  
 mə=bahis-ya ang saha-ya runay-Ø mabo minkay-ya  
 some=day-LOC AT come-3SG.M fox-TOP hungry village-LOC  
 ‘Some day a hungry fox came to a village.’
- b. *Ang naraya aguyanya: Garu, sa ming tangyang kadāre sekay*  
 ang nara=ya.Ø aguyan-ya gara-u sa ming tang=yang kadāre sekay-Ø  
 AT speak=3SG.M.TOP rooster-LOC call-IMP PT can hear=ISG.A so.that voice-TOP  
*veno vana!*  
 veno vana  
 beautiful 2.GEN  
 ‘He spoke to a rooster: “Call, so that I can hear your beautiful voice!”’
- (2) a. *Ang rimaya aguyan viyu nivajas yana nay garayāng babo.*  
 ang rima-ya aguyan-Ø viyu niva-ye-as yana nay gara=yāng baho  
 AT close-3SG.M rooster-TOP proud eye-PL-P 3SG.M.GEN and call=3SG.M.A loudly  
 ‘The proud rooster closed his eyes and crowed loudly.’
- b. *Sa da-kacisaya runayang ya nay sa ninyāng ya manga*  
 sa da=kacisa-ya runay-ang ya.Ø nay sa nin=yāng ya.Ø manga  
 PT so=grab-3SG.M fox-A 3SG.M.TOP and PT carry=3SG.M.A 3SG.M.TOP DIR  
*kong vinimya.*  
 kong vinim-ya  
 in forest-LOC  
 ‘There he was grabbed by the fox and carried to the forest by him.’
- (3) a. *Tadayya si ang kengyan bedangye adaley, ang nimpyan*  
 taday-ya si ang keng-yan bedang-ye-Ø ada-ley ang nimp=yan.Ø  
 time-LOC REL AT notice-3PL.M farmer-PL-TOP that-P.INAN AT run=3PL.M.TOP  
*manga pang runayya nay babatang:*  
 manga pang runay-ya nay nay  
 DIR behind fox-LOC and cry.out=3PL.M.A  
 ‘As the farmers noticed, they ran after the fox and cried out.’
- b. *Ang manga pabya runay aguyanas nana!*  
 ang manga pah-ya runay-Ø aguyan-as nana  
 AT PROG take.away-3SG.M fox-TOP rooster-P ISG.GEN  
 ‘The fox is taking our rooster away!’
- (4) a. *Nay ang naraya aguyan runayya: Ningu cam:*  
 nay ang nara-ya aguyan-Ø runay-ya ning-u cam  
 and AT speak-3SG.M rooster-TOP fox-LOC say-IMP 3PL.M.DAT  
 ‘And the rooster said to the fox: “Tell them:”’

- b. *Sa ninyang aguyan nā; ninoyyang da-vana.*  
 sa nin=yang aguyan-Ø nā nin-oy=yang da=vana  
 PT carry=1SG.A rooster-TOP 1SG.GEN carry-NEG=1SG.A SO=2PL.GEN  
 ‘I am carrying my own rooster; I am not carrying yours.’
- (5) a. *Ang bomya runay aguyanas bantana yana nay garayāng:*  
 ang bom-ya runay-Ø aguyan-as banta-na yana nay gara=yāng  
 AT release-3SG.M fox-TOP rooster-P mouth-GEN 3SG.M.GEN and call=3SG.M.A  
 ‘The fox released the rooster from his mouth and called.’
- b. *Sa ninyang aguyan nā; ninoyyang da-vana.*  
 sa nin=yang aguyan-Ø nā nin-oy=yang da=vana  
 PT carry=1SG.A rooster-TOP 1SG.GEN carry-NEG=1SG.A SO=2PL.GEN  
 ‘I am carrying my own rooster; I am not carrying yours.’
- (6) a. *Ang nunaya para nārya aguyan manga ling mebirya.*  
 ang nuna-ya para nārya aguyan-Ø manga ling mehir-ya  
 AT fly-3SG.M quickly though rooster-TOP DIR on tree-LOC  
 ‘The rooster, though, quickly flew onto a tree.’
- b. *Sitang-gasiya runayang, yāng depangas, nay lampyāng mangasara.*  
 sitang=gasi-ya runay-ang yāng depang-as nay lamp=yāng mangasara  
 REFL=scold-3SG.M fox-A 3SG.M.A fool-P and walk=3SG.M.A away  
 ‘The fox scolded himself, that he were a fool, and walked away.’<sup>1</sup>

## B.3 Ozymandias

(Adapted from Becker 2011b)

### Ozymandias

I met a traveller from an antique land,  
 Who said – “two vast and trunkless legs of stone  
 Stand in the desert ... near them, on the sand,  
 Half sunk a shattered visage lies, whose frown,  
 And wrinkled lips, and sneer of cold command,  
 Tell that its sculptor well those passions read  
 Which yet survive, stamped on these lifeless things,  
 The hand that mocked them, and the heart that fed;  
 And on the pedestal these words appear:

<sup>1</sup> This sentence was translated rather literally from the German *der Fuchs schalt sich einen Narren*, literally ‘the fox scolded himself a fool’, with *einen Narren* ‘a fool’ as an object-predicative nominal.

