

A Grammar of Tawrã Mishmi

Rolf Stephan Hotz Molina

A thesis submitted to fulfill requirements for the degree of Doctor of Philosophy

Department of Linguistics

Faculty of Arts and Social Sciences

The University of Sydney

June 2025

Abstract

This is a comprehensive grammar of Tawrã Mishmi (simply Tawrã [ta¹.ɾwã]), an indigenous Tibeto-Burman language from the Eastern Himalayas. Most speakers reside in Lohit and Anjaw districts, at the easternmost tip of Arunachal Pradesh. The data for this grammar came from four field trips to India between 2022 and 2024. Tawrã is closely related to Kera'a and together they form the Kera'a-Tawrã subgroup within the Tibeto-Burman language family. Tawrã is furthermore in strong contact with Kaman, with which it is related on a higher phylogenetic level. These three speaker communities form the *Mishmi* ethnic group.

Tawrã has 25 consonant phonemes and two marginal phonemes with a three-way distinction among stops and affricates (voiceless non-aspirated, voiced non-aspirated, voiceless aspirated). There are six phonemic vowels and four rhymes. As a tonal language, it makes a four-way distinction between high-falling, high, low and mid tone. The tone- and stress-bearing unit is the word. This means that disyllabic words receive stress/tone on one syllable, thus forming two groups: a) stress/tone may fall on the second syllable. This is the most common pattern in the language and all four tones available. B) stress/tone may fall on the first syllable, although here only high-falling tone is available.

Three major word classes in Tawrã are nouns, verbs and adjectives. These are distinguished by morphosyntactic features, e.g. by the suffixes that are available for each word class. Cross-class membership is an extended feature of the language: many roots exist as verbs, nouns and adjectives.

Many nouns display a root structure that is synchronically derived from an etymological prefix + root collocation. This pattern is most common for body parts, flora and fauna and material culture.

Tawrã has a large number of classifiers. Classifiers trigger a reduced form of the numeral 'one' as opposed to nouns, which are counted with the full numeral. Classifiers are engaged in counting but also have other grammatical functions: they are employed for anaphora and individuation, especially if they occur *bare*, i.e. without a numeral.

While most members are transparently related to nouns, there are some classifiers where this link is no longer evident. This suggests that the system of classifiers is old in the language. Classifiers also participate in word formation, as evidenced in several *classifier compounds*: lexicalized compounds built by a noun followed by a classifier.

A predicate in Tawrã consists minimally of a verbal stem. The predicate may be optionally expanded into a complex predicate by one or more *predicate derivations* that occur between the stem and the final suffixal chain. Predicate derivations are morphologically bound derivational morphemes that occur between the verbal stem (head of the predicate) and the *tense, aspect, mode and number* (TAMN) suffixes. Predicate derivations modify the semantics of the predicate in terms of aspect/Aktionsart, manner and valency. Most predicate derivations are related to independent verbs. The predicate can also take TAMN suffixes. These resemble inflectional suffixes in other languages but, crucially, there is no obligatory agreement/argument cross-reference in the language. Rather, TAMN suffixes exhibit “person-based sensitivities” (cf. Post 2007, 689).

Multi-clause constructions are dominated by the cosubordinator marker *gũ*. Clauses marked by *gũ* modify a final clause (not marked by *gũ*), forming a discourse episode (cf. Givón 2001b, 355). The structure is thus flat, i.e. not nested, as all non-final (*gũ*) clauses modify the final clause. Cosubordinate constructions are exploited for diverse pragmatic effects, e.g. reason-cause, recapitulation or background. Conversely, several non-final markers in Tawrã indicate subordination, here understood as a nested structure in which a subordinated clause can only modify its superordinate clause. Some subordinating constructions are temporal (‘when’), conditional (‘if’) and purposive (‘in order to’). Nominalization in Tawrã is employed for building relative clauses, event nominalization and location nominalization. It should be noted that only the latter has a dedicated morpheme; both the relative clause and event nominalization rely on syntax alone.

This grammar includes a lexicon of approximately 1600 entries, three glossed excerpts from texts and a list of all classifiers found in the language along with etymological notes and comparisons.

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Rolf Hotz, 22.06.2025

The research reported in this thesis was supported by the award of a Faculty of Arts and Social Sciences Postgraduate Research Scholarship

Acknowledgments

This grammar is the product of many helping hands. I would like to thank the Tawrã communities of Sunpura and *Tajap* (Tezu), Arunachal Pradesh for their continuous help and patience. Therein, I would like to thank the members of the Cultural and Literary Society of Mishmi (CALSOM) for their help. A particular thank you goes to those who took the time to sit with me and tried to help me understand their language: Barup Chaitom, M. Ritti, Soling Bellai, Badamso Tayang, Chogum Tayang, Dibru Cathrin Bu, Bisamsong Tayang, Badamso Tayang, Soronso Chaitom, Juhailum Tayang, Thalaso Karap, Wanlu Talampai, Amanso Tayang and many more!

The Prashad/Patel family in Sunpura, Assam gave me shelter for many months during my stay in India. We ate together, celebrated *pujas*, lived through so many things and I always felt at home. आप लोग को बोहोत धन्यवाद।

I am in immense debt to many people from Rajiv Gandhi University (RGU) in Itanagar: Prof. Jumyir Basar and Dr. Zilpha Modi came to my aid in the most desperate of times. I would also like to thank Dr. Yankee Modi, Prof. Marpe Sora, Dr. Dr. Nani Tamang Jose and Dr. Wanglit Mongchan for their help in sorting the ordeals of Indian bureaucracy. Other people at RGU with whom I had the chance to discuss specific issues regarding this grammar include Dr. Lisa Lomdak, Dr. Tarun Mene, Hari Apralo, Tage Umbrey, Dr. Kaling Dabi, Dr. Hermant Gambre, Kulenso Pul and Ometso Ngadong.

This work has benefited from academic exchange in the form of conversations, comments, corrections and encouragement. I would like to thank Prof. Fernando Zúñiga, Prof. Nick Enfield, Dr. Jonathan Evans and Karl Hotz. A particular thank you goes for Agnes Conrad for her help editing this thesis. My dissertation committee provided valuable comments and insights: thank you to Prof. Linda Konnerth and Prof. Guillaume Jacques.

To my fellow students from the University of Sydney, thank you for helping me make my stay there most enjoyable: Chintan Seth, Sharmi Barua, Dr. Manuel David González Pérez, Dr. Carl Bodnaruk and Sharmi Barua. I also want to thank Naomi Peck from the University of Freiburg for our fruitful discussions.

I would like to give a further word of thank you to those speakers of Tawrã who spent considerable time with me working on the language. This often meant for them to spend long hours sitting around a recording machine, enduring endless lists of impossible questions and hearing the same recording again and again. These are: Jabralum Chaitom, Amit Chaitom, Ajitso Ama, Bablin Awe, Sawimso Tayeng, Sahiblu Tayang, Thagnem Tayang, Satolum Talampai and Batamso Talampai. To all of you, *khodemu!*

Adv. Johakso Manyu from Lulyang (Loiliang) deserves a special mention. As one of the first members of the Tawrã community who started documenting his language, Johakso was a strong supporter of this project from the first day. Not only did he spend many hours working with me, he also helped me navigate through local politics and allowed me to stay at his wonderful place. Without his constant help and commitment, this grammar could not have been possible.

I am very grateful to my supervisor Dr. Mark W. Post. Mark's dedication to the documentation of languages and culture has been an inspiration to me, even before I started my PhD. Throughout the multiple drafts of this grammar, Mark was both a meticulous and insightful commenter. Thank you for encouraging me to pursue this work, even at times of uncertainty.

Finally, my most sincere gratitude goes to my wife Isidora, who has supported me through thick and thin. ¡Gracias por ser el faro de este bote a la deriva!

Contents

1	Introduction	24
1.1	Linguistic situation	25
1.1.1	Genealogical affiliation	25
1.1.2	The ‘Mishmi group’ and multilingualism	26
1.1.3	Contact with Austroasiatic	29
1.1.4	Other neighbors	35
1.1.5	Relation to the Macro-Tani group	37
1.2	Previous literature on Tawrã	39
1.3	A sketch of Tawrã culture and society	43
1.4	Theoretical framework	45
1.5	Fieldwork and data	46
1.6	Main typological features	50
1.6.1	Phonology	51
1.6.2	Word classes	51
1.6.3	Morphology	52
1.6.4	Multi-clause constructions	52
1.6.5	Morphosyntax	53
1.6.6	Intersections of language, culture and cognition	53
1.7	Transcription, orthography and conventions	54
1.7.1	Consonants	55
1.7.1.1	Consonant clusters	55
1.7.2	Vowels	56
1.7.3	Tone	57
1.7.4	Other conventions	59

2	Phonology	60
2.1	Overview	60
2.2	Syllable structure and wordhood	63
2.2.1	Moraic structure and monosyllabic wordhood	63
2.2.1.1	Syllabification and sonority	64
2.2.2	Heavy syllable (H)	66
2.2.3	Syllable onset	69
2.2.3.1	Onset consonant clusters	70
2.2.4	Syllable coda	73
2.2.4.1	Alternating rhymes	74
2.2.4.2	Devoiced codas	76
2.2.5	Light syllables (L)	79
2.2.6	Tone and metrical structure: tonal patterns	82
2.2.7	The phonological word	87
2.2.7.1	Phonological dependency or cliticization	92
2.3	Vowels and rhymes	92
2.3.1	Previous work	93
2.3.2	Monophthongs	93
2.3.2.1	Vowel length	98
2.3.2.2	Nasalization of monophthongs	101
2.3.2.3	Vowels and nasal codas	102
2.3.3	Rhymes	104
2.4	Consonants	106
2.4.1	Previous work	106
2.4.2	Overview	106
2.4.3	Oral stops (excluding glottal stops)	107
2.4.3.1	Labial stops	114
2.4.3.2	Alveolar stops	114
2.4.3.3	Velar stops	116
2.4.4	Glottal stop	116

2.4.5	Affricates	119
2.4.5.1	Alveolar affricates	119
2.4.5.2	Alveo-palatal affricates	120
2.4.6	Nasals	121
2.4.7	Fricatives	123
2.4.8	Lateral	123
2.4.9	Approximants	124
2.5	Tone	126
2.5.1	Tone in NEI languages	126
2.5.2	Previous work on tone	128
2.5.3	Tone on monosyllables	130
2.5.3.1	High-falling	131
2.5.3.2	Mid tone	139
2.5.3.3	Low tone	143
2.5.3.4	High tone	144
2.5.4	Toneless suffixes	145
2.5.5	Tonal patterns: disyllabic words	147
2.5.5.1	High-falling+mid (HF-M)	147
2.5.5.2	Mid+high-falling (M-HF)	148
2.5.5.3	Mid+high (MH)	149
2.5.5.4	Mid+mid (MM)	150
2.5.5.5	Mid+low (ML)	151
2.5.6	Tonal patterns: trisyllabic morphemes	152
2.5.6.1	High-falling+mid+mid	153
2.5.6.2	Mid+high-falling+mid	154
2.5.6.3	Mid+high+mid	155
2.5.7	Tonal change in nominal modification	156
2.6	Morpho-phonology	160

3	Morphology and Lexicon	161
3.1	Basic morphological units	161
3.2	Overview of major open classes	163
3.2.1	Cross-class membership	165
3.2.2	Nouns	168
3.2.3	Verbs	169
3.2.4	Adjectives	170
3.3	Semantics and subgroups of open word classes	174
3.3.1	Nouns	174
3.3.1.1	Relator nouns	174
3.3.1.2	Kinship terms	177
3.3.1.3	Class terms	181
3.3.1.4	Collective suffix <i>làng</i>	183
3.3.2	Verbs	186
3.3.2.1	Verb transitivity	186
3.3.2.2	Event semantics	188
3.3.2.3	Humanness and animacy restrictions in verb selection	193
3.3.3	Adjectives	194
3.4	Other open classes	196
3.4.1	Adverbs	196
3.4.1.1	Classification of adverbs	200
3.4.1.2	Ideophones	203
3.4.2	Classifiers	204
3.5	Closed classes	205
3.5.1	Pro-forms	205
3.5.1.1	Personal pronouns	206
3.5.1.2	Interrogative pronouns	207
3.5.1.3	Indefinite pronouns of person <i>mě</i> ‘someone, somebody’ . . .	211
3.5.1.4	Reflexive pronoun <i>ga</i>	212
3.5.1.5	Quantifier <i>mé</i>	214

3.5.1.6	Other indefinite pronouns	215
3.5.2	Particles	216
3.5.2.1	Modal particles	217
3.5.2.2	Filler	218
3.5.2.3	Interjections	218
3.5.2.4	Postpositions	218
3.5.2.5	Non-final markers	219
3.5.3	Numerals	220
3.5.4	Demonstrative <i>we</i> and complex demonstratives	221
3.6	Reduplication	221
3.6.1	Reduplication of question words and pronouns	222
3.6.2	Reduplication of adjectives	222
3.6.3	Reduplication of adverbs	223
3.6.4	Reduplication of nouns and NPs	223
3.7	Diachronic aspects of root morphology	225
3.7.1	a- root prefix	228
3.7.1.1	A brief note on *a-/*ha- prefixes	229
3.7.1.2	a- root prefix body parts	229
3.7.1.3	a- root prefix flora & fauna	230
3.7.1.4	a- root prefix: kinship terms	231
3.7.1.5	a- root prefix: directions and deictic terms	232
3.7.1.6	a- root prefix: diverse verbs	232
3.7.1.7	a- root prefix: material culture	233
3.7.2	ha- root prefix	233
3.7.2.1	ha- root prefix: flora	233
3.7.2.2	ha- root prefix: body parts	234
3.7.2.3	ha- root prefix: diverse verbs	234
3.7.3	ka- root prefix	235
3.7.3.1	ka- root prefix: small animals and insects	235
3.7.3.2	ka- root prefix: adverbs, classifiers and numerals	237

3.7.4	me- root prefix	238
3.7.5	ma- root prefix	239
3.7.6	ta- root prefix	240
3.7.6.1	ta- root prefix: flora & fauna	240
3.7.6.2	ta- root prefix: bodies of water and related terms	242
3.7.6.3	ta- root prefix: material culture	242
3.7.7	we- root prefix	244
4	Noun phrase	245
4.1	NP structure	245
4.1.1	Maximal structure	245
4.1.1.1	NP head	246
4.1.2	Restricted subtypes	247
4.1.2.1	NP headed by a personal name	247
4.1.2.2	NP headed by a pronoun	248
4.1.3	Headless NP	249
4.2	Nominal compounds and echo formations	251
4.2.1	Hierarchical nominal compounds	251
4.2.2	Symmetrical nominal compounds	254
4.2.3	Echo formations: partial reduplications of nouns	256
4.3	NP-head modifiers	258
4.3.1	Adjectives	258
4.3.1.1	Small/diminutive <i>a(qa)</i>	263
4.3.2	Classifier phrase and numerals	264
4.3.3	Demonstrative <i>we</i>	265
4.3.3.1	Basic features	266
4.3.3.2	Pragmatic functions	268
4.4	Conjoined NPs	273
4.4.1	Additive <i>ma</i>	273
4.4.2	Asyndetic coordination	274

4.5	Nominal affixes	276
4.5.1	‘place of many’ <i>-pa</i>	276
4.5.2	Feminine suffix <i>-yǎ̀</i>	277
4.5.3	Prefix for kinship terms <i>na-</i>	279
5	Classifiers	281
5.1	Basic properties of classifiers	282
5.2	Syntactic constituency	284
5.3	Functions	286
5.4	Origin of classifiers	290
5.5	Classifier compounds	292
6	Predicate complex	295
6.1	Predicate structure	295
6.1.1	Stem types	297
6.1.1.1	Simple stem	297
6.1.1.2	Compound stem	298
6.1.1.3	Discontinuous stems	299
6.1.1.4	Stem reduplication	300
6.2	Predicate derivations	303
6.2.1	Theoretical considerations: inflection, derivation and predicate deriva- tions	304
6.2.1.1	Inflection, derivation	304
6.2.2	Valency operating predicate derivations	308
6.2.2.1	Associative <i>-né/nê</i>	308
6.2.2.2	Causative predicate derivations	309
6.2.2.2.1	Causative <i>-jǐ</i>	310
6.2.2.2.2	Volitional causative/permissive <i>-shyüw</i>	311
6.2.2.3	Bene/malefactive <i>-hàng</i>	312
6.2.2.4	Permissive <i>-shyüw</i>	314
6.2.3	Manner/modality predicate derivations	315

6.2.3.1	Inopinative <i>-ta</i>	315
6.2.3.2	Frustrative <i>-ha</i>	315
6.2.3.3	‘by force’ <i>-mazi</i>	316
6.2.3.4	‘at last’ <i>-ha</i>	317
6.2.3.5	‘in exchange’ <i>-dá</i>	318
6.2.3.6	‘thoroughly’ <i>-makáw</i>	318
6.2.4	Aspect/Aktionsart predicate derivations	319
6.2.4.1	Permansive <i>-(ta)lá</i>	319
6.2.4.2	Ambulative <i>-wi</i>	320
6.2.4.3	Completive <i>-dâng</i>	321
6.3	Tense, aspect, mode and number (TAMN) suffixes	323
6.3.1	Relevant concepts	325
6.3.1.1	Egophoricity (and allophoricity)	325
6.3.1.2	Tense and aspect	332
6.3.1.3	Suffixal chain	333
6.3.2	Number	335
6.3.2.0.1	Plural suffixes <i>-ga</i> and <i>-ring/-rẽ</i>	335
6.3.2.0.2	Dual marker <i>-nega</i>	339
6.3.3	Associated motion suffixes	340
6.3.3.1	Directive <i>-zâ</i>	340
6.3.3.2	Go (and come back) <i>-ge</i>	343
6.3.4	Aspect/Aktionsart	344
6.3.4.1	‘ahead’ <i>gadyô</i>	344
6.3.4.2	Continuative <i>-q</i>	344
6.3.4.3	Completive <i>-kũ</i>	345
6.3.4.4	Iterative <i>-mâ</i>	347
6.3.4.5	Stative <i>-ho</i>	349
6.3.4.6	Intentional (egophoric) <i>-ne</i>	350
6.3.4.7	Egophoric habitual <i>-de/di/d</i>	351
6.3.4.8	The <i>bô</i> group	352

6.3.4.8.1	Perfective <i>-dyibô/-bô</i>	353
6.3.4.8.2	‘and go/go to’ suffix	355
6.3.4.8.3	Typological and diachronic overview	355
6.3.4.9	Recent past <i>-li/lü</i>	358
6.3.4.10	Certain future (egophoric) <i>-(vi)dyô</i>	359
6.3.4.11	‘permanently’ <i>-tá</i>	360
6.3.5	Command mood suffixes	360
6.3.5.0.1	Imperatives <i>-na</i> and <i>-tè</i>	361
6.3.5.0.2	Prohibitive <i>-yâq</i>	362
6.3.5.0.3	Hortatives	363
6.3.5.0.3.1	Singular hortative <i>-ta</i>	363
6.3.5.0.3.2	Hortative <i>-kéq</i>	363
6.3.5.0.4	Jussive <i>-tyêyu</i>	364
6.3.6	Other TAMN suffixes	365
6.3.6.1	Allophoric suffix <i>-ya</i>	365
6.3.6.2	Factitive <i>-na</i>	366
6.3.6.2.1	The term <i>factitive</i> and its use in related languages	373
6.4	Negation	375
6.4.1	General negator <i>yüm</i>	375
6.4.2	Observational verbal negator <i>shyüw</i>	378
6.4.3	Obsolesced negator <i>-ma</i>	381
6.5	Nominal incorporation	381
7	Clause types	385
7.1	Verbless clause	385
7.1.1	Locative construction	390
7.2	Predicate clauses	392
7.2.1	Constituents of a predicate clause	393
7.2.1.1	Arguments	393
7.2.1.2	Adjuncts	394

7.2.2	Predicate clause based on argument structure	396
7.2.2.1	Atransitive clause	397
7.2.2.2	Intransitive clauses	397
7.2.2.3	Transitive clauses	399
7.2.3	Predicative possessive construction	400
7.3	Non-declarative clauses	401
7.3.1	Interrogative clauses	401
7.3.1.1	Polar tag questions	402
7.3.1.1.1	Disjunction construction with <i>kya</i>	404
7.3.1.2	Content question	405
7.3.1.2.1	Rhetorical and wondering particle <i>palang/(sa)mlang</i>	406
7.3.2	Command moods	408
8	Semantics and pragmatics of clause constituent core arguments	409
8.1	Key concepts	410
8.2	Argument selectors	412
8.2.1	Predicate marking	413
8.2.2	Addressee in imperatives	416
8.2.3	Constituent order	417
8.2.4	Marking of core argument NPs	421
8.2.5	Omitted arguments in multi-clause constructions	422
8.3	Topic and focus	426
8.3.1	Topic construction	428
8.3.1.1	Topic construction with <i>bye</i>	428
8.3.1.2	Contrastive topic <i>ma</i>	432
8.3.2	Focus particle <i>nyu</i>	434
9	Particles	438
9.1	Preliminaries	438
9.2	Modal particles	440
9.2.1	‘maybe’ <i>matyô</i>	441

9.2.2	Inferential evidential <i>de</i>	444
9.2.3	‘unfortunately’ <i>pà</i>	445
9.2.4	Non-visual evidential <i>kh râk</i>	445
9.2.5	‘sadly/ for all I care’ <i>thâ</i>	448
9.2.6	‘it seems’ <i>delé</i>	449
9.2.7	‘unexpected/controversial’ <i>kho</i>	450
9.2.8	‘counter-expectation’ <i>ki/kale/kalai/ke</i>	454
9.2.9	Advisative <i>ta</i>	455
9.3	Fillers	456
9.4	Interjections	457
9.5	Postpositions	458
9.5.1	General locative/instrumental <i>gũ</i>	460
9.5.2	‘inside’ locative <i>ke</i>	462
9.5.3	‘atop’ <i>(ta)kyã</i>	464
9.5.4	‘with’ <i>dũ</i>	464
9.5.5	‘for’ <i>(a)ba</i>	465
9.5.6	Comparative ‘like’ <i>dahi</i>	465
9.5.7	NP quantifiers ‘each, every’	466
9.6	Non-final markers	467
10	Multi-clause constructions	469
10.1	Introduction	469
10.2	Conjoined clauses	470
10.2.1	Semantic functions of conjoined clauses	471
10.3	Cosubordination	472
10.3.1	Sequential interpretation	477
10.3.1.1	Reason-cause	478
10.3.1.2	Recapitulation	480
10.3.1.3	Non-canonical use of cosubordinate clauses	482
10.3.2	Simultaneous interpretation	484

10.3.3	Negation in cosubordination	486
10.3.4	Double marker with <i>gũ</i>	488
10.3.4.1	Co-temporality marker <i>dũ</i>	489
10.3.4.2	Concomitant action <i>gũ thĩ</i>	490
10.3.5	Grammaticalization path	492
10.4	Subordinating constructions	493
10.4.1	Temporal and conditional subordination <i>bô</i> , <i>bobyẽ</i> and <i>byẽ</i>	494
10.4.1.1	Conditional subordination	494
10.4.1.2	Temporal subordination	496
10.4.2	Other subordinating constructions	499
10.4.2.1	Lest/until NEG <i>tapwi</i> construction	499
10.4.2.2	Purposive <i>kanüm</i>	500
10.4.2.3	Unrealized alternative and contra-factual conditional with <i>(go)ji</i> and <i>ji</i>	501
10.4.2.3.1	Unrealized alternative <i>(go)ji</i>	501
10.4.2.3.2	Contra-factual alternative <i>ji(-na) yüm</i>	502
10.4.3	Secondary non-final markers: the <i>thĩ</i> group	503
10.4.3.1	Temporal and hypothetical subordination <i>bũ thĩ</i>	503
10.4.3.2	‘As if’ <i>da thĩ</i>	504
10.5	Nominalization based constructions	505
10.5.1	Theoretical considerations	505
10.5.2	Relative clause	506
10.5.3	Event nominalization	511
10.5.4	Locative nominalizer <i>blúm</i>	512
10.6	Complement clause and embedded discourse	514
10.6.1	General strategy for complementation	514
10.6.1.1	Different-subject restriction of <i>wê</i> ‘think, want to’	516
10.6.1.2	Same subject restriction of complementation modal verb <i>dũ</i> ‘need, must’	517
10.6.2	Direct speech complement	518

10.6.3	Embedded cited discourse or clause subordinated to a negative verb of cognition	519
10.7	Extended functions of the verb <i>la</i> ‘say’	521
10.7.1	Hearsay particle	521
10.7.2	Name labeling	523
10.7.3	Discourse continuity with <i>la</i> ‘say’	524
10.8	Complex demonstratives (discourse operators)	526
10.8.1	Discourse deictics <i>weli</i> and <i>(we) cuwe</i>	527
10.8.2	Sentence connector <i>webo</i> ‘and then’	531
11	Lexicon	533
12	Appendix I: Texts	564
13	Appendix II : Classifiers	619
14	References	627

Abbreviations and conventions

A	Agent; most actor-like argument in a transitive verb	COND	conditional
AA	Austroasiatic	CONT	continuative (- <i>q</i>)
ABL	ablative	CONTR.TOP	contrastive topic (<i>ma</i>)
ADD	additive	COUNT.EXP	counter-expectational (<i>ki/kale/kalai</i>)
adju	adjunct	DIR	directive (- <i>zâ</i>)
AFF	affirmative (<i>hâ</i>)	dir.speech	direct speech
ALAS	‘unfortunately’ (<i>pà</i>)	D	dual person
ALLO	allophoric	DISC.DEIC.DEM	discourse deictic demonstrative
AMB	ambulative (- <i>wi</i>)	DU	dual (- <i>nega</i>)
Asm	Assamese	DEM	demonstrative (<i>we</i>)
ASSO	associative (<i>né/nê</i>)	DUB	dubitative
BEN	benefactive (- <i>hàng</i>)	ECHO	echo formation
CAUS	causative (- <i>ji</i>)	E	English
CAUS.V	volitional causative/permissive (- <i>shyüw</i>)	EGO	egophoric
CERT.FUT	certain future (egophoric) (- <i>(vi)dyô</i>)	EMPH	emphatic
cl	clause	EVI	non-visual evidential (<i>kh râk</i>)
CLF	classifier	FACT	factitive (- <i>na</i>)
CLP	classifier phrase	FILL	filler
CNT	continuative (-(<i>ta</i>) <i>lá</i>)	FEM	feminine (- <i>yă</i>)
CO.TEMP	co-temporality marker (<i>dũ</i>)	FOC	focus (<i>nyu</i>)
COLL	collective (<i>làng</i>)	FRUS	frustrative (- <i>hă</i>)
COMPL.CL	complement clause	FS	false start
COMPL.I	completive I (- <i>dâng</i>)	G	Goal; most goal-like or ground-like
COMPL.II	completive II (- <i>kũ</i>)	(H)	high tone
CONC	concomitant (<i>thĩ</i>)	HAB	habitual
		(HF)	high-falling tone
		HORT	hortative

HSAY	hearsay (<i>la</i>)	PERM	permissive (<i>-shyüw</i>)
IA	Indo-Aryan	PERMA	permansive (<i>-(ta)lá</i>)
IMP	imperative (<i>-na</i>)	PFV	perfective
IMP.NARR	narrative imperative (<i>-tè</i>)	PLAC	pluriactionality
INFR	inferential evidential (<i>dě</i>)	PL.I	plural I (<i>-ga</i>)
INSIST	insistive (<i>na/ne</i>)	PL.II	plural II (<i>-ring/-rě</i>)
INS	instrumental (<i>gũ</i>)	POSS	possessive
INOP	inopinative (<i>-ta</i>)	POT	potential
INT	intentional (egophoric) (<i>-ne</i>)	PROH	prohibitive
IRR	irreal (<i>((go)ji</i>)	PRP	purposive (<i>kanüm</i>)
ITER	iterative (<i>-mâ</i>)	prsp	perspective
JUS	jussive	PST	past
(L)	low tone	PTB	Proto-Tibeto-Burman
LOC	locative (<i>gũ</i>)	Q	question tag (<i>kya/ye</i>)
LOC.INS	locative inside (<i>kě</i>)	REC.PST	recent past (<i>-li/lii</i>)
LOC.NMLZ	locative nominalizer (<i>blüm</i>)	REDUP	reduplication
LOC.ON	locative atop (<i>kyã</i>)	rel.cl	relative clause
(M)	mid tone & toneless syllable	SELF	reflexive (<i>ga</i>)
MAL	malefactive (<i>-hàng</i>)	S	singular person
MOD	modifier form	S	subject; only argument of an intransitive verb
MSEA	Mainland Southeast Asia	STAT	stative (<i>-ho</i>)
NE	Northeast (India)	TBU	tone-bearing unit
NEG	general negative (<i>yüm</i>)	Tib	Tibetan
NF	cosubordination marker (<i>gũ</i>)	T	Theme; not most goal-like, not most actor-like
NP	noun phrase	TAMN	Tense, aspect, mode and number
O	object	TB	Tibeto-Burman
OBS.NEG	observational negative (<i>shyüw</i>)	TEMP.SUB	temporal subordinator
OPT	optional	TOP	topic (<i>byě</i>)
P	Patient; not most actor-like argument in a transitive verb	V	verb, verbal
P	plural person		

vi intransitive
 VLC verbless clause
 vt transitive
 WND wonder particle (*palang/(sa)mlang*)
 X not attested, not occurrence
 xxx not audible, not clear
 - morphological concatenation &
 compounds with a non-compositional
 meaning
 = clitic boundary
 ~ reduplication of verbal stem
 [] phonetic transcriptions
 / / phonemic transcriptions
 * reconstruction & unacceptable example
 ? marginal/uncertain
 <.> syllable boundary
 <'> stress mark
 <◌◌> (unmarked) mid tone & toneless syl-
 lable
 <◌◌◌> high tone
 <◌◌◌> high-falling tone
 <◌◌◌> low tone
 ∅ zero/covert realization (morphosyntax)
 1 first person
 1-D one dimension
 2 second person
 2-D two dimensions
 3 third person
 3-D three dimensions

1 Introduction

This is the first comprehensive description of Tawrã Mishmi, or simply *Tawrã* [ta¹ɽwã], a Tibeto-Burman language spoken in Lohit and Anjaw districts of Arunachal Pradesh, Northeast India. This region corresponds to the easternmost tip of the Eastern Himalayas, a hotspot of biological and linguistic diversity. Politically, the Tawrã area lies at the intersection of three countries: (Northeast) India, China (Tibet) and Myanmar; the largest part of the Tawrã area lies in the former. Linguistically, Tawrã lies at the convergence between South Asian and Mainland Southeast Asian linguistic areas, sharing characteristics of both.

Speakers of Tawrã generally prefer the term *Tawrã* to refer to their ethnic identity and language; other neutral terms like *Digaru* and *Taraon* are mostly employed by non-Tawrã. The Tawrã also recognize a larger ethnic alliance with the Kera'a (also known as *Idu*) and the Kaman (also known as *Miju*) under the umbrella *Mishmi*.

The aim of this chapter is threefold: i) to situate the Tawrã in their socio-historical context for the layman, ii) to provide extra-linguistic information that I deem relevant for this grammar and iii) to provide an overview of phylogeny and contact effects. §1.1 is an introduction to the language in its linguistic setting. §1.1.1 provides information on the genealogical affiliation of the language within the Tibeto-Burman family. §1.1.2 discusses the term 'Mishmi' in detail, a topic which gives way to discussing multilingualism. §1.1.3 presents an overview of the potential Austroasiatic influence in Tawrã. *Other neighbors*, in §1.1.4, refers to Indo-Aryan (henceforth IA) and Tibeto-Burman languages that live in proximity to speakers of Tawrã. This section also has a brief discussion on language vitality. §1.1.5 reviews the work of Sun (1993) regarding the phylogenetic connection between Tawrã and the Tani languages. §1.2 provides an overview of previous literature on Tawrã. §1.3 is a sketch of two aspects of Tawrã society and culture: clan or-

ganization and religion. §1.4 discusses the theoretical framework of this grammar. §1.5 lays out the details of my fieldwork in the Tawrã area and of the corpus that supports my analysis. §1.6 is an overview of the main typological features of the language. Finally, §1.7 discusses transcription and orthographic conventions.

1.1 Linguistic situation

1.1.1 Genealogical affiliation

Tawrã and Kera'a form a subgroup within the Tibeto-Burman (henceforth TB) language family, as they share a large part of their core lexicon (cf. §1.1.2). This subgroup is here called *Kera'a-Tawrã*. Kaman, a neighboring TB language, is related on a higher phylogenetic level within the larger TB group¹.

The literature on TB phylogeny recognized the constituency of the Kera'a-Tawrã subgroup early on, although different labels have been proposed. Shafer (1955, 102) employed the label *Digarish*, which included *Taying (Digaro)*² (i.e. Tawrã) and *Midu* (i.e. Kera'a). Shafer placed *Digarish* under the *Bodic Division* of his *Sino-Tibetan* model. That division also included *Midžuish* (i.e. Kaman), *Hrusish* (i.e. Hrusso or Aka), *Dhimalish* (i.e. Toto and Dhimal), *Mišingish* (i.e. Tani) and *Dzorgaish* (i.e. Lepcha, Magar and Qiangic, cf. van Driem 2001, 389). Benedict (1972, 5) tentatively placed Tawrã (spelled *Digaro*) along with Aka, Kaman (Miju) and Dhimal in the *Abor-Miri-Dafla (Mirish)* (i.e. Tani) 'nucleus'. van Driem (2001, 496f) provides a literature review of the *Mishmi clusters* (which includes Zaiwa, also known as Meyor or Tulong, cf. §1.1.4). Within his agnostic *fallen leaves* model, which does not presuppose any macro sub-grouping other than low-level and otherwise well-established groups, van Driem (2014) reproduced Shafer's model and placed, on the one hand, Tawrã and Kera'a under the label *Digarish* and, on the other, Kaman (and Meyor/Tulong) under the *Midžuish* group (sometimes *Midzhuhish*, cf. van Driem 2012, 188; also cf. van Driem 2011). Blench and Post (2014, 95), while keeping

¹Sastry (1984, 2) refers misleadingly to Idu (Kera'a), Digaru (Tawrã) and Miju (Kaman) as the "three principal dialects of Mishmi". He, however, ultimately follows Shafer (1955, 102) in separating Kera'a-Tawrã from Kaman within the larger macro-family.

²This term might be related to a Tawrã clan named *Tayang*. I have not observed the term *Taying* being used.

Tawrã and Kera’a together, rechristened the *Digarish* group *Mishmic*. This is unfortunate, as it does not do justice to Kaman, which is also Mishmi in the political sense. Kaman, furthermore, falls under the singleton branch *Mijuic/Mijuish*, which is unclassified within the language family. In a similar approach, Post and Burling (2017, 223-224) further rename the *Mishmic* subgroup as *Idu-Tawrã*, a more agnostic term, while Kaman is tentatively placed along with Meyor (Tulong) in the *Kaman-Meyor, Lohitic or [Midzuish]* group.

Tawrã was included in a language sample subjected to Bayesian phylogeny by Sagart et al. (2019). The results give a posterior probability of 100 % for a clade comprising Tawrã, Kera’a and Tani. Similarly, Wu et al. (2022) performed a Bayesian phylolinguistic analysis that included the data from Sagart et al. (2019) but was expanded with data from varieties from Kho-Bwa, Hrusish, Tshangla, Bodish, Tani and Mishmic (which included all three Mishmi languages). The results suggest that Tani and Mishmic (i.e. including Kaman) are genealogically the closest relatives of Kho-Bwa and Hrusish (also cf. Jacques and Stevens 2024).

In sum, the close relationship between Tawrã and Kera’a has been established early on in the literature. Sun (1993) has provided evidence for a potential relationship between Tawrã-Kera’a and the Tani languages. This is further discussed in §1.1.5. Finally, recent phylogenetic analysis employing large datasets has pointed to a close relation between Tawrã-Kera’a, Kaman, Tani, Kho-Bwa and Hrusish.

1.1.2 The ‘Mishmi group’ and multilingualism

As mentioned in §1, the tribes³ Tawrã, Kaman (Miju) and Kera’a (Idu) recognize a shared *Mishmi* identity, which serves as an umbrella for these three groups, especially regarding state-level politics. People more acquainted with these tribes often refer to them by the tribe’s name plus the epithet *Mishmi*, hence *Tawrã Mishmi* (also *Digaru Mishmi*). I am not aware of any sound etymology regarding the term *Mishmi*⁴ (also cf. Mene and Chaudhuri 2019, 7-9). The term, however, does seem to be old, as it appears in the Snake Pillar erected by the Ahoms in the XVIth century in Sadiya⁵ (cf. Luthra 1971, 1145).

³In the context of Arunachal Pradesh, the term ‘tribe’ is the preferred form to address indigenous groups.

⁴The initial *mi-* segment could be related to a widespread etymon for ‘man’ **mi* (Matisoff 2003, 658). I have heard the term *Mishmi* being linked humorously to English *miss me!*

⁵Transliterated from the Ahom script as <misimi>.

The inherited Kera’a-Tawrã vocabulary is clearly TB. Table 1.1 displays several Tibeto-Burman roots shared by the Mishmi group. The data for Kera’a and Kaman comes from my own field notes⁶. TB *comparanda* is from Proto-Tani (Sun 1993) and Chintang (Rai et al. 2011)⁷ (if available).

Table 1.1: Shared TB vocabulary in Kera’a-Tawrã

	Tawrã	Kera’a (Midu)	Kaman	TB
year	<i>(ka)náng</i>	<i>inu</i>	<i>law</i>	Proto-Tani *ñiŋ ‘year’
road	<i>alyüm</i>	<i>ayuci, aluci</i>	<i>blogn</i>	Proto-Tani *lam ‘road’, Chintang <i>lambu</i> ‘road’
bone	<i>rûbrung</i>	<i>ruci</i>	<i>rak</i>	Proto-Tani *loŋ, Chintang <i>saru?wa</i> ‘bone’
leaf	<i>na</i>	<i>na, napra</i>	<i>lap</i>	Proto-Tani *nə ‘leaf’ and *lap ‘wild green onion’, Chintang <i>lapho?</i> ‘big leaf used for wrapping and covering food and for making leaf plates’
nest	<i>asûw</i>	<i>su</i>	<i>sap</i>	Proto-Tani *sup ‘nest, lair’, Chintang <i>hup</i> ‘nest’

Despite the robust TB stock of the lexicon, the Kera’a-Tawrã subgroup is characterized by several *aberrant* words, i.e. words that do not have a clear TB origin. Some of these words are shown in Table 1.2. In many cases, Kaman aligns with widespread TB cognates. The TB *comparanda* is again from Proto-Tani (Sun 1993) and Chintang (Rai et al. 2011) (if available). The presence of potential Proto-Austroasiatic cognates in Tawrã will be discussed in §1.1.3.

⁶I am thankful to Hari Apralo and Tage Umbrey for their help in Kera’a (Midu variety) and Krulenso Pul for his help with Kaman. I have used an orthography that is in line with Tawrã for the sake of comparison. It should be noted, however, that tonal differences have been collapsed since the phonology of these languages is not well understood.

⁷These language were chosen to reflect two different groups within the language family.

Table 1.2: Aberrant inherited vocabulary in Kera'a-Tawrã

	Tawrã	Kera'a (Midu)	Kaman	notes
eye	bülüm	elobra	mik	Proto-Tani *mik 'eye', Chintang <i>mik</i> 'eye'
flesh	thô	tho	shitn	Proto-Tani *duun 'meat, flesh'
give	hàng	hã	phit	Proto-Tani *bi 'give', Chintang <i>pid-</i> 'give'
hand	hatyô	ato	rok	Proto-Tani *lak 'hand, arm', Chintang <i>chuk, muk</i> 'hand'

Tawrã and Kaman share a large part of their territory and there are many areas where bilingualism is commonplace, especially around Hayuliang, Wakro and Dav valley. The latter is a particular case since the valley was difficult to access until relatively recently and hence there is a higher degree of inter-mixing between Tawrã and Kaman families and clans so that bilingualism is exceptionally widespread. Although the languages share only a small part of their core vocabulary, they display striking syntactic, morphological and semantic similarities, a topic that has not yet been explored. Figure 1.1 is a map of the approximate distribution of Tawrã and Kaman in Lohit and Anjaw districts. This distribution is based on my observations and inquiry and, as such, I do not claim completeness - especially since I was never allowed to enter Anjaw. If we follow the map, the heartland of the Kaman is the stretch between Hayuliang and the border with China. The Tawrã area begins roughly at the Dav and Dalai valleys and continues down to the plains into Lohit. The agglomeration of villages around Tajaw (Tezu) constitutes the largest concentration of Tawrã speakers. A TB language Meyor or Tulong is spoken in one village near Walong (further discussed in §1.1.4). We know that there are both Tawrã and Kaman speakers across the border in China (Tibet) due to the work of Jiang et al. (2013) and Sun (1991) hailing from that area (as discussed in §1.2) but I have no updated information to give an assessment.

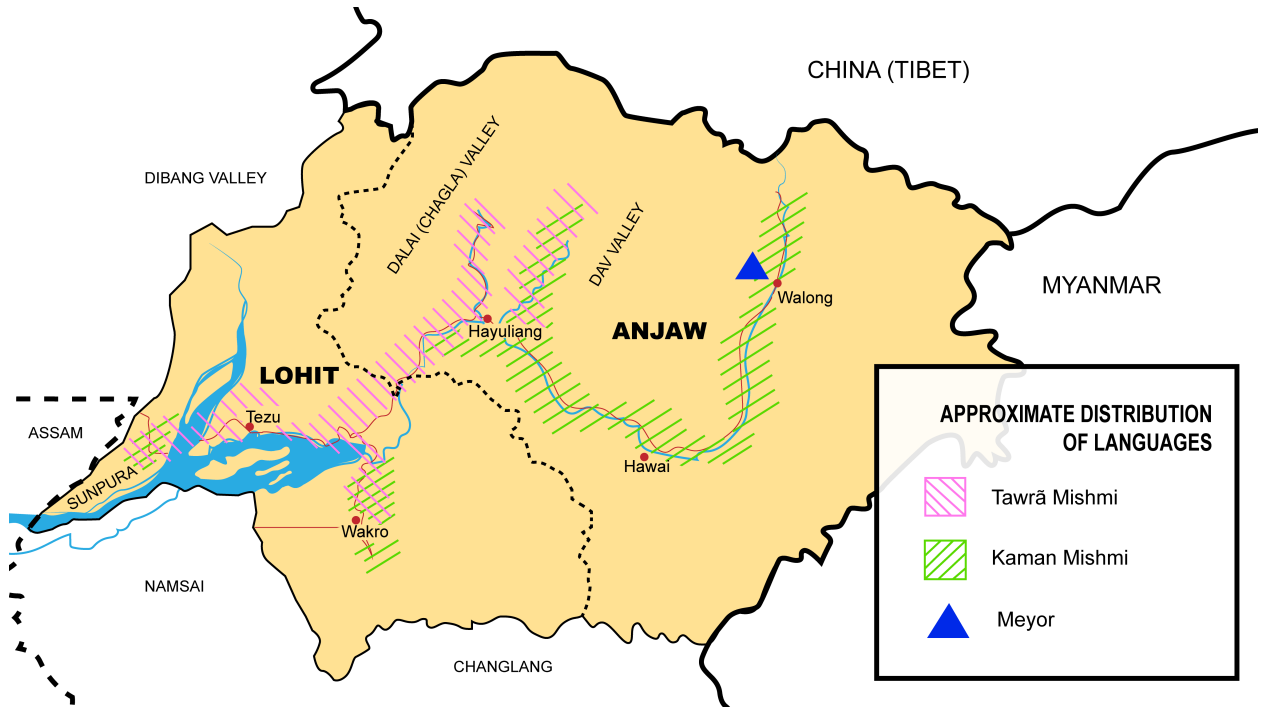


Figure 1.1: Approximate distribution of native TB languages in Lohit and Anjaw districts

Based on clan stories of the Tawrā, the plains (Lohit district) have seen a surge of populations coming down from the mountains (Anjaw district) at least since the nineteen fifties. Therefore, many clans recognize their original place to be up in Anjaw. The situation of the Kaman around Wakro is probably similar but more research in their oral literature is required to understand their migration paths. The villages around Sunpura have also been populated recently, or at least there have been important migration waves in recent times from the mountains of both Tawrā and Kaman populations. This area has traditionally been in contact with Assamese and is ethnically diverse. Tajaw (Tezu) is the district capital and the center of commerce and education for both Lohit and Anjaw districts. As such, Tajaw hosts people from many parts of India.

1.1.3 Contact with Austroasiatic

Some lexemes in Tawrā may have come through contact with an extinct Austroasiatic (henceforth AA) language. This section will offer an overview of the research on AA influence in TB, it will present a list of potential AA loanwords in Tawrā and it will also include my hypothesis of how this language contact unfolded, bearing in mind that

linguistic evidence is, in this case, the sole evidence available.

Compared to TB, AA is considered an older linguistic layer in Mainland Southeast Asia (henceforth MSEA) based on its scattered modern distribution: from East India (Munda) to the western fringe of SEA (Vietic) and to the southern tip of the Malay Peninsula (Aslian) (Sidwell 2014, 144f). The homeland or center of the spread of the language family is a matter of contention and several areas have been suggested between MSEA and NE India (for the *Southeastern Riverine Hypothesis* cf. Sidwell and Blench 2011, 338f; for an overview cf. van Driem 2012, 193). In any event, what does seem certain is that the expansion of AA is associated with rice agriculture (dry and later wet), as there is archaeological evidence as well as a reconstruction of vocabulary associated with rice (Diffloth 2005; Bellwood 2021, 37f; Bellwood 1993, 91, 109).

The early expansion of AA could have reached modern Northeast India and there might have been extant AA populations when TB speakers started migrating into the area⁸. It has been suggested that there might have been several stages of contact with AA throughout the history of TB expansion (Norman and Mei 1976; Benedict 1967; Benedict 1992; Matisoff 2010; Sidwell 2014, 146). Furthermore, the directionality of borrowing is in some cases not clear, particularly regarding *Wandelwörter*, i.e. loanwords found across different language families, often acquired through trade or transfer of knowledge, as well as areal phenomena where different linguistic phyla are involved (cf. Matisoff 1991, 414; Enfield and Comrie 2015; van Driem 2001, 362).

The influence of AA on the TB languages of MSEA has been well described for Burmic (Bradley 2021, 306f; Jenny and Hnin Tun 2017, 29f, 94f; also cf. Jenny 2011, 48f) and Karen (cf. Kato 2021, 339f).

In the case of NE India, the influence of AA on TB languages can be tentatively separated into two interconnected groups: in the south, the influence of Khasi over Karbi has been noted by some authors (cf. Joseph 2009; also cf. Konnerth 2014, 10). A second area is found north of the Brahmaputra. This group includes the Tani and the Mishmi languages (including Tawrã Mishmi) and it may possibly extend to some languages of northern Burma (cf. Button 2011, 23,70). Crucially, it should be noted that there are no

⁸There is little evidence to assess when the Khasian group reached Meghalaya (Sidwell 2018, 5-7; Bradley and Mohanty 2023, 192; also cf. Diffloth 2008). Zoller (2023, 245f) has furthermore argued that AA reached up to North India at some point.

extant AA populations in a radius of 500 km from the Tawrã area. Image 1.2 illustrates the approximate distance between the Tawrã/Kaman area and Khasi and Palaung/Wa, the closest AA languages.

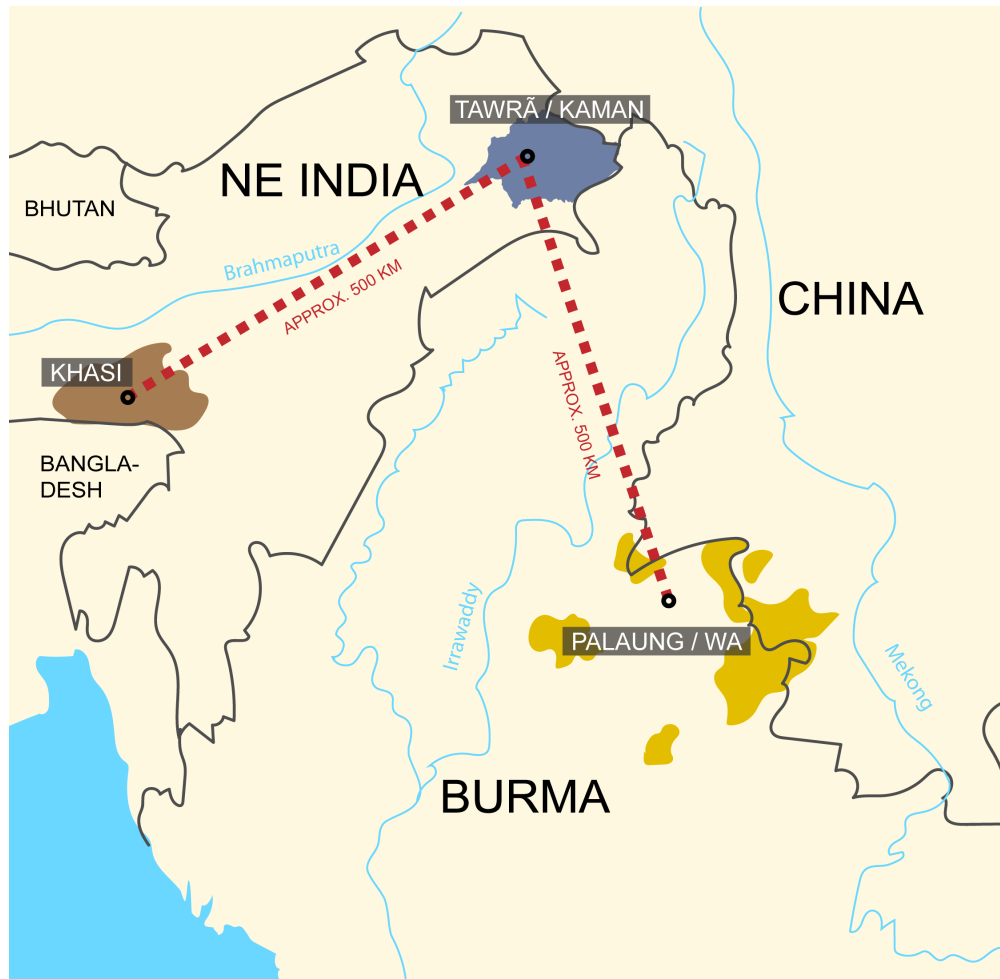


Figure 1.2: Approximate distance between the Tawrã/Kaman area and two AA phyla

Sun (1993) was probably the first to note potential AA words in Proto-Tani. He included potential loanwords like Proto-Tani **granj* ‘hornbill’ and **par* ‘mortar’. Considering that Tani and Mishmi languages are in close proximity and are related at some level, the existence of AA loanwords in Tawrã is to be expected.