(Shelley 2003)

[illegible]

*Sa pengalyang asano similena tado, ang naraya: Namāng sam kāryo nay taryankay bengyon adābalya. Ya hemayong kiyisa nasay adany', abalya, marinas avanu-ngas. Ang ningyon igān nay nanding dijisu yona nosānas kilisarya nay sagoyamanas: Sa layaya ban-ikan tiyanyāng da-dikun si telujong tarela, ya saprayos linyaye:<sup>2</sup>*

<sup>2</sup> Originally ᮊᮥᮒᮥᮒᮥ ᮊᮥᮒᮥᮒᮥ *ᮊᮥᮒᮥᮒᮥ* ‘survive’ (survive=3PL.N.A), which at least presently is a mistake since ᮊᮥᮒᮥ *dikun* ‘passion’ is listed in the dictionary as a *singulare tantum*. The same goes for



*sapayas si sagoyong; padangas si kondis'yong.*  
*Nay sa tahanyo eda-narān bengyamanya:*  
*Garanang nā SIMANJAS, baybiang baybiyena:*  
*Sa silvu gumo nā, nay prisu, vāng si lita!*  
*Hangara ranyareng palung. Le apanisareng*  
*ahal-nama kebay, pray, soya, litoya kayvay,*  
*miday nernanyēa eda-kiyanena nake.*

- (1) *Sa pengalyang asano similena tado, ang*  
 sa pengal=yang asano-Ø simil-ena tado ang  
 PT meet=ISG traveler-TOP country-GEN old AT  
 'I met a traveler from an old country,'
- (2) *naraya: Namāng sam kāryo nay taryankay*  
 nara=ya.Ø nama-ang sam kāryo nay taryan-kay  
 say=3SG.M.TOP leg-A two big and torso-less  
 'he said: Two big and torsoless legs'
- (3) *bengyon adābalya. Ya hemayong kiyisa*  
 beng-yon ada=ahal-ya ya hema=yong kiyisa  
 stand-3PL.N that=desert-LOC LOCT lie-3SG.N.A shattered  
 'stand in that desert. There lies shattered'
- (4) *nasay adany', abalya, marinas avanu-ngas.*  
 nasay adanya-Ø ahal-ya marin-as avanu=ngas  
 near.of that.one-TOP sand-LOC face-P sunken=almost  
 'close to there, in the sand, an almost-sunken face.'
- (5) *Ang ningyon igān nay nanding dijisu yona*  
 ang ning-yon igān-Ø nay nanding-Ø dijisu yona  
 AT tell-3PL.N frown-TOP and lips-TOP twisted 3SG.N.GEN  
 'Its frown and twisted lips tell'
- (6) *nosānas kilisarya nay sagoyamanas:*  
 nosān-as kilisarya nay sagoyaman-as  
 command-P strict and mocking-P  
 'of strict command and mockery'

ᑭᑭᑭᑭᑭ *saprayos*, which was plural ᑭᑭᑭᑭᑭ *sapratos* 'they are stamped' (stamp=3PL.N.P) before. One word of the line in the original poem was not translated as it did not fit the meter in Ayeri, "lifeless," which may be translated as ᑭᑭᑭᑭᑭ *si tenarya* 'which are unalive' (REL unalive).