Table 1.3 displays several potential AA loanwords in Tawrã. The data from PAA (Proto-Austroasiatic) comes from Shorto (2006), Proto-Khasian from Sidwell (2018) and Proto-Palaungic from Sidwell (2015)⁹. Data from Proto-Waic (a member of the Palungic

⁹The orthography employed by Sidwell (2015) has been slightly adapted to match the one of Shorto (2006). This involves vowel length, which is given as two vowels (i.e. *aa*) by Shorto (2006) while Sidwell

group) is from Diffloth (1979).

Table 1.3: Some potential AA loanwords in Tawrã

Tawrã	AA
<i>tapũ</i> ‘bamboo shoots (generic)’	PAA *t ₁ ɓaŋ ‘bamboo shoots’, Proto-Palaungic *ɓaŋ (no Proto-Khasian form)
<i>pla</i> ‘CLF:blades, large, flat objects’	PAA *plaaʔ ‘blade, edge’, Proto-Palaungic *plaaʔ, cf. Proto-Khasian *blaa ‘shoulder blade’
<i>katyàw</i> ‘type of crab’	PAA *ktaa ₁ m ‘crab’, Proto-Khasian *thaam, Proto-Palaungic *kəthaam
<i>apũ</i> ‘fence’	Proto-Palaungic *-ɓuum ‘fenced garden’ (neither PAA nor Proto-Khasian form)
<i>alũ</i> ‘big mortar’	Proto-Khasian *t ^h lɔŋ ‘mortar’ (neither PAA nor Proto-Palaungic form)
<i>bülik</i> ‘pig’	PAA *lik, *liik ‘pig’, Proto-Waic *lik, Proto-Palaungic *-leek (no Proto-Khasian form)
<i>pé</i> ‘winnowing tray’	Proto-Palaungic *piɛr ‘winnowing tray’ (neither PAA nor Proto-Khasian form)

Before discussing each word individually, it should be noted that the reconstructions presented by Shorto (2006) display astonishing lexical stability, particularly within the Palaungic-Khmer subgroup. This is an impressive characteristic of the AA language family given that many subgroups are far apart. We may presume that this is due to a relatively rapid process of early expansion. However, some words are either found in either one phylum or the other, e.g. PAA *lik, *liik ‘pig’, only found in Proto-Waic and Proto-Palaungic but not in Proto-Khasian. This can be explained as a case of independent innovation within AA and/or we may presume a differential effect of language contact with now-extinct language groups from a time before the arrival of TB to the area.

The Tawrã word *tapũ* ‘bamboo shoots (generic)’ exhibits a *ta-* root prefix (cf. §3.7.6) that is common among terms for flora.

(2015) employs a colon (i.e. *α:*).

The word *katyàw* ‘type of crab’ has a *ka-* initial syllable that is noteworthy, as it matches both the ‘animal prefix’ found in modern AA (cf. Matisoff 2003, 138-139) and the Tawrã root prefix *ta-* associated with small animals (cf. *kacêk* ‘type of mole’ or *kame* ‘type of squirrel’). These matching *ka-* prefixes seem to point to morphological convergence brought about by repeated cases of language contact in ancient times. This prefix is further discussed in §3.7.3. The final nasal *-m*, present for both Proto-Palaungic (also Proto-Wa, cf. Diffloth 1979, 144) and Proto-Khasian (Sidwell 2015), remains unexplained¹⁰.

In the case of *apû* ‘fence’ and *alû* ‘big mortar’, the *a-* segment goes back to an *a-* root prefix (cf. §3.7.1), which is in many semantic groups, including tools (material culture). The complex onset of Proto-Khasian **tʰlɔŋ* seems to have been simplified to **lɔŋ* in Tawrã form (note that there is no *tʰ-l* onset cluster in Tawrã, cf. §2.2.3.1). Kaman *glogn* ‘big mortar’ seems also to have adapted this AA word, although with a different onset cluster.

The case of *bülik* ‘pig’ is remarkable, as it is likely the product of a fusion between a TB element (cf. PTB **pʷak* ‘pig’ (Matisoff 2003, 62)) and an AA root (cf. PAA **lik*, **liik* ‘pig’), something on the lines of **p-lik* (for a similar case in ‘otter’, cf. Matisoff 2010, 583). It is also noteworthy that this PAA etymon has no reflex in Sidwell’s Proto-Khasian, although it is reconstructed for Proto-Waic and Proto-Palaungic. Sun (1993, 217) suggests that Proto-Tani **rjek* ‘pig’ could also be a AA loanword.

Lastly, Tawrã *pé* and Proto-Palaungic **pier* ‘winnowing tray’ resembles PAA **tpal* ‘mortar’, which has been noted as an AA borrowing in Proto-Tani (**par* ‘mortar’ cf. Sun 1993, 224). In any case, an AA origin (**pier*) for Tawrã *pé* ‘winnowing tray’ has to assume a drop of final alveolar *-r* and a simplification of the diphthong.

Table 1.4 displays three potential *Wandelwörter*: ‘iron’, ‘cow (cattle)’ and ‘salt’. The data from AA (PAA and Palaung) comes from Shorto (2006), data from Proto-Palaung from Sidwell (2015), data from Pnar (Khasian) from Ring (2015) the data from PTB from Matisoff (2003), the data from Rabha from Joseph (2006) and the data from Proto-Tani from Sun (1993). These words in Tawrã could have been adapted along with the

¹⁰Coda *-m* is certainly allowed in Tawrã (§2.2.4). There is, however, a similar case of coda *-m* loss regarding loanwords: Adi (tribe and language) is called *pada* in Tawrã. This term goes back to the *Padam*, a group within the Adi tribe. In this case, the final nasal was also lost when adapted to Tawrã.

knowledge associated with the use of iron, the domestication of cattle and the salt trade (cf. Blench and Post 2014, 83f).

Table 1.4: Two potential *Wandelwörter* in Tawrã with AA correspondences

Wandelwörter	Tawrã	AA	Notes
iron	<i>sâi</i>	Palaung <i>hir</i> ‘iron’	PTB *syi:r *sya:l
cattle, cow	<i>matsâp</i>	PAA *mæk, *mæək ‘ox, cattle’	Rabha <i>má-su</i> ‘cattle’
salt	<i>plà</i>	Proto-Palaungic *blu:ʔ, Pnar <i>bloʔ</i>	Proto-Tani *lo ‘salt’

Benedict (1972, 33, 80) was probably the first to point out that ‘iron’ in TB could be a (Austro-Tai) loanword due to the alternation *syi:r *sya:l (echoed by Matisoff 2003, 395). In the case of ‘cattle, cow’, potential cognates with a structure #mVCV are widespread among AA and TB languages of NEI and a relation to IA languages cannot be discarded, cf. Assamese *māṃsa* ‘meat, flesh’ (cf. Blench and Post 2014, 83f; Diffloth 2008, 103). Finally, the word for ‘salt’ could have been adopted through trade, although more information is required to assess this claim (cf. Diffloth 2008, 100). The relation to Proto-Tani *lo ‘salt’ is not clear.

The study of borrowed material should consider extra-linguistic aspects, such as technological innovations (e.g. cultivation, domestication of animals, use of materials and weapons) and ontological changes (e.g. religious vocabulary) since linguistic data are just one (albeit important) part of the complex social process behind the early contact between speakers of distinct linguistic phyla (cf. Blench and Post 2014, 74; also discussed by Enfield 2021, 95f). In this light, the words presented in Table 1.3 suggest a form of technological transfer from an AA group to the Kera’a-Tawrã, which included tools, some of which are associated with rice cultivation (mortar, blade, winnowing tray), the domestication of pigs and the consumption of crabs and bamboo shoots, all potentially new sources of food for the incoming TB groups. The presence of AA words in Proto-Tani, as noted by Sun (1993), also suggests the possibility of early contact between now-extinct AA group(s) and perhaps a common ancestor of Tani and Kera’a-Tawrã. Conversely, the presence of *Wandelwörter* (Table 1.4) paints a complex picture of language contact and influence. Lastly, the potential loanwords in Tawrã (presented in Table 1.3) further

suggest an affinity with the Palaungic and Khasian subgroup of the AA language family. These two phyla are generally considered to form a subgroup within the language family (Sidwell 2014, 163). However, it cannot be discarded that this AA lexifier belonged to one (or more) now-extinct AA subgroup(s).

1.1.4 Other neighbors

This section will provide an overview of the languages spoken in the Tawrã area other than Kaman. It will also discuss language vitality in conjunction with contact with IA languages.

In the Walong and Kibithoo Circle of Anjaw district, Arunachal Pradesh, there is a Tibeto-Burman language spoken by a small community of less than 1000 people named *Meyor* or *Tulong* (previously referred to as *Zakhring*) (cf. Zeite 2022; van Driem 2001, 496). Meyor is in direct contact with Kaman. According to Blench (n.d.), Meyor displays lexical affinities with Kaman, but this might be due to language contact rather than genetic affiliation. There is not much information available on this language.

Going further south in Sadiya, Assam, there are speakers of *Deori* (also spelled *Deuri*), a Bodo-Garo language. Given that some Tawrã claim that they used to have land even further south than Sadiya, before the establishment of the modern border between Assam and Arunachal Pradesh, there is a chance that there might have been some contact between Tawrã and Deori-speaking people. However, without historical data and sufficient description of the Deori language, these statements remain conjecture.

English is commonly known among younger people since it is taught at school. People who have had access to higher education tend to be fluent in the language. Since Hindi-English code-switching is very common (not only for educated classes), some English words have entered Tawrã language after being adapted to the phonology of (Arunachali) Hindi, e.g. *boder* ‘border’ or *daireksan* ‘direction’.

Tawrã has been in close contact with IA languages. Assamese used to be a strong language in the Tawrã area (and through most of Arunachal Pradesh as well), serving as *lingua franca* and as the language of the government, including education. Most elder speakers that I have met are proficient in Assamese. During the 50s and 60s, Assamese

was replaced by Hindi as the main language of education, amidst the wave of *hindization* that swamped Arunachal Pradesh (cf. Sharma 2014, 196f).

Nepali is also a widespread language in the Tawrã area, although it has had, to my knowledge, no impact in terms of loanwords. Command of Nepali is common for Tawrã speakers who live and/or work among Nepali workers, as they are often hired in construction or agricultural work. Table 1.5 displays several IA loanwords in Tawrã. It should be noted that for many of these loanwords, it is difficult to establish which IA language is the source, as Hindi (H) and Assamese (Asm) share an important portion of their vocabulary. Thus the overarching label *IA* is employed either when the source language is unknown or when it is the same term for Hindi and Assamese.

Table 1.5: Some IA loanwords in Tawrã

IA loanwords in Tawrã	Meaning	Source
<i>châgulyei</i>	goat	Asm
<i>gângkhi</i>	milk	Asm
<i>nyêm</i>	rule	Asm
<i>jorâ, jura</i>	pair	Asm
<i>aram</i>	ease, rest	IA
<i>baji</i>	time, o'clock	IA
<i>bawã, bawan</i>	god	IA
<i>bera</i>	fence	IA
<i>daram, durum</i>	religion	IA
<i>garêi</i>	car	IA (also cf. E <i>car</i>)
<i>dimag</i>	brain	IA
<i>juto</i>	shoe	IA
<i>accha</i>	I see (lit. 'good')	H

A last note on language vitality is due before finishing this section. As has been noted by several authors (van Driem 2001, 1140-1142; Turin 2007; Post 2020), many TB languages from the Himalayas are losing ground at an accelerated pace amid a widespread language shift, mainly to IA languages. Tawrã is by no means exempted from this trend

and Arunchali Hindi has emerged in the last decade as the main language for inter-tribe communication (*lingua franca*) and education (cf. Modi 2006). During my fieldwork in *Lulyang* (Loiliang), I observed that Tawrã was still the preferred language among adults. However, younger people below 20 years old often had Arunachali Hindi as their first language. Further sociolinguistic surveys would be required to assess the vitality of the language in other parts of the Tawrã area. My impression remains that further transformation of Tawrã society will undoubtedly have a toll on the transmission of the language; the integration of larger parts of Tawrã society into the Indian state might see a continuous homogenization of their culture and an inexorable loss of the language.

1.1.5 Relation to the Macro-Tani group

Some authors have pointed out the relation between Tawrã and the Macro-Tani group. This section will overview some of the data that has been presented in the literature.

As part of the reconstruction of Proto-Tani, Sun (1993, 398f) undertook a lexicographic test to assess the degree of lexical affinities between Tani and four potentially close groups: Tawrã (spelled *Taraon*), Kaman, Lepcha and Dhammai (and the control groups Written Tibetan, Written Burmese and Garo). The test had 200 words from a list that included body parts and common verbs. There, Tawrã received the highest average score: Tawrã 33.75%, Dhammai 27.5%, Kaman 23.3% and Lepcha 24%. For the author, this result does not paint a clear picture and he further warns the reader that, without a comparative study between these languages, alternative explanations for the high percentage of shared material cannot be ruled out. The author does recognize two shared phonological innovations between Proto-Tani and Tawrã (idem, 405f): One is the parallel development of PTB *dz to *d in Proto-Tani and *th in Tawrã, while other members of the TB group either have retained *dz or a fricative (e.g. [ʃ]). Table 1.6 displays part of the data employed by Sun (1993, 430f) for his lexicographic test that is relevant to this presumed parallel development. I have adapted the Tawrã data presented by the author to my own orthography. This also applies to Table 1.7. In the case of Kaman, I have added my own field notes in parentheses after the data presented by the author¹¹.

¹¹The author employs Chao tone numerals for Kaman. I am agnostic about the tones of Kaman until

Although the **d* to */th/* correspondence is evident in the first three examples, the case of ‘sit’ suggests that this correspondence has either been partly obfuscated by semantic change or a further layer of phonological change, as both the Tani and Tawrã forms have the same alveolar consonant. Furthermore, I have not been able to find more examples of this correspondence.

Table 1.6: Parallel development of PTB **dz* to **d* in Proto-Tani and **th* in Tawrã (based on Sun 1993)

Word	Proto-Tani	Tawrã	Kaman	Notes
eat	<i>*do</i>	<i>thâ</i>	<i>tɕa⁵³, ɕa⁵³ (sha)</i>	cf. Bantawa <i>ca</i> ‘eat’ (Winter 2013, s.v.)
mountain	<i>*di</i>	<i>thîya</i>	<i>a³¹ dzau³⁵ (anjaw)</i>	
poison, hot (spicy)	<i>*duuk/ *tuuk</i>	<i>thâi</i>	<i>tau⁵³ (la)</i>	
sit	<i>*duŋ</i>	<i>dî</i>	<i>lăp⁵³ (lap)</i>	

The second shared feature of Tawrã and Proto-Tani according to Sun (1993, 405f) is the palatalization of certain initials. Table 1.7 displays the data presented by the author. Note that the data from Proto-Tani has been adapted to match the Tawrã orthography, keeping the original in parentheses. The Tawrã data has also been adapted to the orthography employed in this work.

Table 1.7: Palatalization of initials (based on Sun 1993)

Word	PTB	Proto-Tani	Tawrã
fathom	<i>*la(:)m</i>	<i>*ryam <*rjam></i>	<i>lyâng</i>
dusk, evening	<i>*rum~*rim</i>	<i>*ryum<*rjum></i>	<i>talyâng</i>
nose	<i>*s-na</i>	<i>*nya- <*ña-></i>	<i>hanyâgum</i>

Perhaps more relevant than the lexical similarities and the parallel phonological developments between Proto-Tani and Tawrã is a list of potentially cognate grammatical

more research is conducted.

morphemes presented by Sun (1993, 406). Table 1.8 includes all four grammatical markers proposed by the author. The leftmost column, including Proto-Tani and Tawrã, displays the author’s data adapted to the current orthography. The rightmost column displays my data. Note that only one of the four morphemes matches my data. This can be due to errors in the transcription or dialectal variation (although this is unlikely). I have also listed similar-looking morphemes when available. While the prohibitive is clearly cognate, the relation with the comparative auxiliary, the imperative¹² and the experiential aspect is unclear.

Table 1.8: Comparison of gram. morphemes between Proto-Tani and Tawrã (based on Sun 1993)

gram. morpheme	PT	Tawrã by Sun	Tawrã by Hotz
comparative auxiliary	*yang <*jaŋ>	yôŋ <joŋ ⁵³ >	unknown
imperative	*to <*to>	tyô <tio ⁵³ >	unknown, cf. jussive -tyêyuke
Prohibitive	*yo <*jo>	yâ <ja ⁵³ >	-yâq
experiential aspect	*kuu <*kuu>	kóng <koŋ ³⁵ >	unknown, cf. completive -kũ

The relation between Tawrã and Proto-Tani, based on the work of Sun (1993), consists of lexical similarities, two parallel phonological developments and one cognate morpheme. As remarked by the author, no conclusive phylogenetic tie can be established without historical work in Tawrã and Kera’a. Based on the evidence presented, however, it seems safe to assume that Kera’a-Tawrã is related to the Tani group on a higher taxonomic level, as has been suggested by Post and Sun (2017, 323), although the details of this relation are unclear.

1.2 Previous literature on Tawrã

This section includes the main works completely or partly devoted to the documentation and description of Tawrã. Other works with a phylogenetic character, which may have included some data of Tawrã, are discussed in §1.1.1. The work of Sun (1993),

¹²Jiang et al. (2013, 233, 235) also list an imperative suffix *ɬa*⁵³ along with *ku*⁵³.

which includes some data from Tawrã, has been commented in §1.1.5.

The work of Robinson (1855) is likely the first description of Tawrã. Although it is a short overview of the language, the author provides data on morphology, syntax and lexicon of both Tawrã and Kaman. It also includes a short vocabulary in both languages with a translation in English.

A short grammatical description of Tawrã is found in the monumental work *Linguistic Survey of India*, by Grierson (1909, 616-618). This is largely a repetition of the work of Robinson (1855) with some notes and comparisons.

The STEDT database¹³ directed by Professor James A. Matisoff from the University of California at Berkeley includes several entries in Tawrã. Most notable are the wordlists by Bufan and Qingxia (1992) (under the term *Darang [Taraon]* and the subgroup *Deng*, along with Kera'a and Kaman) with more than one thousand entries, the wordlist with more than nine hundred entries by Sun (1991) and one with 172 entries by Sun et al. (1980). All three wordlists (along with the work of Jiang et al. 2013) belong to a Tawrã variety spoken in China. Furthermore, all wordlists are exceptional in their tonal notation, which is mostly in line with the findings of Evans and Manyu (2021) and my own analysis. They do, however, indicate tone on every syllable in polysyllabic words, which reflects a relevant difference between East Asian languages on the one hand and MSEA languages and the situation described for some languages of Northeast India (including Tawrã) on the other. While the former languages often have a tone-bearing unit (TBU) at the level of the syllable, the latter have tonal patterns that are lexically determined, i.e. at a word level, instead of the syllable level (cf. Post 2015a, 184f; Kirby and Brunelle 2017, 711f), cf. §2.5.1. Apart from their value for comparative and historical linguistics, these word lists do not have much grammatical information and some of the entries include roots along with suffixes and numerals — I presume that the authors were not able to distinguish the basic morphological units of the language.

Jiang et al. (2013) have produced a short grammar of Tawrã based on the varieties of the language spoken in Tibet, China¹⁴. In most cases, my analysis deviates from their work to the point that I have found it to be of little use, except for the comparative notes of

¹³Available at <https://stedt.berkeley.edu/~stedt-cgi/rootcanal.pl>.

¹⁴I am thankful to Prof. David Bradley for pointing this work out to me.

dialectal variation. One example may illustrate this point: regarding the phonology, the authors suggest a three-way distinction in the place of articulation for voiceless affricates, i.e. alveolar affricate /ts/, vs. retroflex affricate /tʂ/ vs. palatal affricate /tɕ/. The authors provide no minimal pairs. Such an analysis is untenable for Tawrã and clearly reflects the phonology of Mandarin Chinese (and some Bodic languages).

From the Indian side, there are some word lists and phrasebooks as well. The North-East India Frontier Agency Administration (NEFA nd) produced a dictionary that includes some grammatical notes and an interesting introduction about the Tawrã. This dictionary was conceived as a help to non-native troops stationed in the Tawrã area and, as such, it appeals to the layman and is of little help to researchers. The dictionary employs the (non-modified) Devanagari script, which constrains certain phonological oppositions of the language. Most notably, it does not indicate tone (although tone is mentioned in the introduction), it does not indicate the vowel *ü* [i~u~u] and it collapses the distinction between palatal and alveolar affricates (i.e. च <c> for both *ts* [ts] and *c* [tɕ]).

Pulu (1991) published a phrasebook in Tawrã (written *Taraon*). The book includes a brief introduction to the phonology of the language although, similar to the dictionary published by NEFA (nd), it neither marks the vowel *ü* nor does it distinguish between palatal and alveolar affricates. The book includes a short comparison of vocabulary between Tawrã, Kera'a and Kaman, as well as a list of clan names that exist in two or all three languages.

Singh Grewal (1997, 188-213) published two volumes termed “The Tribes of Arunachal Pradesh” which include information about the cultures of many tribes. It also has short grammatical descriptions of many languages, including Tawrã (which he calls simply *Mishmi*). There is a description of the phonology that makes use of an IPA-inspired phonological notation that is hard to follow and not consistent. The section on morphology and syntax seems to be informed by the author's knowledge of other languages (mainly Indo-European, I presume), as it discusses case marking and verbal inflection. As such, I found this description of the language to be of little help.

Tawsik (2014) has self-published a book that includes a comparison of Tawrã and Kaman lexicon and short texts. Sadly, the book does not follow a clear orthography and is hard to read, even for native speakers.

There is an unpublished, short dictionary of Tawṛā by Blench (2017). The author employs an orthography that includes IPA symbols, e.g. the sign <i>, but this is not done consistently. The placement of tones is also inconsistent. Although the author does have different spellings for palatal and alveolar affricates (<c> and <ts>), he is again inconsistent in their use, e.g. the 3rd person pronoun 's/he' occurs both as <cha> and <tsha> (cya in the orthography used in this work). We may presume that this is still a work in progress.

Evans and Manyu (2022) have uploaded a short online dictionary with entries in practical orthography and phonetic transcription. Their work is still in progress.

Ama (2024) has recently self-published a description of the death ritual *taló*. Although the work is not linguistic in nature, it is a remarkable accomplishment that includes technical vocabulary in both Tawṛā and Kaman, based on the author's research in the area. As such, it is of immense value for the lexicography of the Mishmi languages.

Devi Prasada Sastry has published both a grammar (1984) and a short dictionary (1991), which also includes some notes on grammar. Although the author must be given credit for being the first to publish a lengthy manuscript with grammatical information on Tawṛā, I have found his work to be of little use due to inconsistencies in the orthography and inaccurate grammatical notes. The grammar (1984) has a description of vowels which is reasonable but the discussion of tone is hard to follow and the placement of tone markers over syllables seems to be random at best. The discussion on morphology and syntax provides some insight into the language but one soon encounters grave confusion (not to say, mistakes) that make it clear to the reader that the author a) did not spend much time analyzing the data and b) attempted to homogenize the grammar of Tawṛā to that of Indo-European languages.

Evans and Manyu (2021) published a sketch phonology of the language which includes a description of vowels, consonants and tone. This paper is of enormous value and it has influenced the phonological description in this grammar. Their approach, and the main difference to my analysis, are discussed in §2.

Finally, Mya-Tu et al. (1967) published the results of their expedition to the *Taron* of Northern Burma. Based on the name and the proximity to Arunachal Pradesh, there is a chance that these people were related to the Tawṛā. The report includes a short wordlist

of substandard quality that is certainly not Tawrã. Further research might be able to point out whether there is some connection between these people and the Tawrã.

1.3 A sketch of Tawrã culture and society

This section will provide extra-linguistic information about the speakers of Tawrã and their socio-political context by addressing two issues: the structure of a clan-based society and religion. This should be taken, at best, as a sketch of Tawrã culture and society. A longer period of fieldwork would be required for me to give more detailed and comprehensive accounts.

In order to address these two topics, I have decided to lay down my notes from interviews with two friends from the Tawrã area that I think can explain these issues with clarity. The first notes come from an interview with Shri Johakso Manyu from *Lulyang* (Loiliang) about the place of clans in Tawrã society.

The Mishmi are a tribe from Arunachal Pradesh, Northeast India. They live in the districts of Lohit, Anjaw, Dibang Valley and Lower Dibang Valley. The Mishmi are composed of three subgroups: Tawrã (Digaru), Kaman (Miju) and Idu (Kera'a). Their traditional way of life, before modern amenities, was in remote hills, without roads, electricity or access to education. They used to depend on the forest; went hunting, fishing, practiced *jhum* cultivation (i.e. *swidden*) to cultivate rice, maize and other crops.

Clans are very important. Under each tribe there are clans, usually termed *surnames* or *titles*. This system is similar to other clan-based societies in the area: a person's identity can be known through his/her clan. There are more than one hundred clans of the Tawrã and Kaman Mishmi, often divided in many sub-clans. Some of these clans have become extinct. A person's family tree can be drawn and identified through his/her clan lineage. We can know one's relatives, relations and origin. Marriage within a sub-clan is not allowed. A sub-clan will maintain close ties: they will help each other, work for each other, and live together. They are the root of a person (and his generation). Clans are the base of Mishmi society.

Tawrã (and Kaman) clans know their history and relations through clan stories. These stories tell of four main migration routes: One group migrated from Burma, one group from Assam, one group came from the Idu side and one group does not have a story of migration; they say they are the original inhabitants ('sons of the soil').

The names of certain clans and sub-clans reveal part of their history: they are often the names of streams and hills. A common pattern goes as follows: a group of brothers could no longer live together and had to migrate. One settled by a stream and named his clan after that stream, another brother settled on a hill and got the name of that hill.

The second point I wish to address is that of religion. The Tawrã have traditionally practiced a belief system that echoes similar practices found across the Himalayas and which have been termed *animism* or *shamanism*. These are my notes from an interview with Shri Ajitso Ama from Thowang, near Wakro.

Religion is mixed with all aspects of life: kinship, house building and language. This is a vast topic and I will address some points with which I am acquainted. In most *pujas*, the *gwôk* 'shaman, priest' needs to address every clan guardian. Clan guardians (called *dēmē*) are powerful, important and can be numerous as well. If we fail to make an offering to a *dēmē*, these guardians can punish us. A *dēmē* enters a clan's pantheon through marriage (i.e. through women): if my son were to marry a woman from the *Tawsik* clan, he and his descendants will have to give offerings to the *Tawsik dēmē*, which would be incorporated into the household's pantheon. We thus have to make offerings (*namüing*) for all our (female) ancestors. If a couple has no children, the *dēmē* has to be returned to the wife's parent's house upon the death of the wife.

A further spiritual being to consider is the *kapã*. They are associated with particular trees and hills. If disturbed or offended, e.g. if its tree is felled, they cause damage and ruination. It is the task of the shaman to identify the particular *kapã* as the source. Thereafter, the *kapã* can be integrated to the pantheon of a clan as *dēmē kasâk* or *bá-mê kasâk*. A *kasâk* (i.e. a tamed *kapã*) will officiate

as the messenger between the *gwôk* ‘shaman, priest’ and the major Gods like *jâp-maliü* or *ringyã*.

Some people say that we worship nature because we name water bodies and mountains during *puja*. This is imprecise because we make offerings to the *kapã* that are associated with a particular forest or lake.

A last point concerns the *gwôk* ‘shaman, priest’. We say that a *gwôk* is *zâng* i.e. a magical or powerful being. The *dēmē* ‘clan guardians’ help people who want to become priests. During *puja*, the priest has to narrate the clan genealogy and the details of the offerings (e.g. number of types of animals offered). He then mediates with the divine dimension. The priest speaks using a special form of the language called *khûnyim-tahwî* ‘teachings of the Gods’.

1.4 Theoretical framework

I adhere to the belief that serious linguistic analysis must approach language without prejudice and seek a description that is based on its own structure and (if possible emic) terms (cf. Haspelmath 2009). It should neither be primed by the linguist’s native language, his/her language of expertise nor a language of prestige but rather be driven by data itself. In this sense, I subscribe to a naive, if not opportunistic approach regarding linguistic theory.

This being said, I further recognize that no academic pursuit starts *ex nihilo* and the decisions that a scholar makes regarding his/her data are, more often than not, grounded in existing theories and traditions, albeit unknowingly. This is particularly poignant for the study of language and communication, here understood as eminently social phenomena. Furthermore, the aim of this grammar is not to be purely descriptive; it seeks to uncover, whenever possible, diachronic developments, cognitive and functional explanations as well as phylogenetic ties. These attempts are grounded both in theory and in areal commonalities, as will be explained presently.

During my formative years and especially during the process of writing this grammar, I have been influenced by two, often overlapping traditions: the South and Mainland Southeast Asian grammatical tradition on the one hand and the typological-functionalist

approach on the other. Under the South and Mainland Southeast Asian grammatical tradition I understand the convergence of an ancient tradition, best exemplified in the works of Pāṇini, with the school of grammatical description that started in the 70s and 80s, particularly that which has focused on Tibeto-Burman languages. By a typological-functional approach, I understand the conception of grammar as coding “simultaneously, both propositional semantics and discourse coherence (pragmatics)” (Givón 2001a, 13). Applied to descriptive linguistics, a functional approach is translated into an attempt to identify structures, categories and relations, make sense of them and describe their function in their communicative context. Furthermore, a typological approach looks at the “functional domain” (Givón 2001a, 23) of a structure, assuming that there is often more than one structural means to fulfill the same communicative function.

Structures and functions, in turn, are pinned down through labels and this grammar intends to employ well-established terminology whenever possible. Labels, however, are often the source of dispute as they can mean varied things for different authors. Although this inextricable malaise of the Humanities won’t be solved in this grammar, it can be aided through hermeneutic comments, particularly on broad terms (e.g. *aspect*) but also on those terms that have not yet been fully established in the literature (e.g. *egophoricity*).

Finally, linguists working on endangered or minority languages have unique access to (extra-linguistic) knowledge, such as native forms of classification, ecological knowledge of fauna and flora, inter-clan politics, oral literature, historical knowledge (e.g. clan lineages) or more practical knowledge of house building, agriculture, etc. Due to the rapid and irreversible transformations of most societies in Arunachal Pradesh, to name one example, much of it is seriously endangered if not already lost (cf. Post et al. 2022, 1-3; also cf. Post 2012, 113-115). I believe that descriptive linguistics and grammar-writing should, to the extent possible, pay attention to extra-linguistic information and try and include it in the hermeneutic process and analysis.

1.5 Fieldwork and data

The data used in this grammar was collected during four field trips to NE India: From January to June 2022, from November 2022 to January 2023, from October to November

2023 and from September to November 2024. My main abode during these field trips was Sadiya, Assam. My stay there started as a temporary solution to the intransigent and unreasonable no-visit policy implemented by the local police of Lohit (and Anjaw) district, Arunachal Pradesh. By 2023, it became clear that Sadiya would become my permanent abode while in NE India. Sadiya is an incredible hotspot of linguistic diversity, with speakers of Deuri and Mising (both TB) as well as Assamese, Hindi, Marwari and Nepali (all IA). The latter is the *lingua franca*, although Hindi is taken over at an accelerated pace. Sunpura, the village where I stayed in Sadiya, is split between Assam and Arunachal Pradesh. On the Assamese side, there are no Tawrã speakers, although they often come to the Assamese side to buy things at the bazaar. During my first stay in Sunpura, I was dependent on speakers coming all the way from *Tajaw* (Tezu), approximately one hour drive, to work with me on the language. Luckily, by the end of 2022 we found two speakers of Tawrã who lived on the Arunachali side of Sunpura and were willing to come work with me. This arrangement yielded the largest portion of my data. I stayed two times in Lohit, Arunachal Pradesh: one month in *Tajaw* (Tezu) in June 2022 and three weeks in October 2024 in *Lulyang* (Loiliang). During my first stay in *Tajaw* (Tezu), I was strictly prohibited from leaving the town. This was inconvenient as *Tajaw* (Tezu), the capital of Lohit district, is mostly populated by outsiders and Tawrã speakers are surprisingly hard to come by. During my second stay in Arunachal Pradesh, I had finally obtained a research permit that allowed me to travel freely within Lohit (although not Anjaw) and I stayed in *Lulyang* (Loiliang), which is a cluster of Tawrã villages.

All sessions were recorder using a **Zoom H4n** audio recorder. Sometimes I used a **Rode NT4** stereo microphone. My workflow consisted in transcribing the audios using the **Transcriber App**, which were then converted to .txt files. These were glossed and translated using **Toolbox**. Doubts about the phonology were cleared using **Praat** (Boersma and Weenink 2024).

The type of data that builds this grammar can be divided into three: a) standard elicitation, b) stories/interviews and c) field notes. By A) standard elicitation I mean a setting in which I sat together with a native speaker(s). This was employed to translate and analyze recordings and address questions (e.g. tone, lexicon, etc.). The language of elicitation was, depending on the speaker, Nepali or English and it progressed to Tawrã

as I became more acquainted with the language. In my case, this accounts for the overwhelming majority of my data. In this grammar, the data from such elicitation sessions is indicated by a tag that starts with a date, e.g. [291024_STBT_eli1 161].

B) Stories/interviews refer to instances in which one or more speakers talked freely, either telling a story or as an interview. I recorded several traditional stories both in Sunpura and in Lohit. As of the interviews, I was fortunate enough to work with Shri Johakso Manyu and Shri Jabralum Chaitom, who are committed to the preservation of their language and culture. They conducted interviews and introduced me to virtuous speakers in order to record them. Shri Johakso Manyu also conducted several interviews related to clan stories by himself and I was able to work with those recordings. Excerpts from these recordings are presented in §12. Some of these recordings have been archived in the PARADISEC catalog¹⁵. In this grammar, examples from such recordings take a tag that starts with **TA** and a year, e.g. [TA22_origin_112 326].

C) I took some notes during my stay at *Lulyang* (Loiliang) of naturalistic speech or remarks that speakers said to me, e.g. corrections or potential misunderstandings. These notes are referenced by a tag that says **field notes**, e.g. [field notes 261024].

One of the consequences of having spent comparatively little time in a Tawrã village is dialectal inconsistency, as the linguistic background of my main consultants is very diverse. It should be stated that this is also a feature of the linguistic landscape of the *ashâ*, the Tawrã plains; Lohit, as opposed to Anjaw, has villages that have been populated recently by the Tawrã. Meanwhile, the migration from mountains to plains has not stopped and *Tajaw* (Tezu) is experiencing explosive growth, incorporating the surrounding clusters of villages. Speakers of remote valleys in Anjaw have come and keep coming to Lohit, creating a mixture of dialects that obfuscates clear isoglosses. This is visible, e.g. in the pronunciation of codas, as discussed in §2.2.4.1. My impression is that the dialectal variation does not go beyond minor phonological processes and some lexemes but this impression can be due to the sample of speakers I have encountered. One exception is the varieties spoken in Chagla, which exhibit a markedly different phonology, lexicon and also have different verbal suffixes¹⁶. Since two of my main informants are from Chagla, there is undoubtedly a dialectal mix in my data. An exploration of the

¹⁵The catalog is available at <https://catalog.paradisec.org.au/collections/RH1>.

¹⁶Chagla speakers who reside in the plains adopt a form of soft-Chagla variety that can be understood

varieties of Tawrã spoken in Anjaw would be of immense value for a historical study of Kera'a-Tawrã. Anjaw district, however, remains inaccessible to foreigners as of now.

Table 1.9 lists several recordings that correspond to the corpus used in this grammar. It indicates the total length of the recordings, the percentage which is glossed and the date and location where it was recorded. *Type* stands informally for the genre of the file; whether it is a story (S), a conversation (C) or both. Under *Notes* I indicate whether somebody else made the recording. The initials *JO* stand for Shri Johakso Manyu. The initials *TR* stand for recordings given to me by Shri Johakso Manyu that were recorded during a language training program TRICL organized by Dr. Mark W. Post and Dr. Yankee Modi.

by non-Chagla speakers. This situation is comparable to speakers of Wallis-Swiss-German, a **Highest** Alemannic variety, residing in places where **High** Alemannic varieties are spoken, like Bern or Zurich.

Table 1.9: Data sources

Name of file	Length	Glossed length	Date	Location	Type	Note
TA22_frog1	12min47s	fully glossed	150122	Sadiya, Asm	S	
TA22_JAstory	3min33s	fully glossed	310122	Sadiya, Asm	S	
TA22_SOstory	1min36s	fully glossed	010422	Sadiya, Asm	S	
TA22_clan	7min38s	fully glossed	050622	Tezu, AP	S, C	
TA22_CLAN2	6min32s	fully glossed	050622	Tezu, AP	S, C	
TA22_origin	9min43s	fully glossed	110622	Tezu, AP	S	
TA22_tiger	5min30s	fully glossed	091222	Sadiya, Asm	S	
TA23_phlakam_clan	2min29s	fully glossed	111222	Tezu, AP	S	JO
TA22_kutyemba_phla	9min47s	fully glossed	121222	Sadiya, Asm	S	
TA22_chicken_thief	12m26s	7min2sec	161222	Sadiya, Asm	S	
TA23_tayang_clan	26min15s	9min45s	020123	Tezu, AP	S, C	JO
TA23_dimha_clan	5min39s	1min13s	080123	Tezu, AP	S	JO
TA23_talyu_clan	8min56s	fully glossed	080123	Tezu, AP	S, C	JO
TA23_pineapple	5min6s	2min57s	100123	Pasighat, AP	S	JO
TA23_chaitom_clan	2min50s	fully glossed	311023	Tezu, AP	S	JO
TA24_land	23min24s	10min14s	201024	Tezu, AP	C	
TA23_FishStory	5min13s	fully glossed	241015	Tezu, AP	S	TR
TA23_newmyth	6min20s	fully glossed	240615	Tezu, AP	S	TR

The text TA24_land deserves a note, since it has been anonymized in order to avoid any potential conflict.

1.6 Main typological features

This section will present the main typological features of the language. As such, this serves as an overview of the grammar. §1.6.1 discusses the phonology of the language. §1.6.2 presents the main points regarding parts of speech (POS). §1.6.3 discusses the main points of morphology. Multi-clause constructions are introduced in §1.6.4. Finally, morphosyntax is discussed in §1.6.5

1.6.1 Phonology

Tawrã has 25 consonant phonemes and two marginal consonant phonemes. These are distributed in five places of articulation (labial, alveolar, palatal, velar and glottal). Stops and affricates make a three-way distinction between voiceless non-aspirated, voiced non-aspirated, voiceless aspirated. There are six vowels and four diphthong rhymes. The central vowel displays a high level of variation between a close and a mid position. Tawrã is a tone language that distinguishes four tonemes: high-falling, high, low and mid tone. As it is common for languages of Northeast India, the tone-bearing unit (TBU) is the (phonological) word. Polysyllabic words are the locus of interaction between tone and stress, which yields tonal patterns. Most relevant is the case of disyllabic words, which have two possible patterns: in an *iambic* pattern, the final syllable receives stress and all four tonemes are available. This is also by far the most common word type in the language. Conversely, in a *trochaic* pattern, the initial syllable is stressed and the only toneme available is high-falling.

1.6.2 Word classes

Major open word classes in Tawrã are nouns, verbs and adjectives. While verbs and nouns can be distinguished due to their syntax and the affixes available to them, adjectives share properties with both: they can occur as heads of NPs without any derivation but they also host two verbal suffixes. Other open word classes include adverbs (often homophonous with adjectives) and classifiers, of which an important number are transparently derived from nouns.

Closed classes include *proforms* (interrogative pronouns, personal pronouns, among others), particles, numerals and demonstratives. *Particles* stand for a group of syntactically bound, uninflected words that often have a clause/phrase-final position as well as clausal/phrasal scope and provide pragmatic and epistemic information.

A feature of the word classes in Tawrã is cross-class membership: lexical roots often occur in more than one word class, e.g. a root may occur as a noun, as a verb and as an adjective without any form of derivation.

1.6.3 Morphology

Tawrã is an agglutinative, predominantly suffixing language. *Words* are here distinguished from *affixes*; while the former can stand alone without any form of affixation or derivation, the latter is morphologically bound to a word and cannot head a phrase on its own. Both verbal and nominal heads are words and can stand alone without suffixing. Nouns have three affixes and a wealth of postpositions available. Nominal modification of the type N-N triggers a tonal change that affects mid and low tones making them high. Predicates have a vast array of *tense, aspect, mode and number* (TAMN) suffixes available and they can furthermore be expanded through *predicate derivations* and reduplication. TAMN suffixes are not inflectional in the syntactic sense. Although TAMN suffixes and predicate derivations correspond to two different slot positions within the predicate complex, they often have overlapping functions that indicate aspect/Aktionsart, manner, modality, valency change, number and egophoricity. Negation has its own slot position after all other verbal suffixes.

1.6.4 Multi-clause constructions

There are several ways in which more than one clause can be combined in Tawrã. Two independent clauses can be conjoined without conjunction; clauses can be simply apposed to each other. A more prevalent strategy is cosubordination with the morpheme *gũ*. Non-final clauses (marked with *gũ*) conclude in a final clause (without *gũ*). The structure of these clauses is flat (i.e. not nested), and thus all non-final clauses modify the final clause. Conversely, subordination stands for a nested structure in which a subordinate clause can only modify the clause where it is embedded. Tawrã has a wealth of subordinating devices that indicate the type of subordination: temporal, causal, conditional, ‘as if’, etc.

An areal and typologically interesting feature of Tawrã is its dearth of nominalizers, having only one device that turns verbs into places where the action of the verb may take place (e.g. ‘place of sleep’, thus locative nominalizer). Relative clauses and event nominalization do not have their own nominalizing device and rely on the syntax of the construction.

1.6.5 Morphosyntax

There is no evidence for a strong grammaticalization of morphosyntactic properties that relate an argument to a clause (e.g. as a Subject or Object) in Tawrã. Common argument selectors like case, predicate marking or constituent order do not yield a clear alignment (e.g. accusative-dative, ergative, etc.) and are rather overridden by semantico-pragmatic considerations. Conversely, topic-focus seems a more apt theory to account for the disambiguation of role and reference of arguments. Although the topic can be readily associated with S/A arguments, they do not have to coincide.

1.6.6 Intersections of language, culture and cognition

There are two aspects of the description of the language that touch on larger questions of culture and cognition and that are of typological interest: ‘coherent group sensibility’ and classification.

By ‘coherent group sensibility’ I mean a tendency in the language to mark plural referents that show internal cohesion, e.g. a clan, a family or group of peers. This is evident in: a special plural verbal suffix *-rẽ/-ring* (opposed to *-ga*, cf. §6.3.2.0.1) and the selective use of egophoricity-sensitive verbal suffixes in conjunction with first person plural referents (§6.3.1.1). This tendency in the language is consistent with the findings of Modi (2017, 212-216) regarding the *ideocclusive* pronoun set in Milang. There seems to be an areal feature of socio-epistemic diversity which surfaces as a distinct grammatical marking for coherent groups.

The second point relates to the role of classifiers (§5) in the language. As devices of nominal classification, they reveal aspects about cognition and how native speakers organize the world. The semantic distribution of classifiers is driven by perceptual features (size, shape, etc.), functionality and the use of metaphors that expand the range of classified nouns. They also have grammatical functions related to referentiality and individuation.

1.7 Transcription, orthography and conventions

Tawṛā Mishmi is eminently a spoken language with no written tradition. In India, there are on-going attempt to write Tawṛā using roman letters. There has been, to my knowledge, no serious attempt the write down the language in a different script, at least not on the Indian side of the border. The present work makes use of a modified form of the orthography that is employed by the *Cultural & Literary Society of the Mishmi* (CALSOM) as well as by Evans and Manyu (2021) in their latest publication. The orthography is mostly phonological and it reflects my analysis of the phonology of the language but also includes some simplifications (e.g. regarding aspiration). Table 1.10 illustrates a list of consonants along with the IPA value in square brackets. The main rationale behind the orthography is to employ symbols from the Roman alphabet instead of IPA symbols, thus [ŋ] is written as <ng> and [ɦ] as <h>. This renders this orthography accessible for both native and non-native speakers with a knowledge of written English or transliterated IA languages.

Table 1.10: Consonants in Tawṛā. The sounds in parentheses are marginal phonemes.

manner\place	Labial		Alveolar		Palatal		Velar		Glottal
	-v	+v	-v	+v	-v	+v	-v	+v	
Stop -asp.	p [p]	b [b]	t [t]	d [d]			k [k]	g [g]	(q [ʔ])
+asp.	ph [p ^h]		th [t ^h]				kh [k ^h]		
Affricate -asp.			ts [ts]	z [dz]	c [tɕ]	j [dʒ]			
+asp.			tsh [ts ^h]		ch [tɕ ^h]				
Nasal		m [m]		n [n]				ng [ŋ]	
Fricatives			s [s]		sh [ɕ]				
Lateral			l [l]						
Approx.		w [w], (v [v])	r [ɹ]		y [j]				h [ɦ]

Examples from IA is here rendered through a macron <̄> (thus <tā>)¹⁷. Aspirated consonants take an <h>, thus ɸ [ɸ^h] is written as *dh*. Retroflex consonants are rendered with an underdot, thus ɽ [ɽ̣] is written <ṭ>. Note, however, that this does not apply to

¹⁷It should be noted that this is not maintained outside of the examples, i.e. in the running text: IA loanwords that are common in (Indian) English follow the English spelling and are in cursive, e.g. *beedi* (and not <bīḍī>) and *pūja* (and not <pūjā>).

IA loanwords in Tawrã, which follow the general orthography of the language proposed here. The remainder of this section will discuss particular issues of the orthography employed for Tawrã.

1.7.1 Consonants

This orthography follows the widespread practice of using an <h> after the consonant to indicate aspiration. Thus [t_ɕ] and [t_ɕ^h] are written <c> and <ch> (and not <ch> and <chh>¹⁸).

The glottal stop (a defective phoneme, cf. §2.4.4) is written as <q> except when it occurs in a predictable word-initial position before a vowel. Table 1.11 illustrates four positions where the glottal stop occurs: word-initially before a vowel and as part of an onset consonant cluster (*qr*-, *qy*- and *qw*-), word internally between two vowels and as a constitutive aspect in coda position. Only in the first case, word-initially before a vowel, is the glottal stop not written.

Table 1.11: The orthography of the glottal stop

	word-initial before vowel	word-initial cluster	word-internal between vowels	coda position (continuative aspect)
orthography	<i>abe</i> ‘morning’	<i>qwîng</i> ‘old’	<i>kaqîng</i> ‘two’	<i>thâ-q</i> ‘he is eating’ (eat-CONT)
IPA	[ʔə.ˈbe]	[ʔwîŋ]	[kə.ˈʔîŋ]	[t ^h âʔ]

1.7.1.1 Consonant clusters

Tawrã has several onset consonantal clusters. The decision to treat these as clusters and not as phonemes is discussed in §2.15 and follows the work of Evans and Manyu (2021). Table 1.12 illustrates all clusters¹⁹ found in the language. Most CC combinations

¹⁸This is a common transliteration of IA languages.

¹⁹When Tawrã names were first written down for the integration of their community into the incipient Indian state, the officials were presumably Assamese and could not distinguish between alveolar and palatal

render a predictable cluster (e.g. [k^h] + [ɿ] gives [k^hɿ]) and, therefore, there is no need to further comment on their orthographic depiction except for five clusters which have an unpredictable pronunciation: *wr* [ɿ^w], *hr* [ɿ̥], *hl* [l̥] or [ɬ], *pw* [pu̯] and *bw* [bu̯].

Table 1.12: Attested onset clusters.

	Cr-				Cl-			
Velar	kr	khr	gr		kl		gl	
Alveolar	tr	thr	dr					
Affricate								
Labial	pr	phr	br	mr	pl	phl	bl	ml
Lateral								
Approx.	wr [ɿ ^w]	hr [ɿ̥]			hl [l̥] or [ɬ]			
Fricative								
Glottal	qr							
	Cy-				Cw-			
Velar	ky		gy		kw		gw	
Alveolar	ty		dy	ny				
Affricate						tshw		
Labial	py		by	my	pw [pu̯]		bw [bu̯]	
Lateral			ly					
Approx.	hy				hw			
Fricative	shy							
Glottal	qy				qw			

1.7.2 Vowels

Tawrã has six phonemic vowels, written <a, e, i, o, u, ü>. Their approximate IPA values are illustrated in Table 1.13. Note that these phonetic values hold for *heavy* syllables; light syllables are often phonologically eroded, e.g. /a/ is often pronounced [ə] in a

stops (note that Assamese has a palatal stop as an allophone of /s/ in certain dialects). This is why certain personal names have a <ch> although they are not pronounced as [tɕ], but with an alveolar affricate [ts] as in the case of *Tsâi.tûm* ‘a clan’ which is pronounced [tsâi.tûm] but written down as <Chaitom> for all official purposes.

light, non-stressed syllable (cf. §2.2.5). The orthography employed here does not make a difference between heavy and light vowels, mainly because the language does not have phonological vowel length (cf. §2.3.2.1). There are four vocalic rhymes in the language, as depicted in Table 1.14.

Table 1.13: Orthography of vowels

Orthography	Approximate IPA
a	[a]
e	[e]
i	[i]
o	[o]
u	[u]
ü	[ʊ~i]

Table 1.14: Orthography of rhymes

Orthography	Approximate IPA
ai	[ai]
ei	[ei]
aw	[aw]
üw	[iʊ]

Nasalization is marked with a tilde atop the nasalized vowel: <ẽ ã ü õ> (Vowels /ü/ and /i/ are never nasalized, cf. §2.3.2.2).

1.7.3 Tone

Tawrã is a tonal language and tone is indicated in the present orthography. For monosyllabic morphemes, tone is marked by a diacritic: <ô> for high-falling (HF), <ó> for high tone (H), <ò> for low tone (L) and no marking for mid tone (M). Toneless monosyllabic morphemes are also unmarked, like prefix *na-* 1s.POSS. Table 1.15 presents a four-way minimal pair based on tone.