- (7) *Sa layaya ban-ikan tiyanyāng da-dikun*  
 sa laya-ya ban=ikan tiyanya-ang da=dikun-Ø  
 PT read-3SG.M well=very creator-A such=passion-TOP  
 ‘Very well did the creator read such passion’
- (8) *si telujong tarela, ya saprayos linyaye:*  
 si telug=yong tarela ya sapra=yos linya-ye-Ø  
 REL survive=3SG.N.A still LOCT stamp=3SG.N.P thing-PL-Top  
 ‘which still survives, stamped into the things.’
- (9) *sapayas si sagoyong; padangas si kondis’yong.*  
 sapay-as si sago=yong padang-as si kondisa=yong  
 hand-P REL mock=3SG.N.A heart-P REL feed=3SG.N.A  
 ‘the hand that mocks; the heart that feeds.’
- (10) *Nay sa tabanyo eda-narān bengyamanya:*  
 nay sa tahan-yo eda=narān-Ø bengyaman-ya  
 and PT write-3SG.N this=word-TOP pedestal-LOC  
 ‘And this word is written on the pedestal.’
- (11) *Garanang nā Simanjas, baybiang baybiyena:*  
 garan-ang nā Simanjas bayhi-ang bayhi-ye-na  
 name-A ISG.GEN Ozymandias ruler-A ruler-PL-GEN  
 ‘My name is Ozymandias, the king of kings.’
- (12) *Sa silvu gumo nā, nay prisu, vāng si lita!*  
 sa silv-u gumo-Ø nā nay pris-u vāng si lita  
 PT see-IMP work-TOP ISG.GEN and tremble-IMP 2.A REL mighty  
 ‘Behold my work and tremble, you who are mighty!’
- (13) *Hangara ranyareng palung. Le apanisareng*  
 hang-ara ranya-reng palung le apanisa=reng  
 remain-3SG.INAN nothing-A.INAN else PT.INAN stretch=3SG.INAN.A  
 ‘Nothing else remains. It stretches’
- (14) *ahal-nama kebak, pray, soya, litoya kayvay,*  
 ahal-Ø=nama kebak pray soya lito-ya kayvay  
 sand-TOP=only lonely smooth empty border-TOP without  
 ‘only the lonely, smooth, empty sand, without borders,’

- (15) *miday nernanyēa eda-kiyanena nake.*  
 miday nernan-ye-ya eda=kiyan-ena nake  
 around part-PL-LOC this=wreckage-GEN large  
 ‘around the pieces of this large wreckage.’



## Bibliography

---

- Aikhenvald, Alexandra Y. 2006. Serial verb constructions in typological perspective. In *Serial verb constructions: A cross-linguistic typology*, edited by Alexandra Y. Aikhenvald and Robert M. W. Dixon, 1–69. Explorations in Linguistic Typology 2. Oxford: Oxford University Press.
- Bauer, Laurie. 2001. Compounding. In *Language typology and language universals: An international handbook*, edited by Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible, 1:695–707. Handbooks of Linguistics and Communication Science 20. Berlin: Walter de Gruyter.
- Becker, Carsten. 2010. The syllable structure of Ayeri's stems: A statistical analysis. Draft. April 3. Accessed June 20, 2016. <https://ayeri.de/files/frequency.pdf>.
- . 2011a. Article 1 of the United Nations 'Universal declaration of human rights'. November 22. Accessed July 23, 2016. [https://ayeri.de/wp-content/uploads/2011/12/xmp\\_humanrights.pdf](https://ayeri.de/wp-content/uploads/2011/12/xmp_humanrights.pdf).
- . 2011b. Simanjas. March 9. Accessed May 17, 2017. [https://ayeri.de/wp-content/uploads/2011/06/xmp\\_ozymandias.pdf](https://ayeri.de/wp-content/uploads/2011/06/xmp_ozymandias.pdf).
- . 2012. Imperial messages: On translating Franz Kafka's "Eine kaiserliche Botschaft" into Ayeri. August 8. Accessed July 4, 2016. <https://ayeri.de/wp-content/uploads/2012/04/2012-04-07-imperial-messages.pdf>.
- . 2013. Happy 10th anniversary, Ayeri! *Benung: The Ayeri Language Resource* (December 1). Accessed July 3, 2016. <https://ayeri.de/archives/3817>.
- . 2014. Translation challenge: The scientific method. *Benung: The Ayeri Language Resource* (April 25). Accessed August 16, 2016. <https://ayeri.de/archives/4092>.