Table 1.15: Tonal minimal pairs

word	tone	translation
<i>pá</i>	H	yeast
<i>pâ</i>	HF	go across
<i>pà</i>	L	axe
<i>pa</i>	M	draw (water)

Disyllabic morphemes require a brief explanation: the tone-bearing unit (TBU) in Tawrã is the phonological word. This means that each phonological word can have up to one tone and that, on native disyllabic words, tone will always fall on only one syllable. The choice of which syllable exhibits tone is lexically determined and isomorphic with stress. For a disyllabic word, there are two options: tone/stress can fall on either the first or the second syllable. Although all four tones can occur on any vowel, there is a restriction on which tone can occur on the first syllable of a disyllabic word: only a high-falling tone (HF) can occur in a first syllable (termed *trochaic* pattern) whereas all four tones can occur on the second syllable (termed *iambic* pattern). To schematize the tonal system in disyllabic words, I have organized them in *tonal patterns*, i.e. the tone of the word is notated along with a dummy mid tone (M) that represents the toneless syllable, thus a wordlike *bürì* ‘lower part’ is notated as M-L: the first syllable is toneless (marked M) and the second has a low tone (L), hence **M-L**. Furthermore, a word like *bülik* ‘pig’ has an initial toneless syllable (M) followed by a mid tone (also M), hence M-M. Table 1.16 illustrates these two patterns along with tonal patterns.

Table 1.16: Tone marking on disyllabic words

trochaic pattern	iambic pattern
<i>bûra</i> ‘earlier’ (HF-M)	<i>bülik</i> ‘pig’ (M-M)
	<i>bürì</i> ‘lower part’ (M-L)
	<i>dabyûm</i> ‘jackfruit’ (M-HF)
	<i>daráng</i> ‘big’ (M-H)

1.7.4 Other conventions

Compounds with a non-compositional meaning, e.g. *bǎ-mě* (lit. father's people) 'clan', or where one of the elements is unknown are written with a dash (-) in plain text. In a glossed example, such compounds take a dot (i.e. *bǎ.mě*) in order to avoid confusion. Compositional (i.e. transparent) compounds are merely apposed, e.g. *ya dabû* (night middle) 'midnight'.

Morphological concatenation (affixation or compounding) is indicated with a dash (-), e.g. *na-tya* (1S.POSS-grandfather) 'my grandfather'. An equal sign stands for clitization, cf. the clitic form of the negative marker *yüm: bô = m* (go = NEG) 's/he won't go'. A tilde (~) is reserved for the reduplication of verbal stems.

Square brackets ([]) stand for the phonetic transcription, whereas slashes (/ /) stand for the phonological transcription. Hence a word like *bwì* 'dance' is transcribed phonemically as [bɯɪ] but phonologically as /bwì/. The orthography employed in this work aims at being phonological, as described in §1.7. Both phonological and phonemic transcription indicate stress with the stress mark ('), as in *tarüm* 'flock', transcribed as /ta'rüm/ and [ta'ɪm]. Syllable boundary is sometimes indicated with a dot <.>, e.g. /'kûm.po/ 'door frame, entrance'.

Most examples in this grammar are presented with interlinear morpheme-by-morpheme glosses and a translation. Morpheme labels are congruent with the Leipzig Glossing Rules²⁰ although many categories have been crafted to fit the Tawrã system. The glosses also include etymological information regarding loanwords: thus (IA) stands for IA²¹, (Asm) for Assamese, (Tib) for Tibetan (only one case) and (E) for English. All abbreviations and conventions are listed in *Abbreviations and conventions*. Glosses are intended for Tawrã and the few instances where a language different from Tawrã is glossed are always marked with the name of the language.

²⁰ Available at <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>.

²¹ In many cases, it cannot be established whether a loanword has been borrowed from Assamese or Arunachali Hindi, cf. *gantā* 'hour'. Such cases are signaled as 'IA'.

2 Phonology

This chapter presents a comprehensive description of the phonology of Tawrã. §2.2 describes the syllable structure of the language and discusses the concept of the phonological word (*pword*). §2.3 describes the vowels and rhymes (diphthongs) in the language. §2.4 presents an overview of the consonants organized according to their place of articulation. §2.5 discusses tone and introduces the tonal patterns of disyllabic words.

2.1 Overview

Tawrã has a moderately large inventory of phonemes with a three-way distinction among stops and affricates (voiceless non-aspirated, voiced non-aspirated, voiceless aspirated). Since consonants and vowels are not discussed until §2.4 and §2.3 respectively, a summary overview of Tawrã consonants, vowels and rhymes are presented in Tables 2.1, 2.2 and 2.3 to aid the readability of this chapter, although all relevant discussions and details are found in the corresponding sections. Symbols in parentheses are defective phonemes, i.e. there are no instances where the glottal stop would contrast with another phoneme.

Table 2.1: Consonants in Tawrã. The sounds in parentheses are marginal phonemes.

manner\place	Labial		Alveolar		Palatal		Velar		Glottal
	-v	+v	-v	+v	-v	+v	-v	+v	
Stop -asp.	p [p]	b [b]	t [t]	d [d]			k [k]	g [g]	(q [ʔ])
+asp.	ph [pʰ]		th [tʰ]				kh [kʰ]		
Affricate -asp.			ts [ts]	z [dz]	c [tɕ]	j [dʒ]			
+asp.			tsh [tsʰ]		ch [tɕʰ]				
Nasal		m [m]		n [n]				ŋ [ŋ]	
Fricatives			s [s]		sh [ɕ]				
Lateral			l [l]						
Approx.		w [w], (v [v])	r [ɹ]		y [j]				h [ɦ]

Table 2.2: Vowels in Tawrã

	Front	Central	Back
High	i [i]	ü [i~u]	u [u], ũ [ũ]
Middle	e [e], ě [ě]		o [o], õ [õ]
Low			a [a], ã [ã]

Table 2.3: Rhymes in Tawrã

-ai [ai]	-ei [ei]	-aw [aw]	-üw [i̯u]

Tawrã is a tone language with a four-way distinction between mid, high, high-falling and low tone. Table 2.4 displays a list of minimal pairs of monosyllabic words. Several suffixes are toneless (§2.5.4).

Table 2.4: Tonal minimal pair

word	tone	translation
<i>pá</i>	H	yeast
<i>pâ</i>	HF	go across
<i>pà</i>	L	axe
<i>pa</i>	M	draw (water)

§2.2 introduces an important distinction between *heavy* and *light* syllables that helps explain other phonological phenomena like tone at the word level and phonological erosion. What is relevant for this overview is that a phonological word consists of minimally and maximally one heavy syllable, which attracts both stress and tone. Although stress and tone may fall on either syllable of a polysyllabic word, there are certain restrictions on the type of tone that can surface in each syllable. To explain this distribution, §2.5.5 introduces the concept of *tonal patterns*: while a disyllabic word with tone/stress on the **initial** syllable (termed *trochaic foot*, cf. §2.2.1.1) may only have a high-falling tone (termed here *M-H*; no other tone is allowed), a disyllabic word with tone/stress on the **final** syllable may have any of the four tones available (*M-HF*, *M-H*, *M-M*, *M-L*). This is illustrated in Table 2.5.

Table 2.5: Tonal patterns in disyllabic words

tone/stress on initial syllable (termed <i>trochaic foot</i>)		tone/stress on second syllable (termed <i>iambic foot</i>)	
HF-M	<i>mâlai</i> ‘scatter’	M-HF	<i>malâ</i> ‘search’
		M-H	<i>malû</i> ‘sickness’
		M-M	<i>malai</i> ‘spring’
		M-L	<i>malà</i> ‘tip, edge’

2.2 Syllable structure and wordhood

This section describes the nature and structure of both Tawrã syllables and phonological words. A distinction will be introduced here between *heavy* and *light* syllables, which is relevant to understanding other areas of the phonology such as non-phonemic vowel length and tone/stress in polysyllabic words.

This section is organized as follows: §2.2.1 introduces the distinction between *heavy* and *light* syllables based on syllable weight measured in *moras*. §2.2.2 describes the key properties of the heavy syllable. §2.2.3 and §2.2.4 describe syllable onsets and codas respectively. §2.2.5 describes the properties of light syllables, which only occur in a polysyllabic word. §2.2.6 introduces the notion of stress and tone in polysyllabic words. §2.2.7 discusses the phonological word based on moraic structure.

2.2.1 Moraic structure and monosyllabic wordhood

Much of Tawrã phonology depends on distinguishing between *heavy* (H) and *light* syllables (L). A *heavy* syllable (H) has either a consonantal coda (§2.2.4) or a long vowel/rhyme. Only a heavy syllable may carry both tone and the primary stress, which are coterminous (cf. §2.5). By contrast, a *light* syllable (L) has a short vowel, which is often at the mercy of varied forms of erosion (cf. §2.2.5, note the discussion on *sesquisyllable*), it has no coda and attracts neither stress nor tone. The difference in weight between these two syllables (H and L) can be conceptualized in the number of *moras* (symbolized by μ), i.e. a unit of syllable weight that is defined in language-internal terms but generally refers to a combination of CV or simply a V (Gordon 2016, 187f; Ladefoged and Johnson 2006, 261; also cf. Hyman 1985). While the consonantal onset is disregarded, both the nucleus and the coda (which can be a vowel) are given a moraic value of *one*. In other words, a *pword* is minimally bimoraic. The distinction between H and L is illustrated in Table 2.6, which depicts the possible syllable shapes of the heavy syllable (counting two moras) on the left and the light syllable (counting one mora) on the right. Two VV moras in Tawrã can either be a rhyme (§2.3.3), a nasalized vowel (§2.3.2.2) or an unreduced vowel, which is structurally longer than a monomoraic vowel and fulfills the bimoraic

minimality requirement (cf. McCarthy and Prince 1986, 6, 19). Note that vowel length is not phonemic in the language in the sense that it is a structural feature, a product of the moraic structure of the language; it reflects differences in syllable weight and it is not a contrastive feature (cf. Green 2005, 7). The length of a monomoraic vowel and all its other properties are described in detail in §2.2.5.

Table 2.6: Shapes of heavy and light syllables

Heavy syllable (two moras)	Light syllable (one mora)
VV or $\tilde{V}$	V
CVV or $C\tilde{V}$	CV
CVC	

2.2.1.1 Syllabification and sonority

Syllabification refers to the separation of words into syllables. In Tawrã, it follows the widespread principle that onsets are preferred over codas, i.e. a consonant immediately before a vowel belongs to the same syllable as the vowel (Gordon 2016, 83f). Thus VCV is broken down to V.CV rather than VC.V, as with the example *abâ* ‘type of bamboo’: /a.^hbâ/ rather than */ab.^hâ/. VCCV combinations require a brief explanation, as they are broken down according to sonority rules. Sonority is commonly associated with acoustic loudness (intensity), although the phonetic phenomena are less well understood (cf. Gordon 2016, 1f; cf. Steriade 2003; also Parker 2002, 295f). Crucially for us, sonority is usually displayed as a scale where vowels are most sonorous and stops are least sonorous, here reproduced as Table 2.7.

Table 2.7: Sonority scale (adapted from Steriade (2003))

vowels	liquids	nasals	fricatives	stops
most sonorous				least sonorous

There are two relevant implications from this scale: a) sonority is a defining factor in

the syllabification process of VCCV combinations and b) consonantal clusters commonly have the member with higher sonority closer to the syllable nucleus, a feature that will be further discussed in §2.15. Concerning the former, VCCV combinations are syllabified as V.CCV unless this violates the notion of *sonority peak* (Steriade 2003; also cf. Gordon 2016, 34f), which states that there cannot be non-adjacent high sonority segments within the same syllable. In other words, there cannot be more than one tautosyllabic sonority peak. The word *ablá* ‘flame’ is syllabified as /a.'blá/ as the *blá* syllable forms one sonority peak: stops have little sonority while liquids and vowels have higher sonority. Since -l- and -a- occur adjacently, they form one peak. On the other hand, *kûmpo* ‘door frame, entrance’ is syllabified /'kûm.po/ (and not */'kû.mpo/), for -m- has high sonority and may only be syllabified to the vowel of the initial syllable. The impossible syllable **mpo* would generate two sonority peaks, one on -m- and one on the non-adjacent vowel -o-. Incidentally, only those consonantal clusters that conform to the sonority peak described here are valid onset clusters in the language (cf. Hayes 2009, 252f.), thus *bl-* is an allowed onset cluster whereas **mp-* is not.

The preference for onset rather than coda consonants in syllabification creates an imbalance between syllables in polysyllabic words as the initial syllable won’t normally display a coda while the final syllable has no restrictions. This has the practical implication that an initial heavy syllable (i.e. a trochaic pattern with both stress and tone falling on the initial syllable, cf. §2.2.6) will never display its coda consonant other than a nasal (no other high sonority consonant is allowed in that position in the language). To explain this, Table 2.8 illustrates four disyllabic words along their syllabic breaks and meanings. The first two examples correspond to the iambic pattern (tone and stress fall on the final syllable, cf. §2.2.6) while the last three words correspond to the trochaic pattern (stress and tone fall on the initial syllable). In the case of *tabrê* /ta.'brê/ ‘meat’ and *taglû* /ta.'glû/ ‘provisions’, the consonants that make the intervocalic clusters fall on the onset of the final syllable since a light syllable has per definition no coda (cf. §2.18). However, the trochaic nouns *tûpla* /'tû.pla/ ‘fold’ and *rûbrung* /'rû.brung/ ‘bone’ do not have a coda on the initial syllable due to the preference of onsets before codas, although their initial, heavy syllable would, in theory, allow /-b/ and /-g/ codas (cf. §2.2.2 on heavy syllables and §2.2.4 on codas). Lastly, as discussed above, *kûmpo* ‘door frame, en-

trance' is syllabified /^lkûm.po/ i.e. /^lCVC.CV/ with the initial, heavy syllable allowing a coda because **mp-* is not an allowed onset cluster in Tawrã. In sum, it is argued here that the heavy initial syllable (i.e. in a trochaic pattern) does not show a defective structure in disallowing most codas but this is rather the effect of the preference of onsets over codas and the sonority peak rule that disallows codas in that position and syllabifies them to the following tautomorphemic onset.

Table 2.8: Examples of syllabification

Example	Syllable breaks	Meaning
<i>tabrê</i>	ta. ^l brê	meat
<i>taglû</i>	ta. ^l glû	provisions
<i>tûpla</i>	^l tû.pla	fold
<i>rûbrung</i>	^l rû.brung	bone
<i>kûmpo</i>	^l kûm.po	door frame, entrance

2.2.2 Heavy syllable (H)

A phonological word in Tawrã must have maximally and minimally one heavy syllable. The structure of a heavy syllable is represented in Table 2.9. C₁ stands for the optional initial consonant or syllable onset, both word-initial and medial. The precise consonants that can fill that position are listed in §2.2.3. The optional C₂ can only occur if there is a C₁. Following the economic approach to the phonology of the language proposed by Evans and Manyu (2021), features of secondary articulation such as labialization and palatalization are treated as consonant clusters and further discussed in §2.2.3.1. If no consonant precedes the syllable nucleus, a prothetic glottal stop occurs automatically before the vowel (cf. §2.4.4). The nucleus (V) is the only obligatory element in the syllable. Since it is stressed, it must take one of the four tones (cf. §2.5). Vowels are further discussed in detail in §2.3. Lastly, X stands for the optional coda, which may either correspond to a) vowel length of the nucleus (V:)¹ or a rhyme (VV),

¹As mentioned in §2.2.1.1, vowel length is not phonemic in the language in the sense that it does not contrast e.g. in a monosyllabic word [ba] vs [ba:]. Rather, vowel length is used here to represent the

b) nasalization of the nucleus vowel ($\tilde{V}$) or consonantal coda (C_3). The pool of possible codas in the language is relatively small with only seven consonants (/ -p, -b, -k, -g, -m, -n, -ng/) and no clusters allowed. Codas are further discussed in §2.2.4.

Table 2.9: Maximal structure of a syllable

		+ tone	
(C ₁)	(C ₂)	V	(X)
Onset		nucleus	coda

Table 2.10 illustrates all types of monosyllabic syllables found in the language. It also includes the moraic structure (note that VV is depicted as two vowels).

Table 2.10: Examples of all syllable structures

syllable shape	example	moraic str.	translation	comment
V	<i>a</i>	aa	child	No-onset syllables take a prothetic glottal stop cf. §2.4.4
VC ₃	<i>am</i>	am	mineral lick	idem.
C ₁ V	<i>bi</i>	bii	elder sister	
C ₁ $\tilde{V}$	<i>tã</i>	tã	fish	
C ₁ VC ₃	<i>bôm</i>	bôm	pour down	
C ₁ C ₂ V	<i>pra</i>	praa	good	
C ₁ C ₂ VC ₃	<i>dyûm</i>	dyûm	be full	

Two heavy syllables may co-occur in a trochaic word (H.H). In these cases, stress (and tone) will fall only on the initial syllable, for stress falls on the first heavy syllable from left to right, as will be discussed in §2.2.6. An initial heavy syllable may not have a coda other than a nasal due to syllabification rules described in §2.2.1.1. However, I have not found a H.H trochaic word with a nasal coda on the first syllable. Table 2.11 illustrates the structure of a H.H (heavy.heavy) trochaic word in which each syllable

structural opposition between the vowels in a *light* syllable (§2.2.5) and a *heavy* syllable. This is further discussed in §2.3.2.1.

weighs 2 moras ($\mu\mu$). All elements in parentheses are optional. Furthermore, no word in my corpus would have all elements of this table. This, however, may also be because trochaic words are *per se* rare in the language compared to iambic words. Table 2.12 displays a number of H.H trochaic words in the language along with different syllabic structures.

Table 2.11: Structure of a H.H trochaic word

+ stress/tone	
(C ₁)(C ₂)V(X).	(C ₁)(C ₂)VV(X)
$\mu\mu$.	$\mu\mu$

Table 2.12: Examples of H.H trochaic words

Example	syllabic structure	Translation
<i>côcang</i> /'cô.cang/	C ₁ VV.C ₁ VC ₃	type of bird
<i>nyâhing</i> /'nyâ.hing /	C ₁ C ₂ VV.C ₁ VC ₃	epilepsy
<i>rûbrung</i> /'rû.brung/	C ₁ VV.C ₁ C ₂ VC ₃	bone
<i>sâkwã</i> /'sâ.kwã/	C ₁ VV.C ₁ C ₂ $\tilde{V}$	anus
<i>klâipom</i> /'klâi.pom/	C ₁ C ₂ VV.C ₁ VC ₃	intestines
<i>nâshik</i> /'nâ.shik/	C ₁ VV.C ₁ VC ₃	hunger
<i>lâhlak</i> /'lâ.hlak/	C ₁ VV.C ₁ C ₂ VC ₃	trick, cheat

A final note on trochaic words is in order before continuing: once stress has been assigned (from left to right) the metrical system does not discriminate between a heavy and a light syllable. The main consequence of this is that we may henceforth collapse the distinction between trochaic **H.H** (heavy-heavy syllable) and trochaic **H.L** (heavy-light syllable), for it bears little repercussion in the phonology. As an example, in a word like *sûmbe* /'sûm.be/ 'type of earring' the second syllable *be* is open and not stressed. Without the cue of stress, a light syllable CV simply does not contrast with a heavy syllable CVV (cf. §2.3.2.1). We may therefore write the syllable shape of this word as C₁ VN.C₁ V?, noting that we cannot know whether it is a heavy or a light syllable. This is illustrated

in Table 2.13.

Table 2.13: Structure of the trochaic word *sûmbe*

Example	syllabic structure	Translation
<i>sûmbe</i> / ^l sûm.be/	C ₁ VN.C ₁ V?	type of earring

2.2.3 Syllable onset

This section describes syllabic onsets, i.e. the consonants that precede the nucleus, including onset clusters. A syllable onset is not mandatory in the language. Some words without an onset include *agã* ‘set up (a trap)’ and *asûm* ‘taro’ (*colocasia esculenta*) (the glottal cover of no-onset words is discussed in §2.4.4). Table 2.14 displays the consonantal phonemes allowed in onset position. This includes almost all consonant phonemes in the language except for the glottal stop, a marginal phoneme that occurs a) predictably before a vowel (a syllable with no onset) and b) in the (marginal) clusters *qy*, *qr* and *qw* (§2.2.3.1) but does not occur as a contrastive consonant in onset position. The glottal stop is discussed in detail in §2.4.4 while the orthographic rules have been discussed in §1.7.

Table 2.14: Syllable onsets

		labial	alveolar	palatal	velar	glottal
stop	-asp.	p-, b-	t-, d-		k-, g-	
	+ asp.	ph-	th-		kh-	
affricate	-asp.		ts-, z-	c-, j-		
	+ asp.		tsh-	ch-		
nasal		m-	n-		ng-	
fricative			s-	sh-		
lateral			l-			
approx.		w-	r-	y-		h-

2.2.3.1 Onset consonant clusters

Tawrã consonant clusters occur in the syllable onset position while they are disallowed in coda position, a typologically common configuration (Steriade 2003) (cf. §2.2.4). Consonant clusters follow a widespread preference for having the member with the highest sonority closer to the syllable nucleus, or, in other words, onset clusters display a rising sonority profile (cf. Gordon 2016, 99) (sonority has been introduced in §2.2.1.1). Some examples are /kr-/ as in *krâng* ‘bite (of a person)’ or *plà* ‘salt’. This analysis of Tawrã clusters also includes CC- combinations that are discussed as palatalizations² and labializations in other languages.

Table 2.15 displays all known combinations of C₁C₂: On the leftmost column is the place of articulation of the C₁. Within each of the four other columns (Cr-, Cl-, Cy- and Cw-) each line stands for different modes of C₁ (for each place of articulation). The asterisk before the individual cluster indicates that it has been attested only once.

²The Cy- group in Tawrã bears a striking similarity with onset clusters in Tangam (Post 2017, 19f) and Milang (Modi 2017, 96). As described by Post (2017, 19), the Proto-Tani clusters are usually well preserved in Tangam. In the case of Cy- clusters (which includes *Cr- clusters that became Cy-), these did not palatalize, as in most Western Tani varieties, nor were they simplified to C-, as in some Eastern Tani varieties. The Tawrã *tamyá* and Tangam *myo* both ‘tiger’ (Proto Tani *mro /myo?, cf. Sun 1993, 462) would thus be a case of shared retention, either from a PTB or a common Tawrã-Kera’a-Macro-Tani form.

Table 2.15: Attested onset clusters. Asterisk <*> indicates that the cluster has been attested in only one word.

	Cr-	Cl-
Velar	kr khr gr	kl gl
Alveolar	tr thr dr	
Affricate		
Labial	pr phr br mr	pl phl bl ml
Lateral		
Approx.	wr [ɹw] hr [ɹ]	hl [l̥] or [ɬ]
Fricative		
Glottal	*qr	
	Cy-	Cw-
Velar	ky gy	kw gw
Alveolar	ty dy ny	
Affricate		*tshw
Labial	py by my	pw [pɥ] *bw [bɥ]
Lateral	ly	
Approx.	hy	hw
Fricative	shy	
Glottal	*qy	qw, *qwr

The C₂ /w/ has an allophone [ɥ] in clusters *pw* and *bw*, thus *bwì* ‘dance’ is pronounced [bɥì] and *pwi* ‘other, different’ [pɥì] (Evans and Manyu 2021, 5).

As discussed in §2.4.9, the approximant /ɹ/ in C₂ position is realized as [ɹ] after a velar or alveolar C₁ (cf. Evans and Manyu 2021, 10).

The decision to analyze these as clusters rather than as extra phonemes was proposed by Evans and Manyu (cf. 2021, 6f) and follows an economic approach to avoid a large number of phonemes. Based on an independent analysis, the present study has also adopted this decision for three reasons: a) [j] (written <y>) and [w] are independent phonemes (cf. §2.4.9), b) there are no instances in which Cy-, Cw- would behave differently from Cr-, Cl- regarding the first element in the cluster. Table 2.15 includes

a number of phonological gaps, such as *phw* or *tsh*. There could be phonological constraints that hinder those precise combinations but I have not detected them. Rather, I think that these gaps are mostly the result of a normal Zipfian distribution in which certain phonemes (in this case e.g. the alveolar affricate group) are rarer than others. Cf. the cluster *tshw* which has only one attestation: *tshwî* ‘flea’. C) Clusters are associated with heavy syllables; in an iambic pattern³ (i.e. tone/stress falling on the second syllable of a disyllabic word, cf. §2.2.1.1) clusters can only occur on the second (heavy) syllable, e.g. *matyâ* ‘release, set free’ or *tabri* ‘bamboo basket’. Conversely, a word like **myasá* **[mja'sá]* would not be allowed, as the syllable with the cluster onset is not stressed. In sum, there are three arguments to treat these clusters as C-C combinations: the second element in these clusters is always a full phoneme, all C-C combinations have the same behavior and, morpho-phonologically, C-C combinations can only occur on heavy syllables.

As discussed in §2.18, unstressed syllables can drop their vowel nucleus altogether and fuse their onset consonant with the following syllable, thus giving rise to *secondary* clusters, as with *kurû* ‘head’, which can be heard as *[k.ɹû]* (as well as *[kɪ.'ɹû]*). These *secondary* clusters are the product of morpho-phonological processes since they originate from a CV- light syllable but they do not create new CC- combinations that are not attested in Table 2.15, thus a noun like *katsûm* ‘chest’ does not surface as **[ktsûm]* (it rather loses the onset, as *[a'tsûm]* or *[a'sûm]* cf. §2.18). Secondary clusters can be further distinguished from *lexically specified* clusters based on their separability, while *[k.ɹû]* ‘head’ can be separated to *[ku.'ɹû]*, *krî* ‘ruminate’ cannot be separated to **[kɪ.'ɹi]* or the like, suggesting a direction in sound change from unstressed syllable towards clustering (cf. Morey 2012, 153f). Whether this is the origin of these lexically specified onset clusters remains to be established.

The cluster combinations *qy-*, *qw-* and *qr-* have been described by Evans and Manyu (2021, 7f). These initial clusters are rare in the language, with only a handful of attested forms. Some examples are *qwîng* ‘be old, be ruined’, *qwûw* ‘second wife’, *qyá-raw* ‘young woman’ and *qrû* ‘pry’. An interesting case is the complex onset cluster *qwrû* ‘lever’, which

³In the case of trochaic words this is more complicated, as they often seem to go back to a collocation of two heavy syllables and thus the second, non-stressed syllable does not have to be light.

has a cluster with three members: q-w-r.

A note is also required for the *hl* cluster. In careful speech, it is pronounced [həl] but [l̥] or [ɬ] is casual speech. The phoneme [l̥] is widespread in the language family, although it often surfaces with reduced distribution. As an example, in Kurtöp (East-Bodish, Bhutan) the voiceless lateral [l̥] only occurs in syllable onset position before a high tone (Hyslop 2017, 137f. also cf. Burmese Jenny and Hnin Tun 2017, 17; cf. an equivalent evaluation of the cluster/phoneme problem by Handel 2003, 564). In Tawrã, the analysis as a cluster rather than an independent phoneme is preferred for the reasons listed above (i.e. both C₁ and C₂ are independent phonemes in the language, all C₁C₂-combinations behave similarly and *hl* only occurs in heavy syllables) and because of the structural C₁h- cluster found in all four CC combinations (*hl hr hy* and *hw*). However, there is one analytical issue found in the Chagla dialect of Tawrã. The noun *kahlâi* ‘guts’ is sometimes assimilated to *khhlâi* [k^hl̥âi] by speakers from Chagla, making an onset cluster of three members /kh-h-l/. This development has not been detected in the varieties of Tawrã taken as the basis for this investigation and should be further explored.

2.2.4 Syllable coda

Codas are not obligatory in the language and words without a coda are plentiful, cf. *maro* ‘talk’ or *tahrô* ‘ten’. Consonant clusters are not allowed in coda position. Table 2.16 displays all consonants allowed in coda position. A first overview reveals that Tawrã is restrictive in this syllable position, for only six consonants are allowed (cf. Evans and Manyu 2021, 11)⁴. As described by Evans and Manyu (2021, 10f), obstruent codas are actively devoiced and non-released. This is discussed in §2.2.4.2.

⁴As remarked by Evans and Manyu (2021, 11), the lack of alveolar codas is particularly striking but a similar distribution is found in the phonology of Northern Sangtam (Aoic): while the onset may be occupied by any of the 27 phonemic consonants, coda position may only be filled by /p, k, m, ŋ, ɬ/ (Coupe 2020, 160f). Also note that Galo and most Western Tani languages do not allow alveolar stop codas (Post 2007, 891).

Table 2.16: Syllable codas

		labial	alveolar	palatal	velar	glottal
stop	non-asp.	-p, -b			-k, -g	
	asp.					
affricate	non-asp.					
	asp.					
nasal		-m			-ng	
fricative						
lateral						
approx.						

2.2.4.1 Alternating rhymes

As described by Evans and Manyu (2021, 10f), a series of rhymes have a lexicalized alternative pronunciation that is either a /-ei/, /-üw/ or /-aw/ rhyme or a /-i/ monophthong. Note that these rhymes are the product of a morpho-phonological process and should not be confused with lexically specified rhymes *ei*, *ai* and *üw* discussed in §2.3.3. Table 2.17 displays all rhyme alternations. While some rhymes have one alternation, others like *-ik* have two or more possible variants. It seems like these variations are lexicalized although the potential relation between tone and shape of the alternations has not been explored yet. Note that the devoicing of codas is discussed in §2.2.4.2.

Table 2.17: Alternating codas

Coda	Example	Variant I	Variant II	Variant III	Meaning
-ab	<i>kalyab</i>	[ka'ljap]	[ka'ljaw]		field
-ap	<i>ráp</i>	[ráp]	[ráw]		CLF (tributaries of a river)
-ak	<i>thâk</i>	[thâk]	[thûw]	[thûw]	bite
-ag	<i>kwâg</i>	[kwâk]	[kwûw]	[kwûw]	bite
-ek	<i>tapyêk</i>	[ta'pjêk]	[ta'pî]	[ta'pjêj]	sand
-ek	<i>rêk</i>	[rêk]		[rêj]	be afraid
-ik	<i>macîk</i>	[ma'cîk]	[ma'cî]	[ma'cêj]	water
-ik	<i>brík</i>	[brík]	[brûw]		CLF (bamboo-shaped)
-ik	<i>bülík</i>	[bə'lík]	[bə'lûw]	[bə'léj]	pig

The rhyme *-ak* has both a [w] and a [aw] variant. The rhyme *-ek* has an [-ei] variant while only *tapyêk* ‘sand’ has a further [-i] variant. The rhyme *-ik* has a variant [w] (cf. the clusters *bw-* and *pw-* in §2.2.3.1, where the allophone [w] also occurs) and some lexemes have a variant [-ej] and [-i]. Note that rhymes ending in *-[Vw]* are written *<Vw>* due to the similarity between the allophone [w] and the glide [w] (cf. §1.7).

The difference between one or the other rhyme is dialectal and several informants have remarked that *vocal codas* (i.e. *-aw*, *-ei*, etc. instead of *-ab*, *-ek*, etc.) are more widespread in the plains whereas consonant codas are prominent towards the Western Hills, e.g. Dav valley. In the Northern hills, e.g. around the Chagla valley, *vocal codas* are also prevalent. The word lists hailing from China by Sun (1991) (971 records) and Bufan and Qingxia (1992) (1138 records) suggest that *vocal codas* are also prominent there. According to Johakso Manyu (2022, p.c.), the presence of consonant codas in the Western Hills coincides with contact phenomena with Kaman (Miju) where consonant codas, including final */-t/*, are widespread⁵. Conversely, a tendency to further reduce the already limited pool of consonant codas (cf. §2.2.4) seems at least to converge with

⁵Consonant codas in Kaman are attested in two independent word lists: the work of Blench and Kri (n.d.) is based on one or more unspecified Indian dialects while Sun (1991) (962 records) is from China. Further research in Kaman is a desideratum.

Kera'a (Idu) spoken to the West of the Tawrã region. In Kera'a, only nasal consonants are allowed in coda position and open syllables are widespread (cf. Reinöhl 2022).

2.2.4.2 Devoiced codas

As described by Evans and Manyu (2021, 10f), coda consonants are regularly devoiced in coda position. The underlying voicing is, however, detectable before a voiced consonant. Image 2.1 illustrates the waveform and spectrogram of *klüb* 'CLF:sip' as uttered by AJ. The yellow arrow points to a voiceless segment in the spectrogram that corresponds to the coda [-p]. However, the underlying voiced consonant may show signs of voicing when succeeded by a heteromorphemic voiced consonant. This is illustrated in Image 2.2, which depicts the waveform and spectrogram of *klüb gê* 'one sip' as uttered by AJ. Here the yellow arrow points to a voiced segment that corresponds to the coda of the classifier before the following voiced consonant in *gê* 'one'.

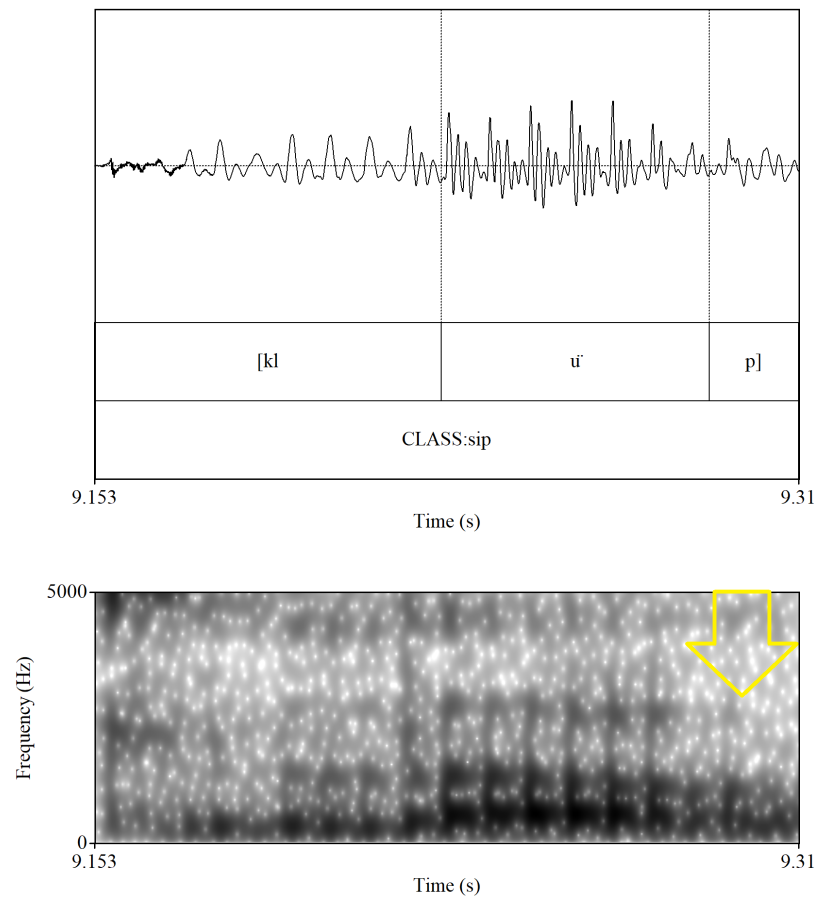


Figure 2.1: Waveform and spectrogram of *klüb* 'CLF:sip' by AJ

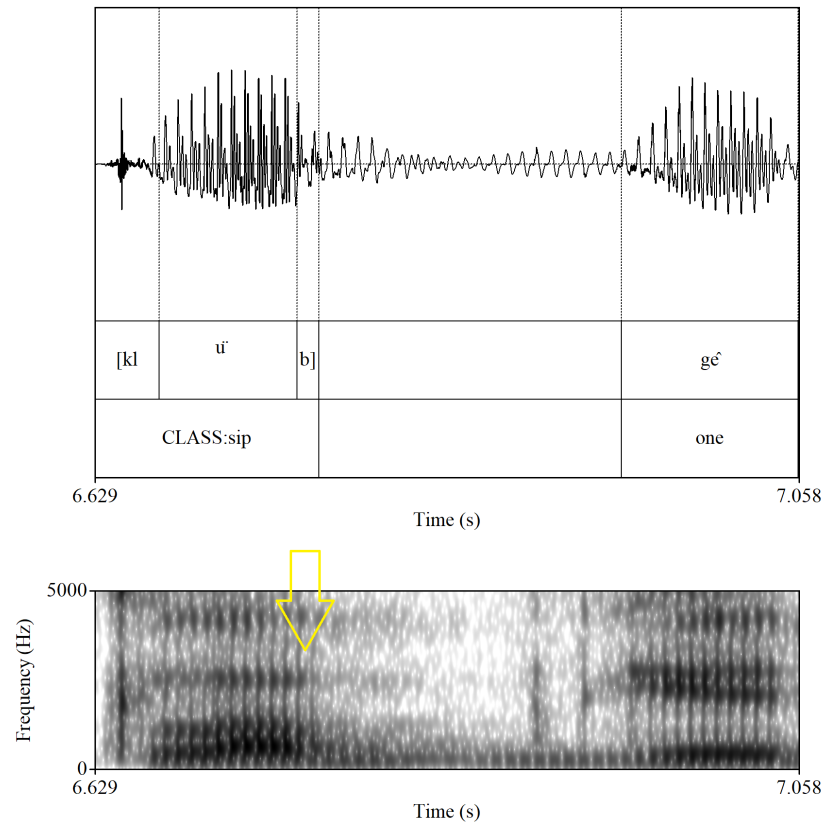


Figure 2.2: Waveform and spectrogram of *klüb gē* ‘one sip’ by AJ

It should be mentioned that voicing of codas is not automatic but lexically determined. Thus, an inherent voiceless consonant coda will keep its voicelessness regardless of the next consonant. This is illustrated in Image 2.3, which depicts the waveform and spectrogram of *thâk gũ* (bite NF) ‘by biting’ as uttered by JO. The yellow line points to a gap in the voicing that corresponds to voiceless [-k].

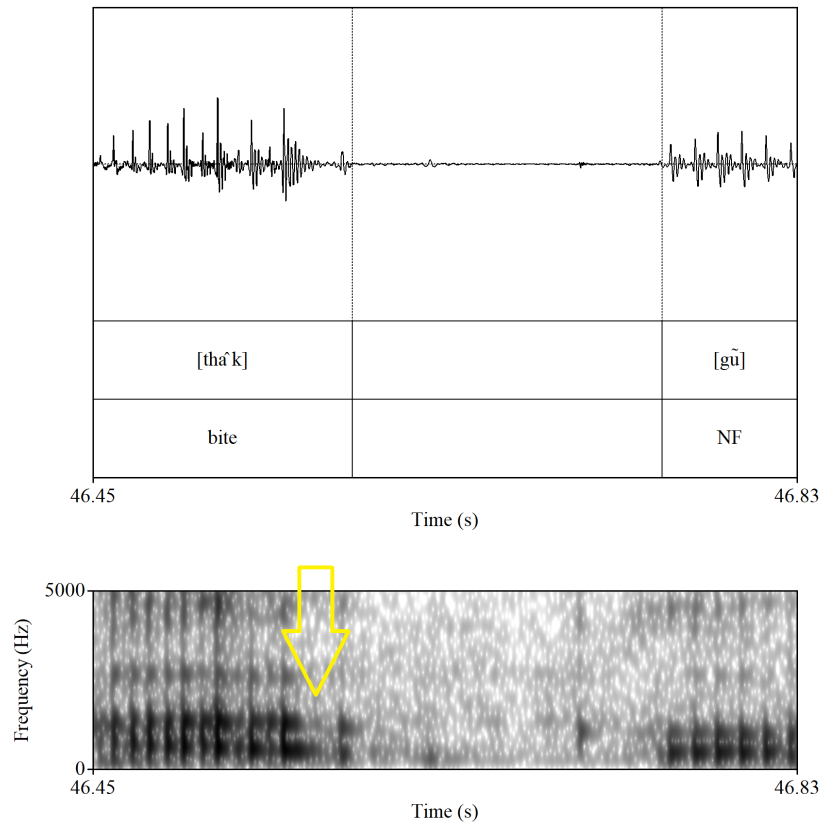


Figure 2.3: Waveform and spectrogram of *thâk gũ* ‘by biting’ by JO

A contrast in voicing in coda position is uncommon in the language family. A historical explanation for this contrast should be further studied.

2.2.5 Light syllables (L)

The term *sesquisyllable* or *unstressed presyllable* is often encountered in the literature to describe a phonological word consisting of an initial light syllable with varying types of phonological erosion followed by a final heavy syllable, or “a syllable and a half” as termed by Matisoff (1973b, 86) while discussing the structure of Proto-Austroasiatic. This particular configuration is also known as iambic foot (Gordon 2016, 201). Although there is little agreement on the details of what a *sesquisyllable* is, there are some general characteristics: a) the light syllable often has a reduced consonant inventory, at least compared to the major syllable, b) it has a reduced vowel inventory to the extent that the light syllable can often only have a schwa or a syllabic consonant and c) it carries neither

stress nor tone (Butler 2015, 443f; Green 2005, 2f; Pittayaporn 2015, 502f; Thurgood 1999, 60f; Van Dam 2019, 94f; Watkins 2002, 32-35). This phonological pattern is considered a characteristic of the languages of Mainland Southeast Asia (cf. Enfield 2005, 186; Sidwell and Rau 2014, 227) and it is also common in languages of NE India that border Myanmar (cf. Morey 2014, 660).

A light syllable in Tawrã contains exactly *one* mora and stands in opposition to a heavy syllable (cf. §2.2.1). The syllable structure of Tawrã makes a distinction between light and heavy syllables, which resonates with the discussion on sesquisyllabicity. The rhythmic structure of disyllabic words has been modeled, in this analysis, as consisting of two main patterns: iambic (light + heavy syllable) and trochaic (heavy + heavy/light syllable) (cf. Hayes 1985). The former pattern is by far more common in the language. The interplay of tone and stress is discussed in §2.2.6.

Image 2.18 illustrates the structure of a light syllable, which can either occur before the heavy syllable (iambic pattern) or after (trochaic pattern, cf. §2.2.6): C stands for an optional onset. Onsets are described in §2.2.3 (this applies to both heavy and light syllables). V stands for the nucleus, which can be any of the six vowels available in the language. The vowel of the light syllable consists of one mora (as discussed in 2.2.1), which again stands in opposition to the bimoraic nucleus of a heavy syllable (VV cf. §2.2.2). In practical terms, this means that the unstressed vowel is shorter than the stressed syllable, as discussed in §2.3.2.1, and it cannot be nasalized (cf. §2.3.2.2). Lastly, several phonological processes only affect light syllables as will be described presently.

Table 2.18: Structure of a light syllable in Tawrã

(C)	V
onset	nucleus

Although the metrical structure of the language will be discussed in detail in §2.2.6, it suffices for now to say that the light syllable is an *unfooted* or extra-metrical element, i.e. it does not carry the main stress of the phonological word (and therefore no tone). Rather, as mentioned above, in polysyllabic words the light syllable only occurs along with a heavy syllable, either preceding it (iambic pattern) or following it (trochaic pattern)⁶.

⁶As was discussed in §2.2.2, the distinction between a trochaic H.L (heavy.light) and trochaic H.H

This is illustrated in 2.19.

Table 2.19: Trochaic and iambic pattern

trochaic	iambic
heavy-light/heavy	light-heavy
μμ-μ/μ	μ-μμ

As in many languages, and particularly the languages of MSEA as discussed above, unstressed syllables in Tawrã are subject to different forms of phonological erosion (cf. Ladefoged and Johnson 2006, 104f, 259; Gordon 2016, 175). For Tawrã, these processes have only been attested in iambic words, i.e. in non-stressed light syllables (cf. §2.2.6). The most relevant forms of erosion are a) vowel centralization, b) loss of onset, c) onset alternation and d) loss of the nucleus of the syllable. In sum, a general erosion of all non-stressed segments. The combination of b) and d) leads to a complete loss of the non-stressed syllable in some cases, as in *(ka)hlâi* ‘guts’. By a) vowel centralization is meant that unstressed vowels /a/ and /ü/ (by far the most common in light syllables) often tend towards a schwa [ə] (non-stressed vowels are treated in §2.3). B) the loss of onsets refers to the loss of syllable initial consonants in non-stressed syllables. This phenomenon has been attested with consonants to /k-, h-, m-, t-/, which are arguably the most common onsets in light syllables in the language. Other onset consonants like /kh-/ or /n-/ are never lost in that position but they seldom occur as light syllables in the lexicon.

Table 2.20: Lost initial consonants in non-stressed syllables

affected C	example	translation
k-	<i>(k)acu</i>	judge
h-	<i>(h)atûm</i>	sky
m-	<i>(m)atyô</i>	probably
t-	<i>(t)adyâw</i>	season

Onset consonants /p-, s-, d-/ are also lost in certain words along with the nucleus of the syllable CV-. For the latter two (/s-, d-/ I have only found one example: *(de)lé* (heavy.heavy) does not bear any repercussion in the phonological system of the language and can thus be neglected.

‘it seems’ and *(si)mthâ* [m.^hâ] ‘what-to-say’. These are very common words in the language, which perhaps explains their tendency to be reduced.

C) Onset alternation refers to changes in word-initial light syllables. These changes are i) *n-t* alternation, cf. *tabù/nabù* ‘maize’ or *tadi/nadi* ‘mocus’ ii) *m-h* alternation (only one attestation) cf. *mamyú/hamyú* and iii) *n-l* alternation, cf. *lümúng/nümúng* ‘tail’⁷. Note that several other alternations are not affected by stress patterns, e.g. voicing and aspiration (both discussed in §2.4.3).

D) Loss of the nucleus of the syllable is a process by which the non-stressed syllable CV- becomes C- and is resyllabified with the following heavy syllable as CV.VC > CVC or CV.CVC > CCVC, thus *kalà* ‘lord, officer’ often surfaces as *klà*. There is a restriction in this development in that it does not yield onset clusters that are not otherwise attested in the language, thus *talyâng* ‘evening’ could not surface as *[tljân] (cf. §2.15)⁸.

The distribution of the most common light syllables in the lexicon hints towards a particular semantic value which seems to go back to a once productive Prefix + Root structure (also cf. Morey 2012, 158f). These *prefixes* are *a-*, *ha-*, *ka-*, *me-*, *ma-* *ta-* and *we-* and they are discussed in §3.7.

2.2.6 Tone and metrical structure: tonal patterns

Polysyllabic phonological words in Tawrà are the locus of interaction between tone and stress; word-level pitch features are determined, or rather, are isomorphic with the abstract metrical structure of stress. These are here condensed in *tonal patterns*. Individual *tonal* patterns will be discussed in detail in §2.5.5 but the overall *stress* system can be defined thus: Stress falls of the first *heavy* syllable from left to right. The weight of a syllable is lexically determined and does not change. Stress on the initial syllable (*trochaic* pattern) will define a syllable as heavy⁹. Trochaic words can only have one tonal pattern: [high-falling + toneless (superficially low)] (HF-L), which will be introduced in

⁷An equivalent process is found in some Western Tani languages.

⁸A similar case has been noted in Turung, although in that language *impossible* clusters do surface (cf. Morey 2012, 158f).

⁹The second syllable may again be heavy (based on the syllable structure, cf. §2.2.2) or light but this is not relevant for the system; once stress has been placed, the syllable weight of the second syllable can be neglected - in practical terms, this means that there are no clues to tell whether an open, second syllable is heavy or light in a trochaic word, as was discussed in §2.2.2.

§2.5.5.1. On the contrary, in an iambic pattern, the first syllable is light and the stress falls on the second, heavy syllable. The tonal patterns of the iambic pattern are [toneless (noted as M) + any of the four tones (i.e. surfacing as M-HF, M-H, M-M and M-L)]. Table 2.21 illustrates this feature based on syllable weight: while trochaic words can have a heavy + heavy (H.H) or heavy + light (H.L) weight, iambic words can only have a light + heavy (L.H) structure.

Table 2.21: Stress structure and syllable weight

pattern	syllable weight	Example	stress and syllable
trochaic	H.H/L	<i>côcang</i> ‘bird spc.’	/ˈcô.cang/
iambic	L.H	<i>hatûm</i> ‘sky’	/ha.ˈtûm/

Now that the basic stress system has been sketched, and before continuing with the details of the patterns, a brief discussion on stress is required. Stress refers to the relative intensity with which a syllable is uttered, although it often involves phonetic elements controlled by other phonetic properties (Hayes 2009, 271; an *organizing framework* for Hayes 1985, 429). Stress is often represented in a metrical structure: *feet*, consisting of a stressed syllable and an optionally non-stressed syllable make up an intermediate structure between the syllable and the word. Feet are controlled by rules and parameters which create a structural hierarchy: strong vs. weak. Strong syllables commonly have one or more of the following properties: they have higher pitch levels, a longer duration and loudness and they tend to attract tone (cf. Kager 2007, 195; Hyman 2009, 215f). As was discussed in §2.2.1, syllable weight is described in *moras*, thus heavy syllables have *two moras* whereas a light syllable only has one. In a polysyllabic word, the foot falls on the heavy syllable¹⁰ while the unstressed, light syllable is *unfooted* (cf. 2.18 on light syllables). Although there are no syllable-internal contrasts, i.e. within a syllable there can be no more (or less) than one stress, languages can be divided by the predictability of stress. Languages like French and Polish have phonological rules that determine where the stress will fall (e.g. final syllable) while in other languages, like Spanish, Russian or

¹⁰Only one syllable can have the highest degree of metrical prominence. This attribute of stress has been termed *culminativity* (cf. Hyman 2009, 217; Hayes 2009, 271). There are, however, often exceptions in grammatical words such as articles or adpositions which do not carry a main stress for they are clitics and depend phonologically on a host-word.

English, the placement of stress is phonemic, i.e. it cannot be predicted. There are often stress-based minimal pairs, such as Spanish [ˈkɾi.ʈi.ko] ‘critic’, [kɾi.ʈi.ko] ‘I criticize’ and [kɾi.ʈi.ko] ‘s/he criticized’ (note the placement of the stress mark <ˈ>). However, even in languages where stress is phonemic there are rules and general tendencies (cf. Hayes 2009, 271f). Tawṛā falls into the second group of languages where stress is unpredictable (*free stress* languages) as evidenced by a number of stressed-based minimal pairs. Table 2.22 displays three segmental minimal pairs based solely on stress (and disregarding tonal differences). The first column from left to right has words with the stress falling on the initial syllable (trochaic foot type) while the second column has words with tone on the second syllable (iambic foot type) (Gordon 2016, 201). As a general tendency, iambic words are much more common than trochaic words in the language.

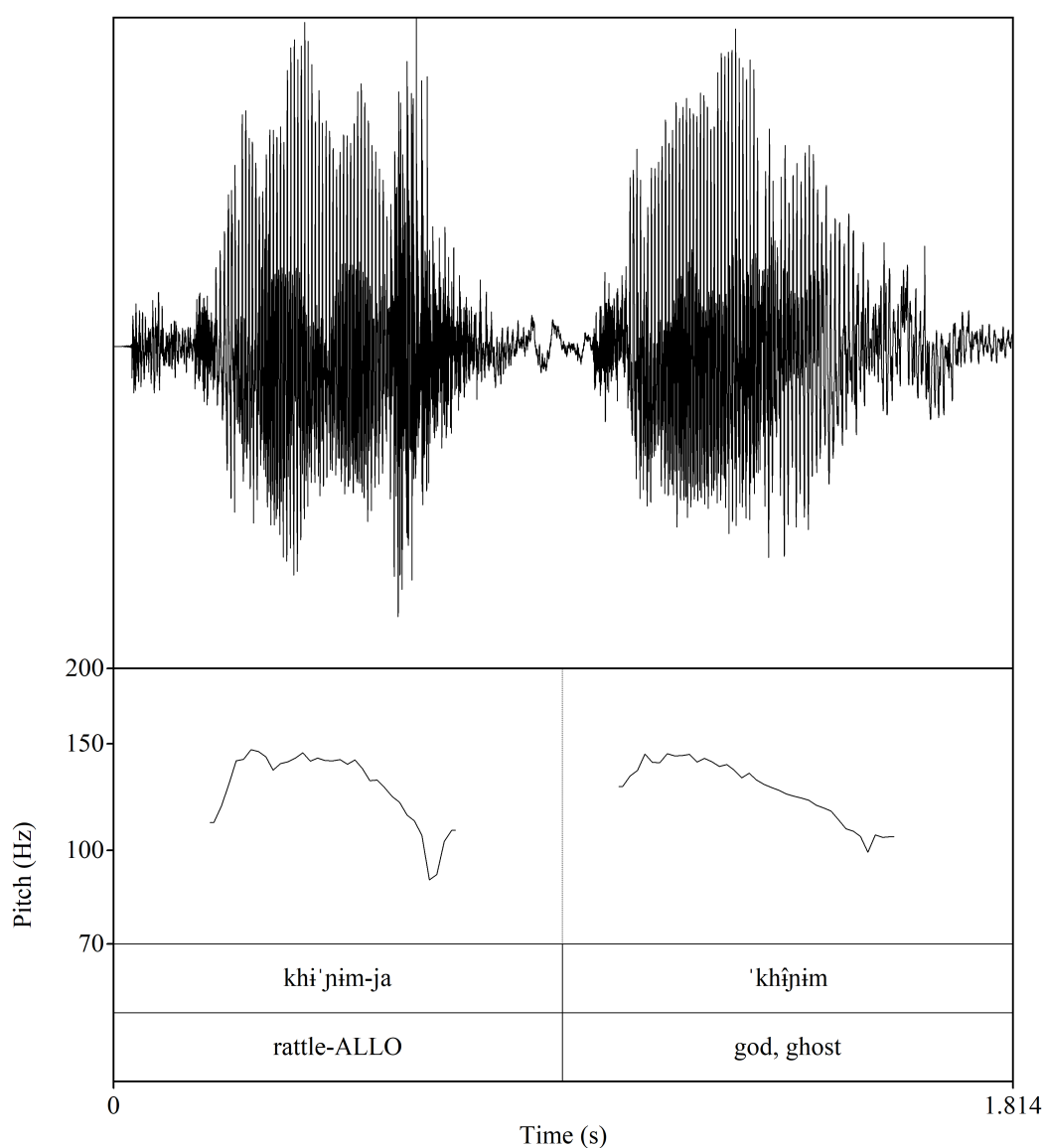
Table 2.22: List of segmental minimal pairs based on stress (disregarding tonal differences)

trochaic foot (stress on initial syllable)	iambic foot (stress on last syllable)
ˈkhû.nyüm ‘supernatural being’	khü.ˈnyüm ‘rattle’
ˈhâ.wa ‘make sound’	ha.ˈwa ‘waste time, delay’
ˈtâ.mẽ ‘boar’	ta.ˈmẽ ‘be crazy’

Stress has a linguistic reality for native speakers as they can assign prominence to one syllable over others in a clapping exercise and thus distinguish iambic from trochaic words. However, the phonological correlation of stress/accentuated syllable in Tawṛā cannot be teased apart from the effect of tone since pitch, intensity and duration are also features of the individual tonemes and their phonotactic effect on surrounding syllables (Hayes 1995, 7f; cf. Hayes 1985, 429; Hyslop 2021, 553), in other words, there is an isomorphism between the tone-bearing unit (henceforth TBU) and stress. It is therefore presently not possible to pinpoint phonetic features or cues that would only apply to stress and not to tone. Image 2.4 illustrates the waveform and pitch of a stress-based minimal pair *khünyüm-ya* (rattle-ALLO) ‘s/he rattles’ [khi.ɲim-ja] and *khûnyüm* [ˈkhî.ɲim] ‘god, ghost, supernatural being’ uttered by BT. The pitch on *khünyüm-ya* [khi.ɲim-ja] ‘s/he rattles’, as we will see in §2.5.5.4, corresponds to the Mid-Mid pattern (plus the descending

pitch on the (toneless) non-past marker *-ya*). The waveform depicts an even amplitude between the two initial syllables, although the second syllable has a slightly higher amplitude. Conversely, the pitch of *khûnyüm* ‘supernatural being’ is also characteristic of the HF-M pattern (tonal patterns are discussed in §2.5.5) with an initial rise followed by a fall that spans across the two syllables. The initial rise in pitch is accompanied by a higher amplitude at the waveform that is reduced progressively. These two features (pitch and intensity) are characteristic of both high-falling tone as well as stress.

Figure 2.4: Waveform and pitch contrast of *khünyüm-ya* ‘s/he rattles’ and *khûnyüm* ‘supernatural being’ by BT



One way to make sense of this interplay between tone and stress at the word level and to account for their lexicalized nature is to arrange all available combinations of stress and tone in word *patterns*. The first clear-cut distinction was illustrated in Image 2.22 between having stress falling on the initial syllable (trochaic) and on the second syllable (iambic). The second step is to list the range of tonal distinctions available in each group. This is illustrated in Table 2.23, which depicts two groups: the left (trochaic) has only one member: HF-M¹¹ (high-falling + mid). Having one possible pattern (HF-M) means that tonal contrasts are neutralized if the stress falls on the first syllable. In practical terms, a trochaic word (i.e. stress on the first syllable) like *khûnyûm* ‘supernatural being’ [k^hî.njîm] never contrasts with *[k^hî.njîm] in a *L-HF pattern (note the different tone on the stressed (initial) syllable). On the other hand, the iambic group has all four tonemes available on the second syllable while the first syllable does not contrast tone (the lack of tonal contrast on the first syllable in the iambic foot is a common pattern, at least regionally, cf. Morey 2014, 660f).

Table 2.23: Tonal patterns divided by stress

trochaic foot	iambic foot
HF-M	M-HF
	M-H
	M-M
	M-L

The tonal range of iambic words can be demonstrated with minimal pairs. Table 2.24 includes two sets of minimal pairs: The first set demonstrates a 3-way contrast between MH, M-HF and ML. The second set demonstrates the contrast between MM and M-HF. A four-way minimal pair has not been found. Conversely, the trochaic group does not show tonal contrast on the initial, unstressed syllable (cf. §2.5.5.1).

¹¹Note that in both groups the letter M stands for mid tone (cf. §2.5.3.2) but in the context of tonal patterns, it should rather be thought of as a toneless syllable (except for the second M in M-M which does stand for the mid toneme (§2.5.5.4). M is employed to mean ‘toneless syllable in a tonal pattern’ due to its superficial similarity with a toneless syllable. Light syllables are thoroughly described in §2.18.

Table 2.24: Minimal pairs among iambic words

example	translation	pattern
<i>hamá</i>	aura	MH
<i>hamâ</i>	get lost	M-HF
<i>hamà</i>	smoky fragrance	ML
<i>hana</i>	come	MM
<i>hanâ</i>	squeeze	M-HF

The following section will relate these patterns to the phonological word.

2.2.7 The phonological word

The phonetic word in Tawrã is defined in metrical parlance as having maximally and minimally one heavy syllable or *one foot*. Notice, however, that in metrical calculation one plus one does not always equal two: two light (monomoraic) syllables do not make a heavy syllable. Rather, a phonological word in Tawrã requires a heavy syllable and not merely two moras. Image 2.25 depicts three impossible words. The first one is a monomoraic, non-stressed **a*, the second a monomoraic segment **ba* and the third a combination of monomoraic, non-stressed segments **a.ba*.

Table 2.25: Three impossible words

Syllables	σ	σ	σ	σ
Moras	μ	μ	μ	μ
Fictitious example	a	b a	a b a	
Syllable shape	(V)	(C V)	(V) (C V)	

Conversely, Image 2.26 depicts four different models of heavy syllables and one combination of light plus heavy syllables. The first example is bimoraic *a* ‘child’, a heavy syllable and as such suffices to be considered a phonological word. Note that the underlying,

moraic structure is *aa*, as opposed to the monomoraic example **a* discussed above. As mentioned above, it is crucial to note that vowel length is not phonemic i.e. monomoraic **a* and bimoraic *aa* ‘child’ never contrast (cf. §2.3.2.1). Rather, their differences in length and vowel quality are epiphenomena of stress and related phonological processes that are discussed in detail in §2.18. Their underlying moraic depiction as **a* vs. *aa* is to model crucial and complex differences in the syllable structure of the language. The second example *bà* ‘father’ has a CVV structure (again with a full vowel). The third example has a rhyme which, in metric terms, has the same weight as a long or full vowel. The fourth example has a consonantal coda, which again, in metrical terms, has the same weight as a long or full vowel or a rhyme (cf. McCarthy and Prince 1986, 19f). Finally, the fifth example has two syllables, an initial light (monomoraic) syllable and a final heavy (bimoraic) syllable. While the light syllable has no tone, the final, heavy syllable carries a HF tone. As mentioned in §2.3.2.1, only the heavy syllable carries tone and stress. The interaction between stress and tone within a phonological word is discussed in §2.2.6.

Table 2.26: Five phonological words

Syllables	σ		σ		σ			σ			σ	σ			
	/		/		/		/		/			/			
Moras	μ	μ	μ	μ	μ	μ	μ	μ	μ	μ	μ	μ	μ		
underlying structure	a	a	b	a	a	br	e	i	bl	ü	m	a	l	o	o
Syllable shape	(V	V)	(C	V	V)	(CC	V	V)	(CC	V	C)	(V)	(C	V	V)
Example	<i>a</i> ‘child’		<i>bà</i> ‘father’		<i>bréi</i> ‘buy’			<i>blúm</i> ‘place’			<i>alô</i> ‘garlic’				

It is relevant to note that stress can have a demarcative function, as it is usually attracted to syllables located at the edge of phonological words (Kager 2007, 196). In Tawrã, the tonal patterns discussed in §2.2.6 and listed in §2.5.5 are diagnostic of both the presence and shape of a phonological word. Any deviation therefrom is indicative of two phonological words. An example is the lexeme *bá-mě* ‘clan’. There is no tonal

pattern high + low (*HL) and the formative elements can be etymologized as *bà* ‘father’¹² + *mề* ‘person’, thus yielding ‘father’s people’ (a *true compound*, cf. §4.2).

A further case worth mentioning here is that of the classifier phrase. It is typologically common for classifiers to have a certain degree of phonological fusion with the numeral¹³. In Tawrã this is not the case as classifiers and numerals are distinct phonological words even though there is a morphological process in which a truncated form of the numerals one to three show up only with classifiers (more details on this feature in §3.5.3). Superficially, however, the classifier-numeral collocation may resemble a disyllabic phonological word if the tonemes happen to match a tonal pattern, as in the case of *tyàng gê* ‘one body’: a low tone followed by a high-falling tone. This is illustrated in Figure 2.5, which depicts this precise example as uttered by JO. Here the pitch goes slightly down (because of the low tone) before the characteristic curve of HF (cf. §2.5.3.1). It superficially resembles a M-HF pattern but these are two phonological words with independent pitch. Furthermore, a classifier with a different toneme would produce a different pattern. This is illustrated in Figure 2.6, which depicts the classifier phrase *tyâng gê* ‘one plot of land’ as uttered by JO. Crucially, the classifier (the first element) has a high falling (HF) tone, thus the figure depicts two almost symmetrical HF curves, i.e. two distinct phonological words, as there is no high falling + high falling (*HF-HF) tonal pattern.

¹²The change in tone from L to H in the possessive construction has been described in §2.5.3.3.

¹³Cf. Nepali (Indo-Aryan) *ek* ‘one’ but *euṭa* ‘one (of something)’.

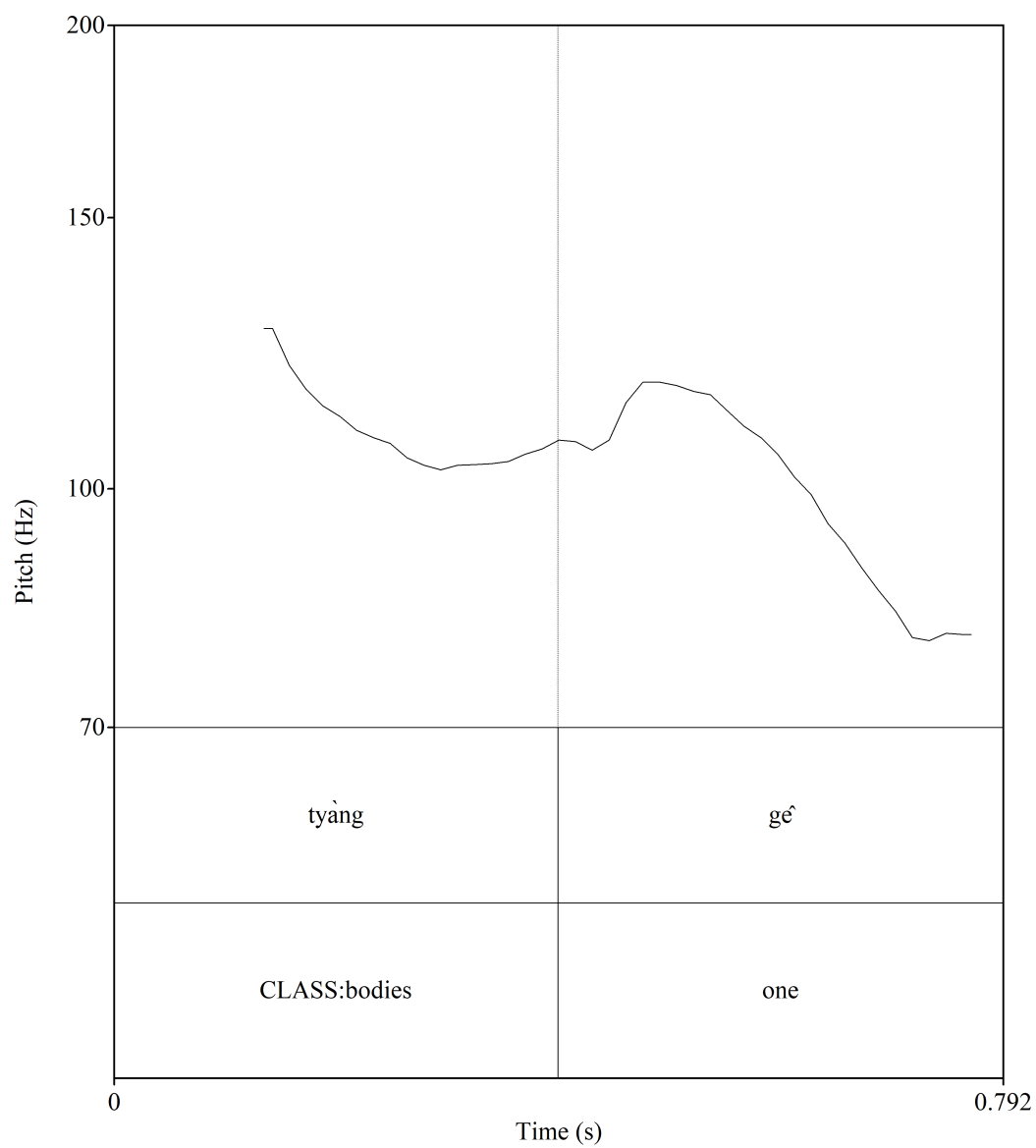


Figure 2.5: Pitch of *tyàng gē* ‘one body’ by JO

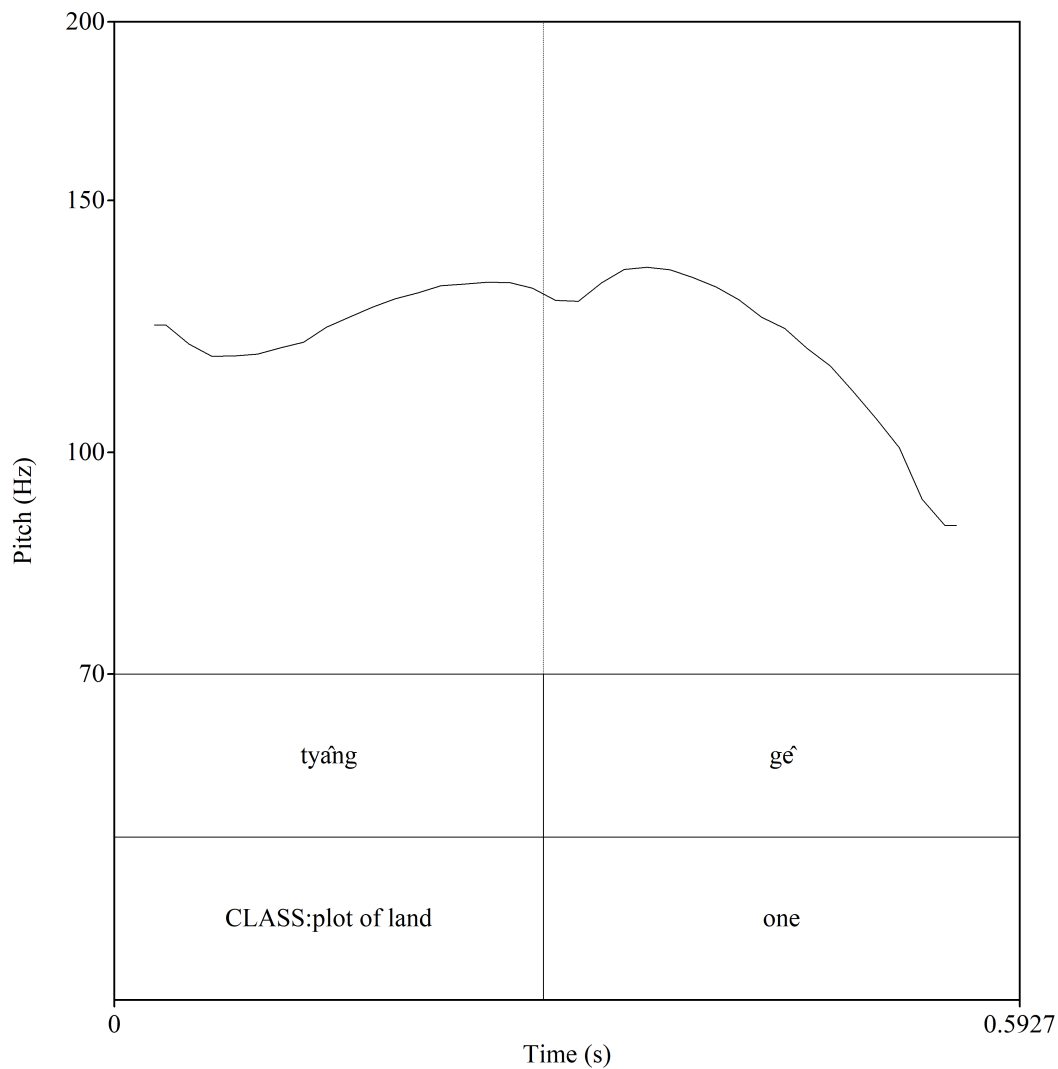


Figure 2.6: Pitch of *tyâng gê* ‘one plot of land’ by JO

Lastly, there are a number of lexemes, including many personal names which are composed of two phonological words even though the formants that compose them are not always known. As an example, *jâp-malu* ‘a god’ is probably composed of *jâp* ‘a tribe, Kaman’ and an unknown element *malu*. Furthermore, trochaic words of the type H.H (heavy.heavy) are composed of two phonological words although they only have one grammatical meaning (cf. §2.2.2).

2.2.7.1 Phonological dependency or clitization

§2.2.7 has described the characteristics of the phonological word. Some morphemes exhibit a defective phonological configuration in certain grammatical environments. This behavior is known as *clitization* (cf. Aikhenvald 2003, 47f; Gerlach and Grijzenhout 2000). The concept of clitization does not necessarily align with the grammatical concept of clitic (cf. Bickel and Nichols 2007, 174f). Clitization is here understood as a phonological behavior and not the inherent property of certain morphological units. A defective phonological configuration in Tawrã means that a morpheme may lose its tonal/stress properties and/or fuse with adjacent phonological units (cf. Zwicky 1985, 287f). One example is the negative marker *yüm*, which has a cliticized form = (*ü*)*m* (cf. §6.4.1). The forms *yüm* and = (*ü*)*m* seem to be in free variation, as they are not triggered by morphophonological factors. A further example is the focus particle *nyu* ‘FOC’ (§8.3.2). When used with interrogative pronouns, it may fuse to the pronoun and constitute one phonological word, in the sense that it has one main stress and one tone, cf. *kadényu* ‘how’ (< *kadé* = *nyu*; §3.5.1.2). It should be noted that determining the loss of tonal properties is not unproblematic, since there is a fine distinction between a toneless syllable and a syllable with a mid tone (cf. §2.5.3.2); in careful speech, a mid tone prompts a slightly longer vowel. However, this distinction can be easily overridden in fast speech. Therefore, a minimal pair between *nyu* ‘FOC’ (toneless syllable) and *nyu* ‘2s’ (mid tone) cannot be established, as these forms are indistinguishable in isolation.¹⁴

2.3 Vowels and rhymes

This section describes the vowels in Tawrã. Non-phonemic vowel length is also discussed.

¹⁴The isomorphism between the focus marker and the 2nd person pronoun is further discussed in §8.3.2. The same situation is found in Milang, where *ni* is both a focus marker and a personal pronoun for 2nd person singular (cf. Modi 2017, 211, 272f.).

2.3.1 Previous work

Previous works on Tawrã generally recognized six phonemic vowels /a~ɑ, e, i, o, u, i/ (Evans and Manyu 2021, 12; Jiang et al. 2013, 29; Sastry 1984, 9). An exception is Blench (2015), who proposed 10 phonemic vowels, including /ɛ, ə, ʌ, ɔ/. The author does not provide minimal pairs supporting his claim. Evans and Manyu (2021, 12) note that the vowel /i/ varies between [u] and a featureless [ə], especially when it is not stressed. Vowel length is described by Evans and Manyu (2021, 14) and Blench (2015) along with minimal pairs. Evans and Manyu (2021, 15) and Sastry (1984, 9) recognize three nasal vowels, one for each level of backness: /ẽ, ã, ü/. Evans and Manyu (2021, 15) include some examples. Blench (2015) recognizes six phonemic nasal vowels, adding /ẽ, ã, õ/. The author presents some examples. The work of Evans and Manyu (2021, 17) is the only work to mention the existence of diphthongs/rhymes. They analyze V-V combinations as glide-vowel and vowel-glide, since a) all potential V-V combinations include a close vowel and b) final stops and nasals never occur tautomorphemically after final glides.

The present work, which is based on an independent collection of data and analysis, has reached similar conclusions, e.g. regarding the amount of vowels, but differs in other aspects: I present four nasal vowels and do not recognize phonemic vowel length. Regarding rhymes/diphthongs, I recognize four rhymes: *ai* [ai], *ei* [ei], *aw* [aw] and *üw* [iɯ]. The remainder of this chapter will look carefully at these issues.

2.3.2 Monophthongs

Vowels in Tawrã differentiate between stressed and non-stressed. This distinction is informed by a larger dichotomy between heavy and light syllables (discussed in detail in §2.9). Tawrã has thus six phonemic vowels: /a, e, i, o, u, u~i~u/ <a e i o u ü> that occur in heavy syllables. This is demonstrated by the segmental minimal pairs based on vowel quality illustrated in Table 2.27 (cf. Evans and Manyu 2021).

Table 2.27: Segmental minimal pairs based on vowel quality (heavy syllables)

vowel	IPA	example	translation
a	[a]	<i>la</i>	say
e	[e]	<i>lê</i>	rub, roll
i	[i]	<i>lî</i>	escape, rescue
o	[o]	<i>lô</i>	pierce
u	[u]	<i>lu</i>	poke
ü	[i~u~o]	<i>lû</i>	save, spare, survive

As discussed by Evans and Manyu (2021, 12), the vowel /a/ is much more common than all other vowels. To illustrate this, I analyzed a list of 241 words pertaining to core vocabulary recorded in April 2022. The results are displayed in Table 2.28. The tokens are divided between heavy and light syllables: words have minimally one heavy syllable (cf. §2.2.7) regardless of the total number of syllables. Heavy and light syllables are defined in §2.2.1. This table does not include values from trisyllabic words and non-accentuated heavy syllables (cf. §2.2.2) for the sake of simplicity. Nasal values are included in parentheses and summed to the total. Note that rhymes cannot occur on light syllables. The most important point is that /a/ is the nucleus of the heavy syllable of almost 40% of the lexemes, followed at a distance by /u/. In the case of light syllables, the pool is much smaller, as it only includes disyllabic words. The percentage occurrence of the vowel /a/ in a light syllable is very high (more than 80%). Although this is but a small sample, it helps convey a rough estimate of the distribution of vowels in the lexicon and the high frequency of /a/ therein.

Table 2.28: Distribution of vowels in heavy and light syllables (sample)

	heavy syllable	%		light syllable	%
a	87 (+ 4ã)	38%	a	97	82%
u	32 (+ 5ũ)	17%	ü	4	3%
ü	22	10%	i	3	3%
o	18 (+ 1õ)	8%	u	3	3%
i	18	7%	e	2	2%
e	11 (+ 6ẽ)	7	o	1	1%
ei	14	6%		Total: 110	
ai	5	2%			
üw	1	0.4%			
	Total: 224				

Central vowel /ü/ displays an important amount of speaker-internal variation between the parameters of [i~u~o] (also discussed by Evans and Manyu 2021). On top of that, there is also a dialectal variation. Images 2.7 and 2.8 depict the (stressed) vowel space of two male speakers aged approximately 40 years old. While both speakers reside in Sunpura (AP), speaker ST has always lived there while speaker BT came originally from Chagla and preserves features from that dialectal zone. In both images, the letters correspond to the mean F1/F2 of 16 attestations (making 96 data points per plot). Ellipses represent confidence intervals set at 60%. Individual data points have been omitted for the sake of clarity. The data points have been extracted using *Praat* (2024), while the plots have been created using *R* (2015).

Perhaps the most striking difference is the distribution of /ü/ and /u/; while for speaker ST there is little overlap between these phonemes, they overlap to a great extent for speaker BT. In the case of the relation between /ü/ and /o/, the overlap is pronounced for both speakers, although /o/ shows little variation for both speakers. Furthermore, it is remarkable that for BT the central vowel /ü/ has a close-mid to close position between /i/ and /u/, whereas for ST the central vowel /ü/ seems close to /o/, almost resembling a close-mid back unrounded vowel [ɤ]. This pronunciation is consistent with my

impressions from other speakers from Chagla.

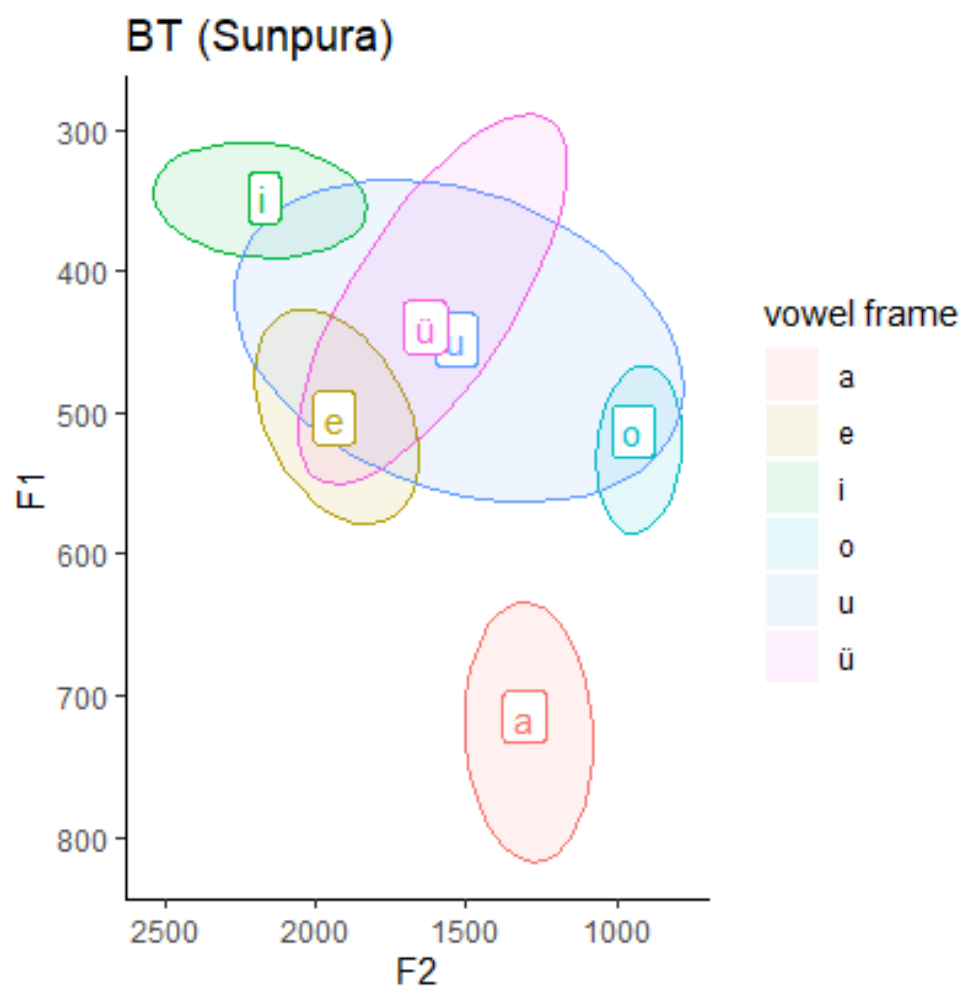


Figure 2.7: Mean and standard error of F1 and F2 (vowel space) by BT

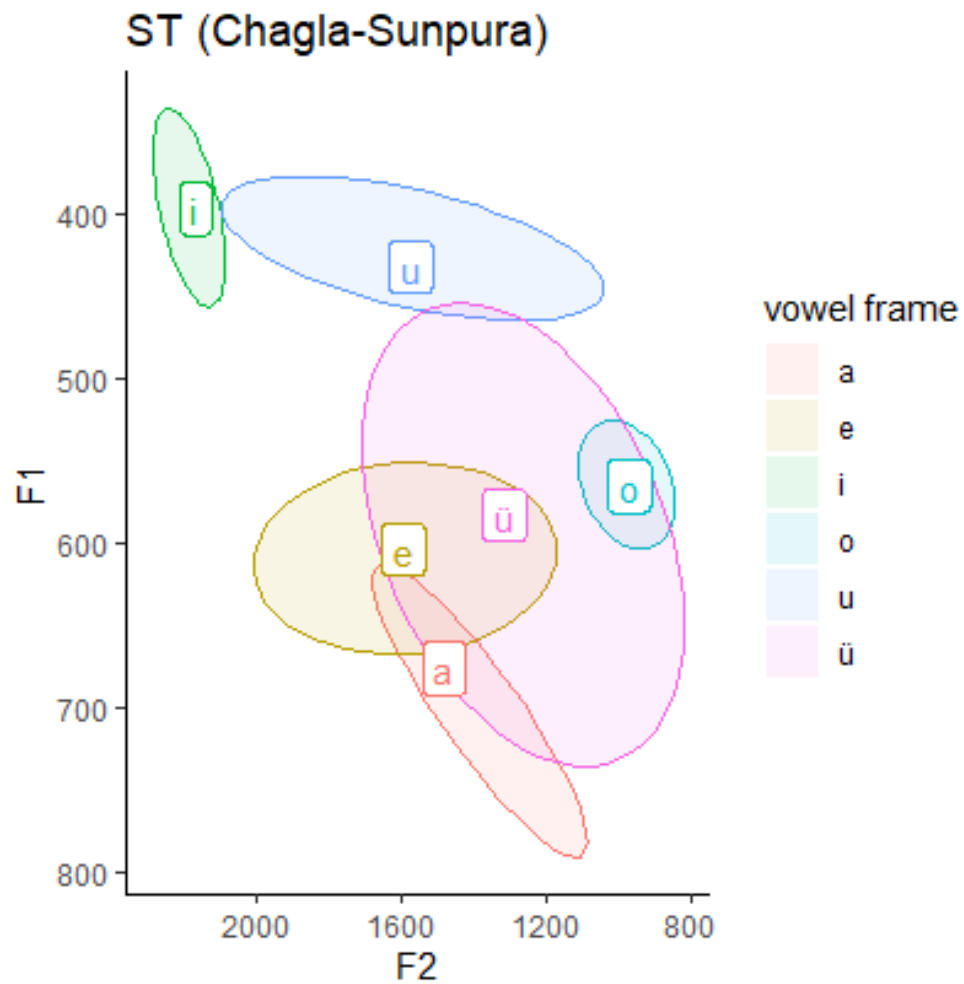


Figure 2.8: Mean and standard error of F1 and F2 (vowel space) by ST

Non-stressed vowels in Tawrã do not form minimal pairs with stressed vowels. Although their precise phonetic value in terms of F1/F2 requires more study, they do impressionistically resemble full vowels. A list of non-stressed vowels is displayed in Table 2.29.

Table 2.29: Examples of non-stressed vowels

non-stressed vowel	example	translation
a	ha.'bá	say, hearsay particle
e	we.'sâ	forget
i	tsi.'kâ	tore open
o	pho.'ko	stomach
u	pu.'tyû	stump
ü	züm.'pâ	together

Non-stressed vowels /a/ and /ü/ are often centralized as compared with their stressed counterparts, thus often surfacing as [ə], e.g. as in *karô* ‘fast’ [kə.'ɪô].

There are cases of vocalic alternation on stressed syllables that involve the vowel /ü/ cf. *lûmkû*, *lâmkû* ‘inside’ or *myûm*, *myûm* ‘elder brother’. This could be due to dialect mixing but more research is required.

Cases of vocalic alternation in non-stressed syllables are likely part of a general phonological erosion along with centralization and vowel loss (cf. §2.18). Some examples are *hayâ*, *hiyâ* ‘cook’ and *powan*, *pawün*, *pûwün* ‘silver’ (the change in stress/tone is unexpected).

2.3.2.1 Vowel length

Vowel length was described by Evans and Manyu (2021, 14f) and they provide two minimal pairs. Vowel length is further described as being related to referential marking, although the exact changes in tone are not further explored. Lastly, the authors indicate that the negative marker *yüm* ‘NEG’ triggers vowel length on the preceding verb.

For the present study, vowel length is not considered phonemic but rather an epiphenomenon of syllable weight, as discussed in §2.2.1¹⁵. For Hayes (1995, 7), languages

¹⁵In a practical sense, the bimoraic requirement of a heavy syllable is similar to phonemic vocalic length. As a case of historical development, a full/reduced vowel distinction has been described for Proto-Athabaskan verbal stems, in which full vowels occurred in CV and CVC stems whereas reduced vowels only occur in CVC stems, i.e. a bimoraic syllable. This system developed differently in the daughter languages but crucially Ahtna (Athabaskan, Canada) has been described as having a phonemic distinction

with phonemic vowel length contrasts disprefer the use of duration as a correlate of stress¹⁶. Conversely, in a system where syllable weight is relevant, heavy syllables are often described as having a full or long vowel as opposed to light syllables (Hayes 2009, 279f; Hyman 2007, 18), along with a set of properties such as tone. It is precisely the intertwined nature of these phonological features that makes it problematic to isolate vowel length. An attempt to measure vowel length could look for the correlates of tone and length on heavy syllables, as studied by Morey (2020) in Tai Phake, but it would not tell us anything about unstressed (often toneless) vowels. Conversely, measuring the vowel length of a stressed vs. non-stressed vowel is also problematic since it would be, as mentioned above, partly influenced by the phonetic nature of the different tones. Comparing the vowel length of heavy and light syllables does, however, shed light on this imbalance.

A further point to consider is the difference between the iambic and trochaic feet since both patterns are attested in the language (cf. §2.2.6). As discussed by Hayes (1985, 430f), the difference between the iambic and trochaic groups is not only a matter of intensity (e.g. stress) but also of duration: “prominence contrasts based on duration lend themselves to iambic grouping, while prominence contrasts based on intensity lend themselves to trochaic grouping” (idem, 430). What this means for the present investigation is that we might expect a different result when measuring light syllables in trochaic and iambic words.

As a small demonstration, I measured the vowel length of 10 attestations of the iambic verb *wesâ* ‘forget’ in an elicitation recording using Praat (Boersma and Weenink 2024). The session was on verbal morphology and JO pronounced the verb alone, with a verbal suffix *wesâ-dyibô* (forget-PFT.EGO) ‘I’ve forgotten’ and in the phrase ‘some say *wesâ-dyibo*’. The varied realizations of this verb should not affect the results since I wanted to measure it in context. Image 2.9 illustrates an example of how the vowels were measured: the vowels /e/ and /â/ were selected based on their waveform and pattern in the spec-

in vowel length, although maintaining the requirement for stem minimality and the distinction between full/reduced vowels in verbal stems that mark, e.g. mode and aspect (cf. Rice and Hargus 2005, 6; Kari 1990, 12f).

¹⁶Note that there are well-described counter-examples, as in the case of Tetsôt’iné (Dëne, Canada). As described by Jaker and Howson (cf. 2022) duration is the main correlate of stress on top of a system that distinguishes vocalic length and tone.

trogram. The results of the 10 attestations are displayed in Table 2.30. In all cases, the stressed vowel is longer than the non-stressed vowel. The average values reveal that this difference in time is almost twice as long. A larger study that includes trochaic words would provide a comprehensive view of the difference in vowel length.

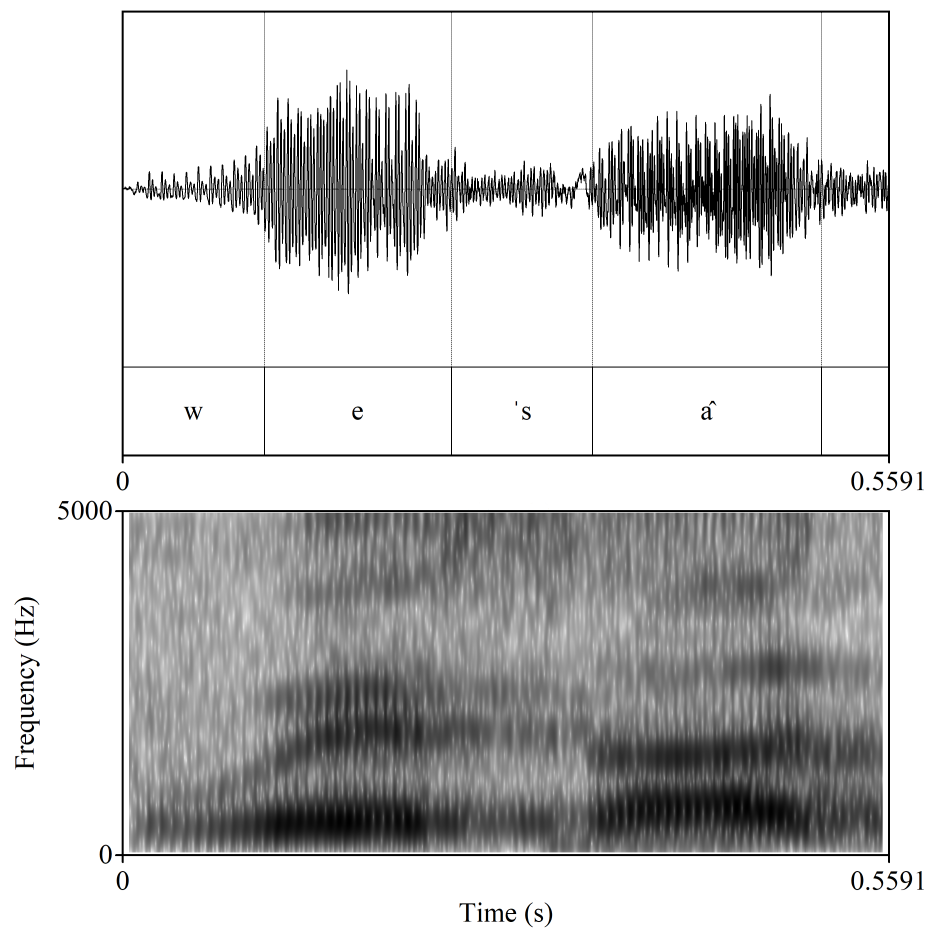


Figure 2.9: Spectrogram and waveform of *wesâ* [we'sâ] 'forget' by JO

Table 2.30: Vowel length of 10 attestations of *wesâ* ‘forget’ by JO and the average length

nr. of attestation	length V1 (non-stressed)	length of V2 (stressed)
1	0.1	0.26
2	0.1	0.16
3	0.06	0.09
4	0.07	0.08
5	0.1	0.14
6	0.08	0.1
7	0.02	0.06
8	0.05	0.1
9	0.1	0.16
10	0.05	0.1
Average:	0.073	0.125

2.3.2.2 Nasalization of monophthongs

Nasalization is a widespread suprasegmental feature of vowels in the language. Evans and Manyu (2021, 15) reported three nasal vowels /ẽ ã û/, which is in line with the reduced set of vowels that occur before a nasal final (§2.3.2.3). The present work further includes the nasal vowel /õ/. Vowels /ü/ and /i/ are never nasalized. Table 2.31 displays a set of structural minimal pairs for nasalized vowels.

Table 2.31: Structural minimal pairs for nasalized vowels

Nasalized vowel	Example	Non-nasalized vowel	Example
ã	yã grandmother	a	ya night
ẽ	zè defeat	e	zè be exhausted
õ	zõ expand	o	zô bud
û	zû cold (felt on body)	u	zû write

Nasalization is contrastive. There are no active phonological processes in which nasalization would spread from a nasal coda or onset to the vowel. Nasalized vowels may not occur before a tautomorphic bilabial stop, and there are no words like **tãb* in my data. Across morpheme boundaries, a nasalized vowel undergoes nasal unpacking and inserts a homorganic nasal consonant in running speech, cf. *gũ bô* (NF TEMP.SUB) occurring as [gum.'bô].

Nasalized vowels are important to syllable weight and can define a heavy syllable (§2.2.2). This claim is supported by the fact that there are no nasal vowels on light syllables in iambic words (i.e. L-H, cf. §2.2), thus a word like */tã.'sê/ would not be possible. Iambic words provide the best testing ground for syllable weight as the initial, non-stressed syllable is surely light. For trochaic words, where the second, non-stressed syllable can be heavy, i.e. trochaic words can have a H-H structure, like *hûcê* /'hû.cê/ ‘type of bamboo’ or *tâplã* /'tâ.plã/ ‘hammer’. Furthermore, the nucleus of a heavy syllable of CVC type can not be nasalized (*CVC̃), for it would constitute a 3-mora heavy syllable¹⁷.

We may assume that nasalization originated from an eroded nasal stop which retained its moraic value, as this is a common development cross-linguistically (cf. Campbell 1998, 41). There is one hint that this process occurred in the language, cf. the dialectal variation of *tâ* ‘fish (generic)’ as *tangâ* in the Chagla variety (among others). We may assume that the form *tâ* is more innovative as the final nasal syllable was lost and the nasalization (as well as tone) was transferred to the initial syllable (*tangâ* > *tâ*). A similar case is the plural suffix *-ring/rê* (cf. 6.3.2.0.1, note the change in vowel quality), although in this case, both forms occur in free variation.

2.3.2.3 Vowels and nasal codas

Evans and Manyu (2021, 15) mention that not all vowels occur before a nasal coda. Table 2.32 displays the combinations of vowels and nasal codas described by the authors (as V-N). Vowels /i/ and /o/ would not occur before nasal coda /-m/ and vowels /e/ and /u/ would not occur before nasal coda /-ng/.

¹⁷This is in line with the phonological description of other TB languages of NE India, e.g. Apatani (cf. Post and Kanno 2013, 36) and Bulu Puroik (Lieberherr 2017, 38).

Table 2.32: Vowels and nasal codas (adapted from Evans & Manyu 2021, 15). ‘X’ stands for no occurrence.

	-m	-ng
a	pam ^H ‘wasp’	pang ^H ya ^{HL} ‘finish’
e	pem ‘anyway’	X
i	X	taping ‘residue of grain’
o	X	pong ‘bomb’
u	tamyum ^{HL} ~ tamyüm ^{HL} ‘monkey’	X
ü	tapüm ^H ~ tapum ^H ‘insect’	kapüng ^{HL} ‘bat’

Based on an independent survey of the phonology of the language, I have come to similar results to those of Evans and Manyu (2021) regarding vowels and nasal codas. There are some small differences regarding possible V-N combinations: regarding nasal coda /-m/, I came across the combination O-M. For nasal coda /-ng/ I agree with the lack of /e/ and /u/ but I also included a lack of /o/. Table 2.33 illustrates all V-N combinations found in my survey.

Table 2.33: Vowels and nasal codas. ‘X’ stands for no occurrence.

	-m	-ng
a	<i>phalâm</i> ‘tea’	<i>ruzàng</i> ‘in between’
e	<i>kadyêsyem</i> ‘hardly’	X
i	X	<i>hîng</i> ‘day’
o	<i>bôm</i> ‘edge’	X
u	<i>hatûm</i> ‘sky’	X
ü	<i>katsûm</i> ‘chest, breast’	<i>kashûng</i> ‘front yard’

As described by Evans and Manyu (2021, 15f), a nuclear vowel /e/ is lowered before a nasal coda /-m/ to [ɛ~æ], cf. *chûweng* ‘louse’ [t^hî.wæŋ]. This process also affects nuclear vowel /a/, e.g. before a cliticized negative marker =*m* (§2.2.7.1): *thâ=m kya?* (eat=NEG Q) ‘haven’t you eaten?’ [t^hâm kja].

2.3.3 Rhymes

The analysis of rhymes is not without complications, as combinations of vowels often have complex and diverse historical origins that complicate their classification (Gordon 2016, 108). In their description of the phonology of Tawrã, Evans and Manyu (2021, 17) analyze V-V combinations as glide-vowel or vowel-glide based on two arguments: all V-V combinations have at least one close vowel (e.g. *ei* or *üw*). Furthermore, rhymes cannot occur tautomorphemically before final consonantal stops (i.e. there are no words like **peik*). Their analysis is also consistent with their economic approach to consonant combinations, which treats sounds like *wr* [ɹw] or *hl* [l̥]~[ɬ] as clusters rather than independent phonemes (cf. §2.4.1 and §2.2.3.1).

The present work recognizes four falling rhyme/diphthongs *ei* [ei], *ai* [ai], *aw* [aw] and *üw* [iɯ]. Table 2.34 displays an overview of these four rhymes with examples.

Table 2.34: Four rhymes

type of rhyme	Example
-ai	<i>râi</i> ‘be scared’
-ei	<i>sêi</i> ‘beg’
-aw	<i>khaw</i> ‘smoke’
-üw	<i>prûw</i> ‘set ablaze, burn’

There are four arguments for considering these V-V collocations as phonemic: a) they are common in the language (minimal pairs are provided below), b) they can occur with all four tones. C) Morpho-phonologically, rhymes/diphthongs have a weight of two moras (like a *long vowel* (VV), i.e. bimoraic). Therefore, they are only found on heavy syllables and cannot co-occur with a consonantal coda (cf. §2.2.1). Finally, d) they are distinct from the alternative codas discussed in §2.17 that match the rhymes *ei* and *ai*. In the case of *üw*, note that the rhyme is pronounced as [iɯ] while the rhyme alternation has a coda version [ɯɯ]. It cannot be ruled out that all rhymes in the modern language go back to an eroded coda. However, there is a difference: while the rhymes discussed in §2.17 are *secondary* diphthongs, i.e the product of a morpho-phonological process,

the rhymes discussed in this section are lexically specified, i.e. they do not have a coda counterpart; there is no alternation **arâk* for *arâi* ‘snow’. Further historical work can try and explain the origin of rhymes/diphthongs in the language and their relation to secondary diphthongs /ai/, /ei/ /üw/ and /aw/. Table 2.35 displays minimal pairs for the rhymes /ai/ and /ei/. Table 2.36 displays minimal pairs of the rhyme and /aw/ and /üw/.

Table 2.35: Minimal pairs of rhymes /ai/ and /ei/

Phoneme	Tawrã	Translation
/ai/	<i>lâi</i>	be half grown
/ei/	<i>lêi</i>	slither
/e/	<i>lê</i>	rub, roll
/i/	<i>lî</i>	escape, survive
/o/	<i>lô</i>	pierce
/u/	<i>lû</i>	sweat, pour

Table 2.36: Minimal pairs of rhymes /aw/ and /üw/

Phoneme	Tawrã	Translation
/aw/	<i>yâw</i>	mother-in-law
/üw/	<i>yüw</i>	CLF: side of face
/a/	<i>ya</i>	habit, addiction
/o/	<i>yo</i>	cross (water)
/u/	<i>yu</i>	beer (particularly rice beer)

It should be noted that /w/ has an allophone [ɰ] in clusters with a bilabial plosive as C₁. Therefore *bwî* ‘dance’ is pronounced [bɰî] and *pwî* ‘other, different’ [pɰî] (Evans and Manyu 2021, 5). This is considered a consonant cluster in this grammar. Clusters have been discussed in §2.2.3.1.

2.4 Consonants

2.4.1 Previous work

The description of the phonology of Tawrã proposed by Evans and Manyu (2021, 4f) has twenty-six consonant phonemes, all of which occur as onsets. The authors emphasize the defective distribution of the glottal stop; although it occurs as the first element in three clusters (*qy*-, *qr*-, *qw*-) it does not contrast as a simple onset. However, it occurs intervocalically and in coda position as a marker of continuative aspect (*-q*). The authors also describe the allophone realization of the rhymes *-Vk*, *-Vg* as [-au], e.g. /kwag/ [kwaɯ] *kwaü* ‘dog’. They also described consonants and syllable structure but this has been discussed in §2.2.3 and §2.2.4.

After my own independent survey and analysis, I agree with the results proposed by Evans and Manyu (2021) in most aspects of their description of consonants. I also agree with their distinction between consonants and consonant clusters and present my arguments in §2.2.3.1.

2.4.2 Overview

Tawrã has 25 consonant phonemes and two marginal phonemes (fricative [v] and glottal stop, cf. §2.4.4) in five places and five manners of articulation. Stops and affricates (excluding the glottal stop) have a series of voiceless unaspirated, voiceless aspirated and voiced unaspirated. These are displayed in Table 2.37, in which the practical orthography (discussed in §1.7) is followed by the IPA symbol in square brackets.