- Becker, Carsten. 2015 [2013]. Translation challenge: The first two chapters from Saint-Exupéry's "Le Petit Prince". Accessed August 16, 2016. <https://ayeri.de/wp-content/uploads/2015/05/2013-05-25-stexupery-petitprince-ayeri.pdf>.
- . 2016a. Benung: The Ayeri language resource. Accessed June 9, 2016. <https://ayeri.de>.
- . 2016b. SE Kunst- und Plansprachen – von Esperanto bis Dothraki: Übersetzungsaufgabe (October 10). Accessed May 6, 2017. <https://ayeri.de/wp-content/uploads/2017/05/2016-10-10-uebersetzungsaufgabe.pdf>.
- . 2016c. The north wind and the sun: An Ayeri translation, revisited. Revision 1. April 26. Accessed July 16, 2016. <https://ayeri.de/wp-content/uploads/2016/04/2016-04-26-northwind-and-sun.pdf>.
- . 2016d. Verb agreement in Ayeri: Bound, clitic, or both? *Benung: The Ayeri Language Resource* (June 1). Accessed June 14, 2016. <https://ayeri.de/archives/5909>.
- . 2017. Clitics in Ayeri: Thoughts and notes. *Benung: The Ayeri Language Resource* (April 16). Accessed April 18, 2016. <https://ayeri.de/archives/6170>.
- Behaghel, Otto. 1932. *Wortstellung. Periodenbau*. Vol. 4 of *Deutsche Syntax: Eine geschichtliche Darstellung*. Germanische Bibliothek, I. Abteilung: Sammlung germanischer Elementar- und Handbücher, I. Reihe: Grammatiken 10. Heidelberg: Carl Winters Universitätsbuchhandlung. Accessed November 15, 2016. <http://digisam.ub.uni-giessen.de/diglit/deutsche-syntax-4>.
- Beyer, Stephan V. 1992. *The Classical Tibetan language*. SUNY Series in Buddhist Studies. Albany, NY: State University of New York Press.
- Bittner, Maria. 1995. Quantification in Eskimo: A challenge for compositional semantics. In *Quantification in natural languages*, edited by Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee, 59–80. Studies in Linguistics and Philosophy 54. Dordrecht: Springer. doi:10.1007/978-94-017-2817-1\_4.
- Bluhme, Friedrich, ed. 1858. Burgundionum leges: Gundobada et Papianus vulgo dictae. In *Monumenta Germaniae historica inde ab anno Christi quingentesimo usque ad annum millesimum et quingentesimum: Legum*, 3:497–630. Monumenta Germaniae historica 15. Hanover: Hahn. Accessed July 25, 2017. [http://www.dmgh.de/de/fs1/object/display/bsb00000877\\_00508.html](http://www.dmgh.de/de/fs1/object/display/bsb00000877_00508.html).
- Boga, Bella, Madita Breuninger, Thora Daneyko, and Martina Stama-Kirr. 2016. Ayeri. Universität Tübingen. June 30. Accessed July 4, 2016. <http://www.sfs.uni-tuebingen.de/-abuch/16ss/Referate/Ayeri.pdf>.

- Borsley, Robert D., Maggie Tallerman, and David Willis. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Botha, Rudolf P. 1984. *Morphological mechanisms*. Oxford: Pergamon Press.
- Bresnan, Joan, ed. 1982. *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
- Bresnan, Joan, Ash Asudeh, Ida Toivonen, and Stephen Wechsler. 2016. *Lexical-functional syntax*. 2nd ed. Blackwell Textbooks in Linguistics 16. Chichester: Wiley-Blackwell.
- Conlang mailing list: Constructed languages list. 2017. Brown University. September 14, 1998–. Accessed August 20. <https://listserv.brown.edu/conlang.html>.
- Buch, Armin. 2016. Kunst- und Plansprachen – von Esperanto bis Dothraki. Universität Tübingen. Accessed July 2. <http://www.sfs.uni-tuebingen.de/~abuch/16ss/conlang.html>.
- Butt, Miriam. 2005. Toy German grammar. Accessed July 16, 2017. <http://ling.uni-konstanz.de/pages/home/butt/main/material/german-toy.html>.
- Butt, Miriam, and Tracy Holloway King. 2015. Lexical functional grammar. In *Syntax—theory and analysis: An international handbook*, edited by Tibor Kiss and Artemis Alexiadou, 2:839–874. Handbooks of Linguistics and Communication Science 42. Berlin: De Gruyter Mouton. doi:10.1515/9783110363708-002.
- Bybee, Joan, and Paul Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. In *Frequency and the emergence of linguistic structure*, edited by Joan Bybee and Paul Hopper, 1–24. Typological Studies in Language 45. Amsterdam: John Benjamins.
- Cain, Bruce D., and James W. Gair. 2000. *Divehi (Maldivian)*. Munich: Lincom Europa.
- Carnie, Andrew. 2013. *Syntax: A generative introduction*. 3rd ed. Introducing Linguistics 4. Chichester: Wiley-Blackwell.
- Christian, Jon. 2017. When did “Neuromancer” actually take place? *Motherboard* (January 5). Accessed February 7, 2017. [https://motherboard.vice.com/en\\_us/article/when-did-neuromancer-actually-take-place](https://motherboard.vice.com/en_us/article/when-did-neuromancer-actually-take-place).
- Chung, Karen Steffen, Nathan W. Hill, and Jackson T.-S. Sung. 2014. Sino-Tibetan. In *The Oxford handbook of derivational morphology*, edited by Rochelle Lieber and Pavol Štekauer, 609–650. Oxford Handbooks in Linguistics. Oxford: Oxford University Press.