Table 2.37: Consonants in Tawrä. The sounds in parentheses are marginal phonemes.

manner\place	Labial		Alveolar		Palatal		Velar		Glottal
	-v	+v	-v	+v	-v	+v	-v	+v	
Stop -asp.	p [p]	b [b]	t [t]	d [d]			k [k]	g [g]	(q [ʔ])
+ asp.	ph [p ^h]		th [t ^h]				kh [k ^h]		
Affricate -asp.			ts [ts]	z [dz]	c [tɕ]	j [dʒ]			
+ asp.			tsh [ts ^h]		ch [tɕ ^h]				
Nasal		m [m]		n [n]				ng [ŋ]	
Fricatives			s [s]		sh [ɕ]				
Lateral			l [l]						
Approx.		w [w], (v [v])	r [ɹ]		y [j]				h [ɦ]

The remainder of this section will describe all relevant places of articulation individually as well as other critical aspects of the consonants. For each sub-section there will be two tables to help determine their phonemic value: one with minimal pairs and a further table depicting its allowed position in the syllable structure (cf. §2.2), i.e. whether it can occur in onset, intervocalic and/or coda position. If a phoneme has not been attested in a particular syllable position, this will be indicated by an ‘X’. For these lists, only native, monomorphemic lexemes were used. It will be noted when relevant assimilation processes spread over morpheme boundaries.

Other aspects of syllable structure relevant to consonants are described in §2.2.3.

2.4.3 Oral stops (excluding glottal stops)

Oral stops occur at four places of articulation: labial, alveolar, velar and glottal (cf. Evans and Manyu 2021, 4). The glottal stop will be discussed in §2.4.4. The remaining stops (labial, alveolar and palatal) have three manners: voiceless unaspirated, voiceless aspirated and voiced unaspirated. The place of articulation of plosives (labial, alveolar, velar and glottal) is reminiscent of related Tibeto-Burman languages from the Hi-

malayas¹⁸ such as Kera'a (Idu) and Kaman (Miju) as well as Assamese¹⁹.

The contrasting qualities of voiceless and voiced stops can be analyzed by looking at the waveform in terms of Voice Onset Time (VOT), which is the measurable interval between the release of a consonant closure and the start of the voicing. The VOT has a negative value if the voicing begins before the release of the stop closure, as is normally the case with voiced stops (cf. Ladefoged and Johnson 2006, 159f). The three-way distinction in mode (non-aspirated voiced, non-aspirated voiceless and aspirated voiceless) in Tawrã is illustrated for dental plosives in Image 2.10. On the first /d/ (in the word *dû* 'jump') the voicing spans from the closure to the vowel, yielding a negative VOT. For the second alveolar /t/ (here in the example of *tã* 'type of bird') there is no sign of voicing before the closure until the vowel starts, i.e. 17 ms after the closure. The last phoneme /th/ is also voiceless but it has a larger VOT of 68 ms due to the opening of the glottis that produces the aspiration (here illustrated as a slightly darker black line before the bigger oscillations of the vowel).

¹⁸Retroflex and alveolar stops are seldom contrastive in the macro Tibeto-Burman family (cf. Matisoff 2003, 21f) and their advent in certain Bodic languages of Nepal is attributed to language contact with Nepali (cf. Noonan 2003, 70; LaPolla 2001, 234), although the retroflex contrast seems very entrenched in some languages, e.g. Western Tamang (Regmi and Regmi 2018, 8,10,124). Some Kiranti languages in Nepal have a retroflex series instead of the alveolar stop, e.g. Chulung or Chintang.

¹⁹Note that the collapse of the retroflex-(lamino-dental) alveolar contrast in Assamese is ascribable to long-standing language contact phenomena (cf. Goswami and Tamuli 2004, 431), particularly with Tibeto-Burman languages.

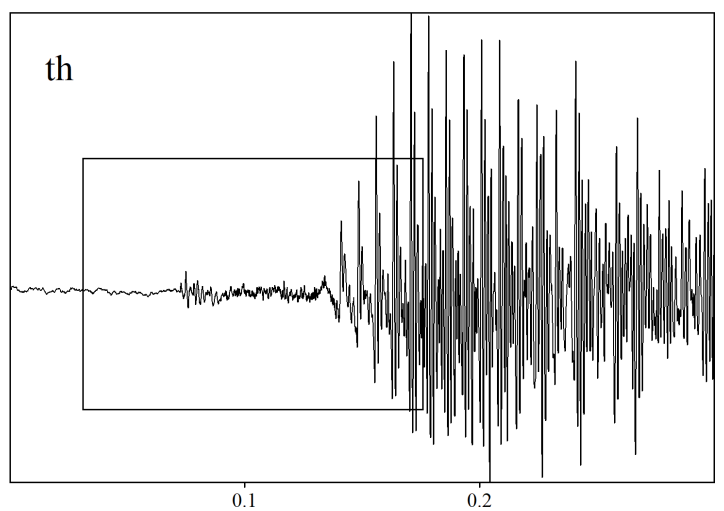
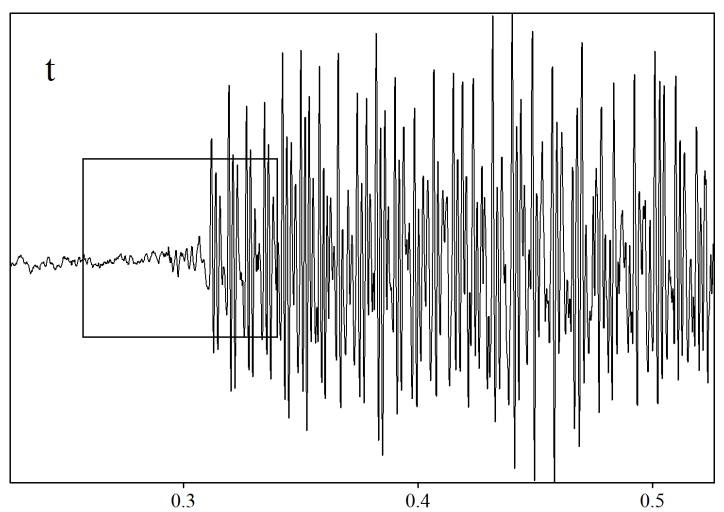
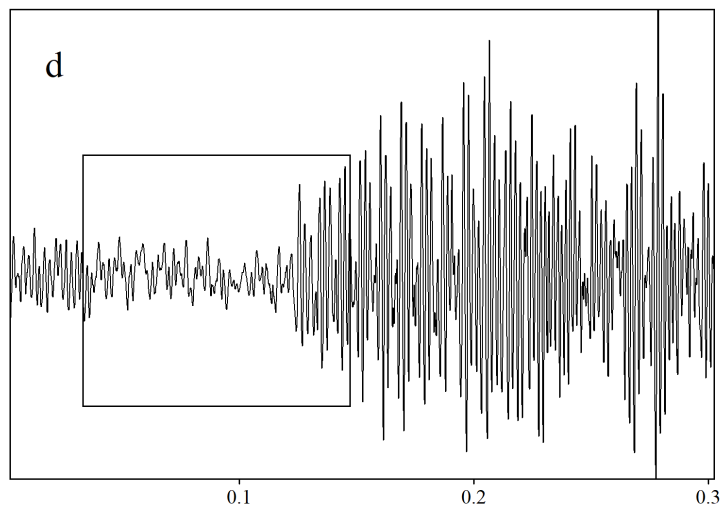


Figure 2.10: VOT of /d, t, th/ by JO

The values for VOT also vary depending on the place of articulation and its distance to the glottis. The further away the higher the values (Ladefoged and Johnson 2006, 164). Thus in Tawrã, as illustrated in Image 2.11, the VOT of /k/ (illustrated by the word *kà* ‘gharial’) is 11 ms while that of /kh/ in the word *khünyüm* ‘rattle’ is 42. Finally, Image 2.12 illustrates the labial group /b, p, ph/ in the words *bà* ‘father’, *pà* ‘axe’ and *phâbum* ‘knee’. In the case of *pà* ‘axe’ the VOT is 10 ms (actually less than that of /k/ and /t/; this requires more research) and for *phâbum* ‘knee’ is 84 ms.

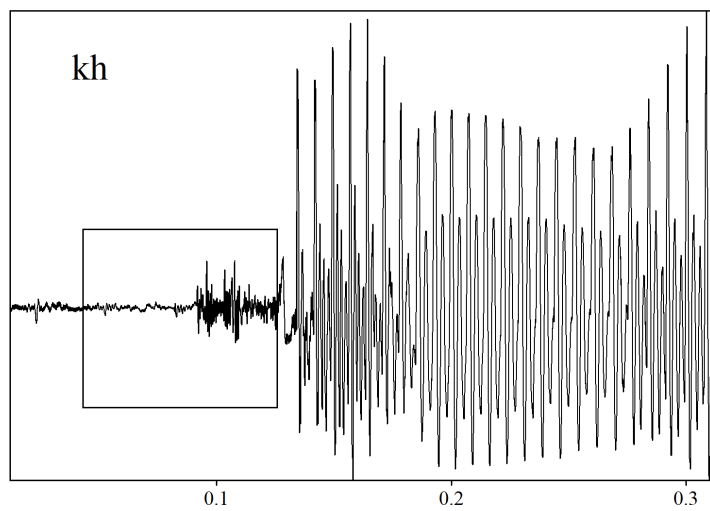
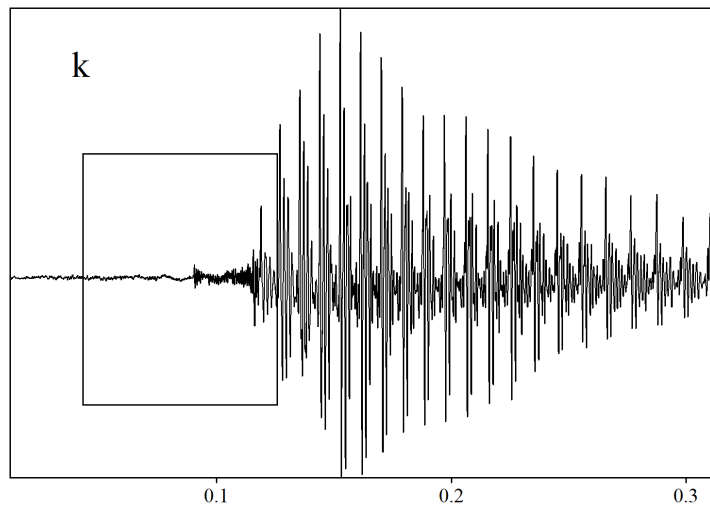
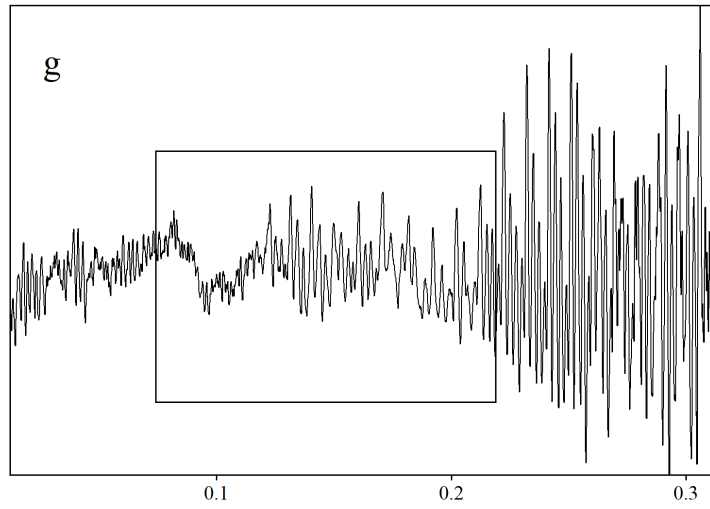


Figure 2.11: VOT of /g, k, kh/ by JO

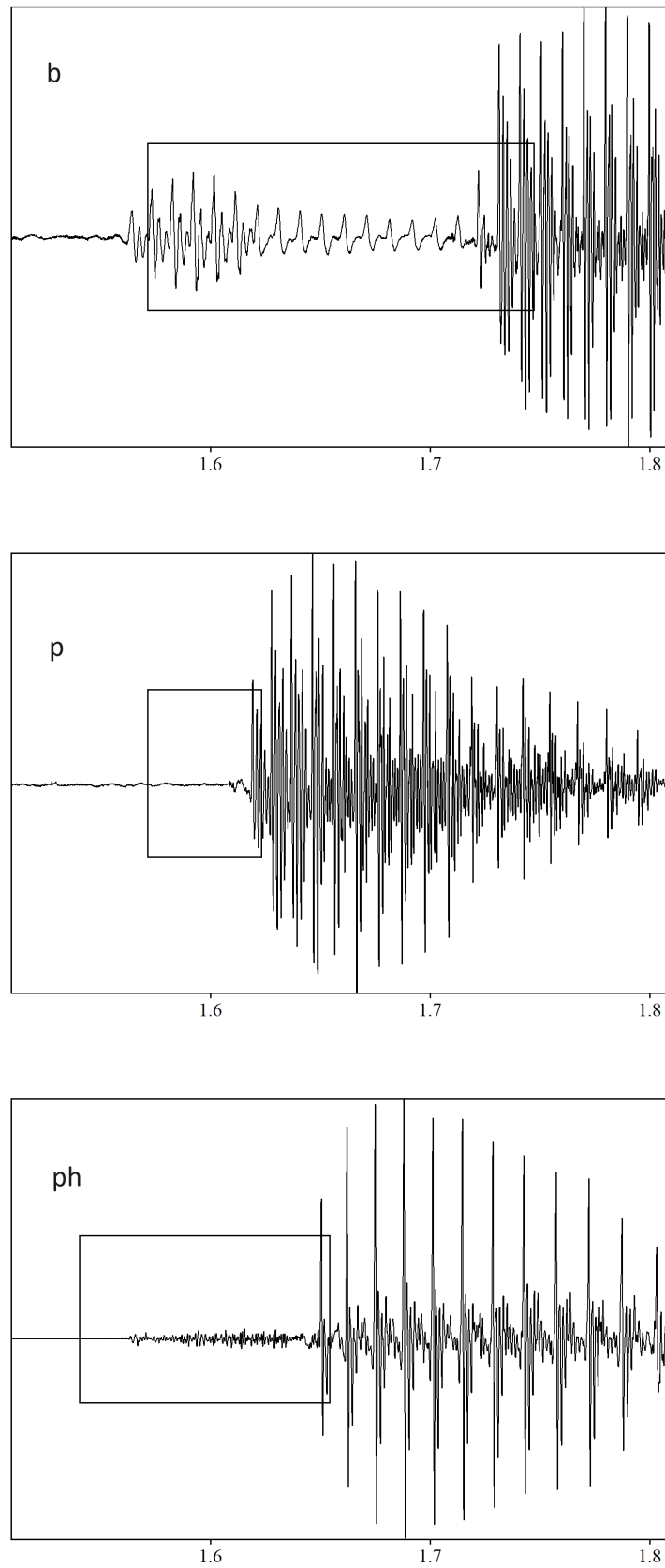


Figure 2.12: VOT of /b, p, ph/ by JO

Plosives are devoiced in coda position, as described by Evans and Manyu (2021, 10f). The underlying voicing is only detectable before a voiced consonant. This will be described in §2.2.4 on codas.

Certain voiceless aspirated stops in word-initial position are in free alternation with non-aspirated counterparts. This phenomenon is lexicalized (reduced to specific lexemes), non-productive and I haven't found that it can be linked to a specific dialect of the language. The most common initial plosive to display this alternation is velar *kh-/k-*: *khaláng/kaláng* 'door', *khû/kû* 'feud, revenge' or *khumpha/kumpha* 'bribe'. For labial *ph/p* there are fewer examples: *phrü/prü* 'yet', *phûdu/pôdu*²⁰ 'type of eggplant'. Voiceless alveolar plosive /t/ does not partake in this alternation. The word list of Tawrã from China by Sun (1991) includes some words that differ in terms of aspiration from my data. As an example, cf. *khɑ³¹dun⁵⁵* 'star' (in the current orthography *kadün*) while I have an unaspirated *kading* 'star'; the aspirated version is rejected by my informants. This difference in aspiration suggests dialectal variation.

The voiceless alveolar plosive /t/ has a voiced alternation in word-initial position in a reduced number of lexemes²¹. Some examples are: *tajaw/dajaw* 'Tezu, a city', *tawó/dawó* 'hoe' or *takhrêi/dakhrêi* 'weaver'²².

Concerning manner in particular, aspiration seems to be an innovation in the language family, as it is not reconstructed for the proto-language (cf. Matisoff 2003, 15). Its uneven distribution among sub-groups suggests independent innovations. In the Tawrã area, aspiration is found in its sister-language Kera'a (Idu) and the more distantly related Kaman (Miju). The neighboring Macro-Tani language group generally does not contrast aspiration and the reconstruction proposed by Sun (cf. 1993, 55f) does not contemplate aspiration to be contrastive²³.

²⁰The change in vowel quality is unexpected; it could be the effect of language contact, cf. Kaman *phundu* 'type of eggplant'.

²¹This feature could partly explain the origin of the name of the river *Digarū* (*tatyû* in Tawrã), perhaps a dialectal alternation of *Tawrã*. Note that the name of the river *Digarū* has been extended as an exonym for the Tawrã tribe.

²²The word *kaqîng* 'two' can be voiced but only when counting, i.e. *ekhing gâqîng* 'one, two'. This is, however, likely a different phenomenon in which a particular prosody associated with counting affects only this word (also note the change in stress/tone *kaqîng* > *gâqîng* 'two').

²³Neseng (Western-Tani, NE India) has developed a secondary aspiration only found on the velar plosive /k-/, especially in /k^hj-/ and low tone (Post, 2024, p.c.).

2.4.3.1 Labial stops

The phonological status of the three members of the labial stop group is demonstrated in Table 2.38.

Table 2.38: Minimal pairs of labial stops

Phoneme	IPA	Example	Translation
/p/	[p]	<i>pa</i>	yeast
/b/	[b]	<i>ba</i>	carry (child)
/ph/	[p ^h]	<i>pha</i>	CLF(bunch)

The aspirated labial stop /ph/ alternates freely with a labio-dental fricative [f] and bilabial fricative [ɸ]. This feature is widespread among Indo-Aryan and Tibeto-Burman languages of upper Assam.

Table 2.39 shows the distribution of the labial stops in the syllable structure. As a rule, aspirated consonants cannot occur in coda position (cf. §2.2.4). They are also rare inter-vocally, although consonant clusters with a /ph/ do occur between vowels, cf. *kaphrûlyu* ‘butterfly’ (§2.15). ‘X’ stands for no occurrence.

Table 2.39: Position in syllable structure of labial stops

phoneme	Onset	Intervocalic	Coda
p	<i>pâ</i> ‘go across’	<i>kapa</i> ‘fool, dumb’	<i>tshâp</i> ‘punch’
b	<i>bô</i> ‘go’	<i>habang</i> ‘outside’	<i>rab</i> ‘stretch (vt)’
ph	<i>phák</i> ‘shake (vi)’	X	X

2.4.3.2 Alveolar stops

A list of segmental minimal pairs for alveolar stops is displayed in Table 2.40.

Table 2.40: Minimal pairs of alveolar stops

Phoneme	IPA	Example	Translation
/t/	[t]	<i>tài</i>	hot
/d/	[d]	<i>dài</i>	Idu, a tribe
/th/	[t ^h]	<i>thài</i>	poison

There is a lexicalized phonological free alternation between the alveolar stop /t/ and the nasal /n/ that affects light syllables *ta-/na-*, as in *tabũ/nabũ* ‘maize’ or *tamüŋ/namüŋ* ‘fire’. Other phonological alternations that affect light syllables are mentioned in §2.2.5.

The distribution of alveolar stops in the syllables is shown in Table 2.41. While all three alveolar stops can occur as syllable onsets, they cannot occur in coda position. Intervocalic position is restricted for the voiced labial stop. ‘X’ stands for no occurrence.

Table 2.41: Labial stops as onset, coda and intervocalic position

phoneme	Onset	Intervocalic	Coda
t	<i>tã</i> ‘type of bird’	<i>atĩ</i> ‘leave, be lost’	X
d	<i>dablang</i> ‘flood’	<i>kading</i> ‘star’	X
th	<i>thâ</i> ‘eat’	<i>athê</i> ‘hail’	X

Indo-Aryan loanwords with an aspirated (or rather breathy voice) voiced alveolar [d^h] stop are consistently reproduced as non-aspirated voiced alveolar [d], such as *dharma* ‘religion’, which is rendered as *daram/durum* in Tawrã. This is perhaps partly due to the fact that the local variety of Hindi (the local lingua-franca cf. §1.1.4) employed in the Tawrã region has merged the aspirated voiced alveolar with its non-aspirated counterpart. The occurrence of *aspirated* voiced alveolar [d^h] stops in other Tibeto-Burman languages has been attributed to the long-standing influence of Indo-Aryan (Matisoff 2003, 15; for a concrete case in Bantawa, cf. Doornenbal 2009, 18).

2.4.3.3 Velar stops

Table 2.42 displays segmental minimal pairs for the three members of the velar group /k, g, kh/.

Table 2.42: Minimal pairs of velar stops

Phoneme	IPA	Example	Translation
/k/	[k]	<i>kâ</i>	be bitter
/g/	[g]	<i>gá</i>	crack open (vi)
/kh/	[k ^h]	<i>kha</i>	lie down

Their distribution in the syllable structure is displayed in Table 2.43. Similar to the labial stops, the aspirated velar stop [kh] cannot occur in coda position. ‘X’ stands for no occurrence.

Table 2.43: Velar stops as onset, coda and intervocalic position

phoneme	Onset	Intervocalic	Coda
k	<i>kû</i> ‘plateau’	<i>takang</i> ‘takin, gnu goat’	<i>nâshik</i> ‘appetite’
g	<i>gû</i> ‘ashes’	<i>agã</i> ‘set up (a trap)’	<i>píg</i> ‘squeeze’
kh	<i>khú</i> ‘be enough’	<i>akhà</i> ‘type of fishing trap’	X

2.4.4 Glottal stop

The phonological status of the glottal stop [ʔ] <q> in Tawrã is inconspicuous for it displays a defective distribution pattern. The glottal stop does not form minimal pairs, i.e. there are no instances where the glottal stop would contrast with another phoneme. However, the glottal stop is relevant for other areas of phonology and verbal morphology. This section will review the arguments for considering the glottal stop a marginal phoneme.

The glottal stop occurs predictably covering non-onset syllables, e.g. *a* ‘child’ [ʔa] (cf. Evans and Manyu 2021, 4; Jiang et al. 2013, 27). This is illustrated in Figure 2.13,

which depicts the waveform and spectrogram of the noun *ám* as uttered by JO. Note a slight oscillation at the beginning of the word that corresponds in the spectrogram to a slightly shaded area followed by a white area (where the yellow arrow is pointing). This stands for a closure in the vocal folds, characteristic of the glottal stop (cf. Hayes 2009, 3). However, the non-onset insertion of the glottal stop is often dropped in less careful and fast speech. This is illustrated in Figure 2.14 which represents the same noun for the same speaker but without the glottal closure: note the absence of both the initial oscillation at the waveform and the shaded area followed by a white space before the pattern associated with the vowel (the yellow arrow).

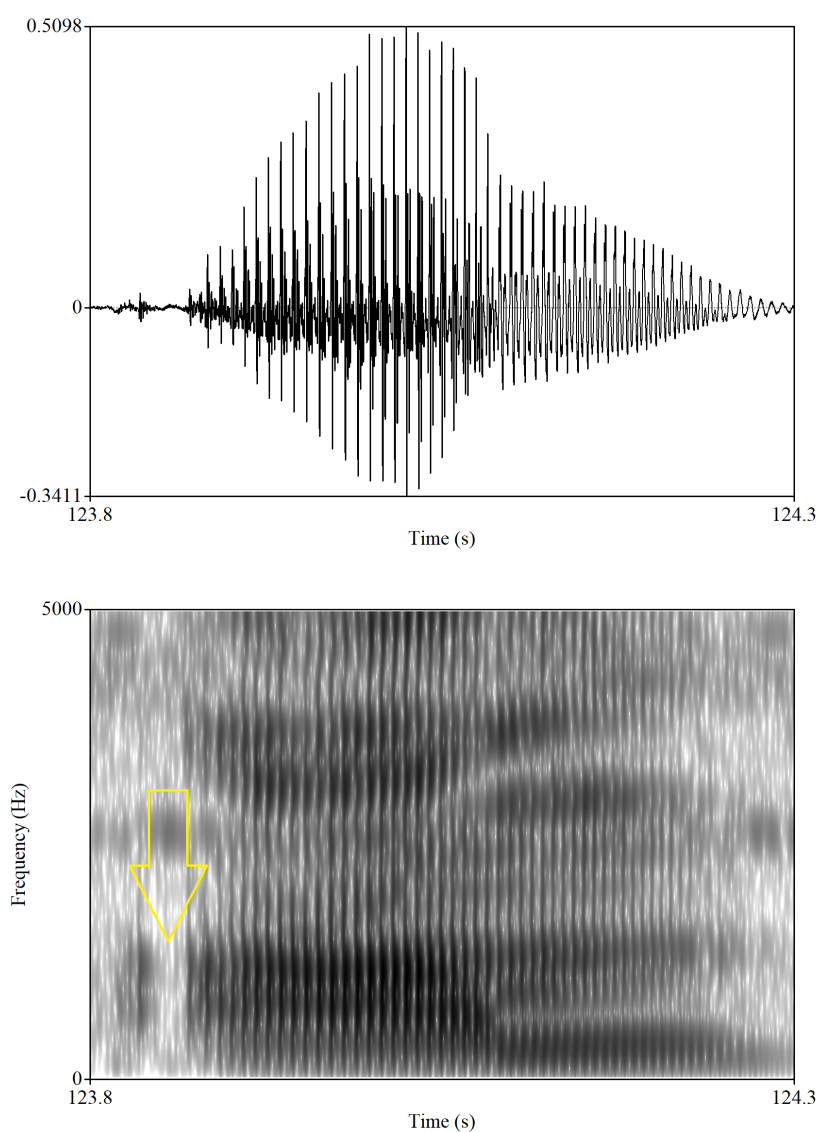


Figure 2.13: Waveform and spectrogram of *ám* by JO

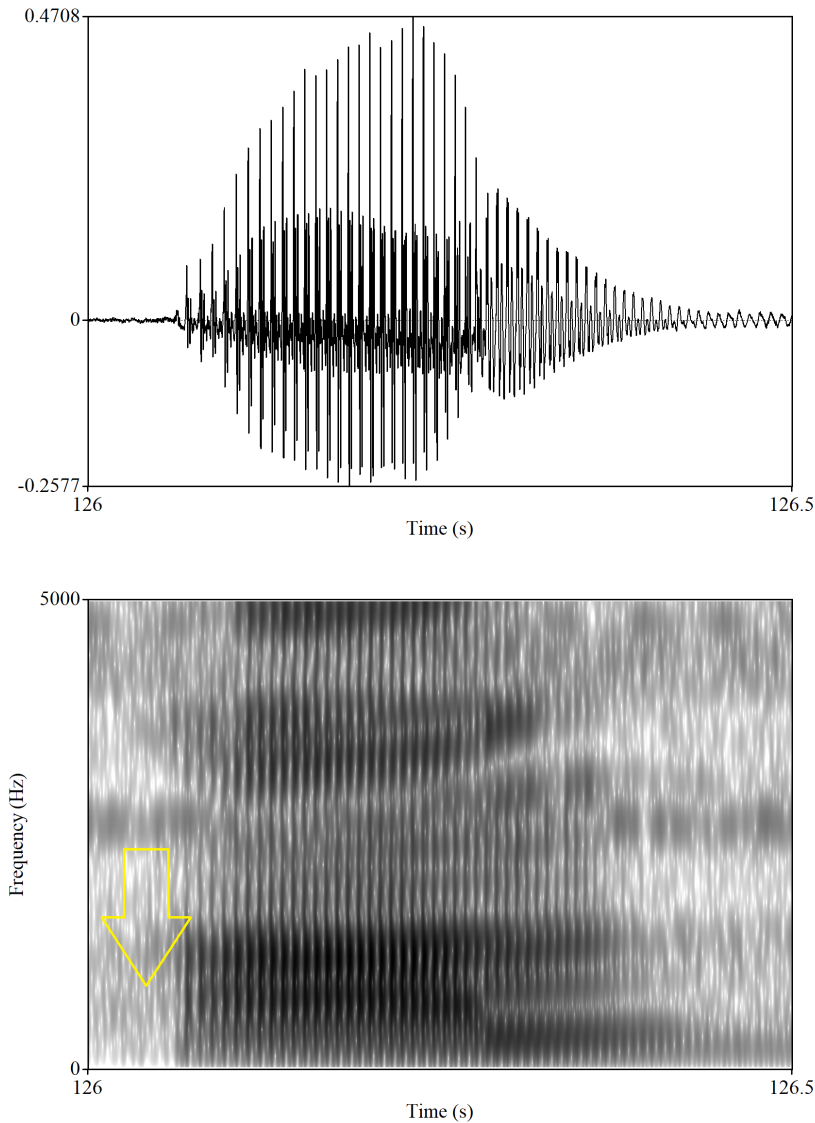


Figure 2.14: Second realization of waveform and spectrogram of *ám* by JO

Syllable initial glottal stops also surface in tautomorphemic intervocalic position, e.g. in the numeral *kaqîng* ‘two’ (although it is often dropped, thus [ka.ʔîŋ]).

The glottal stop further occurs as the first element in CC onset clusters *qy*, *qw* and *qr* (cf. §2.15) (cf. Evans and Manyu 2021, 4, 9). A further point to consider is that the glottal stop does not occur in tautomorphemic coda position (cf. Evans and Manyu 2021, 11), which is in line with the general restrictiveness of the language regarding consonants in that position (cf. §2.2.4). Heteromorphic, however, the continuative aspect is expressed by a glottal stop suffix *-q* (§6.3.4.2), as in *thâ-q* (eat-CONT) ‘s/he is eating’. Crucially,

the glottal stop does not seem to provide a mora, i.e. it does not affect syllable weight (cf. §2.2.1), since it occurs heteromorphically on bimoraic syllables, like *hám-q* (rest-CONT) ‘s/he is resting’. This further suggests that the glottal stop is a defective phoneme in the language.

In sum, the glottal stop does not have a robust set of properties to consider it a *full* phoneme: it has a restricted distribution in the $C_1C_2VC_3$ syllable canon, compared to other segments, it does not form minimal pairs and it does not provide a mora (in terms of syllable weight §2.2.1). However, its heteromorphemic coda position in the continuative *-q* added to its occurrence as C_1 in the clusters *qy*, *qw* and *qr* does suffice to consider it a *marginal* phoneme (cf. Doornenbal 2009, 26; Coupe 2008, 77f).

2.4.5 Affricates

Affricates have two places of articulation: alveolar and alveo-palatal. As with the stops, there is a contrast in voicing and only voiceless affricates contrast in aspiration (Evans and Manyu 2021, 4).

2.4.5.1 Alveolar affricates

The alveolar affricates group comprises the phonemes /ts/, /z/ and /tsh/. Their status as phonemes is described in Table 2.44.

Table 2.44: Minimal pairs of alveolar affricates

Phoneme	IPA	Tawrã	Translation
/ts/	[ts]	<i>tsú</i>	blister
/z/	[dz]	<i>zû</i>	write
/tsh/	[ts ^h]	<i>tshû</i>	type of pine tree

The distribution of alveolar affricates in the syllable structure is displayed in Table 2.45. While alveolar affricates can occur in onset and intervocalic positions, they are barred from coda position. ‘X’ stands for no occurrence.

Table 2.45: Alveolar affricates as onset, coda and intervocalic position

Phoneme	Onset	Intervocalic	Coda
ts	<i>tsú</i> ‘blister’	<i>atsâw</i> ‘appoint’	X
z	<i>zû</i> ‘write’	<i>azâ</i> ‘drop’	X
tsh	<i>tshû</i> ‘type of pine tree’	<i>katshã</i> ‘clear (e.g. a jungle)’	X

The decision to keep the alveolar affricate group as phonemes rather than clusters rests on an economic point made by Evans and Manyu (cf. 2021, 6f), since onset cluster *tshw-*, as in *tshwîng* ‘flea’, would otherwise be considered a cluster composed of three elements: *t^h-s-w* - the only one of its kind. Concerning the distribution of the alveolar affricate group, they are disallowed in coda position, a feature shared with the affricate, fricative and lateral groups and the alveolar stops (cf. §2.2.4).

Although neighboring Kera’a does have alveolar affricates (Culhane et al. n.d., 46f) and they also seem to exist in Kaman, a brief inspection of the languages of the larger area suggests that the alveolar affricate group is not very common, as it does not exist in Turung (Morey 2012, 154), Hakhun Tangsa (Boro 2017, 46f) and Proto Tani (Sun 1993, 122). Elsewhere, the alveolar affricate group is much more common, cf. Trung (Perlin 2020, 45) to the east or Tshangla (Andvik 2010, 8), closer to the Bodish world.

2.4.5.2 Alveo-palatal affricates

The alveo-palatal affricate group comprises three phonemes: the voiceless non-aspirated /c/, the voiced /j/ and the voiceless aspirated /ch/. Their phonemic status is described in Table 2.46.

Table 2.46: Minimal pairs of palatal affricates

Phoneme	IPA	Tawrã	Translation
/c/	[tɕ]	<i>cya</i>	s/he, 3
/j/	[dʒ]	<i>ja</i>	what
/ch/	[tɕ ^h]	<i>cha</i>	Assamese

Their position in the syllable is displayed in Table 2.47. While all three members occur as onsets and in intervocalic position, they are barred from coda position. ‘X’ stands for no occurrence.

Table 2.47: Palatal affricates as onset, coda and intervocalic position

phoneme	Onset	Intervocalic	Coda
c	<i>cya</i> ‘s/he, 3sg’	<i>pacû</i> ‘type of bean’	X
j	<i>ja</i> ‘what’	<i>majéi</i> ‘buffalo’	X
ch	<i>tshû</i> ‘type of pine tree’	<i>tachà</i> ‘type of banana’	X

2.4.6 Nasals

The nasal group comprises three places of articulation (labial, alveolar and velar). As expected, this group makes no phonological distinction in either voicing or aspiration.

Table 2.48: Minimal pairs of nasals

Phoneme	IPA	Tawrã	Translation
/m/	[m]	<i>ma</i>	become dull
/n/	[n]	<i>na</i>	CLF (leaves, fabric)
/ng/	[ŋ]	<i>ngâ</i>	fell (e.g. a tree)

In the phonologically aberrant Chagla dialect, a syllable rhyme with a nasals codas is pronounced as a syllabic nasal, i.e. the vowel is deleted. Null onset syllables with a nasal coda lose the vowel altogether, thus *àng* becomes a syllabic consonant *nğ* [ŋ̥] and *um* ‘feather’ is *m* [m̥]. If the onset is an /h-/, the airflow through the vocal tract is obstructed and channeled through the nose, thus nasalizing the consonant [h̃]. One example is [ta.^{h̃}m̥] ‘bear’ (elsewhere [ta.^{h̃}um]), which is displayed in Image 2.15. Note the yellow arrow that indicates the voicelessness of the consonant. Other syllables with a non-/h/ consonant as onset are not affected by this process.

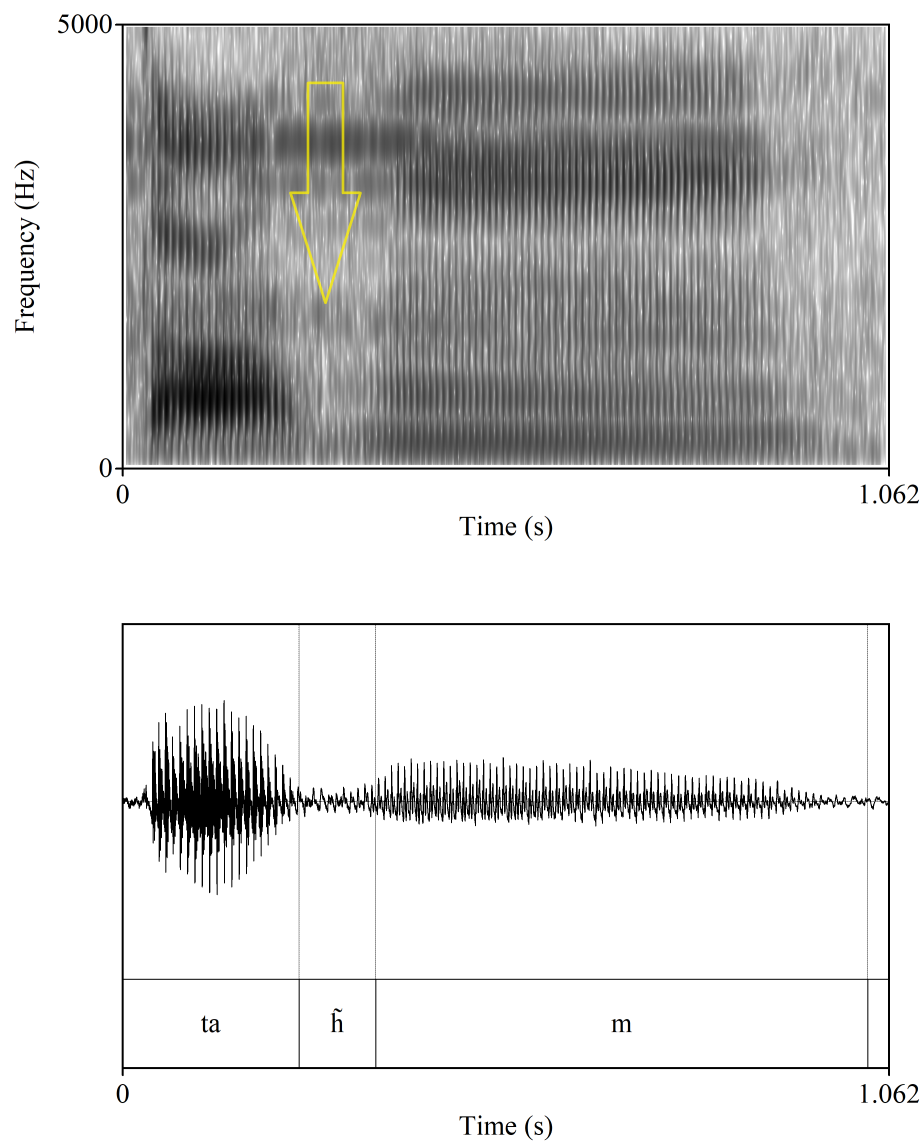


Figure 2.15: Waveform and spectrogram of *tahum* [ta.ʰm̃] ‘bear’ uttered by ST (Chagla dialect)

Table 2.49 describes the syllable position of nasals. ‘X’ stands for no occurrence. All nasals are allowed in onset as well as intervocalic positions. Alveolar nasal is not allowed in coda position.

Table 2.49: Nasals as onset, coda and intervocalic position

phoneme	Onset	Intervocalic	Coda
m	<i>ma</i> ‘become dull’	<i>chûme</i> ‘type of fish’	<i>tyûm</i> ‘drink’
n	<i>na</i> ‘leaf’	<i>ane</i> ‘you, 2P’	X
ng	<i>ngâ</i> ‘fell (e.g. a tree)’	<i>mângâ</i> ‘five’	<i>dâng</i> ‘finish’

In the dialect of Chagla, velar nasal codas correspond to alveolar nasals. Thus, *amang* ‘name’ is pronounced [ə'man].

2.4.7 Fricatives

The fricative group has two members: alveolar /s/ [s] and alveo-palatal /sh/ [ç]. Table 2.50 demonstrates minimal pairs.

Table 2.50: Minimal pairs of fricatives

Phoneme	IPA	Tawrã	Translation
/s/	[s]	<i>sâ</i>	step on (vt)
/sh/	[ç]	<i>shà</i>	wild

The distribution in the syllable is described in Table 2.51. Like alveolar affricates, fricatives cannot occur in coda position. ‘X’ stands for no occurrence.

Table 2.51: Fricatives as onset, coda and intervocalic position

phoneme	Onset	Intervocalic	Coda
s	<i>sâng</i> ‘shadow’	<i>asâ</i> ‘artery, vein’	X
sh	<i>shêi</i> ‘die, get caught’	<i>ashei</i> ‘beg’	X

2.4.8 Lateral

The lateral group has one member: voiceless alveolar [l] <l>. Its status as a phoneme is justified in Table 2.52. Fricative /s/ is shown to form a contrast.

Table 2.52: Minimal pair of the lateral /l/

Phoneme	IPA	Tawrã	Translation
/l/	[l]	<i>lâi</i>	be grown (of the moon)
/s/	[s]	<i>sâi</i>	iron

As hinted by Evans and Manyu (2021, 10), the lateral /l/ has a geminated allophone [l:] when it is the onset of a heavy syllable in intervocalic position. Thus *alô* ‘garlic’ is pronounced [a.ˈl:ô] while *lacrâp* [la.ˈkɾâp̚] ‘elbow’ does not have this alternation. This process is not restricted to tautomorphemic words, cf. [cja.ˈl:ãŋ] (3 COLL) ‘they’.

The voiceless lateral [ɭ] <hl> is considered a cluster and is discussed in §2.2.3.1.

The distribution in the syllable is described in Table 2.53. ‘X’ stands for no occurrence.

Table 2.53: Lateral as onset, coda and intervocalic position

phoneme	Onset	Intervocalic	Coda
l	<i>lê</i> ‘roll, rub’	<i>alô</i> ‘garlic’	X

2.4.9 Approximants

The approximant group comprises three members: voiced labial [w] <w>, voiced alveolar [ɹ] <r> and voiced palatal [j] <y>. Approximants commonly form clusters, such as [kʰr] or [gw]. These are treated separately in §2.2.3.1. The phonological status of approximants is demonstrated in Table 2.54.

Table 2.54: Minimal pairs of approximants

Phoneme	IPA	Example	Translation
/w/	[w]	<i>wê</i>	think
/r/	[ɹ]	<i>rê</i>	tie up, twist
/y/	[j]	<i>yê</i>	question tag
/y/	[j]	<i>yã</i>	grandmother
/h/	[ɦ]	<i>hã</i>	I (1s)

Word-initial /h/ is often dropped in both unstressed and stressed syllables, (*h*)*aprê* [(*h*)a.'prê] ‘full up’, (*h*)*atyû* [(*h*)a.'tjû] ‘fry (with sand)’ or (*h*)*âqa* [(*h*)â.'qa] ‘without’ (cf. §2.18).

The alveolar approximant /ɹ/ has a tap [ɾ] allophone in intervocalic position in free variation, thus *tarim* ‘flock’ can be heard as [ta.'ɹim] and [ta.'ɾim]. In clusters where /ɹ/ is C₂, the alveolar approximant is sometimes realized as an tap [ɾ], although it may have a tap pronunciation in careful speech. The alveolar approximant /ɹ/ should be distinguished from the approximant *Cr* cluster /wr/ [ɹw], cf. the minimal pair *rê* [ɹê] ‘tie, twist (e.g. a rope)’ and *wrê* [ɹwê] ‘take off (clothes)’ (cf. onset clusters on §2.2.3.1). There is one word, however, where these two sounds can be interchanged: *tawra* ‘sword’, pronounced both as [(ta).ɹwa] and [(ta).ɹa].

The alveolar approximant /ɹ/ is found as a phoneme in Assamese. Other Indic (meant non-Assamese) loanwords with intervocalic [ɾ] are readily turned to /ɹ/, as in *aram* [a.'ɹam] ‘easy, to rest’ (from Hindi *aram* ‘rest’).

The distribution of these phonemes in the syllable is displayed in Table 2.55. While approximants can occur in onset and intervocalic positions, they are not found in coda position. Rhymes of the type /-Vk/ have an allophone [-Vw] (this is further discussed in §2.2.4.1). For this reason, the approximant /w/ has a (X) in coda position, for it occurs as an allophone of /-Vk/ rhymes.

Table 2.55: Approximants as onset, coda and intervocalic position

Phoneme	Onset	Intervocalic	Coda
/w/	<i>we</i> ‘demonstrative’	<i>tâwa</i> ‘jaw’	(X)
/r/	<i>ru</i> ‘landslide’	<i>karà</i> ‘wild (only in compounds)’	X
/y/	<i>ya</i> ‘night’	<i>wáyi</i> ‘be upside down’	X
/h/	<i>habá</i> ‘hit, whip’	<i>ahû</i> ‘sweat’	X

Indic loanwords with /r/ in coda position are adapted in different ways to fit the phonology of Tawrá. As an example, the Indic word *kirtan* [kiɾ.'tan] ‘a Hindu devotional practice that involves call-and-response chanting’ is rendered as *kritân* [kɾi.'tân]; a metathesis in which the banned coda /ɹ/ is changed to an onset cluster [kɾ]. In a differ-

ent example, the Indic word *sarkar* ‘government’ has been adapted as *sokarêi* [so.ka.ʼɛi] or *sokâri*²⁴ [so.ʼkâ.ɿ]. In both cases, the first coda [r] is deleted and an epenthetic vowel/diphthong is added to render /ɿ/ intervocalic.

2.5 Tone

Tawrã is a tone language in the sense that “indication of pitch enters into the lexical realization of at least some morphemes” (Hyman 2006, 229f). There are four contrastive tonemes: high tone (H), low tone (L), high falling (HF) and mid tone (M) (absence of contour).

The present section will describe the key features of tone in the language. After a brief discussion of relevant areal features (§2.5.1), I will review the pivotal work of Evans and Manyu (2021), which includes a description of tone in the language. Thereafter I’ll describe tone for monosyllabic words (§2.5.3). Finally, §2.5.5 describes tonal *patterns* for polysyllabic words. Note that §2.2.6, on moraic structure, discusses theoretical points that are crucial to understanding this subsection.

2.5.1 Tone in NEI languages

Tone has been reported for several Tibeto-Burman languages of NEI. In spite of the overwhelming diversity of these languages, researchers have pointed out common properties or areal trends regarding tone. This constitutes a good starting point to approach the tonal system of Tawrã.

The first point is the domain of tonal contrast. While many languages of East and Southeast Asia have the *syllable* as the tone-bearing unit (TBU), i.e. tonal changes operate at the syllable level and are generally stable, Tibeto-Burman languages from in and around NEI often have a TBU at the word or root level. As an example to illustrate this point, Galo contrasts [tu⁴⁴nam³³] ‘to kick’ with [tu⁵⁵nam²¹] ‘to garland’. As described by Post (2015a, 196f), these verbs are composed of a root plus a toneless nominalizer *-nam*:

²⁴The variation *sokâri* ‘governmental’ (instead of *sarkar*) could have been influenced by the Indic adjective *sarkârî* ‘governmental’. An equivalent form also exists in Galo, even though coda [-r] is permitted (Post 2024, p.c.).

tú + *nam* ‘kick + NMLZ’ and *tù* + *nam* ‘garland + NMLZ’. The word-level, superficial tone is determined by the root and predictable phonological rules (also cf. Kirby and Brunelle 2017, 711f; Lin 2012, 629f; Matisoff 2000, 87f; Morey 2014, 660f). Perhaps the difference in the domain of tonal contrast has led scholars more acquainted with a syllabic TBU structure, such as Chinese, to extrapolate that system to languages with a word-level TBU (Post 2015a, 187). This has also been the case of Tawrã, as the works based on Tibetan Tawrã (Jiang et al. 2013, 33-34; Sun 1991) indicate the tonal value of every syllable, cf. *ku*³¹*pau*⁵³*liu*⁵⁵ ‘butterfly’ (*kaprûlyu* in the orthography used for the present publication).

A further point that also applies to the languages of MSEA is the phonetic manifestation of tone, which has been described as a *bundle* of suprasegmental features (also cf. Kirby and Brunelle 2017, 711f; Morey 2014, 638) including intensity, vowel quality and duration and phonation (cf. Mulder 2018) and we should also add stress to that list, as current research points out that these two features (tone and stress) are often inseparable (cf. Hyslop 2021; Evans 2017, 236f; Yip 2002, 31, 255f).

A last point, which is probably related to the latter, is the reported low importance of tone²⁵ for the phonological system. This has been described for individual languages (in Karbi, cf. Konnerth and Teo 2014, 14) as well as for comparisons of different languages and surveys (cf. Post 2015a, 186; Morey 2014, 640f; Van Dam 2019, 9). This is expressed in different manners for each language but, as a general point, it refers to the lack of awareness of tonal categories by speakers and a difficulty in finding minimal pairs for tone.

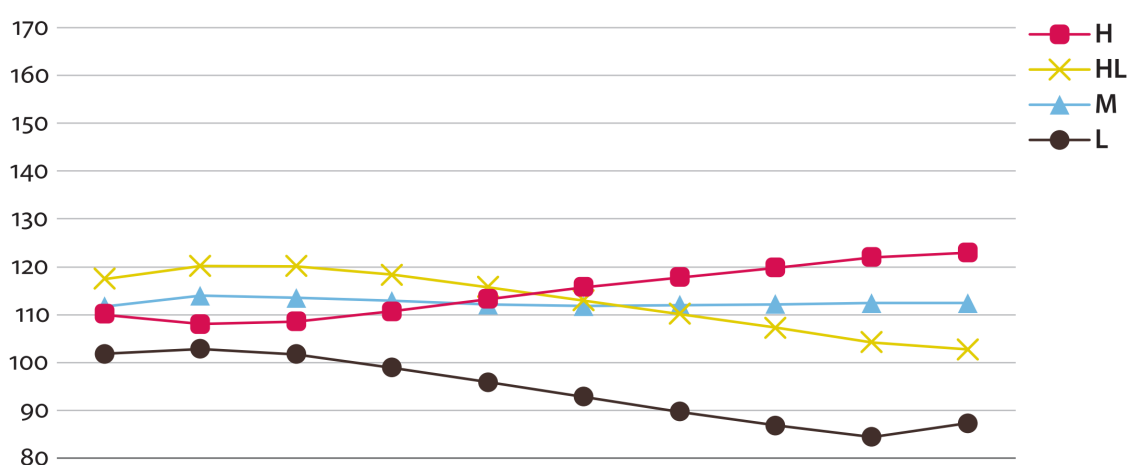
As will become evident in this chapter, several points are relevant for Tawrã, particularly the *word* as the domain of tonal contrast and the different cues, apart from pitch, for the identification of tone.

²⁵This is usually expressed as having a low *functional load*, a term borrowed from information theory that refers to the relative importance of one linguistic contrast (usually a phonetic trait or a segment) over others in relation to the large system (cf. Hockett 1967; Oh et al. 2015), e.g. as a mere guess, merging the vowels /a/ and /ü/ in a large corpus of Tawrã would create much more homophony in the language than merging /e/ and /i/ and thus the /a-ü/ contrast would be considered to have a higher functional load.

2.5.2 Previous work on tone

The analysis of Evans and Manyu (2021, 19f) proposed a tonal system of maximally one tone per morpheme (regardless of the number of syllables) based on three tonemes: high falling (indistinctly referred to as high low), high and low. Morphemes without an assigned tone and non-stressed syllables receive a default mid pitch, which is not a toneme *per se* (also cf. Evans 2017, 234). On monosyllables, the pitch of high and low tones is described as often beginning in the mid-range and then sliding up or downwards, respectively. Creaky voice is a feature of the low tone and it also shows up in high-falling syllables. The pitch of the high-falling tone typically descends to the mid range, particularly in a pre-pausal position, as in *nyâ* ‘thread’ and *tamyûm* ‘monkey (generic)’, but it can fall to low when a toneless syllable follows, for example in *sû-ya* (boil-ALLO). The authors provide a clever plot that traces the change in pitch in Hz on monosyllables containing the vowel /a/ [ɑ], here reproduced as Figure 2.16. The reader should note the representation of the changes in pitch characteristic of each toneme (and the lack of tone). The F0 plot depicts time-normalized averages for three utterances uttered in isolation and individually by the second author (in the plot *JM* but in this study the acronym *JO* is employed). The authors recognize that a four-way minimal pair based on tone alone is hard to find.

Figure 2.16: F0 tracing of contrasting tones in Hz on monosyllables containing [ɑ] (three utterances by JM) (Evans and Manyu 2021, 20)



Concerning polysyllabic words, the authors propose a difference between disyllabic

and trisyllabic words. While the former are unrestrained in the combination of tones on the first and second syllable (i.e. all three tonemes plus the unassigned tone (mid tone) can occur on either the first or the second syllable), trisyllabic words have only three possible tone patterns: a) high-falling on the first or b) second syllable and c) high tone on the third syllable. Words with more than three syllables have not been attested.

Nominal compounds often retain their tonal specifications. The authors take notice of the difference between disyllabic compounds and disyllabic morphemes, which may display “different surface pitch patterns even if the underlying tones are the same” (Evans and Manyu 2021, 22). This is exemplified by mono-morphemic *tâqã* ‘light (weight)’, where the second syllable takes a low tone (which spreads from the initial high-falling tone), while the compound *pyâ tshri* ‘type of bird’ does not display tonal spread like the mono-morphemic example, with the second morpheme retaining its toneless specification and surfacing as mid pitch. This example is represented in Table 2.56. Note that the underlying tonal pattern of the mono-morphemic *tâqã* ‘light (weight)’ HF-M changes to a surface HF-F while this does not occur with the bi-morphemic compound *pyâ tshri* ‘type of bird’

Table 2.56: Tonal difference between disyllabic compounds and disyllabic morphemes (based on Evans and Manyu 2021, 22)

morpho-syllabic structure	example	underlying tonal pattern	>	phonemic script	surface tonal pattern	translation
disyllabic monomorphemic	<i>tâqã</i>	HF-M		<i>tâqã</i>	HF-F	light (weight)
disyllabic bi-morphemic (compound)	<i>pyâ tshri</i>	HF M		<i>pyâ tshri</i>	HF M	type of bird

Lastly, the authors note the difficulty of tracing the underlying tone in some compounds. The example provided is *tù chei* ‘mustard oil’, composed of *tù* ‘mustard’ (cf. IA *tori* ‘mustard’) and *chei* ‘oil’. Crucially, the authors attested the compound with an alternative tonal pattern, namely [tû c^hei] (i.e. high-falling tone on an underlying low tone), which admittedly contravenes their claims about the behavior of tones in nominal

compounds.

The present study has taken a bottom-up approach to the phonology of the language and all features thereof were independently studied and tested. In most of the cases regarding tone, there were no important differences with the foundational work of Evans and Manyu (2021), e.g. regarding the number of tones and their associated features (creaky voice and pitch assimilation). Based on a set of minimal pairs, I've also concluded that there are four contrastive tones, as will be discussed in §2.5.3. As will be described presently, I regard the mid tone as its own toneme based on its phonetic features and the long list of minimal pairs. On top of that, toneless syllables usually behave as a mid tone in lacking an inherent pitch contour. The present analysis offers an explanation for the unexpected tonal change in [tû c^hei] 'mustard oil'; low tone becomes regularly high (however not high-falling!) in nominal modification of the type N-N (i.e. *tù* vs. *tú*). Furthermore, the present analysis proposes an alternative approach to polysyllabic words based on syllable weight that reduces the potential combination of tone on the first and second syllable to five *tonal patterns* (cf. §2.5.5).

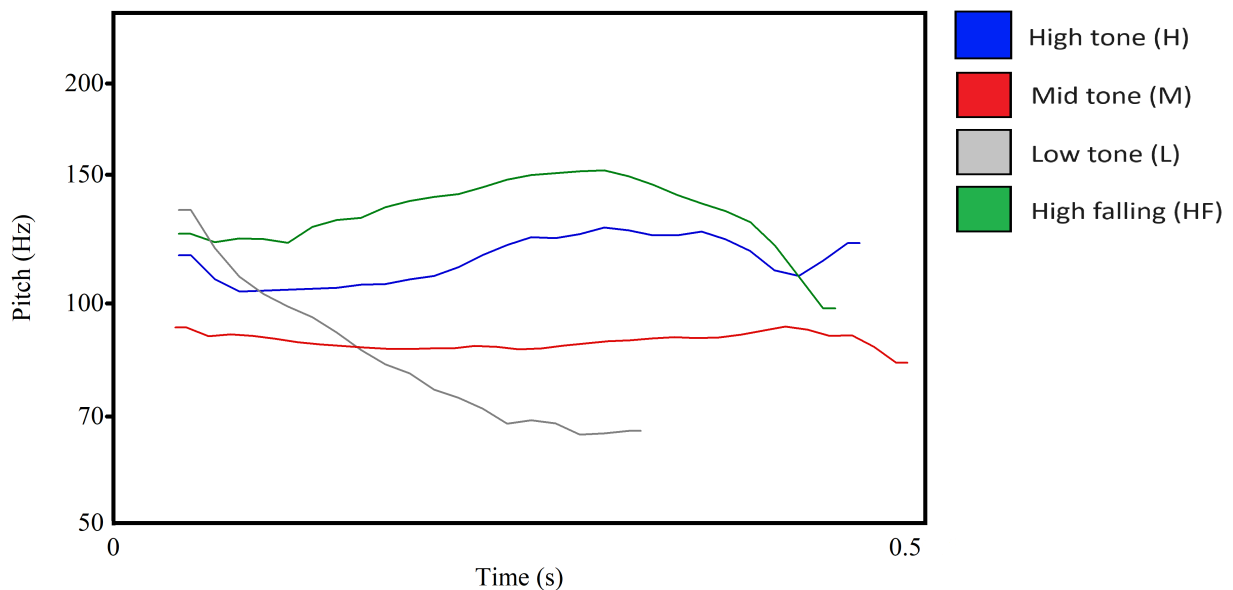
2.5.3 Tone on monosyllables

This section will describe all four tonemes as found in monosyllables. Tone on polysyllabic words is discussed in §2.5.5. Table 2.57 displays a list of segmental minimal pairs based on tone in monosyllabic words. It also includes the Chao tone letters (Chao 2011[1968]) to help indicate the phonetic realization of tonal categories. Figure 2.17 represents the tone of the same 4-way contrast based on one utterance per tone. The pitch of these utterances was superposed (i.e. non-normalized) using four different colors for the sake of comparison. As will become clear in the remainder of this section, apart from pitch, there are other cues for individual tones: creaky voice for low tone with, vowel length for mid tone and high intensity for high-falling tone.

Table 2.57: Tonal minimal pairs

word	tone	translation	tone letter
<i>pá</i>	H	yeast	↑
<i>pâ</i>	HF	go across	↘
<i>pà</i>	L	axe	↓
<i>pa</i>	M	draw (water)	⊥

Figure 2.17: Representation of pitch of all four tonemes based on the examples *pá*, *pâ*, *pà* and *pa* by JO



The remainder of this section will discuss each of the four tones available in the language: high-falling tone (§2.5.3.1), mid tone (§2.5.3.2), low tone (§2.5.3.3) and high tone (§2.5.3.4).

2.5.3.1 High-falling

There are two main cues for high-falling tone: a) change in pitch, with one optional ascending contour followed by a falling contour (Evans and Manyu 2021, 19) and b) a relatively high intensity (also a feature of stress, cf. §2.2.6). Image 2.18 illustrates the

waveform (i.e. pressure over duration) and pitch (i.e. F0 over duration) of the verbs *lû* ‘pour’ and *tyû* ‘sell’ as uttered by BAB and JO respectively. In both cases, the pitch illustrates the characteristic falling contour, although it is much more pronounced in the right panel than in the left panel.

Figure 2.18: High-falling tone: *lû* ‘pour’ by BAB and *tyû* ‘sell’ by JO

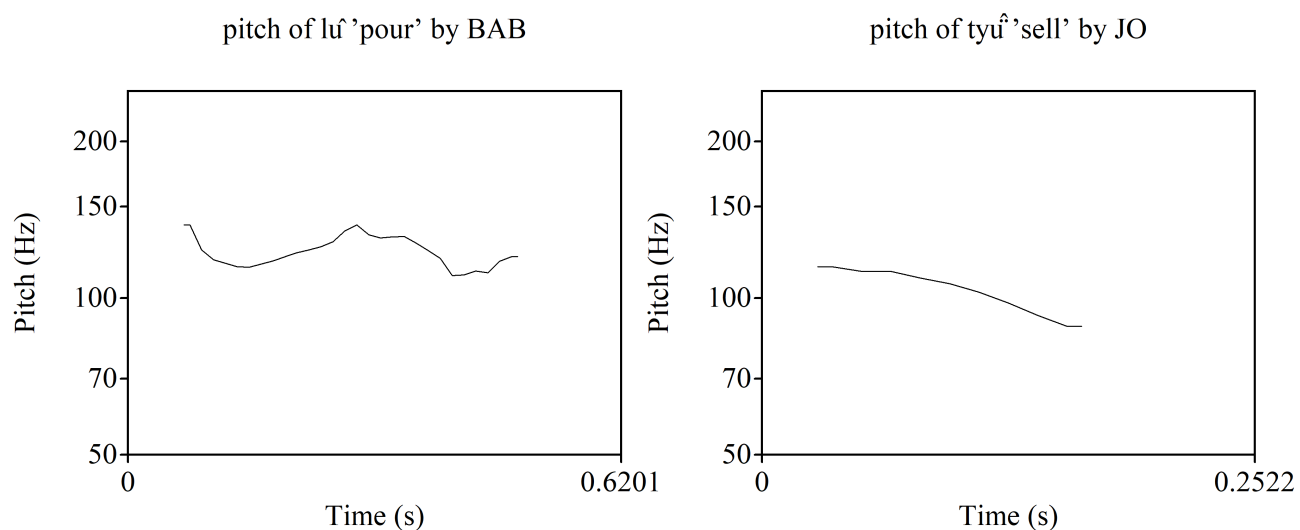
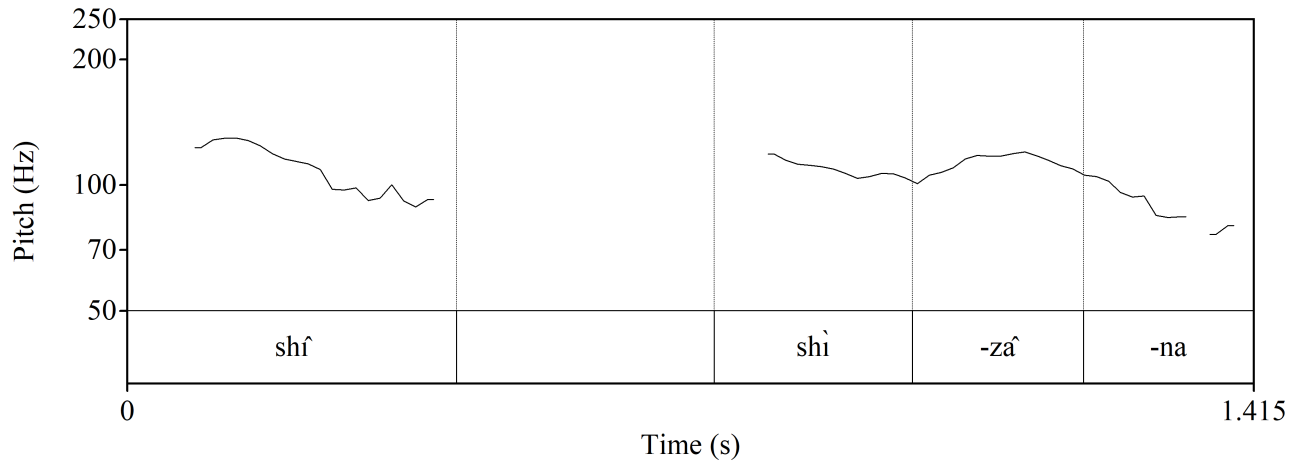


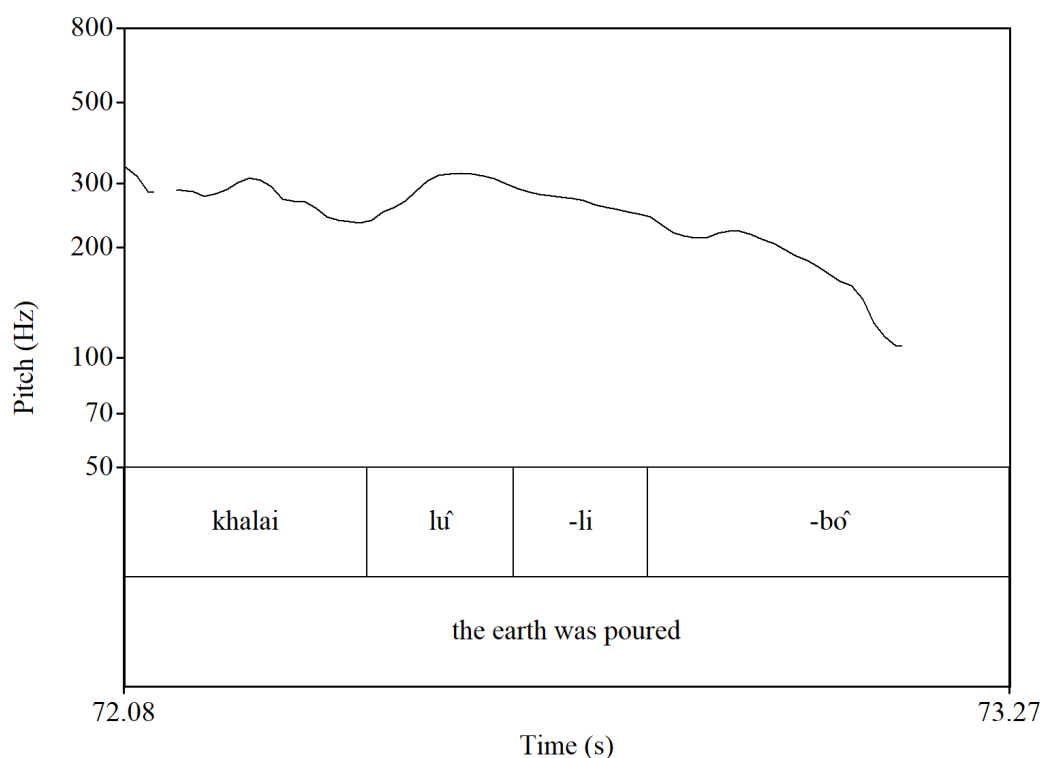
Figure 2.19 illustrates a comparison of pitch between HF *shî* ‘claw’ and L *shî* ‘take’ (in a sentence *shî-zâ-na* ‘s/he took it away’. As for HF *shî* ‘claw’, the pitch undergoes a sharp fall, which is also a common realization of this tone. In disyllabic words of the type HF-M (i.e. trochaic foot, cf. §2.5.5.1) the pitch of the initial syllable might go even lower on the second syllable.

Figure 2.19: Comparison of pitch between HF *shî* ‘claw’ and L *shì* ‘take’ by ST



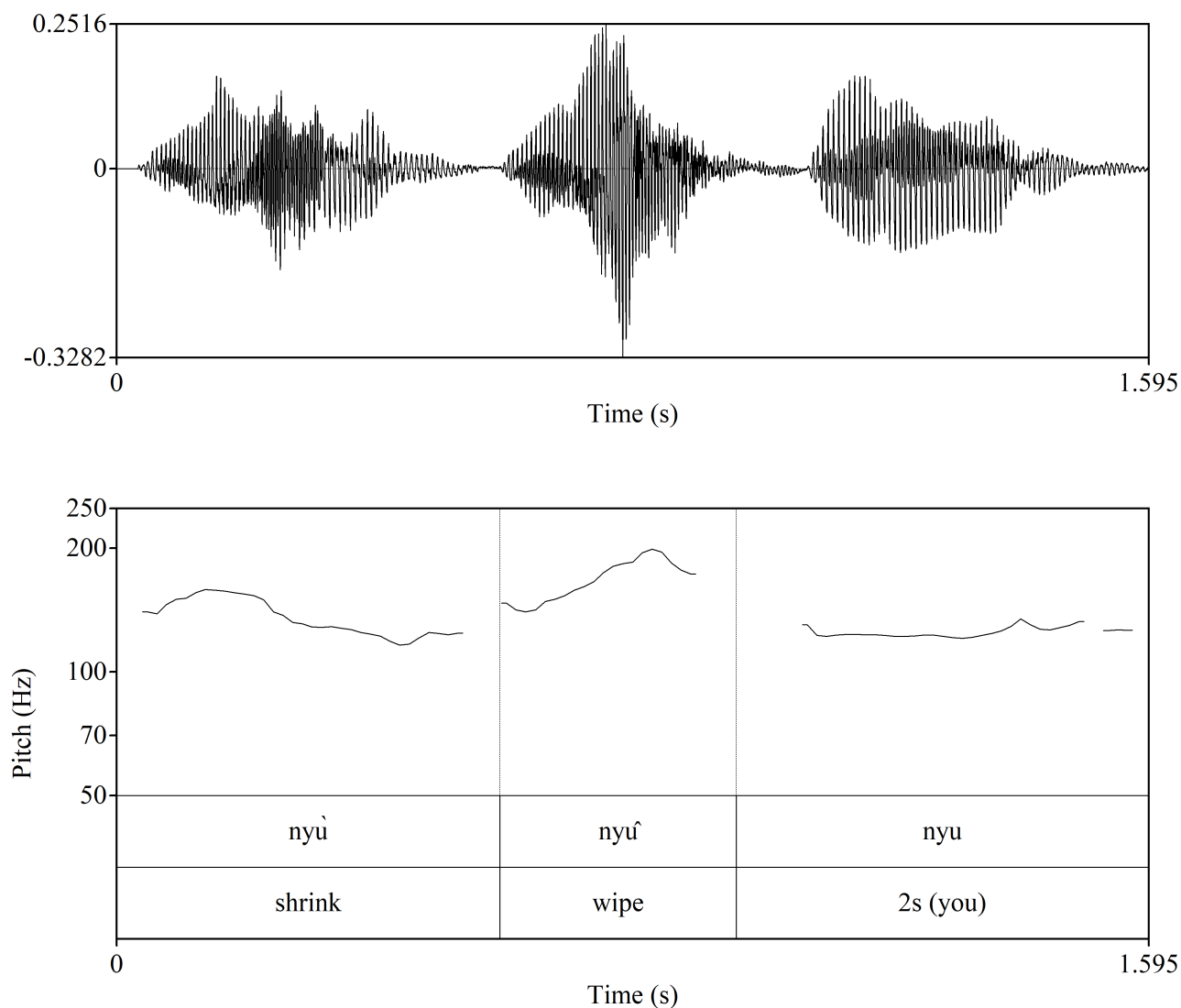
Pitch generally decreases across the utterance. Image 2.20 illustrates the sentence *khalai lûlibô* ‘the earth was created’ uttered by TY as an example in which the pitch goes progressively down. Although the verb *lû* ‘pour’ displays the characteristic high-low change in contour, the marker of recent past *-li* and particularly the perfective marker *-bô* describe a progressive lowering of pitch. This phenomenon is known as *downdrift* and it seems to be a universal tendency in tone languages (cf. Yip 2002, 12). Toneless suffixes are prone to prolong a descending pitch (cf. §2.5.4).

Figure 2.20: Illustration of pitch change in the sentence *khalai lûlibô* ‘the earth was created’ by TY



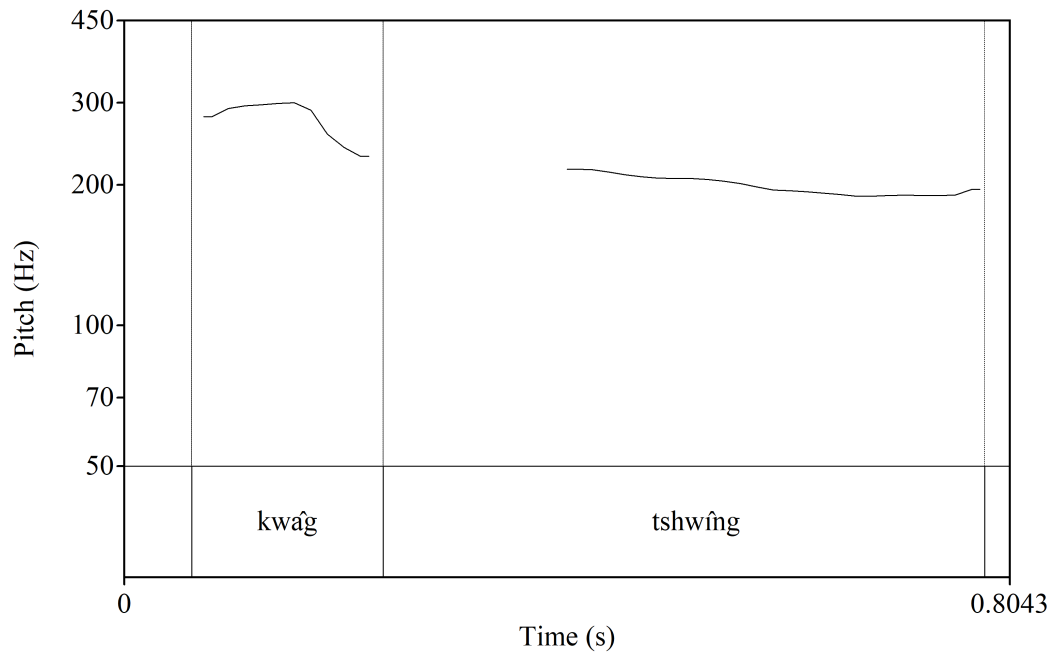
The second cue for high-falling tone is intensity; monosyllables with a HF tone are perceived as *louder* and this can be observed when looking at the amplitude of the waveform of words uttered in isolation. Image 2.21 illustrates three minimal pairs that only differ in tone. These were not uttered in sequence in order to avoid list intonation. The top-level illustrates their relative waveform while the lower level their pitch. These are *nyù* ‘shrink’, *nyû* ‘wipe’ and *nyu* ‘2s’. These words were uttered individually by BAB and merged into one single graph for the sake of comparison. The higher amplitude of the HF tone as compared to a L and M tone is taken here as a sign of its intensity.

Figure 2.21: Comparison between L *nyù* ‘shrink’, HF *nyû* ‘wipe’ and *nyu* ‘2s’ by BAB



The pitch of two adjoined HF tones presents some degree of variation. Similarly to the example presented above, one possible realization of two adjoined HF is one HF tone losing its ascending contour if preceded by another HF tone. This is exemplified in Figure 2.22, where the underlying HF tone in the monosyllabic noun *tshwîng* ‘flea’ surfaces with a descending contour after a HF tone.

Figure 2.22: Pitch of *kwâg tshwîng* ‘dog’s flea’ by DI



A further possibility is to merge two HF tones into one. Image 2.23 depicts the pitch of the verbal compound *zû-lô* (lit. ‘write pierce’) ‘read’ as uttered in *allegro* speech by TY while telling a story. It has an ascending pitch on the first element and a descending pitch on the second. Lastly, a careful pronunciation of the same verbal compound is illustrated in Image 2.24, uttered by JO. Here each element has a HF pitch, without assimilation.

Figure 2.23: Pitch of *zû-lô* ‘write’ by TY

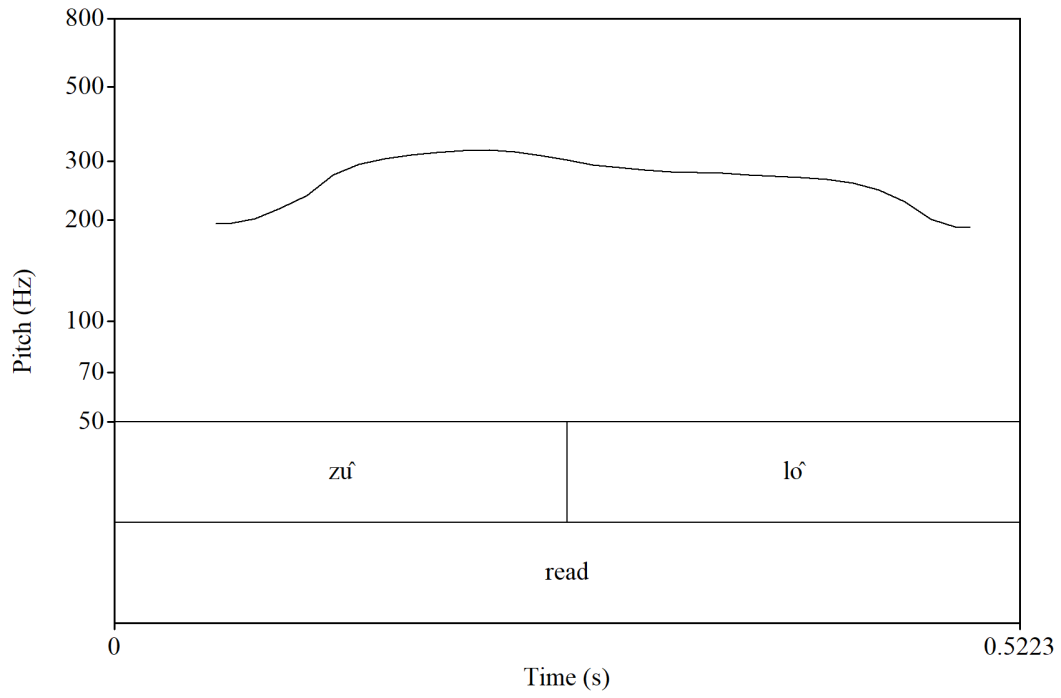
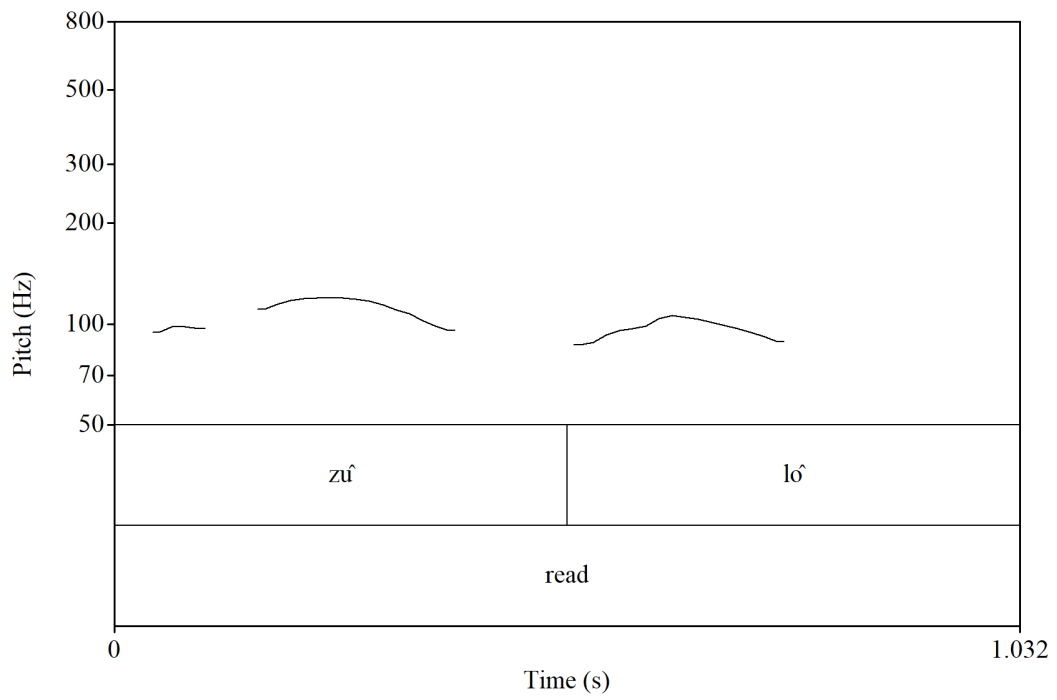


Figure 2.24: Pitch of *zû-lô* ‘write’ by JO



In the tonal patterns of disyllabic words (§2.5.5), HF tone has a prominent role as it is the only tone allowed in the trochaic group (i.e. when an initial syllable is stressed). When HF tone is the second syllable (i.e. iambic foot) it prompts a low contour on the preceding toneless syllable. This is illustrated in Image 2.25 uttered by JA. Here the pitch of the initial syllable goes slightly down before the (interrupted) HF curve. This type of leftward assimilation triggered by a HF does not only occur tautomorphemically and is also found across syllable boundaries before a M syllable, as illustrated in Figure 2.26 uttered by JO, which depicts *cya thâ* ‘s/he eats/ate’. Here the pitch of the M tone word *cya* ‘3’ goes slightly down before the proceeding HF tone.

Figure 2.25: Pitch of *malâ* ‘search’ by JA

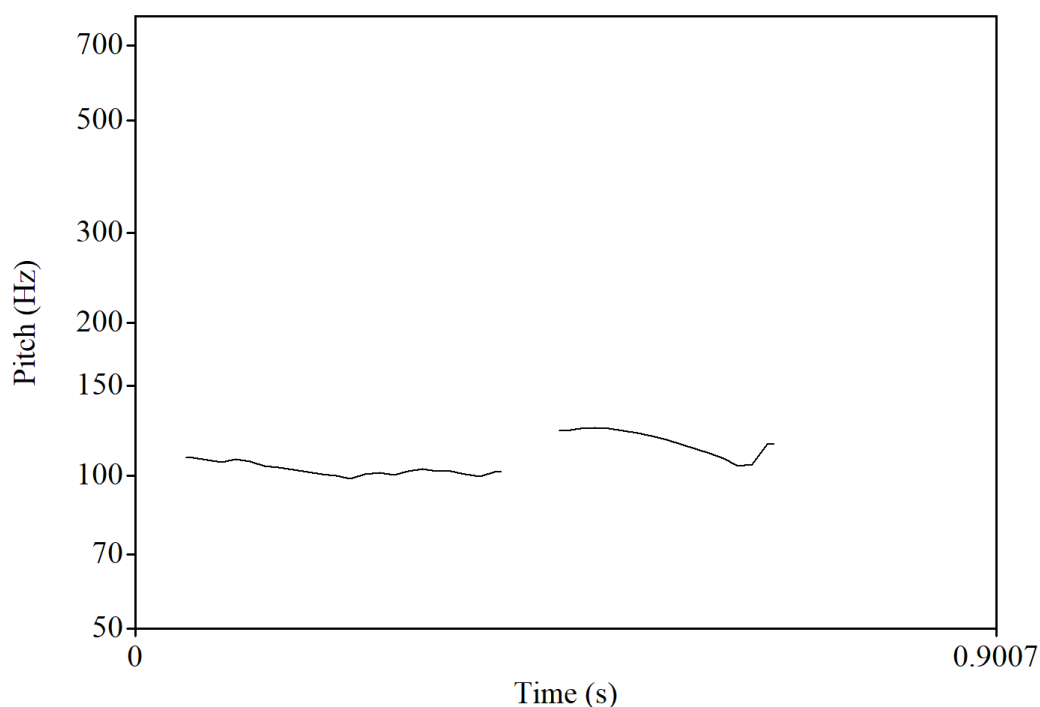
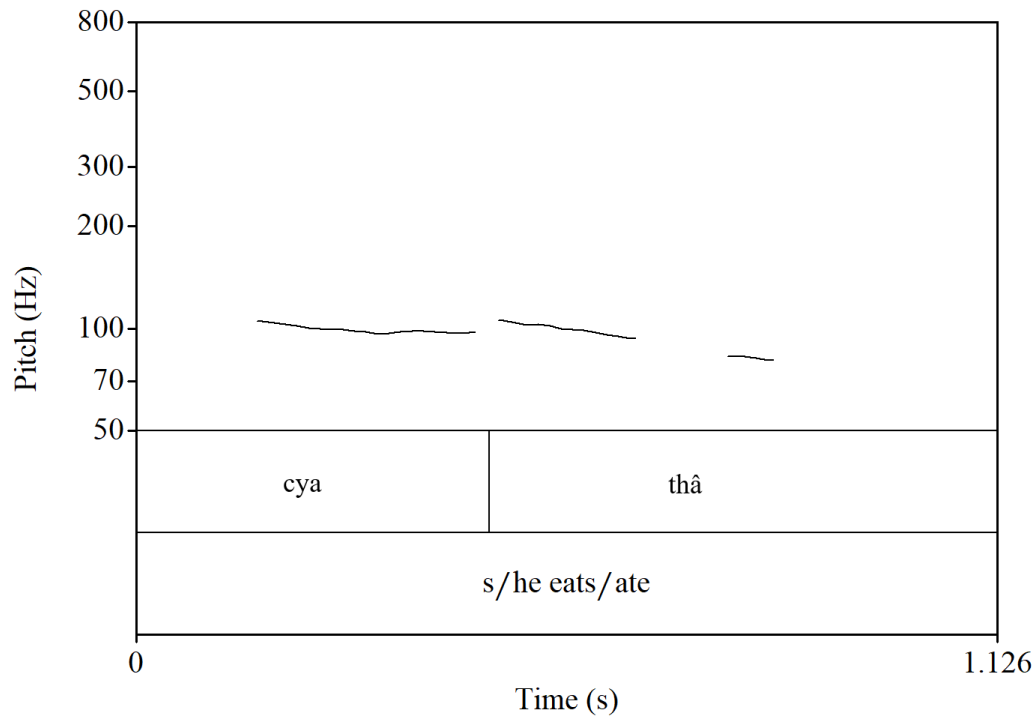


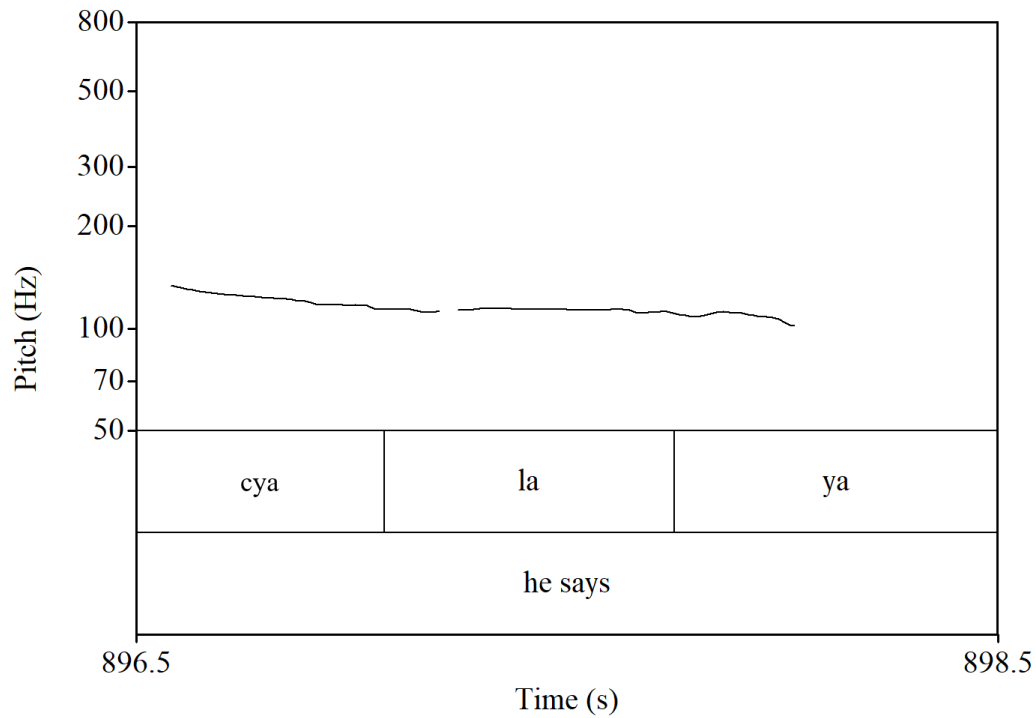
Figure 2.26: Pitch of *cya thâ* ‘s/he eats/ate’ by JO



2.5.3.2 Mid tone

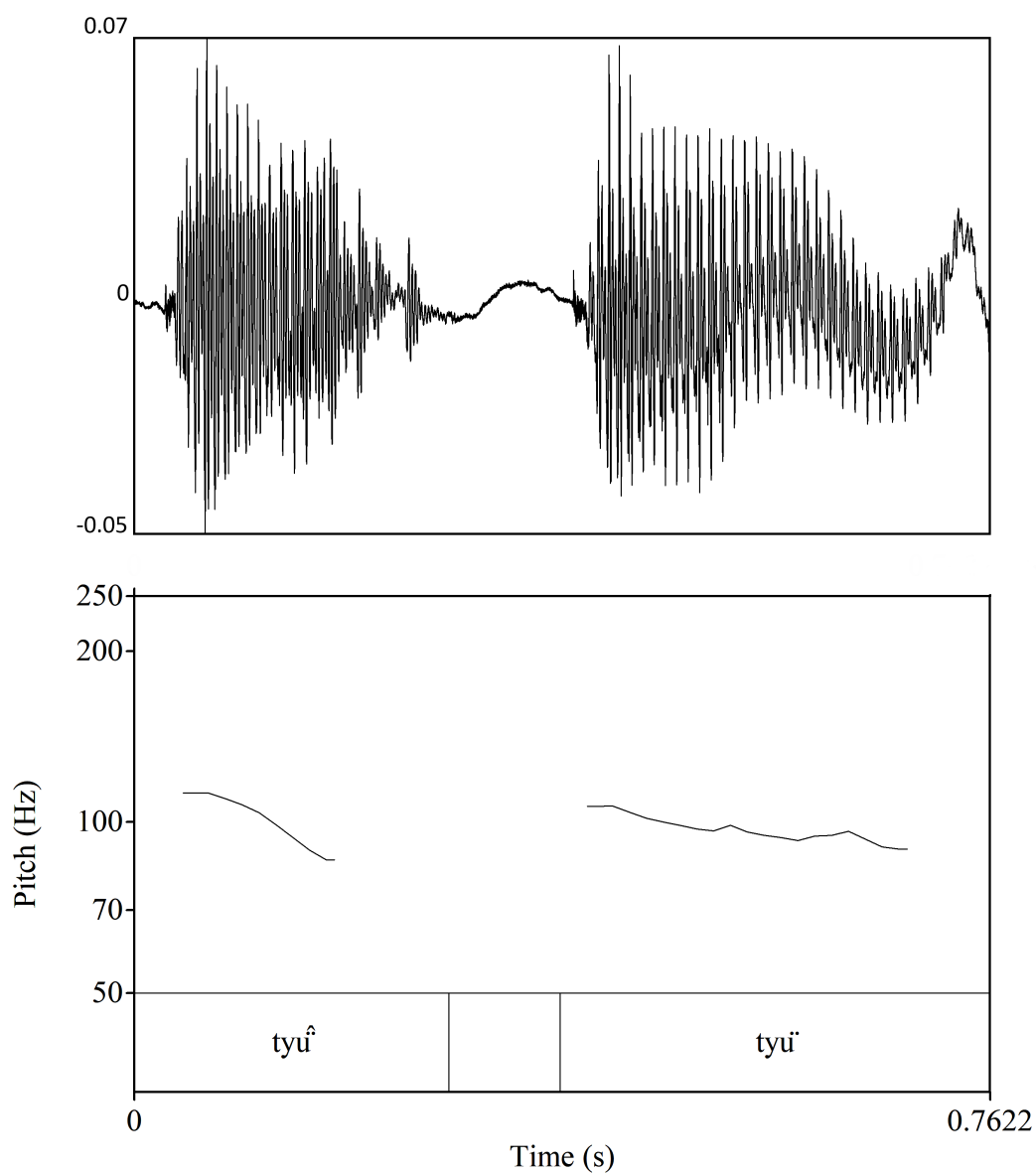
Mid is a contrastive toneme characterized by a continuous (often long) pitch that is located below the pitch associated with high tone and above the low tone with an absence of contour (i.e. change in pitch). This is illustrated in Image 2.27, which corresponds to *cya la-ya* (3 say-ALLO) ‘s/he says’. Here the first two morphemes have a M tone while the last morpheme *-ya* is toneless and effectively gets a M surface tone.

Figure 2.27: Pitch of *cya la-ya* ‘s/he says’ by JA



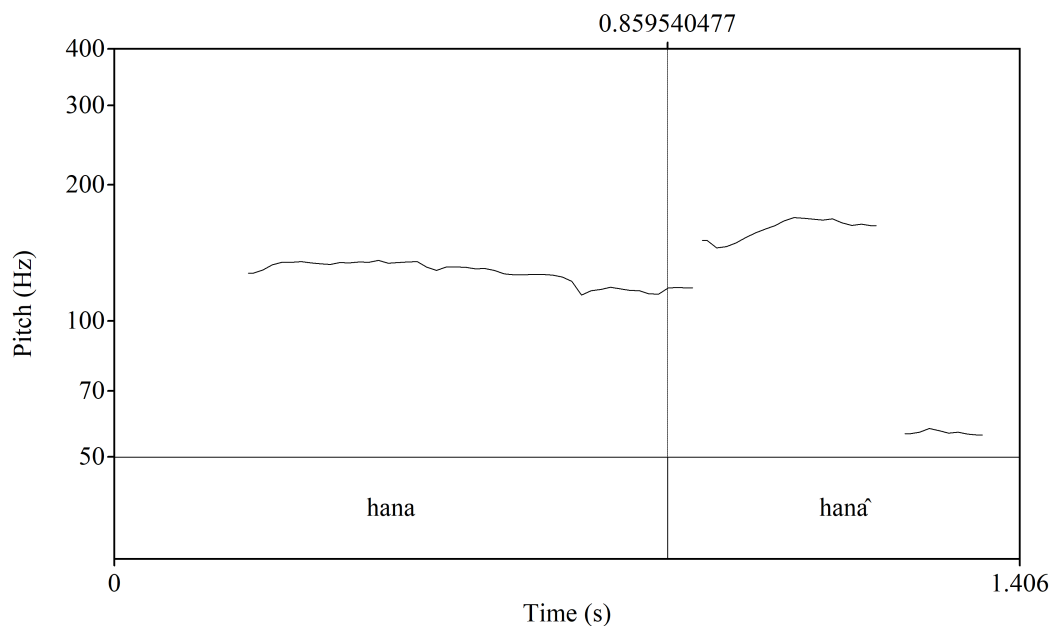
When contrasted with other tones, Mid tone is described by speakers as having a longer duration. This longer duration is often exaggerated in elicitation. This is illustrated in Image 2.28, which displays the spectrogram and pitch of the monosyllabic morphemes *tyû* ‘sell’ and *tyü* ‘snatch’ as uttered by JO. Crucially, the length of the vowel of *tyû* ‘sell’ is 0.19 seconds while that of *tyü* ‘snatch’ is 0.31 seconds.

Figure 2.28: Difference in length (spectrogram and pitch) between *tyû* ‘sell’ and *tyü* ‘snatch’ by JO



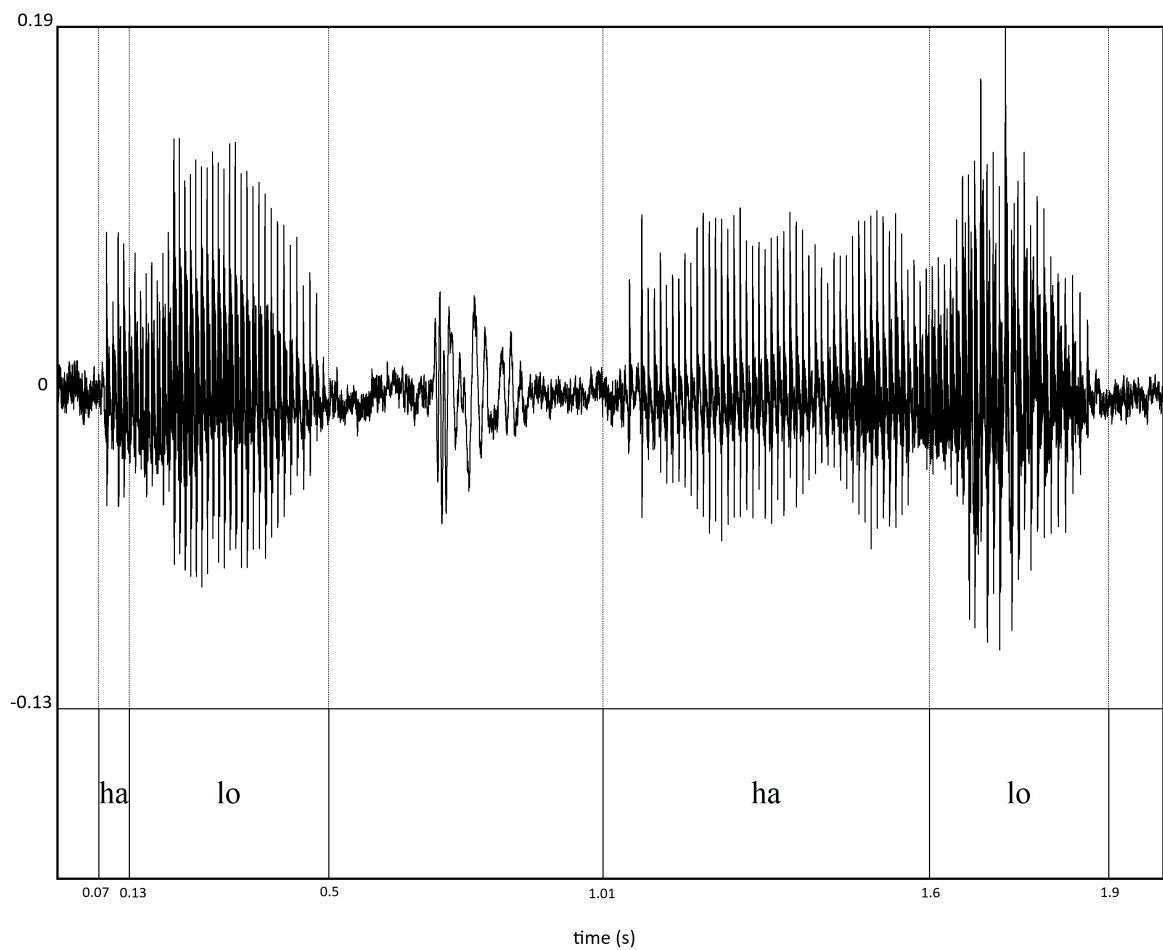
This trait is maintained in disyllabic morphemes as illustrated in 2.29. Here the disyllabic verbs *hana* ‘come’ and *hanâ* ‘squeeze’, uttered by JO, are again portrayed through a spectrogram and pitch. The length of the M-M verb *hana* ‘come’ is 0.56 sec whereas the M-HF verb *hanâ* ‘squeeze’ is 0.42 sec.

Figure 2.29: Difference in length (spectrogram and pitch) between *hana* ‘come’ and *hanâ* ‘squeeze’ by JO



Finally, Image 2.30 depicts two independent realizations of the same noun *halo* [ha'lo] ‘moon’ uttered by ST and merged into one image for clarity. After the first pronunciation by ST, which is 0.44 sec long, I produced an erroneous pronunciation (**hâlo*; not shown here) upon which ST pronounced the second, more careful realization (0.89 sec) with exaggerated length to clarify. The point here is again that vowel length is a cue of mid tone and speakers may even exaggerate vowel length to convey this tone.

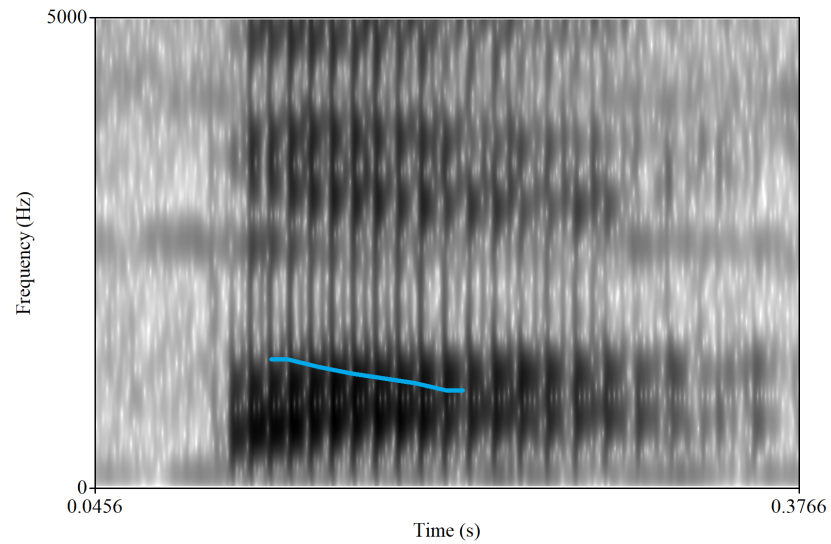
Figure 2.30: Difference in length (waveform) between two realizations of *halo* ‘moon’ by ST



2.5.3.3 Low tone

Low tone is characterized by a lowering of pitch and, commonly for some speakers, a creaky voice. A low tone is depicted in Image 2.31, uttered by JO. The cyan descending line represents the change in pitch over time. The pulses on the spectrogram demonstrate creaky voice.