- Chung, Sandra, and James McCloskey. 1987. Government, barriers, and small clauses in modern Irish. *Linguistic Inquiry* 18 (2): 173–237. Accessed August 11, 2017. <http://www.jstor.org/stable/4178536>.
- Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
- . 1989. *Language universals and linguistic typology: Syntax and morphology*. 2nd ed. London: Blackwell.
- . 2013. Numeral bases. In Dryer and Haspelmath 2013.
- Corbett, Greville G. 1983. Resolution rules: Agreement in person, number, and gender. In *Order, concord and constituency*, edited by Gerald Gazdar, Ewan Klein, and Geoffrey Pullum, 175–206. Linguistic Models 4. Dordrecht: Foris.
- . 2006. *Agreement*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Court, Christopher. 1996. The spread of Brahmi scripts into southeast Asia. In Daniels and Bright 1996, 445–449.
- Cowan, John. 1995. Diff gram lang completed. E-Mail in reply to Christopher Collins. In *Archives of the conlang mailing list*, Archived version, May 7, 2005, in *Internet archive: Digital library of free books, movies, music and Wayback Machine*. December 16. Accessed August 20, 2017. <https://web.archive.org/web/20050507073209/http://mercury.ccil.org:80/~cowan/conlang/>.
- Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Basil Blackwell. Accessed October 5, 2016. <http://www2.ling.su.se/staff/oesten/recycled/Tense&aspectsystems.pdf>.
- Dalrymple, Mary. 2001. *Lexical-functional grammar*. Syntax and Semantics 34. San Diego, CA: Academic Press.
- Daniels, Peter T., and William Bright, eds. 1996. *The world's writing systems*. New York, NY: Oxford University Press.
- Demske, Ulrike. 2001. *Merkmale und Relationen: Diachrone Studien zur Nominalphrase des Deutschen*. Studia Linguistica Germanica 56. Berlin: De Gruyter.
- Dindelegan, Gabriela Pană, ed. 2013. *The grammar of Romanian*. Oxford Linguistics. Oxford: Oxford University Press.
- Dixon, Robert M. W. 2000. A typology of causatives. In *Changing valency: Case studies in transitivity*, edited by Robert M. W. Dixon and Aleksandra Y. Aikhenvald, 30–83. Cambridge: Cambridge University Press.



- . 2010. *Methodology*. Vol. 1 of *Basic linguistic theory*. Oxford: Oxford University Press.
- . 2010–2012. *Basic linguistic theory*. 3 vols. Oxford: Oxford University Press.
- . 2012. *Further grammatical topics*. Vol. 3 of *Basic linguistic theory*. Oxford: Oxford University Press.
- Dryer, Matthew S. 2013a. Order of subject, object and verb. In Dryer and Haspelmath 2013.
- . 2013b. Relationship between the order of object and verb and the order of adposition and noun phrase. In Dryer and Haspelmath 2013.
- Dryer, Matthew S., and Martin Haspelmath, eds. 2013. *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed July 12, 2016. <http://wals.info>.
- Eisenberg, Peter, Jörg Peters, Peter Gallmann, Cathrine Fabricius-Hansen, Damaris Nübling, Irmhild Barz, Thomas A. Fritz, Reinhard Fiehler, and Mathilde Hennig. 2016. *Die Grammatik: Unentbehrlich für richtiges Deutsch*. 9th ed. Edited by Angelika Wöllstein. Duden 4. Berlin: Dudenverlag.
- Emonds, Joseph E. 1988. *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. New York: Academic Press.
- Evans, Nicholas D. 2009. Two plus one makes thirteen: Senary numbers in the Morehead/Maro region. *Linguistic Typology* 13 (2): 321–335. doi:10.1515/LITY.2009.015.
- Everson, Michael. 2008. *Proposal for encoding the Javanese script in the UCS*. International Organization for Standardization, January 28. Accessed July 24, 2016. <http://www.evertype.com/standards/iso10646/pdf/n3319r2-javanese.pdf>.
- Fillmore, Charles J. 2003 [1968]. The case for case. Reprint. In *Papers on semantic roles*, vol. 1 of *Form and meaning in language*, edited by Charles J. Fillmore, 21–119. Chicago, IL: University of Chicago Press. Accessed August 1, 2016. <http://linguistics.berkeley.edu/~syntax-circle/syntax-group/spro8/fillmore.pdf>.
- Fleischer, Jürg. 2012. Grammatische und semantische Kongruenz in der Geschichte des Deutschen: Eine diachrone Studie zu den Kongruenzformen von ahd. *wīb*, nhd. *Weib*. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 134 (2): 163–203. doi:10.1515/pbb-2012-0002.
- Foley, William A., and Robert D. Van Valin Jr. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