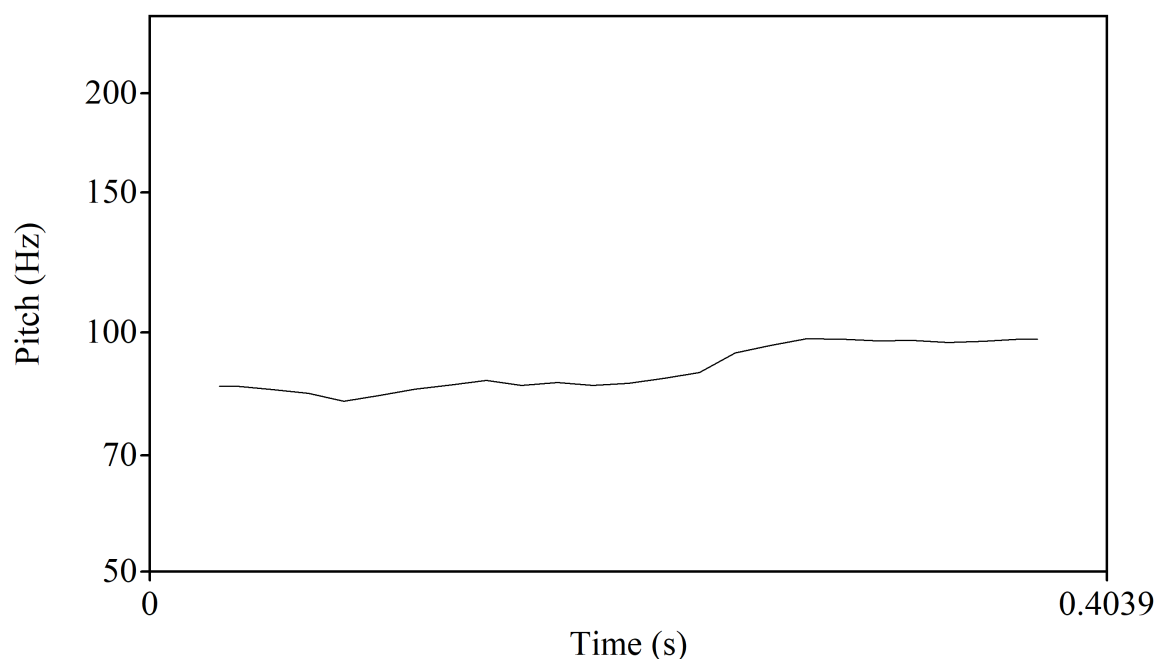
Figure 2.31: Pitch and spectrogram of *pà* ‘ax’ by JO



2.5.3.4 High tone

The high tone is characterized by a gradual heightening of pitch. The change in contour is illustrated by Image 2.32, which corresponds to the noun *wrái* ‘bamboo pattern (used as walls)’ as uttered by JO.

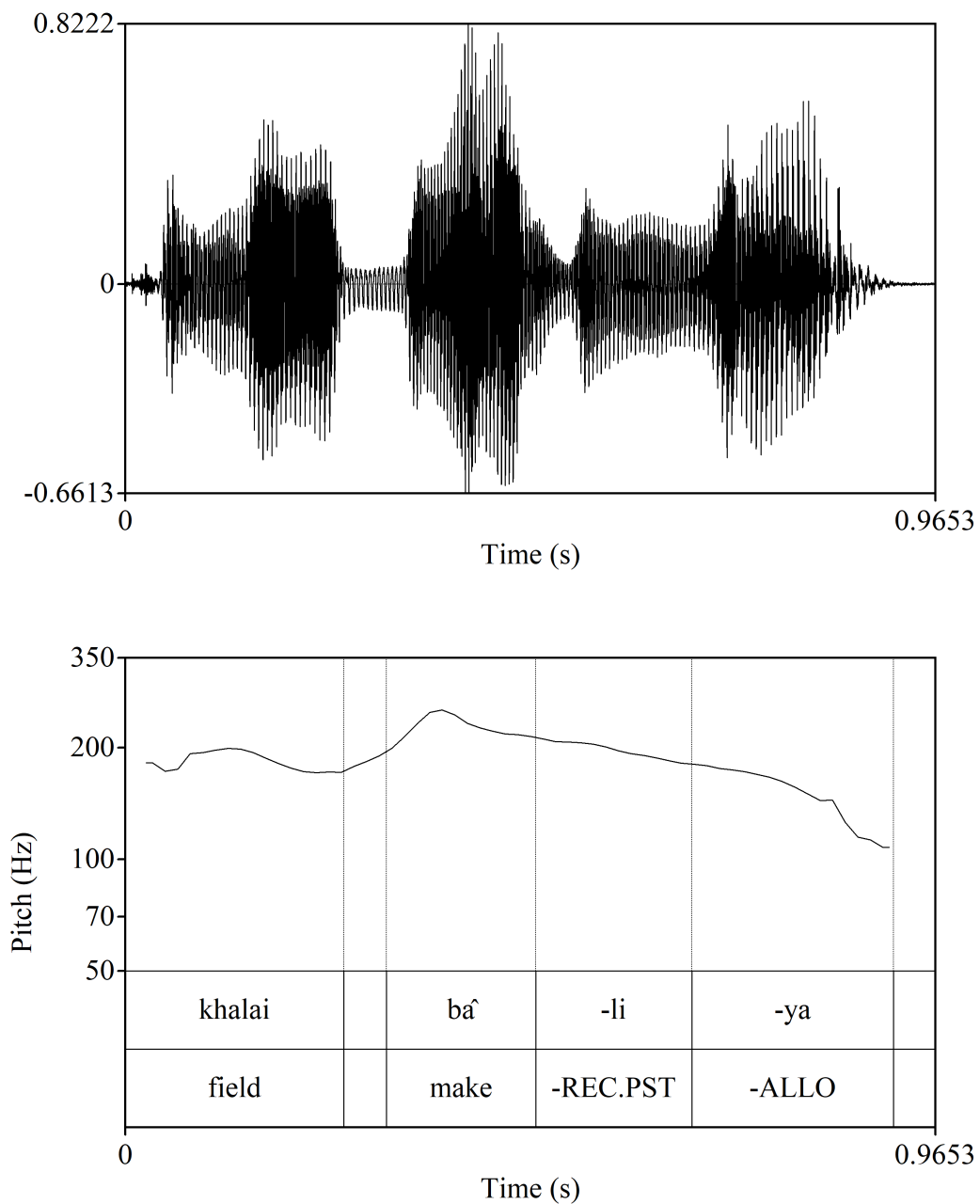
Figure 2.32: Pitch of *wrái* ‘bamboo pattern’ by JO



2.5.4 Toneless suffixes

A number of suffixes in the language are toneless. A toneless morpheme is a light syllable (§2.2.5), i.e. it contains only one mora, a reduced syllable structure (no coda) and a shorter vowel that may not be nasalized. Its pitch is determined by the surrounding tones. Image 2.33 displays the whole sentence *khalai bâ-li-ya* (field make-REC.PST-ALLO) ‘(S/he) (just) made the land’ which includes two toneless suffixes: the recent past marker *li* (§6.3.4.9) and the allophoric marker *-ya* (§6.3.6.1). Note that their amplitude is much smaller than the main verb. Similarly, their pitch contour merely prolongs the descent prompted by the high-falling verb *bâ* ‘make’.

Figure 2.33: Spectrogram and pitch of whole sentence *khalai bâliya* ‘(S/he) (just) made the land’ by TY



We can assume that many toneless suffixes were full words at some point but were stripped of their suprasegmental features upon being recruited as grammatical morphemes.

2.5.5 Tonal patterns: disyllabic words

Tonal patterns refer to the interaction of pitch at the syllable level as determined by word-level stress and tone. As discussed in detail in §2.2.6, stress may fall on either the initial (trochaic pattern) or the second syllable (iambic pattern). The former has only one possible high-falling tone while for the latter all four tonemes are available on the stressed syllable. The toneme on the stressed heavy syllable will determine the surface pitch of the light syllable(s), thus forming tonal *patterns*. Light syllables are toneless (cf. §2.18) in the sense that they do not have an underlying tone. This section describes all five tonal patterns for disyllabic words. Table 2.58 (repeated from Table 2.5) illustrates all five tonal patterns.

Table 2.58: Tonal patterns in disyllabic words

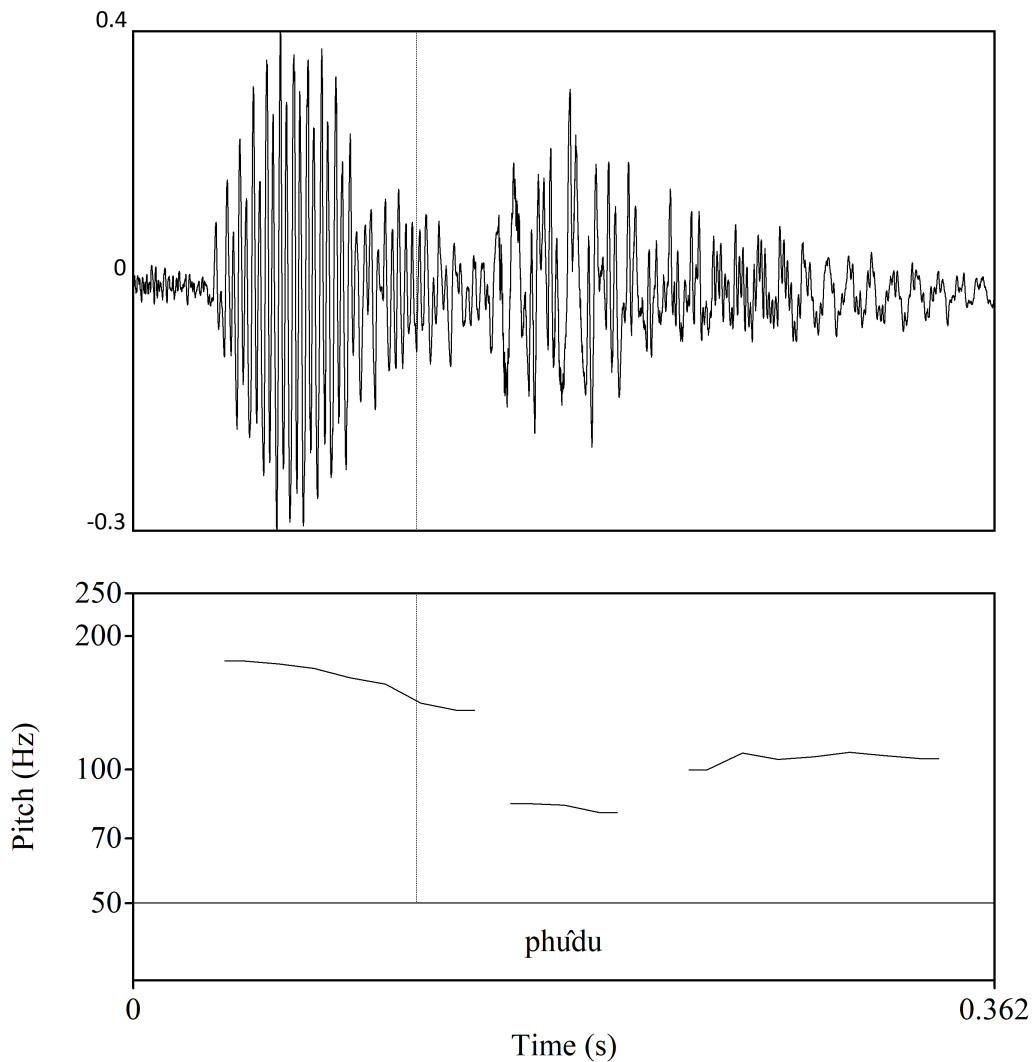
tone/stress on initial syllable (termed <i>trochaic foot</i>)		tone/stress on second syllable (termed <i>iambic foot</i>)	
HF-M	<i>mâlai</i> ‘scatter’	M-HF	<i>malâ</i> ‘search’
		M-H	<i>malû</i> ‘sickness’
		M-M	<i>malai</i> ‘spring’
		M-L	<i>malà</i> ‘tip, edge’

2.5.5.1 High-falling + mid (HF-M)

The pattern High-falling + mid consists of a HF tone on the initial syllable and a toneless second syllable (here marked as M). This is the only trochaic pattern, i.e. where the initial syllable is stressed (tone, therefore, is not contrastive on an initial syllable in a disyllabic word).

Image 2.34 illustrates the pattern of the example of *phûdu* ‘eggplant’ as uttered by AJ. The dotted line represents the syllable break. While the initial syllable shows the contour characteristic of HF (cf. §2.5.3.1) the second, toneless syllable continues the descending pitch of the initial syllable. The amplitude of the waveform also shows a slight decrease, which is consistent with stress falling on the initial syllable and the general prominence of HF as described in §2.5.3.1.

Figure 2.34: Waveform and pitch of *phûdu* ‘eggplant’ by AJ

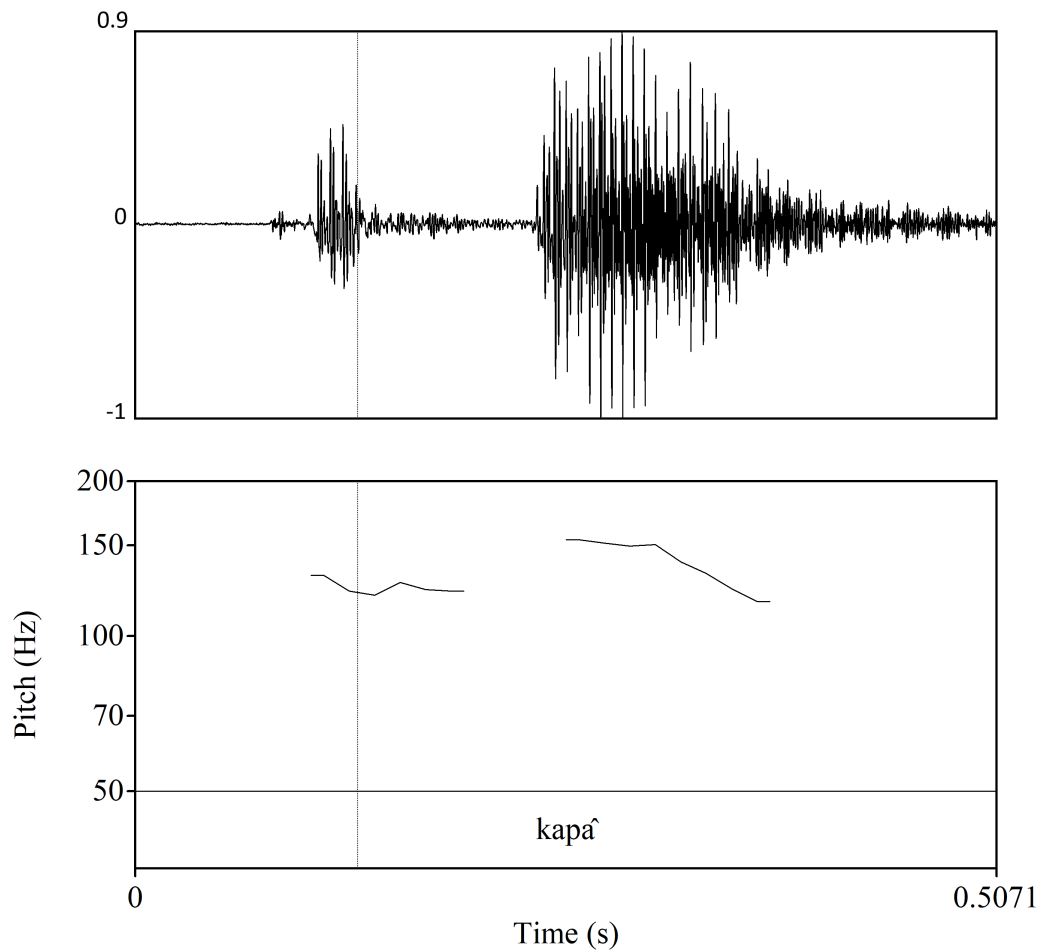


2.5.5.2 Mid + high-falling (M-HF)

The pattern M-HF of disyllabic words has the stress (and tone) on the second syllable (iambic foot). The initial, toneless syllable commonly describes a descending pitch that anticipates the HF contour of the second syllable. My impression is that this is by far the most common pattern in the language although I have not yet made a calculation.

Image 2.35 illustrates the M-HF pattern based on the example *kapâ* ‘alone’ as uttered by AJ. The dotted line represents the syllable break. The pitch of the initial syllable goes slightly down before the HF curve of the second syllable.

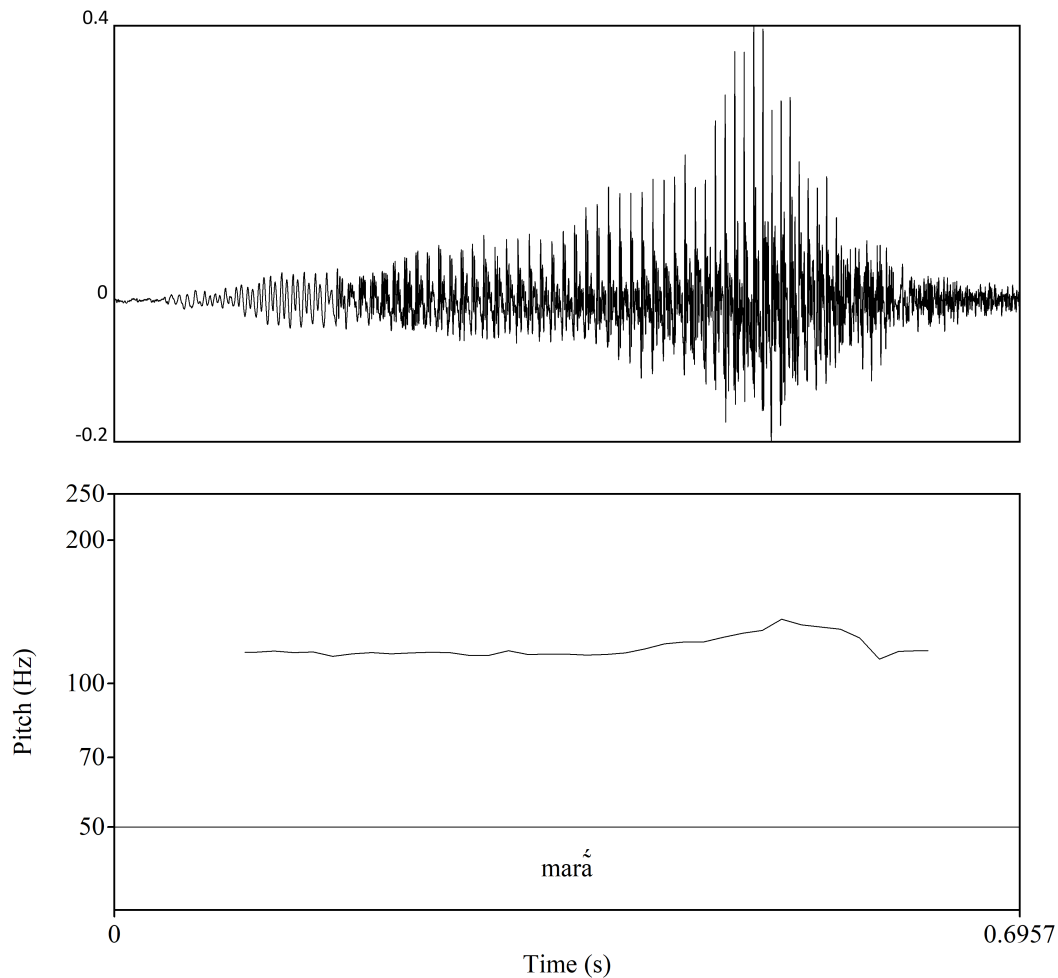
Figure 2.35: Waveform and pitch of *kapâ* ‘alone’ by AJ



2.5.5.3 Mid + high (MH)

The MH pattern belongs to the group where the stress falls on the second syllable (i.e. iambic foot). A toneless syllable is followed by a high tone. This is illustrated in Image 2.36 as uttered by BT. The initial (toneless) syllable shows a M pitch and the waveform shows a low amplitude while the second syllable has a rising pitch and the waveform a bigger amplitude.

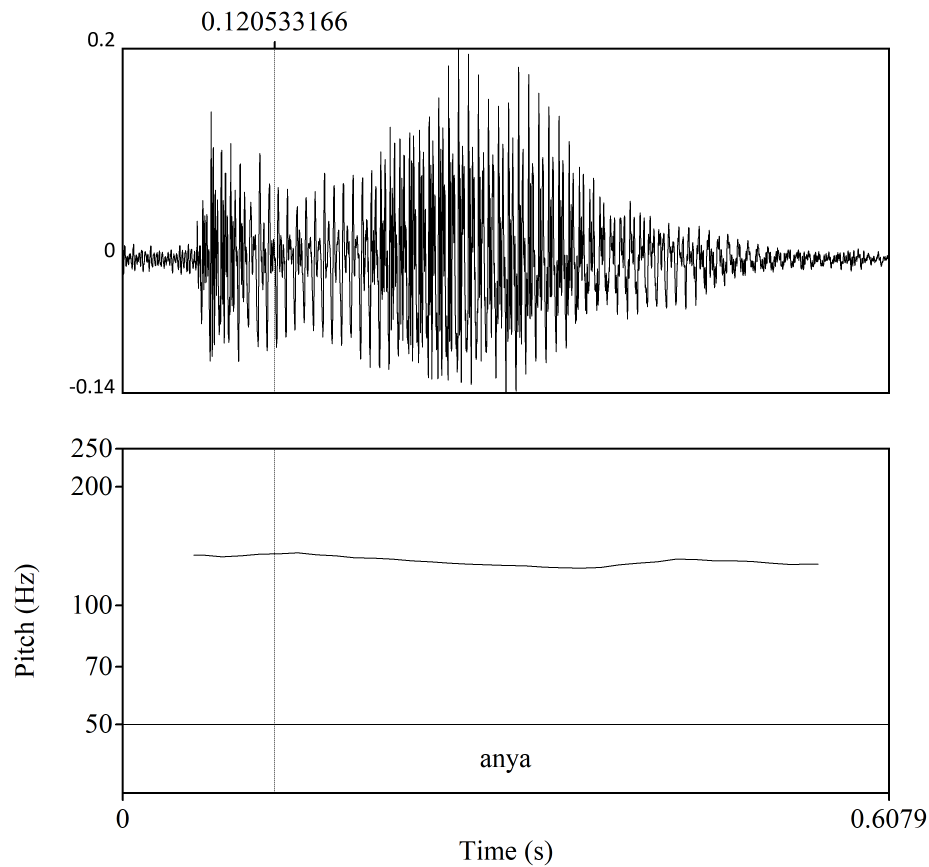
Figure 2.36: Waveform and pitch of *mará* ‘abandon’ by BT



2.5.5.4 Mid + mid (MM)

The MM pattern belongs to the group where the stress falls on the second syllable (iambic foot). It consists of a toneless syllable plus a syllable with a mid tone. Image 2.37 illustrates the waveform and pitch of the noun *anya* ‘face’ uttered by AJ. The dotted line stands for the syllable break. The waveform of the initial, toneless syllable has a shorter amplitude and it also has a shorter pitch (vowel length is discussed in §2.3). The second, full syllable has a markedly stronger amplitude and a mid pitch that finishes in a slight descent as a result of downdrift.

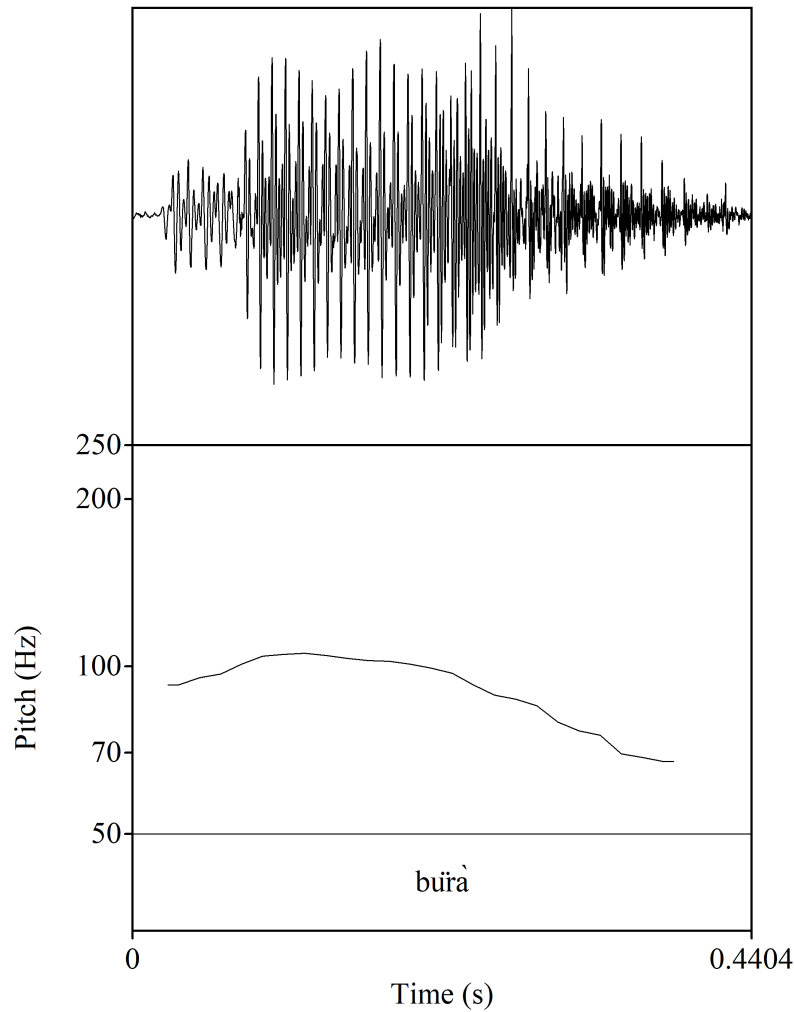
Figure 2.37: Waveform and pitch of *anya* ‘face’ by AJ



2.5.5.5 Mid + low (ML)

The ML pattern belongs to the group where the stress falls on the second syllable (iambic foot). A toneless syllable is followed by a syllable with a low tone. Image 2.38 illustrates the waveform and pitch of the noun *būrà* ‘thunder’ uttered by JO. The dotted line stands for the syllable break. The first, unstressed syllable has a lower amplitude in the waveform along with a mid pitch. The stressed syllable has a higher amplitude and the pitch describes a marked drop downwards. The low tone is often accompanied by a creaky voice (as mentioned in §2.5.3.3 on LT).

Figure 2.38: Waveform and pitch of *būrà* ‘thunder’ by JO



2.5.6 Tonal patterns: trisyllabic morphemes

Native trisyllabic words are exceedingly rare in the language and most attested cases seem to go back to compounds, although the constituent segments are not always transparent. In this sense, a discussion of tone in trisyllabic morphemes must clearly state that there are no stress-based minimal pairs (as there are with disyllabic morphemes) and that the discussed patterns are but an observation based on few existing attestations. It cannot be ruled out that further historical work will reanalyze all patterns as compounds of disyllabic words and toneless syllables.

As noted by Evans and Manyu (2021, 19), trisyllabic words have reduced tonal pat-

terns that are consistent across, albeit few, lexemes. Table 2.59 illustrates all three patterns found in the language. The remainder of this section will explore these patterns individually.

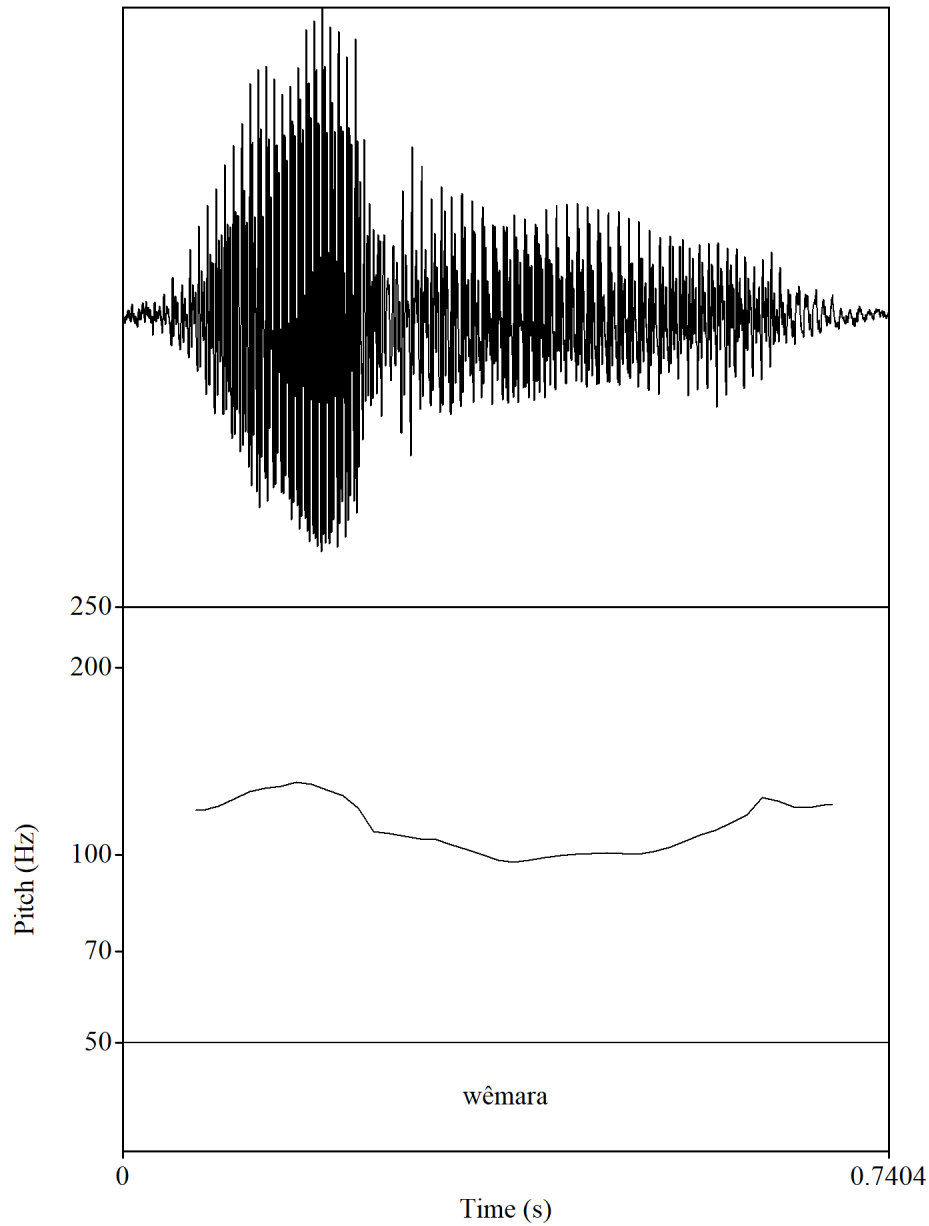
Table 2.59: Tonal patterns in trisyllabic words

Tone pattern	Stress pattern	Example	Reference
high-falling + mid + mid	initial syllable	<i>wêmarā</i> ‘forgive’	§2.5.6.1
mid + high-falling + mid	second syllable	<i>tayîngbum</i> ‘type of bee’	§2.5.6.2
mid + high + mid	second syllable	<i>taméqĩng</i> ‘elephant’	§2.5.6.3

2.5.6.1 High-falling + mid + mid

The high-falling + mid + mid pattern has a high-falling tone on the initial syllable, and thus resembles the trochaic pattern (HF + M), described in §2.5.5.1. It is also the only pattern with a non-mid tone on the initial syllable. Figure 2.39 illustrates the waveform and pitch of *wêmarā* ‘forgive’ as uttered by JO.

Figure 2.39: Waveform and pitch of *wêmara* ‘forgive’ by JO

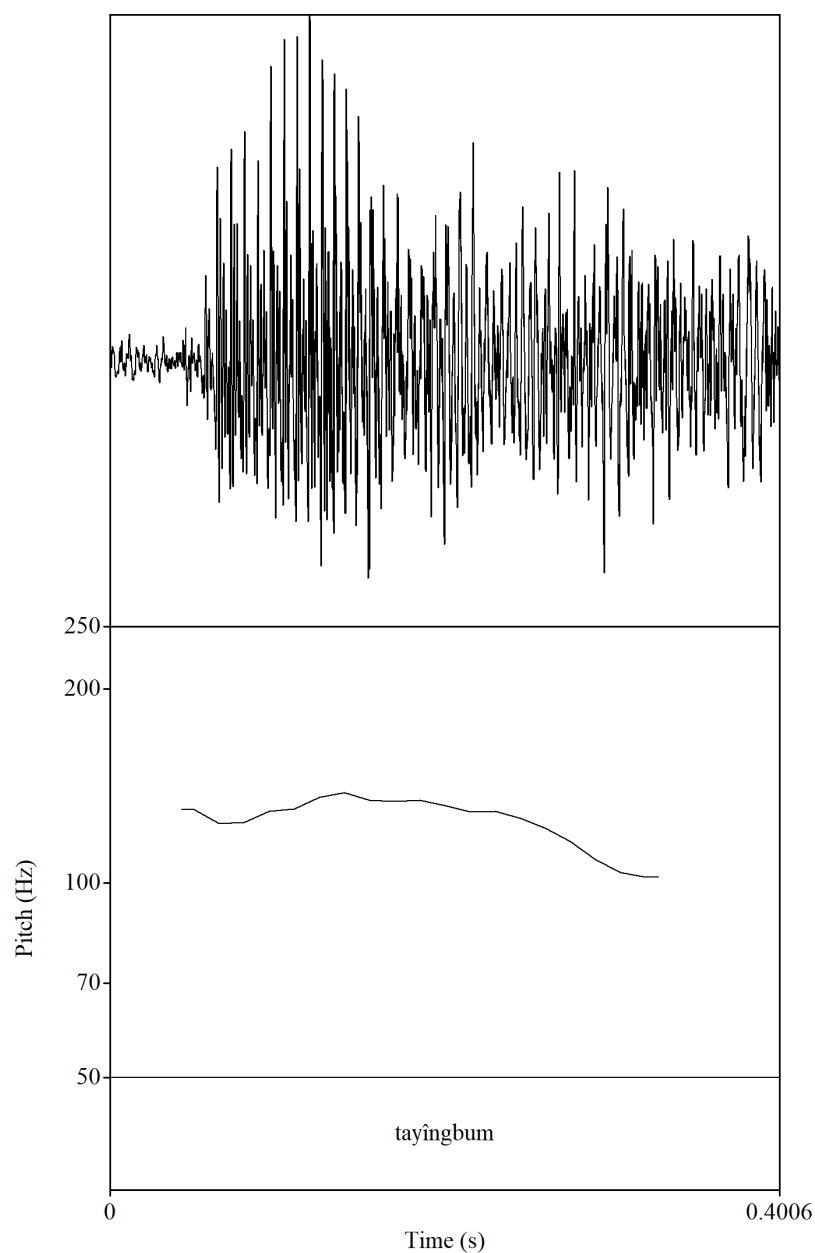


2.5.6.2 Mid + high-falling + mid

The mid + high-falling + mid pattern is dominated by the high-falling tone on the second syllable. Both the first and the last syllable are usually low, anticipating and continuing the change in pitch of the second syllable. This is graphically displayed in Figure 2.40, which depicts the waveform and pitch of the noun *tayîngbum* ‘type of bee’ as uttered

by AC.

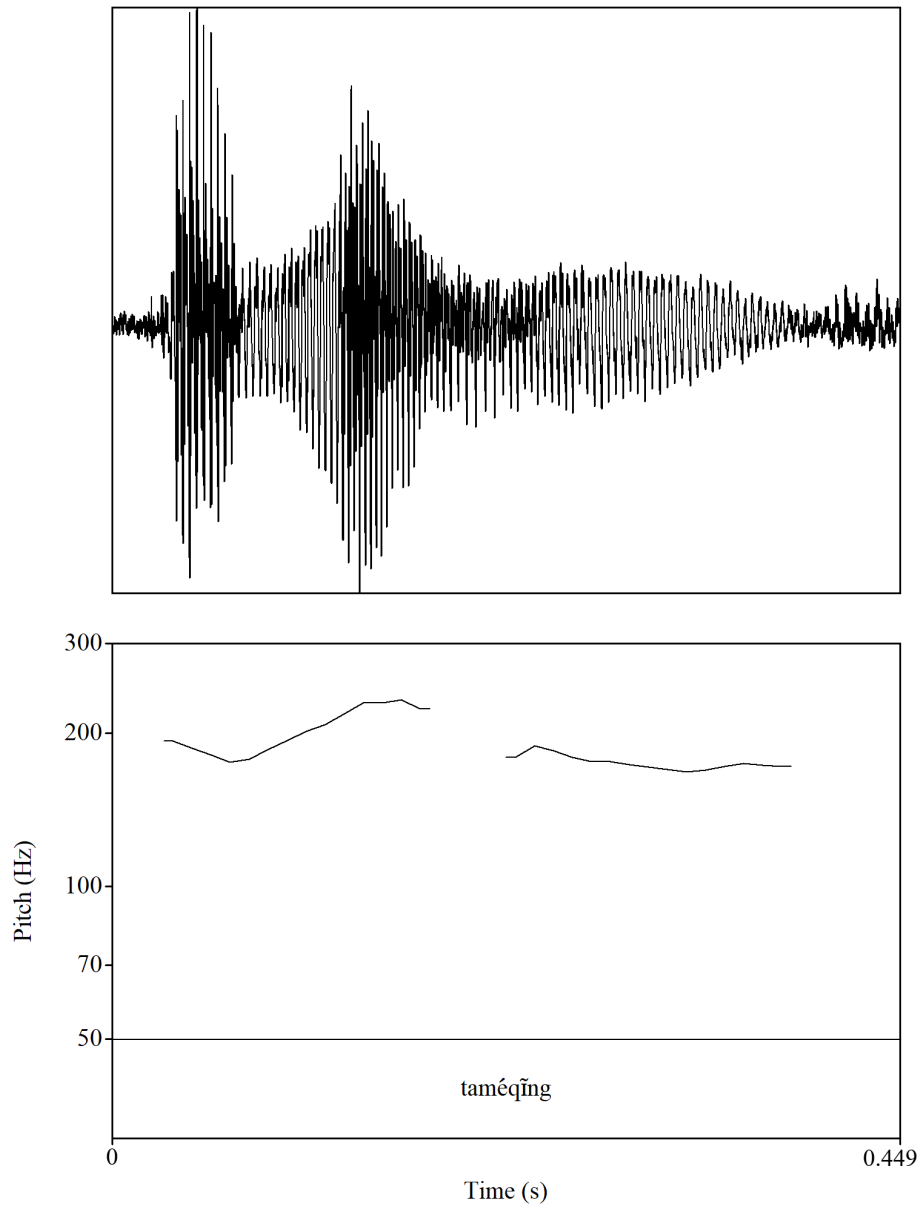
Figure 2.40: Waveform and pitch of *tayĩngbum* ‘type of bee’ by AC



2.5.6.3 Mid + high + mid

The mid + high + mid pattern has a high tone on the second syllable. This is illustrated in figure 2.41, which depicts the waveform and pitch of *taméqĩng* ‘elephant’ as uttered by DI.

Figure 2.41: Waveform and pitch of *taméqĩng* ‘elephant’ by DI



2.5.7 Tonal change in nominal modification

There is a morpho-phonological change that affects syllables with low (L) and mid (M) tone turning them to a high (H). This change is triggered by nominal compounding of the type N-N (noun-noun) and reduplication of nouns, affecting both monosyllables and heavy syllables in disyllabic words (only iambic). Image 2.42 displays the monosyllabic noun *phlǎ̀* ‘stone’ with a low tone in the left panel. When this noun modifies another

noun its tone becomes high, as in the right panel, which shows its *modifier form* (note that *a* is a noun meaning ‘offspring, child, son’). Similarly, in Figure 2.43 the noun *phlǎ̀* ‘stone’ is modified by a postposition *kya* that has some nominal properties (cf. §9.5.3) and triggers the *modifier form*.

Figure 2.42: Pitch of *phlǎ̀* ‘stone’ and *phlǎ́ a* ‘small stone’ by AJ

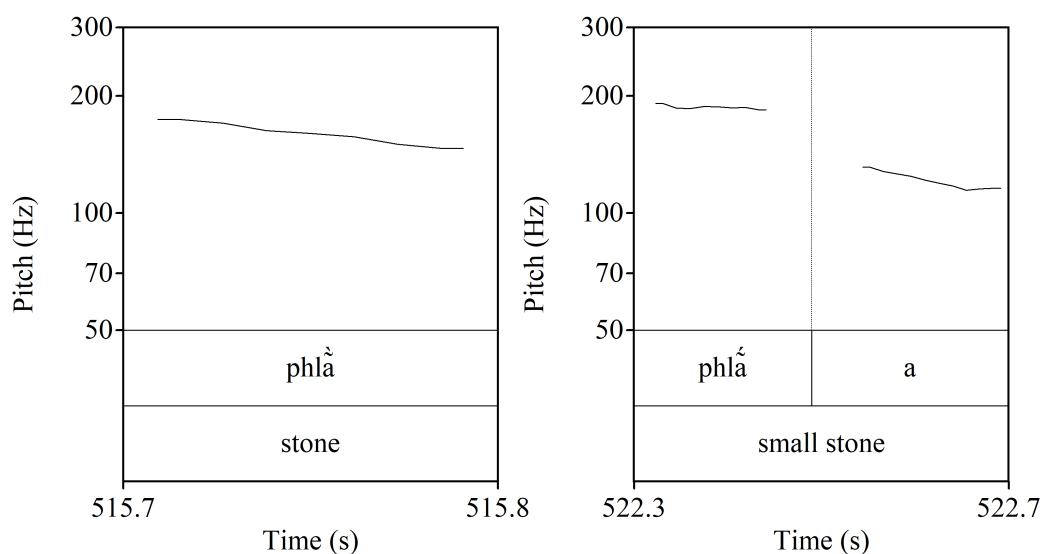
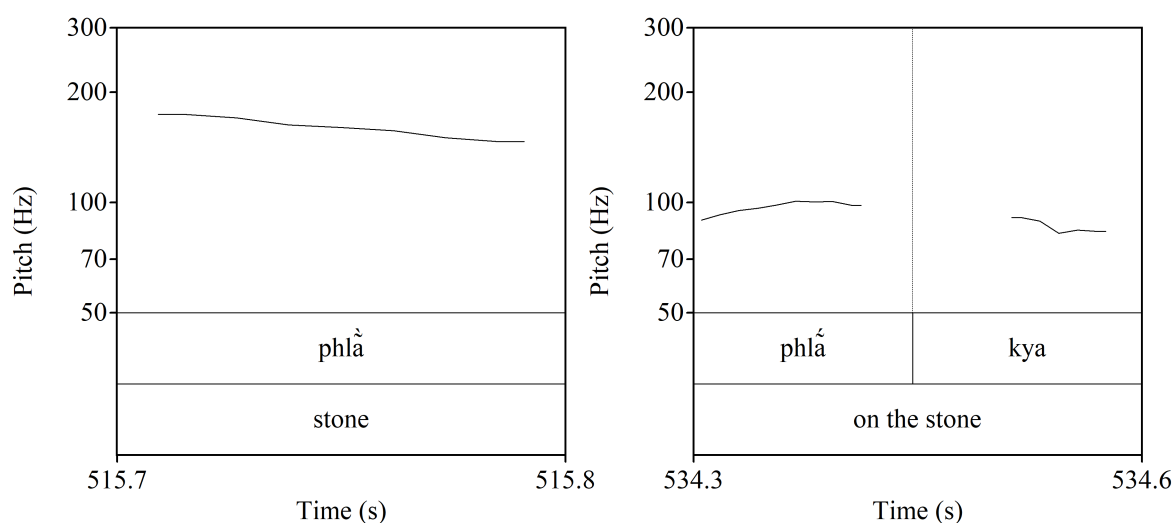


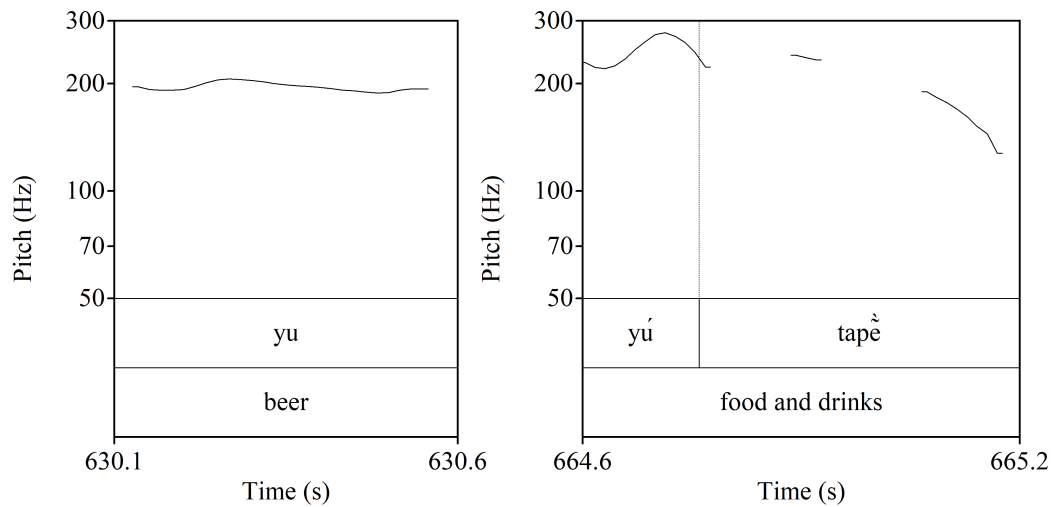
Figure 2.43: Pitch of *phlǎ̀* ‘stone’ and *phlǎ́ kya* ‘on the stone’ by JO



This change affects both hierarchical and symmetrical nominal compounds. In Figure

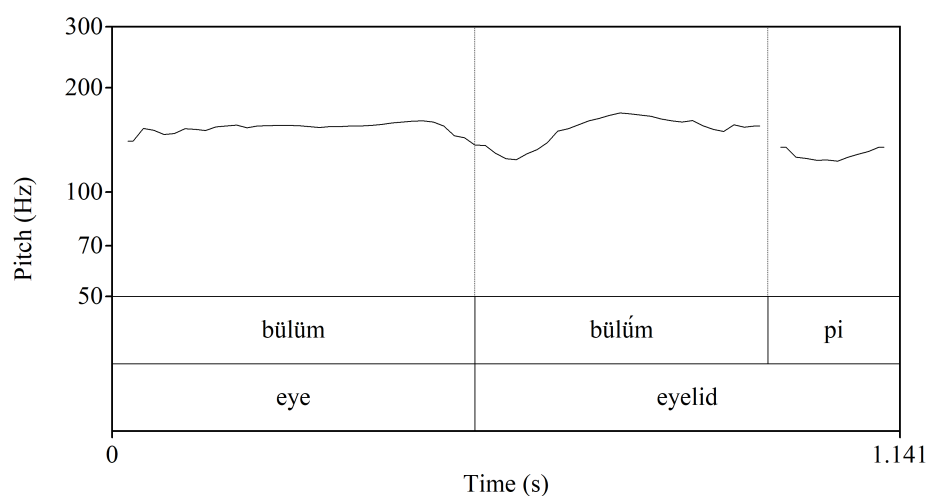
2.44 the noun *yu* ‘beer’ occurs in a symmetrical compound (cf. §4.2), which triggers the tonal change.

Figure 2.44: Pitch of *yu* ‘beer’ and *yú tapè* (beer food) ‘food and drinks’ by DI



In the case of disyllabic nouns, this tonal change only affects the heavy syllable of an iambic word while trochaic words (and light syllables in general) are not affected (§2.2). In Figure 2.45, the polysyllabic MM noun *bülüm* ‘eye’ occurs on the left panel. Its modifier form occurs on the right panel, where it modifies *pi* ‘eyelid’.

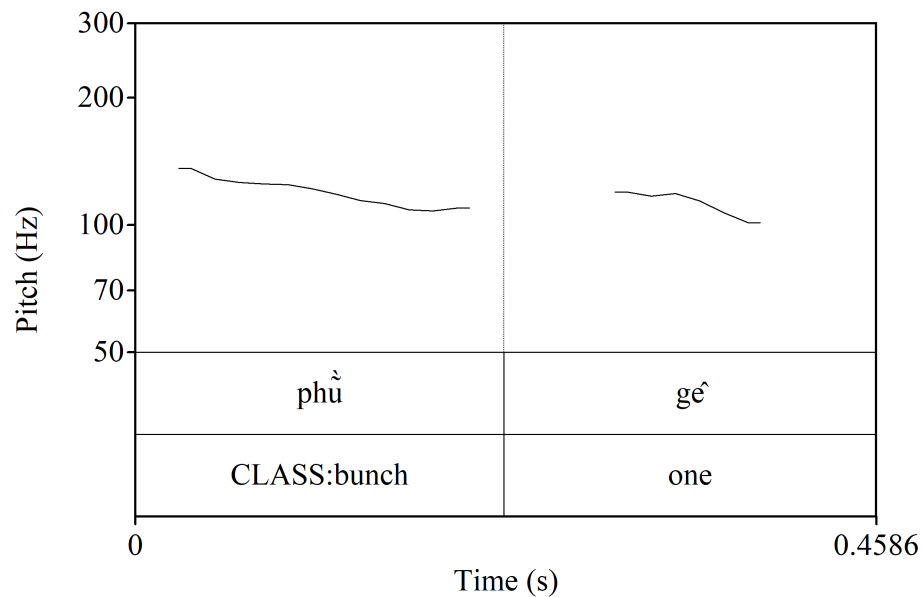
Figure 2.45: Pitch of *bülüm* ‘eye’ and *bülüm pi* by AJ



This tonal change does not affect other forms of modification (i.e. non-N-N). Fig-

ure 2.46 illustrates the pitch of the classifier plus numeral *phũ gê* (CLF:bunch one) ‘one bunch’. Here the ML word keeps its surface pitch for there is no nominal modification.

Figure 2.46: Pitch of *phũ gê* ‘one bunch’ by JA



Reduplication of nouns also triggers this tonal change (cf. §3.6). In Figure 2.47 the noun *mlă* ‘village, place’ is repeated, which renders a distributive sense ‘in every village’. Other POS participate only marginally in this morphological process and it has only been found in lexicalized forms. In Image 2.48, the adjective/adverb *ră* ‘late’ is reduplicated (without a change in meaning, cf. §3.6.3).