- Fortmann, Christian. 2006. Deutsche Syntax in der Lexikalisch Funktionalen Grammatik. Accessed July 12, 2017. <https://www.essex.ac.uk/linguistics/external/lfg/www-lfg.stanford.edu/pubs/lfg-presentations/Fortmann-lfg-06.pdf>.
- Gaeta, Livio. 2008. Constituent order in compounds and syntax: Typology and diachrony. *Morphology* 18 (2): 117–141. doi:10.1007/s11525-009-9125-x.
- Gelderen, Elly van. 2011. The grammaticalization of agreement. In *The Oxford handbook of grammaticalization*, edited by Heiko Narrog and Bernd Heine, 491–501. Oxford Handbooks in Linguistics. Oxford: Oxford University Press.
- Gibson, William. 1995 [1984]. *Neuromancer*. Sprawl Trilogy 1. London: Voyager.
- Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In *Topic continuity in discourse: A quantitative cross-language study*, edited by Talmy Givón, 1–41. Typological Studies in Language 3. Amsterdam: John Benjamins.
- Hagège, Claude. 2010. *Adpositions*. Oxford: Oxford University Press.
- Halle, Morris. 1998. The stress of English words 1968–1998. *Linguistic Inquiry* 29 (4): 539–568.
- Halpern, Aaron. 1998. Clitics. In *The handbook of morphology*, edited by Andrew Spencer and Arnold M. Zwicky, 101–22. Oxford: Blackwell.
- Hammarström, Harald, Robert Forkel, and Martin Haspelmath, eds. 2017. Tagalog. In *Glottolog*, 3rd ed. Leipzig: Max Planck Institute for the Science of Human History. Accessed August 11, 2017. <http://glottolog.org/resource/languoid/id/taga1270>.
- Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford studies in typology and linguistic theory. Oxford: Clarendon Press.
- . 2013. Indefinite pronouns. In Dryer and Haspelmath 2013.
- Healey, Phyllis M. 1960. *An Agta grammar*. Manila: Manila Bureau of Printing. Accessed August 12, 2016. [https://archive.org/details/rosettaproject\\_dgc\\_book-1](https://archive.org/details/rosettaproject_dgc_book-1).
- International Phonetic Association. 2007. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. 9th ed. Cambridge: Cambridge University Press.
- Kager, René. 2007. Feet and metrical stress. In *The Cambridge handbook of phonology*, edited by Paul de Lacy, 195–227. Cambridge: Cambridge University Press.
- Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61 (1): 95–120.

- Kroeger, Paul R. 1991. Phrase structure and grammatical relations in Tagalog. Ph. D. thesis, Stanford University. Accessed December 17, 2016. [http://www.gial.edu/wp-content/uploads/paul\\_kroeger/PK-thesis-revised-all-chapters-readonly.pdf](http://www.gial.edu/wp-content/uploads/paul_kroeger/PK-thesis-revised-all-chapters-readonly.pdf).
- . 1993. Another look at subjecthood in Tagalog. Pre-publication draft, *Philippine Journal of Linguistics* 24 (2): 1–16. <http://www.gial.edu/documents/Kroeger-Subj-PJL.pdf>.
- . 2007. McKaughan's analysis of Philippine voice. In *Piakandatu ami Dr. Howard P. McKaughan, 41–*, edited by Loren Billings and Nelleke Goudswaard. Manila: Linguistic Society of the Philippines / SIL Philippines.
- Kroonen, Guus. 2013. *Etymological dictionary of Proto-Germanic*. Edited by Alexander Lubotsky. Leiden Indo-European Etymological Dictionary Series 11. Leiden: Brill.
- Kuipers, Joel C., and Ray McDermott. 1996. Insular southeast Asian scripts. In Daniels and Bright 1996, 474–484.
- Lehmann, Christian. 2015. *Thoughts on grammaticalization*. 3rd ed. Classics in Linguistics 1. Berlin: Language Science Press. Accessed July 18, 2016. <http://langsci-press.org/catalog/book/88>.
- Li, Charles N., ed. 1976. *Subject and topic*. New York, NY: Academic Press.
- Lieber, Rochelle. 1994. Root compounds and synthetic compounds. In *The encyclopedia of language and linguistics*, edited by Ronald E. Asher, 7:3607–3610. Oxford: Pergamon Press.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Maddieson, Ian. 2013a. Consonant inventories. In Dryer and Haspelmath 2013.
- . 2013b. Consonant-vowel ratio. In Dryer and Haspelmath 2013.
- . 2013c. Vowel quality inventories. In Dryer and Haspelmath 2013.
- McCloskey, James, and Ken Hale. 1984. On the syntax of person-number inflection in modern Irish. *Natural Language and Linguistic Theory* 1 (4): 487–534. doi:10.1007/BF00417057.
- Nordlinger, Rachel. 1998. Constructive case: Evidence from Australian languages. Doctoral dissertation, Stanford University.
- Oxford English dictionary. 2016. Oxford University Press. Accessed August 23, 2016. <http://www.oed.com>.

- ParGram/ParSem: An international collaboration on LFG-based grammar and semantics development. 2009–2016. Accessed January 6, 2017. <http://pargram.b.uib.no>.
- Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Peterson, David J. 2015. *The art of language invention: From Horse-Lords to Dark Elves, the words behind world building*. New York, NY: Penguin.
- Plank, Frans. 1992. Possessives and the distinction between determiners and modifiers (with special reference to German). *Journal of Linguistics* 28 (2): 453–468. Accessed May 24, 2017. doi:10.1017/S0022226700015292. [http://ling.uni-konstanz.de/pages/home/plank/for\\_download/publications/87\\_Plank\\_PossessivesDeterminersModifiers\\_1992.pdf](http://ling.uni-konstanz.de/pages/home/plank/for_download/publications/87_Plank_PossessivesDeterminersModifiers_1992.pdf).
- Ramos, Teresita V., and Resty M. Cena. 1990. *Modern tagalog: Grammatical explanations and exercises for non-native speakers*. Honolulu, HI: University of Hawaii Press.
- Sadler, Louisa. 1997. Clitics and the structure-function mapping. In *Proceedings of the LFG '97 conference*, edited by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications. Accessed August 12, 2017. <https://web.stanford.edu/group/cslipublications/cslipublications/LFG/2/lfg97sadler.pdf>.
- Salomon, Richard G. 1996. Brahmi and Kharoshthi. In Daniels and Bright 1996, 373–383.
- Schachter, Paul. 1976. The subject in Philippine languages: Topic, actor, actor-topic, or none of the above? In Li 1976, 493–518.
- . 2015. Tagalog. In *Syntax—theory and analysis: An international handbook*, edited by Tibor Kiss and Artemis Alexiadou, 3:1658–1676. *Handbooks of Linguistics and Communication Science* 42. Berlin: De Gruyter Mouton. doi:10.1515/9783110363685-007.
- Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog reference grammar*. Berkeley, CA: University of California Press. Accessed August 2, 2016. <http://books.google.com/books?id=E8tApLUNy94C>.
- Shelley, Percy Bysshe. 2003. Ozymandias. In *Representative poetry online*, 3rd ed., edited by Ian Lancashire and J. D. Robins. Toronto, ON: University of Toronto. Accessed May 17, 2017. <https://tspace.library.utoronto.ca/html/1807/4350/poem1904.html>.

- Simpson, Jane H. 1991. *Warlpiri morpho-syntax: A lexicalist approach*. Dordrecht: Kluwer Academic Publishers. doi:10.1007/978-94-011-3204-6.
- Sobin, Nicholas. 2011. *Syntactic analysis: The basics*. Chichester: Wiley Blackwell.
- Speas, Margaret. 1990. *Phrase structure in natural language*. Studies in Natural Language and Linguistic Theory 21. Dordrecht: Kluwer Academic.
- Spencer, Andrew, and Ana R. Luís. 2012. *Clitics: An introduction*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Süddeutsche Zeitung. 2013. Abschied vom “RkReÜAÜG”-Ungetüm: Längstes deutsches Wort verschwindet. *Süddeutsche Zeitung* (June 3). Accessed August 10, 2016. <http://www.sueddeutsche.de/kultur/laengstes-deutsches-wort-verschwindet-abschied-vom-rkreueaug-ungetuem-1.1687160>.
- Teoh, H. S. 2003. Law of ten times worse natlangs. *Conlang Archives* (March 5). Accessed May 20, 2017. <https://listserv.brown.edu/archives/cgi-bin/wa?A2=conlang;5d8981b3.0303A>.
- Department history. 2017. UCLA Department of Linguistics. Accessed September 2, 2017. <http://linguistics.ucla.edu/about/history/>.
- United Nations. 1948. Universal declaration of human rights. United Nations. Accessed July 23, 2016. <http://www.un.org/en/universal-declaration-human-rights/>.
- Wälchli, Bernhard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press.
- Wasow, Thomas. 1997. Remarks on grammatical weight. *Language Variation and Change* 9:81–105.
- Wechsler, Steven. 2009. “Elsewhere” in gender resolution. In *The nature of the word: Studies in honor of Paul Kiparsky*, edited by Kristin Hanson and Sharon Inkelas, 567–586. Current Studies in Linguistics. Cambridge, MA: MIT Press.
- Wiltshire, Caroline, and Alec Marantz. 2000. Reduplication. In *Morphology: An international handbook on inflection and word-formation*, edited by Geert Booij, Christian Lehmann, and Joachim Mugdan, 1:557–567. Handbooks of Linguistics and Communication Science 17. Berlin: Walter de Gruyter.
- Zipf, George K. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Boston, MA: Houghton Mifflin. Accessed August 24, 2016. <http://hdl.handle.net/2027/mdp.39015008729983>.

- Zompist BBoard: A congress of convoluted conworldery. 2017. Accessed August 20. <http://www.incatena.org>.
- Zwicky, Arnold M. 1977. *On clitics*. Bloomington, ID: Indiana University Linguistics Club. Accessed July 22, 2016. [https://web.stanford.edu/~zwicky/on\\_clitics.pdf](https://web.stanford.edu/~zwicky/on_clitics.pdf).
- Zwicky, Arnold M., and Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *N'T*. *Language* 59 (3): 502–513. Accessed July 21, 2016. <https://web.stanford.edu/~zwicky/ZPCliticsInfl.pdf>.

# Index

---

- adjectives, 167–170
- adpositions, 77, 170–179
  - of time, 178–179
- adverbs, 214–223
  - attributive, 214–216
  - degree and quantity, 217–218
  - demonstrative adverbs, 222–223
  - sentence adverbs, 219–222
- agglutination, 63
- agreement
  - number, 180–192
  - person, 180–192
- allomorphy, 193, 195, 204
- allophony
  - of consonants, 7–10, 40
  - of vowels, 10–12, 19
- aspect, 198–200
  - habitual, 200
  - iterative, 201–202
  - progressive, 72, 73, 199–200
- Box script, 45
- Brāhmī scripts, 45
- cases, 75, 113–125
  - agent, 113–115
  - causative, 121–122
  - dative, 9, 117–118
  - genitive, 102, 119–120
  - instrumental, 122–124
  - locative, 120–121
  - patient, 115–117
- clitics, 64, 71–102
- comparison
  - of adjectives, 168
  - of adverbs, 215
- complement clause, 42
- compounds, 38–40, 127–137
  - coordinative, 135
  - endocentric, 129–133
  - exocentric, 136
  - synthetic, 133–134
- conjunctions, 233–237
  - kamo ... nay*, 235–236
  - kamo ... soyang*, 236–237
  - nay*, 234
  - soyang*, 234–235
- consonants, 7–10, 16, 19, 20, 22–24, 26, 28, 47–48
  - clusters, 16, 23, 30–31
- dependent marking, 102
- diacritics, 50–56
- diminutive, 70
- diphthongs, 10, 19, 33, 52
- discourse particles, 72
- gender, 107–110
- gerund, 141–144
- goals, 3, 4

- marking strategies, 102
- modal verbs, 72, 74
- modals, 206–209
- mood, 203–206
  - hortative, 206
  - imperative, 205–206
  - irrealis, 203–204
  - negative, 204–205
- morphophonology
  - of relative pronouns, 9, 11, 71
  - of the plural marker, 7, 9, 111
- negation
  - multiple negation, 205
  - of adjectives, 168–169
  - of verbs, 204–205
- nominalization, 139–144
- nouns, 38, 107–144
- number, 110–113
- numerals, 112, 223–233
  - cardinal, 224–231
  - digits, 56
  - distributive, 232
  - fractions, 227–229
  - integers, 224–227
  - multiplicative, 231
  - ordinal, 229, 229
  - ranges, 232–233
- participle, 209–211
- plural, 7, 9, 111
- postpositions, 176
- prefixes, 64, 66–69
  - on nouns, 125–127
- prepositions, 171–176
- pronouns, 144–167
  - demonstrative, 68, 148–152
  - indefinite, 155–162
  - interrogative, 153–155
  - personal, 65, 144–147
  - possessive, 147–148
  - reciprocal, 167
  - reflexive, 67, 165–167
  - relative, 71, 163–165
- punctuation, 57–59
- quantifiers, 38
- questions
  - closed, 41
  - open, 41
  - tag questions, 215–216
- reduplication, 70, 137–139
- relative clause, 42
- stress, 31–38
- suffixes, 63, 69–70
- suprasegmental, 64, 71
- syllable
  - final, 23–26
  - heavy, 33–38
  - initial, 16–19
  - medial, 20–23
  - monosyllabic words, 26–28, 38
  - superheavy, 34, 37–38
  - types, 14, 33
- Tahano Hikamu, 3, 45–61
  - angular style (ṣūṣ), 60
  - blackletter style, 61
  - hand style (ḥḥ), 60
  - regular/book style, 59
- tense, 67, 192–198
  - future, 195–198
  - past, 10, 193–198
  - present, 193
- topic, 43, 72, 73, 114
- typology
  - of compounds, 127
  - of morphemes, 63–105
- verbs, 179–214
- vowels, 10–12, 16, 20, 23, 28, 48–50