Figure 2.47: Pitch of *mlǎ̀* ‘village, place’ and the reduplicated form *mlǎ́ mlǎ̀* by JA

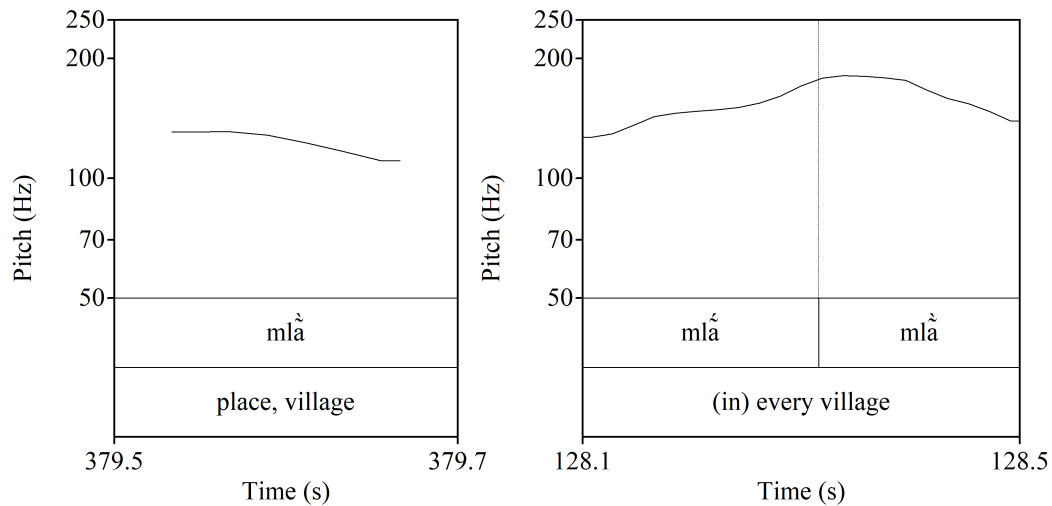
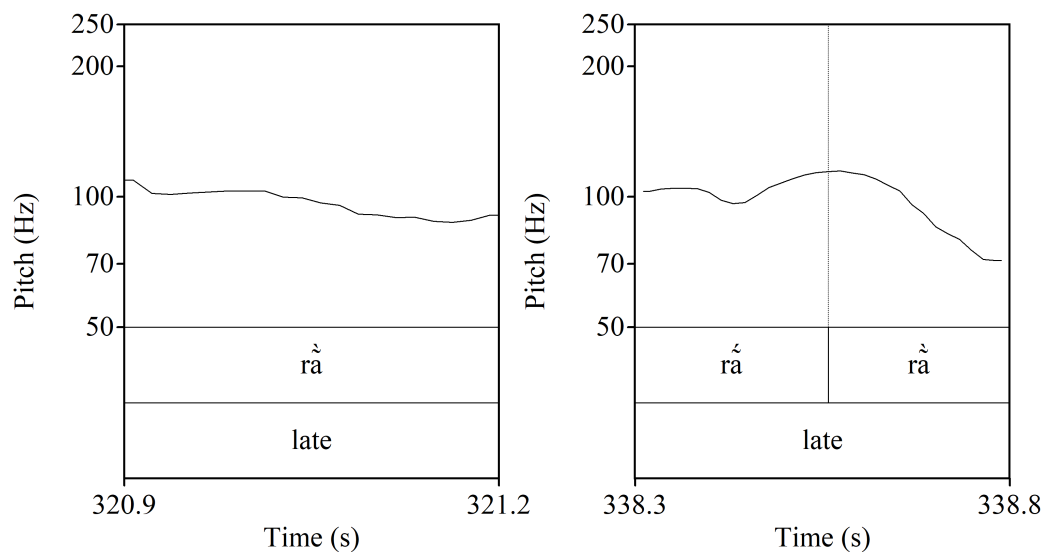


Figure 2.48: Pitch of *rǎ̀* ‘late’ and its reduplicated form *rǎ́ rǎ̀* by JO



2.6 Morpho-phonology

Tawrǎ has a low amount of morpho-phonological processes. The most relevant process is the tonal alternation that occurs in nominal modification and reduplication (thoroughly described in §2.5.7). Cases of assimilation are discussed alongside the phoneme. Similarly, allomorphological alterations are discussed under the pertinent morpheme.

3 Morphology and Lexicon

This chapter encompasses four broad topics: basic morphological units, part of speech (henceforth: POS), reduplication and root morphology: §3.1 will address basic morphological units such as *affixes* and *lexemes*. The discussion on POS is thus presented: the main word classes are discussed in §3.2. Their subgroups are discussed in §3.3. Other open classes, i.e. adverbs and classifiers, are discussed in §3.4. Closed classes (numerals, personal pronouns and demonstratives) are discussed in §3.5. Reduplication is discussed in §3.6. Finally, diachronic aspects of root morphology (i.e. old prefixes) are discussed in §3.7. It should be noted that the noun phrase is discussed in detail in §4 and it will be addressed only tangentially in this section; relevant aspects thereof are cross-referenced.

3.1 Basic morphological units

A basic morphological distinction in Tawrã is that between bound and unbound elements. We can thus distinguish between unbound or free **words** on the one hand and bound **roots** and **affixes** on the other. A **word** in Tawrã is morphologically free, i.e. it can occur independently without derivation or any ancillary morphology and can head phrases. Examples of lexical words are nouns like *tabâw* ‘snake’ or verbs like *shyüw* ‘breath’. Conversely, **roots** and **affixes** cannot be uttered independently and cannot occur in a grammatical phrase without morphological modification (cf. Post 2017, 45): A **root** is a morphologically bound element that heads a word (Booij 2005, 28); it has an independent meaning but it can only occur bound to another morpheme. There is little productive root morphology in the language. As an example, in the numeral ‘two’ *kaqîng*, the root *îng* surfaces alongside classifiers, e.g. *bra îng* (CLF:generic two) ‘two of anything’ (cf. §5.1), with some personal pronouns, e.g. *cya-îng* ‘3DU’ and in the inde-

pendent numeral ‘two’ *kaqîng*. Similarly, an **affix** in Tawrã is morphologically bound, i.e. it cannot occur without a host. Affixes vary in their integration with the host word. As an example, the prefix *ka-* is toneless and is integrated into the phonological word of its host, e.g. in the numeral ‘two’ *kaqîng* (*ka-îng*). Conversely, non-fusing suffixes are distinguished from nouns (a type of *word*) in that they do not trigger the tonal change associated with nominal compounding and modification (termed *modifier form*, cf. §4.2 and §2.5.7). Thus, in Example 1a the noun *mě̀* ‘person’ keeps its underlying tone while in Example 1b the same noun occurs in a modifier form (high tone as part of a nominal compound).

- (1) a. *mě̀-yã*
 person-FEM
 ‘a female person’
- b. *mě́* *tako*
 person.MOD book
 ‘(The) person’s book’

Tawrã is an overwhelmingly suffixing language. The only two prefixes are a 1.SG possessive *na-* found on some close kinship terms (§4.5.3) and *ka-*, found in numerals (although it is fossilized in a wide array of lexemes, cf. §3.7.3). Diachronically, however, the language seems to have had a developed system of nominal and verbal prefixes (reminiscent of Tibeto-Burman languages of the area, cf. Lieberherr 2017, 90; also cf. Post 2006, 45f) that were bleached to the roots where they occurred and today display little syntactic and semantic independence (diachronic aspects of root morphology are treated in §3.7). There are no traces of prefixing verbal agreement morphemes in the language¹. Most affixes in the language are sensitive to POS and can either exclusively occur with nouns/adjectives or exclusively with verbs. The suffixes *-mâ* and *-ho* are an interesting exception as they occur with both verbs and adjectives (cf. §3.2.4). Affixes also

¹Most studies on the reconstruction of the agreement system of the proto-language suggest that ancient prefixes (found today in some languages) were leveled and completely replaced by suffixation at an early stage of the proto-language (cf. DeLancey 2011). I assume that this process also affected Proto-Tawrã-Kera’a.

vary in their level of grammatical and semantic abstraction: while the possessive prefix *na-* ‘1.SG’ (§4.5.3) and the female referent suffix *-yã* (§4.5.2) have both a clearly defined grammatical meaning, the prefix *ka-* does not have semantic or grammatical value and its only function is to allow the root of the numeral ‘two’ *-îng* to occur independently.

A further axis by which to distinguish *words* in a language is by their lexical semantics. In this sense, we may organize the lexicon of a language in *lexical* vs. *non-lexical* vocabulary. The former, also termed *content words*, refers to words that “code stable, culturally-shared concepts” and experiences and can head syntactic phrases (Givón 2001a, 44f.). Members of this category correspond to the open lexical classes of adverbs, nouns, verbs, etc. which will be defined in this section. Two examples are *mê* ‘person’ and *malai* ‘be scattered’. Conversely, the latter group, also termed *function words*, are grammatical and derivational morphemes that modify content words, some examples are pronouns, verbal and nominal suffixes. Table 3.1 summarizes the main characteristics of the basic morphological units **word**, **root** and **affix**. As discussed above, most affixes have a grammatical meaning (termed *function words*) while a small set of inherited prefixes lack a clear grammatical meaning but still have a syntactic function in licensing roots to occur in a construction (termed here *null*).

Table 3.1: Overview of basic morphological units

	word	root	affix
morphemic status	free	bound	bound
lexical semantics	content word	content word	function word/null

3.2 Overview of major open classes

Word classes or part of speech (POS) refer to the main classes of words that can be distinguished grammatically in a language (Schachter and Shopen 2007, 1). These distinct word classes constitute the main building blocks of grammar (cf. LaPolla 2016, 372). Although it is widely held that the diagnosis of POS is language-specific, many scholars agree that there are typological tendencies for basic, major lexical word classes that include *nouns*, *verbs*, *adjectives* and *adverbs* based on semantic and syntactic criteria

(Dixon 2004; Givón 2001a, 49). The study of word classes lies in understanding the limits of linguistic variation, how and to what degree can we speak e.g. of nouns and verbs as distinct classes and what role they play in structuring the grammar of a language (Foley 2023, 109).

The distinction between word classes in Tawrã hinges on a combination of syntactic, morphological and semantic considerations. Formal differences in the behavior of lexical categories are often used as a strong argument for the distinction of word classes, although a purely morphosyntactic approach would result in a high number of word classes, reflecting all possible morphological behaviors (cf. Sasse 2015, 165; also cf. Croft 2023, 219). As discussed by Sasse (2015, 170), the organization of a language's vocabulary in word classes should take notice of both the morphosyntactic dimension (constructional slots), the lexical dimension (lexico-semantic categories) and the emerging interactional pattern. Therefore, a description of POS may reflect at best but a temporary condition, for these are the temporary results of ongoing grammaticalization and lexicalization processes borne out of the said interaction between lexicon and grammar (idem. 173).

In the case of Tawrã, a central criterion to distinguish between nouns and verbs is their different suffixes and derivation: while verbs have a wealth of verbal suffixes (termed TAMN suffixes, cf. §6.3), as well as derivational morphemes (§6.2 and §10.5.4), nouns have a reduced set of affixes (cf. §3.3.1). Adjectives, as will be discussed in §3.3.3, constitute an in-between class: on the one hand they share with verbs the ability to take the stative and iterative suffixes, while, on the other hand, they can head a NP and take all NP modifiers like nouns do (§4.3.1). However, three constructions are specific to adjectives (discussed in §3.3.3). Table 3.2 presents an overview of the main arguments for defining word classes. An 'X' indicates the features that occur for each POS. All three POS may be modified by the postposition/particle *gũ* and all three POS share the ability of occurring unbound in a clause. The negator *yüm* (§6.4.1) can occur with all three POS, although for nouns and adjectives it occurs on a verbless clause (VLC, cf. §7.1). TAMN suffixes are associated with verbs, although two suffixes (stative suffix *-ho* and iterative suffix *-mâ*) also occur with adjectives. The ability to head a NP without derivation is shared between nouns and adjectives. Only adjectives can be compared and intensified. Finally, only nouns can undergo a tonal change (termed *modifier form*

in N-N nominal modification. The remainder of this chapter will put forward the main arguments for three word classes. §3.2.1 will begin by introducing a discussion on cross-class membership.

Table 3.2: Main criteria for word class

	N	V	Adj
Modified by <i>gũ</i>	Yes (postposition)	Yes (non-final particle)	Yes (manner)
Boundness	free	free	free
<i>yüm</i> NEG	only in a VLC	Yes	only in a VLC
TAMN suffixes		Yes	only two suffixes
May head a NP uninflected	Yes		Yes
May take an intensifier, comparative suffix			Yes
Tonal change	only in N-N modification		

3.2.1 Cross-class membership

This section will discuss multiple or cross-class membership as a relevant feature of word classes in the language. Note that nominalization is discussed in §10.5.

A feature found in Tawrã that is also widespread in related languages (Morey 2012, 227-233; and elsewhere cf. Booij 2005, 57; Dixon 1982, 11) is multiple class membership of many lexical roots, i.e. many words occur underived as nouns, verbs and/or adjectives (*zero derivation*, cf. Payne 1997, 224; cf. Manova and Dressler 2005, 71; also termed *ambivalents*, cf. Shimelman 2017, 31). One example of a root occurring in three different POS is *jüm*: a verb ‘be real’, an adjective ‘real’ and a noun ‘truth, direction (only in the expression ‘in all directions’)’. These three meanings are displayed in Table 3.3.

Table 3.3: Different POS of *jüm*

POS	Meaning
V	be real
Adj	real
N	truth, direction (only in ‘eight directions’)

Although such cases of cross-class membership may seem to blur a clear-cut distinction between POS, it must be pointed out that it is not a productive feature of the language as most lexemes exist in only one POS and the distribution of multiple POS membership in the lexicon does not seem to be random (cf. Deutscher 2009, 203f. cf. Bisang 2023, 594f). As recognized by different scholars, certain semantic fields are prone to double categorization, e.g. in the field of human attributes and social roles (cf. Sasse 2015, 161, 185; Post 2007, 207), cf. in Tawrä the cases of *pô* both ‘slave’/‘be a slave’ and *muhwa* both ‘rich’/‘be rich’. In Example 2, the lexeme *pô* is a verb that heads a predicate clause. Conversely, in Example 3 *pô* is a noun that heads a coordinated NP.

(2) *pô gũ matsâp ì yüm*

[*pô* *gũ*]_{cosub.cl} *matsâp ì* *yüm*]_{main.cl}

be.a.slave NF cow be.there NEG

‘Being a slave (*because s/he is a slave*), s/he has no cows.’ [field notes 20102024]

(3) *bá.mě daráng byě bá.mě aqa pô mu kwâg mu bâ la*

[*bá.mě daráng byě*]_{NP} [*bá.mě aqa*]_{NP} [*pô*]_{NP} *mu* [*kwâg*]_{NP} *mu*]_{co-ord NP} *bâ*

clan big TOP clan small slave also dog also make

la]_{main.cl}

HSAY

‘As of big clans, they would enslave small clans (lit. do like slave like dog), it is said.’ [TA22_clan 12 56]

Again, such cases of multiple memberships are lexicalized and other potential members from the semantic group of human and social attributes do not feature double class membership, cf. *hazá* ‘king, officer’ but not *‘be a king, be an officer’. In other words, such cases of cross-class membership are semantically idiosyncratic and not productive as to be considered a form of zero-derivation (cf. Enfield 2006, 315).

Nouns and verbs also exhibit cases of cross-class membership. This is illustrated in Examples 4a and 4b: to wear something and the thing that is worn are homophonous (i.e. argument and predicate). As discussed by Morey (2012, 231), the categorization of lexical roots with multiple class membership as either primarily nouns or primarily verbs often relies on approximate frequency (i.e. the lexeme occurs more often as a N or V) or on semantics, as actions are strongly associated with verbs. However, one must admit that these metrics provide but an educated guess.

(4) a. *tîng tîng-na*

shirt wear.shirt-IMP

‘Wear (your) shirt!’

b. *thaw thaw-na*

skirt wear.skirt-IMP

‘Wear (your) skirt!’ [both 140622_DI_eli1 1022]

One field of the lexicon where we can safely establish a direction in class membership extension is classifiers. As discussed in §5.4, several classifiers are homophonous with nouns. The most accepted direction in grammaticalization is for a noun to be recruited as a classifier (cf. Aikhenvald 2000, 353). As there are no productive morphological devices for this end in the language, classifiers derived from nouns are homophonous with their lexical source², cf. *kahã* both ‘joint, node (in a bamboo)’ and ‘CLF:joints and nodes’.

²Note that several lexemes are composed of a noun and a classifier (termed *classifier compounds*). The evidence suggests that classifiers may feed nominal compounds as regular nominals do cf. §5.5.

3.2.2 Nouns

This short section defines the main parameters of nouns as a distinct word class. Since most features of the noun are discussed elsewhere in this grammar, the relevant sections are always cross-referenced.

Nouns are an open class in Tawrã defined by a series of features: typically, they occur as the head of a NP and define the type of entity involved. NPs have a grammatical role in the clause like Subject or Object (Givón 2001a, 59). The head noun in a NP is commonly omitted if it is retrievable from the context and an adjective or a nominalized verb may formally fill up the rank as an anaphora.

Nouns (and adjectives) can only head a predicate in a verbless clause (cf. §7.1). Nouns can be modified by adjectives, personal pronouns (marking possession) or a further noun (in a nominal compound) (note that this does not apply for personal nouns and, pronouns cf. §4.1.2.1 and §4.1.2.2). Crucially, N-N modification triggers a tonal change on the modifier noun (a *modifier form*), provided that the modifier noun has a mid or low tone as its sole syllable or as its last syllable in a disyllabic word (cf. §2.5.7 and §4.2). This is illustrated in Examples 5a and 5b: the low tone noun *mlà̀* ‘village’ has a *modifier form* when it modifies (in this case, possesses) a further noun (cf. high tone *mlá́*). Nominal compounding exhibits a strict modifier-modified configuration (§4.2).

- (5) a. *mlà̀*
village
‘village’
- b. *mlá́* *mè̀*
village.MOD person
‘village people’

Nouns can be counted and classified by classifiers (§5). Bare classifier constructions have the effect of individualizing nominal referents (except for uncountable nouns like ‘water’) (§5.3). Many nouns in Tawrã display a non-productive root prefix, like *a-* or *ka-* (cf. §3.7).

3.2.3 Verbs

Verbs can head a predicate but cannot occur in a verbless clause (cf. §7.1). Verbs are distinguished from other POS by a wealth of TAMN suffixes and by the ability to be expanded by predicate derivations (cf. §6.2). In Example 6a, a verb takes the allophoric suffix *-ya*, which is neither available for nouns (Example 6b) nor adjectives (Example 6c). Note that adjectives can take the stative suffix *-ho* STAT (§6.3.4.5), as will be discussed in §3.2.4.

(6) a. *thâ-ya*

eat-ALLO

‘S/he eats.’

b. **zô-ya*

embroidery-ALLO

Intended: ‘S/he embroiders.’

c. **pra-ya*

good-ALLO

Intended: ‘S/he is good.’

Regarding the general negator *yüm*, it occurs in both predicative and verbless clauses. On verbless clauses, it negates the identification or attribution of a NP; what in other languages would be to negate the copula³. Verbs can also take one or more derivational morphemes termed *predicate derivations* (cf. §6.2), including causative and other valency-changing devices. Verbs can also be nominalized, as discussed in §10.5.

Verbs occur at the end of the clause in pragmatically neutral contexts (cf. §8.2.3) but before modal particles - this of course does not apply to verbless clauses. Verbs can be nominalized (§10.5) and modify nouns or occur anaphorically if the nominal head of the NP is ellipsed. There is thus a functional overlap with adjectives in modifying nominals. Finally, only verbs can have valency, the number of actants in a clause (cf. §3.3.2.1).

³Cf. Nepali *cha* (be.there.3.S) vs. *chaina* (be.there.3.S.NEG).

3.2.4 Adjectives

The prototypical function of adjectives is that of attribution and specification (Givón 2001a, 84f; Dixon 2004, 10f; cf. Sasse 2015, 183). It is argued here that Tawrã has a large class of adjectives. This is unusual, as many neighboring Tibeto-Burman (henceforth TB) languages are argued to lack a dedicated, non-derived class of adjectives (cf. Coupe 2008, 80; also cf. Modi 2017, 254f.). Furthermore, related languages, where a distinct class of adjectives has been identified, often have few (native) members (for Limbu, cf. van Driem 1987, 21; Kham, cf. Watters 2009, 111; also cf. Dixon 1982, 5f).

The class of adjectives in Tawrã shares properties with both nouns and verbs. Adjectives can be distinguished from verbs in that only the latter can be marked by the full range of aspect, mood and all other categories that pertain to the verbal domain. Adjectives only share with verbs the ability to take the stative suffix *-ho* STAT (§6.3.4.5) and iterative suffix *-mâ* ITER (§6.3.4.4), although with a slightly different meaning, as will be exemplified presently. In Example 7a, the adjective *muhwa* ‘rich’ occurs in a verbless clause modifying a pronoun (as a noun would, cf. §7.1). In Example 7b, the adjective operates as a verb, as it has a stative suffix that confers it a resultative interpretation. Finally, Example 7c demonstrates that adjectives can at best be considered defective verbs, as they cannot take any other TAMN suffix. Any verbal suffix (save for iterative *-mâ* and stative *-ho*) on an adjective is judged ungrammatical.

(7) a. *cya muhwa*

3 rich

‘S/he is rich.’

b. *cya muhwa-ho*

3 rich-STAT

‘S/he became rich’ (and not *‘S/he is rich’).

c. **hã muhwa-de*

1S rich-HAB.EGO

Intended: ‘I am rich’.

Many stative verbs have a descriptive meaning, giving way to functional overlap between these and adjectives (a typologically common situation, cf. Konnerth (2014, 119f)). The semantic distinction between stative verbs and adjectives will be discussed again in §3.3.3. However, stative verbs are distinct from adjectives in syntactic terms since they have the whole palette of verbal suffixes available. Furthermore, the two verbal suffixes available for adjectives have a slightly different meaning, as mentioned above: the stative suffix *-ho* STAT produces a *stative* reading on verbs, while it triggers a resultative reading when attached to an adjective. The former case is illustrated in Example 8a and the latter in Example 8b (repeated from above). A resultative reading for verbs is commonly achieved with the perfective suffix *-bô* (cf. §6.3.4.8), as in Example 8c. Finally, as mentioned above, nouns cannot take the stative suffix or any other verbal suffix, cf. Example 8d.

(8) a. *malai-ho*

be.scattered-STAT

‘It is scattered.’

b. *cya muhwa-ho*

3 rich-STAT

‘S/he became rich’ (and not *‘S/he is rich’).

c. *malai-bô*

be.scattered-PFV.ALLO

‘It became scattered.’

d. **mê-ho*

person-STAT

Intended: ‘S/he is/became a person/human.’

A similar situation is found with the iterative suffix *-mâ* (§6.3.4.4). When used with a verb, it triggers an iterative or continuous reading, as in Example 9a. Conversely, when used with an adjective, it either has an intensifying effect or is interpreted as resultative, as in Example 9b.

(9) a. *agâw wi-mâ*

bend roam-ITER

‘S/he keeps roaming (lit. bend roam).’

b. *muhwa-mâ*

rich-ITER

‘S/he is quite rich’ or ‘S/he is becoming richer’ [both 011024_STBT_eli2 590]

The distinction between adjectives and nouns is more nuanced due to a) a large set of roots that have multiple class membership (as discussed in §3.2.1), b) a dearth of nominal bound morphology and c) pragmatically-driven word order (cf. §8.2.3). Furthermore, both nouns and adjectives can occur predicatively in a verbless clause (cf. §7.1); nouns and unmodified adjectives normally occur as verbal objects in adjectival-noun languages (cf. Schachter and Shopen 2007, 17f; cf. Yauyos Quechua Shimelman 2017, 56). The independent status of adjectives *vis-à-vis* nouns depends on a small number of constructions that either compare or augment the qualities and attributes that are associated with adjectives as a prime means of nominal modification. These *idiosyncrasies* are: the use of the adverb *tyâng* ‘very’ to intensify the adjective, the comparative device *yũ* and the use of reduplication to increase or intensify the meaning. In Examples 10a and 10b an adjective and a verb are intensified with the adverb *tyâng* ‘very’. This is not available for nouns, as illustrated in Example 10c.

- (10) a. *tyâng dyâ hadi*
 very far story
 ‘very old story (lit. very far away)’ [from TA22_clan 15 66]
- b. *tyâng thâ-mâ*
 very eat-ITER
 ‘He eats too much.’ [290924_STBT_eli1 234]
- c. **tyâng taméqĩng*
 very elephant
 -unacceptable

Furthermore, only adjectives can be compared using *yũ*. In Example 11a the adjective in a NP is compared. A similar clause without an adjective is rendered unacceptable, as in 11b.

- (11) a. *we tako aqa yũ*
 DEM book small/DIM COMP
 ‘This book is smaller (than smth else).’ [290924_STBT_eli1 316]
- b. **mề yũ*
 person COMP
 Intended: more like a person, more human

As will be discussed in §3.6.2, the reduplication of adjectives can be employed for exaggeration, although it is an uncommon strategy in the language. Reduplication is not employed for forming the comparative/superlative (as in Purik, cf. Zemp 2018, 144f). In Example 12, an adjective is reduplicated for exaggeration.

- (12) (...) macêi daráng daráng dũ
macêi daráng daráng dũ
 water big REDUP with
 ‘(...) with much, much water (lit. big big water).’ [TA22_origin_69 202]

Many adjectives also occur as adverbs (§3.4.1) and there is even a subset of adverbs/adjectives that share a CVV phonological shape (cf. §3.4.1). Adjectives and adverbs are distinguished on functional and distributional bases: adjectives modify nouns while adverbs modify all other POS (including adjectives) and may even have scope over whole clauses.

3.3 Semantics and subgroups of open word classes

This section presents an overview of the semantics and sub-grouping of verbs, adjectives and nouns found in the language.

3.3.1 Nouns

Nouns prototypically encode concrete entities thought to be bound to a concrete location and to occur independently of an event and not bound in time (i.e. permanent). This is opposed to verbs, which encode transient relations and events that are bound to (concrete) actors (Langacker 2008, 34f, 103f). Nouns include things, personal names and names of places. Nouns are particularly prone to be borrowed from other languages. While Assamese was an important source of loanwords until the recent past, Arunachali Hindi is catching up fast as an important lexifier.

This section will describe the types of nouns found in the language. These include: *relator nouns* (§3.3.1.1), *kinship terms* (§3.3.1.2) and *class terms* (§3.3.1.3).

3.3.1.1 Relator nouns

In the Asian context, *relator nouns* are *function words* (Givón 2001a, 95f) historically derived from nouns that denote both spatial and temporal relations (cf. DeLancey 1997).

They can be considered the functional *pendants* of adpositions in the Indo-European languages. Depending on the language, they exhibit some characteristics of nouns but often have a reduced distribution, e.g. are bound roots (in Purik, cf. Zemp 2018, 145; in Karbi, cf. Konnerth 2014, 149), or show other signs of grammaticalization, e.g. in Lhasa Tibetan, relator nouns occur without the genitive marker, which suggests that these former nouns are undergoing grammatical decategorization (DeLancey 1997, 59). Relator nouns have been described in many languages of South and Southeast Asia (cf. Jenny et al. 2014, 129; cf. Bradley 2017).

In Tawrã, the term *relator noun* is employed in its semantic sense as it refers to spatial and temporal relations. Syntactically, relator nouns do not behave differently from common nouns. In other words, *relator nouns* are versatile nouns in Tawrã that express spatial and temporal relations. The fact that relator nouns may take a locative NP-modifier further suggests that this is an ongoing grammaticalization process. Example 13 displays the relator noun *bürüm* ‘underpart’ modifying a nominal head: *áng bürüm* ‘the lower part of the house’. Nominal compounding (§4.2) is the standard strategy by which relator nouns in Tawrã modify other nouns. In Example 14, *bürüm* ‘lower part’ stands alone (an ellipsis of *áng bürüm* ‘the lower part of the house’) with a locative postposition. Relator nouns can head a NP without restrictions.

- (13) pyâ áng bürüm ìya
 pyâ áng **bürüm** ì-ya
 bird house.MOD underpart be.there-ALLO
 ‘The bird is under the house.’ [010422_SO_eli2 84]

- (14) bürüm kē ìna
 bürüm kē ì-na
 underpart LOC.INS be.there-FACT
 ‘It is in its underpart (i.e. below the house).’ [220222_Ri_eli1]

Taking the ego as the deictic center, relator nouns can be employed using one’s body as a reference (cf. Langacker 2008, 45, 135). The relator noun *plûng* ‘backside’ may

replace *harũ* ‘back (body part), spine’ in the sense of ‘my backside / my back (body part)’⁴, as in Example 15.

- (15) *hã plũng kẽ malapho*
hã plũng kẽ malap-ho
 1s back.side LOC.INS boil-STAT
 ‘My back hurts.’ [290924_STBT_eli1 389]

Some relator nouns also encode a temporal meaning, as in Example 16, where *mamyú* means ‘after, later’ (instead of its locational meaning ‘behind’).

- (16) *hã mamyú ne nyu thãna*
hã mamyú ne nyu thã-na
 1s after only 2s eat-IMP
 ‘Eat only after me!’ [290924_STBT_eli1 406]

Table 3.4 displays a list of all relator nouns so far found in the language.

Table 3.4: Relator nouns

Tawrã	Meaning
<i>habang</i>	outside (also ‘forest/jungle’) ^a
<i>lũmkũ</i>	inside, inner part
<i>nya</i>	front (also ‘face’)
<i>hela</i>	lower part
<i>bürüm</i>	under part (particularly of a house)
<i>tãzi</i>	upper part
<i>tapê</i>	between
<i>lyâng</i>	place in between
<i>ruzàng</i>	place in between
<i>mamyú / hamyú</i>	later, behind

⁴Typologically, the reverse order of grammaticalization seems to be more common (cf. Kuteva et al. 2019, 64-65).

Notes

^a Tawrã seems to have merged the word for ‘forest’ with that for ‘outside’. The semantic shift ‘outside>forest’ seems to be an innovation of Tawrã, as Kera’a has different words for ‘outside’ *amuko* and *ambẽ* ‘forest’ (field notes 2023). Note that Sun (1991) lists *kaŋ⁵⁵baŋ⁵⁵* for ‘outside’ in Kera’a, something that matches the Tawrã form.

An equivalent situation is found in neighboring Kaman, cf. *kanan*, both ‘forest/jungle’ and ‘inside’ (field notes 2024). Language contact with Kaman could have prompted this semantic change in Tawrã, since it is not found in Kera’a.

As a comparative note, Basque *kanpo* ‘land, exterior’ has undergone a metonymic semantic shift that bears some similarities: *kanpo* has been borrowed from Spanish *campo* ‘land, countryside’ (it does not yield the sense of ‘outside’), cf. Basque *kanpo-ra* (exterior-ALLATIVE) ‘outwards’.

Closer to the Tawrã area, the Tani languages offer an inverse picture, as the root of Proto-Tani **raŋ~*roŋ* ‘empty’ (Sun 1993, 187) but also ‘inside’ (Mark W. Post 2025, p.c.) occurs in some words related to ‘forest’, cf. Galo *a-ràa* ‘inside, empty’ and *hiiluu-raaluu* ‘deep, undisturbed jungle’ (Post 2007, 881) or Apatani *ʔuuráa* ‘inside (as of a house)’ and *móoré* ‘jungle, forest’ (Post and Kanno 2013, 52, 240).

3.3.1.2 Kinship terms

Kinship terms are the nouns by which kin are called. Within kinship terms, there is a subgroup of close kind that are morphologically distinguished in that they take a 1st person possessive prefix *na-* (cf. §4.5.3). Table 3.5 indicates those kinship terms for kin older than EGO. Kinship terms younger than the EGO are displayed in Table 3.6. Kinship terms without an inherent difference of age (i.e. husband, wife and co-wife) are displayed in Table 3.7.

Table 3.5: Kinship terms: older than EGO

Term	Meaning
<i>nayã</i>	paternal grandmother
<i>natya</i>	paternal grandfather, ancestor
<i>nabá</i>	father
<i>naba-pâk</i>	father's elder brother
<i>nabâ-a</i>	father's younger brother
<i>nama-pâk</i>	father's elder sister
<i>nama-a</i>	father's younger sister
<i>nama</i>	mother
<i>nakâw</i>	mother's brother (elder and younger)
<i>nabùm-nama</i>	idem.
<i>napû</i>	elder brother, any elder male
<i>nabi</i>	elder sister, any elder female
<i>bra</i>	father-in-law
<i>yâw</i>	mother-in-law

Table 3.6: Kinship terms: younger than EGO

Term	Meaning
<i>pramru/pûmru</i>	‘younger brother (male persp.)’
<i>myúm</i>	‘younger brother (female persp.)’
<i>mathi/pathi</i>	‘EGO’s younger sister’
<i>nanang</i>	‘younger brother-in-law’; any (younger) person from wife’s side
<i>tarap</i>	‘sister-in-law’ (both elder and younger)
<i>ayûwak</i>	‘son’; any male relative below one’s generation
<i>a-yâ</i>	‘daughter (lit. child-FEM)’
<i>kamu</i>	‘son-in-law’
<i>nyàng</i>	‘daughter-in-law’
<i>hadyû</i>	nephew, niece (sister’s side)
<i>anyû-a</i>	nephew, niece (brother’s side)
<i>(a-)ayĩ</i>	grandchild

Table 3.7: Kinship terms: without an inherent difference in age

Term	Meaning
<i>mûwa</i>	‘husband’
<i>mîyã</i>	‘wife’
<i>mîyã qwúw</i>	‘second wife’
<i>akhî</i>	‘co-wife’ (term used by first wife)
<i>kârü</i>	‘brother-in-law’ (and his family; it also means ‘guest’)

The Tawrã kinship system can be described as *patrilineal* since clan membership is inherited through the line of the EGO’s father (cf. Eriksen 2015, 102f.), while the EGO’s mother’s family usually belongs to a different clan or sub-clan. Marriage within one’s sub-clan⁵ is traditionally not permitted. The traditional marriage partner in Tawrã society

⁵I have not detected a specific term for ‘sub-clan’. The term *bá-mě* (sometimes *bá-mě tahũ*) may refer to

is a cross-cousin from the EGO's mother's side, i.e. the child of EGO's mother's brother. Thus, EGO's mother's family plays a key role in Tawrã kinship system. There is also a relevant term *kâw bùm*, which stands for the EGO's mother's family. The term is likely a metaphor for *bùm* 'stem, root, original place, foundation' with the kinship term *kâw* 'maternal uncle'⁶. Hence, ideally, EGO's *kâw* 'maternal uncle' becomes a *bra* 'father-in-law'. Traditionally, a Tawrã husband-to-be was expected to work for his in-laws for months before the wedding⁷. There is no equivalent term for EGO's father's sisters and marriage with their offspring is traditionally not allowed. The EGO's brother-in-law (and his family) is termed *kârü*, which also means 'guest'. The *kârü* is restricted from eating certain foods at the EGO's place, including red meat.

Two kinship terms have a sex-specific perspective: *pramru* stands for 'younger brother' from the point of view of a male EGO while *myúm* has the same meaning but from the point of view of a female EGO.

The Tawrã kinship system does not have special terms for uncle and aunt's spouses; they are also termed 'uncle' and 'aunt'. Thus, the wife of the *nabâ-a* 'father's younger brother' is termed *namâ-a*, which also means 'father's younger sister'.

The terms *napû* 'EGO's elder brother' and *nabi* 'EGO's elder sister' are further employed to any elder male and female, respectively. The EGO's wife's elder brother is always addressed as *napû* (and not *nanang* 'any younger person from the wife's side'), even if he is younger than EGO.

The adjective *tapâk* 'elder' (also a noun meaning 'old man, respected person, elder') can be employed with certain kinship terms, cf. *tarap tapâk* 'brother's elder wife, wife's elder sister'. Some kinship terms take a short form of the adjective *tapâk* 'elder', cf. *nama-pâk* 'father's elder sister' and *naba-pâk* 'father's elder brother'. For other kinship terms, either 'clan', 'sub-clan' or proto-clan.

⁶The practice of taking the maternal uncle's offspring as the ideal marriage partner is also shared among the Kaman Mishmi (cf. Dutta 2006). However, according to some informants, the Kaman observe slightly stricter restrictions: A brother-in-law may not eat any food at the EGO's house. Interestingly, the Idu have a much more complicated marriage system and consanguineal bonds are traditionally only allowed if they have more than 10 (sometimes 15) degrees of consanguinity (cf. Rondo 2006). This means that the maternal uncle's offspring are not ideal marriage partners.

⁷It should be noted that the relevance of in-laws in the kinship system contrasts starkly with other TB languages and rather resembles the situation of matrilineal and matrilineal peoples in Mainland SEA in general and Austroasiatic (henceforth AA) peoples in particular, cf. the discussion on 'in-law' terms in Mlabri (AA, Thailand, Laos) (Jørgen 1995, 123), also cf. the marriage practices among the Khasi (AA, NE India) (Gurdon 1914, 76).

tapâk ‘elder’ is not available, cf. **natyâ tapâk* (intended: ‘grandfather’s elder brother’). Other compositional forms can be employed to specify the order of siblings. Table 3.7 displays the order by birth of EGO’s children.

Table 3.8: Kinship terms: order of EGO’s children

Term	Meaning
<i>ayûwa tazâi</i>	‘eldest child’
<i>ayûwa debû</i>	‘second child’
<i>ayûwa mamyû</i>	‘any non-first child’
<i>ayûwa atyô</i>	‘youngest child’

3.3.1.3 Class terms

In the Asian context, class terms often refer to super-ordinate elements in a nominal compound that express a higher taxon or category commonly found in words related to flora and fauna. A functional equivalent⁸ in English would be *berry* in *strawberry*. They are, in this sense, regular hierarchical compound elements that follow either a [generic-specific] or a [specific-generic] pattern that are confined to taxonomies of varied kinds and have varying levels of obligatoriness.

The term *taxonomy* or *taxa* is often employed in the linguistic literature to describe the paradigm of classification brought about by class terms and other devices of classification. Its use should not convey the idea of a formal scientific classification (e.g. the Linnaean taxonomy) based on morphological traits but rather, we should think of this as a form of *folk classification* that is grammaticalized differently in each language. As such, local folk schemes offer a glimpse into the native classificatory systems that reflect the way that people interact and experience the natural world (a *cultural artifact* for Becker 1975, 111). For example, coherent or *covert* groups that refer to biological clades (e.g. trees) often share characteristics that constitute a network of resemblances sometimes based on morphological, ecological and/or utilitarian considerations (cf. Ellen 2020, 48,90f; Ellen

⁸Functionally equivalent constructions where one noun heads a binomial and classifies the whole compound or collocation are differently labeled e.g. *super-ordinate category-label* in Galo (Post 2007, 217f) or simply *binomial terms* in the literature on ethnobiology (Brown 2020, 76).

2016). Lastly, classifications in general should be thought of as an ongoing process, rather than stable or permanent artifacts, as they “emerge con-substantially with conceptualization” and constitute a transitory consequence between old and new knowledge (cf. Ellen 2017, 246f).

Particularly in the SEA context, *class terms* (or *class nouns*, cf. Beckwith 1993) often form productive phrases or compounds where the head (i.e. the class term) classifies the whole binomial in different classes (e.g. its biological *essence*, cf. Enfield 2007, 147) but also functional groups or other broad categories (Vittrant and Allasonnière-Tang 2021, 748f; Enfield 2005, 300f; these categories can also be employed creatively in metaphors, e.g. in Burmese, cf. Jenny and Hnin Tun 2017, 114). Class terms are often related to nouns and may lexify the class of classifiers (cf. Bisang 1993, 6).

In Tawrã, all class terms are nouns that occur in taxonomies related to some groups of flora and fauna. They do not occur as classifiers in the language. Since class terms organize the lexicon in semantic classes, they are part of the classificatory devices of the language along with classifiers (§5) and verbs that discriminate based on agency (§3.3.2.3). Morphosyntactically, *class terms* in Tawrã are nouns and the class term + noun collocation corresponds to hierarchical compounds (cf. §4.2.1) triggered by semantic considerations. Some examples are *tâ kapá* ‘type of fish (probably banded gourami)’, where *tâ* means ‘fish (generic)’, or *masâng harèi* ‘(tree’s) branch’, where *masâng* is ‘tree (generic)’. These taxonomies may consist of individuals within the taxa (a species) or parts of an individual (e.g. *eyelid*). Class terms are the citation form of semantically subordinate class members but they are not obligatory. Table 3.9 displays some class terms.

Table 3.9: Class terms (excerpt)

Citation form	Translation	Class term	Meaning
<i>bülûm pi</i>	eyelid	bülûm	eye
<i>bülûm dagaw</i>	eye bags	bülûm	eye
<i>bülûm ûm</i>	eyebrow	bülûm	eye
<i>tawà jûm</i>	honey	tawà	bee
<i>tâ chí</i>	type of fish (probably Asian stinging catfish)	tâ	fish (generic)
<i>tâ kapá</i>	type of fish (probably banded gourami)	tâ	fish (generic)
<i>tâ kya</i>	type of fish (Indian flying barb)	tâ	fish (generic)

3.3.1.4 Collective suffix *làng*

The collective suffix *làng* indicates that a referent is composed of a large or indefinite number or a group (cf. Rijkhoff 1991), rather than being merely more than one, as do plural markers in European languages. Equivalent collective suffixes are common for the languages of South Asia and for classifier languages more generally (cf. Bisang 2017, 216f). The collective suffix follows an animacy cline, as it is more likely to be used with human referents rather than animals and it rarely occurs with inert referents (cf. Givón 2001a, 56).

This suffix is likely related to Proto-Tani **lu* ‘group, crowd’ (Post and Sun in prep.), cf. Mising *-lung* ‘a nominal suffix denoting a (usually large) collection of something’ (Taid 2010, 623). This similarity was also noted by Grierson (1909, 569, 617). We may assume that it occurred as an independent noun meaning ‘group’ in Tawrà based on some of its nominal properties, which will be discussed presently.

In Example 17, the speaker employs the collective suffix to indicate an indefinite number of officials. Similarly, in Example 18 the collective suffix implies an indefinite amount of studied people.

(17) tayâng kalà làng bârěya na

[[*tayâng kalà*]_{NP.head} *làng*]_{NP} *bâ-rě-ya* *na*]_{main.cl}

a.clan office COLL make-PL.II-ALLO INSIST

‘(They) made many Tayang officials, right!’ [TA22_clan_34 115]

(18) ning tawrã thĩ mu tako tazo zũho we làng we lamâ

[*ning tawrã thĩ mu*]_{NP} [[[*tako tazo zũ-ho*]_{rel.cl} *we làng we*]_{NP}

1P a.tribe EMPH also book ECHO write-STAT dem COLL DEM

la-mâ]_{main.cl}

say-ITER

‘Even our Tawrã (clan’s men), those who have studied keep saying it.’

[TA23_FishStory_175 305]

The collective suffix may be employed to mean all members of a coherent group, for example, a whole clan, family or lineage. In Example 19, the use of the collective suffix refers to the descendants of *Sokhrelum*.

(19) um.lyang bôya mẽ um.lyang ìga tya Sokhrelum làng

[[*um.lyang bô-ya*]_{rel.cl} *mẽ*]_{NP} *um.lyang ì-ga*]_{main.cl} *tya Sokhrelum làng*

a.place go-ALLO person a.place be.there-PL.I now a.name COLL

‘Those who went to Umlyang are at Umlyang are now the descendants of Sokhrelum.’ [TA23_Chaitom_clan_21 122]

The collective suffix does not commonly occur if the number of referents is specified by a numeral or a classifier phrase. In Example 20, the referent is the leader of seven villagers; the collective suffix would not be allowed. Similarly, in Example 21 there is no collective suffix on *dholes*. It should be noted that the collective suffix, similarly to the plural verbal suffixes (§6.3.2.0.1), commonly does not occur with inanimate and non-human nouns, as evidenced in Example 22.

- (20) *mlǎ̌ mẽ wê kamáng we tîya cya we mane*
 [[*mlǎ̌* *mẽ* *wê*]_{NP} [*kamáng we*]_{NP}]_{VLC} [[*tîya*]_{NP} [*cya we*]_{NP} *mane*]_{VLC}
 village.MOD person seven leader DEM lord 3 DEM FILL(<ASM)
 ‘He was the leader of seven villagers, he was the boss.’ [TA22_CLAN2_61 114]

- (21) *ah meho taprûw bra kaqîng ìya la*
ah [*me-ho*]_{main.cl} [[*taprûw* [*bra* *kaqîng*]_{CLP}]_{NP} *ì-ya* *la*]_{main.cl}
 yes! rear-STAT dhole CLF:generic two be.there-ALLO HSAY
 ‘Yes, (they) reared, there were two dholes, they say.’ [TA22_Talyu_11 60]

- (22) *matsâw ashyâ tyû bülei lângya yüm künüm*
matsâw ashyâ tyû bülei lâng-ya yüm künüm
 cow mithun chicken pig enter-ALLO NEG PRP
 ‘Lest cows, mithuns, chicken (or) pigs (would) come in.’ [TA23_pineapple_16 59]

The morphosyntactic status of *làng* deserves a note. It is a phonological word, as it has tone and stress (cf. §2.2.7). For some speakers, it may also trigger a tonal change characteristic of nominal compounding. In Example 23, the change in tone from *mẽ* ‘person’ to *mế* ‘person.MOD’ suggests that it has nominal elements, as this type of tone change only occurs in N-N modification (§4.2).

- (23) *mế* *làng*
 person.MOD COLL
 ‘people’ [011024_STBT_eli2 634]

Regarding its syntactic features, *làng* may still head a NP if there is no (other) head noun. In Example 24, the collective suffix is flanked by two demonstratives *we*, as nominal heads do. Furthermore, in Example 25 the NP is headed by the collective suffix.

(24) [*we làng we*]_{NP}

DEM COLL DEM

‘they, these’ [011024_STBT_eli2 650]

(25) *caina boder (...) boder nyu shyâzâ làng (...) khiga na*

[*caina boder boder nyu shyâ-zâ*]_{rel.cl} *làng*]_{NP} *khì-ga na*

China(<E) border(<E) border(<E) FOC happen-DIR COLL reach-PL.I INSIST

‘The ones that were at the China border, they reached (there), right?’ [TA22_clan 128 372]

In sum, we may suspect that *làng* was a full-fledged noun in the past and some nominal features remain after it was recruited as a collective suffix.

3.3.2 Verbs

This section will provide an overview of the sub-classification of verbs based on transitivity and semantics. The structure of the verb complex, including stem types, derivation and TAM suffixes, is described separately in §6.

Verbs describe most processes and actions (Schachter and Shopen 2007, 9). Many stative verbs also describe qualities, which correspond to adjectives in other languages. Verbs form an open word class in the sense that there is a seemingly endless number of members. There is no productive process for crafting new verbs e.g. from nouns or another part of speech. However, in rare cases, loanwords may enter the system, as with *khili* ‘play’ (from IA, cf. Hindi *khelnā* ‘play (v.)’ and *khilaunā* ‘toy (n)’). Note that newer loanwords require an auxiliary construction, cf. *kênsel ji* (cancel (<E) do) ‘cancel’ while older, more integrated loanwords (like *khili*) behave just like native verbs.

3.3.2.1 Verb transitivity

This section is an overview of a classification of verbs based on transitivity. Predicate clauses based on argument structure are discussed in §7.2.2. Predicate marking is discussed in §8.2.1.

Verbs in Tawrã express transitivity syntactically and there are neither specialized morphemes for indexing grammatical relations nor are there productive derivational processes at the root level to augment or reduce valency. Furthermore, there is no case marking in the language that would help determine valency classes and referents are regularly dropped if retrievable from context. There are, however, grammaticalized verbal roots and suffixes that alter the valency of a verb, such as the benefactive and the causative predicate derivations (cf. §6.2).

Verb transitivity in Tawrã can be categorized into valency patterns: atransitive, intransitive, transitive and ditransitive. Crucially, verbal roots in Tawrã can encode more than one valency pattern and there are two classes of conventionalized alternations: S/A and S/P ambitransitive patterns. Table 3.10 displays the main valency patterns in the language along with an example.

Table 3.10: Overview of valency patterns

Valency pattern	Verb	Meaning
atransitive	<i>lyâng</i>	be evening
intransitive	<i>bô</i>	go
S/A ambitransitive	<i>lî</i>	survive, rescue
S/P ambitransitive	<i>malai</i>	be scattered, scatter
transitive	<i>dá</i>	return (tr.), finish paying
ditransitive	<i>hàng</i>	give (something)

Atransitive (or zero-transitive) verbs have valency zero. They may not take any argument. Thus, Example 26 is ungrammatical since an atransitive verb cannot have any argument.

- (26) **we lyângho*
we lyâng-ho
 DEM be.evening-STAT
 -unacceptable

Intransitive verbs have a valency value of one; only one argument (S). This includes verbs of movement (*bô* ‘go’, *byúk* ‘run’) of position (*plã* ‘lie down’) and most states (*kâ* ‘be bitter’, *pâ* ‘be curved’) (Payne 1997, 55-58). Transitive verbs have a valency of two, thus two participants in the event (A and P). A verb like *dá* ‘return (tr.), finish paying’ cannot be employed intransitively; *‘return (vi)’. Ditransitive patterns have a valency of three, cf. *hâng* ‘give (something to somebody)’, i.e. A, G and T.

As discussed above, a verbal root in Tawrã may encode more than one *valency pattern*, i.e. the distribution of valency and semantic participants. To illustrate this, cf. the verb *lî* both ‘survive’ (vi, i.e. only S) and ‘rescue’ (vt, i.e. A and P). The alternation between a transitive and an intransitive frame is common in the language to the point that it is difficult to posit a ‘basic sense’ (cf. Malchukov 2015, 77; also Schikowski et al. 2015, 697). In other words, some verbs have only one frame (e.g. intransitive, transitive or ditransitive) whereas other verbs are labile, alternating between two frames. Two *uncoded alternations* are relevant here: a) S/A liability: like in the example of *lî* ‘rescue’ and ‘survive’ from above, the single argument (S) of the intransitive frame corresponds to the A argument in the transitive frame (i.e. S = A). Conversely, b) S/P liability represents verbs that code a S argument in their intransitive frame code that corresponds to the P argument in a transitive frame, as with *malai*, both ‘be scattered’ (vi) and ‘scatter’ (transitive, i.e. S = P) (cf. Haspelmath 1993, 90f; also Malchukov 2015, 108f).

3.3.2.2 Event semantics

A different way of organizing verbs in the language is to clump them in distinguishable semantic groups. One important distinction is that between active and stative verbs. In this case, event semantics correlates with a morphosyntactic feature when the verb is reduplicated: reduplicated stative verbs receive a continuative reading while reduplicated active verbs are interpreted either as continuative or as a willful or otherwise determined action (a similar case in Mona, cf. Dryer 2007a, 260; cf. Comrie 1976, 34f; Malchukov 2015, 115; also *verbal number* as discussed by Mithun 2001, 83-87). In Example 27, the stative verb *tsâi* ‘count’ is reduplicated and it yields an iterative/durative interpretation. This iteration/durative sense is achieved through the reduplication of the stative verb along with an iterative suffix *-mâ* ITER (cf. §6.3.4.4).

- (27) “ekhün gâqing” la tsâitsâi we la gũ thĩ tshâwhamâmâ khrâk
 [e = khün kâqing la tsâi~tsâi we la gũ thĩ]_{cosub.cl} [tshâw-amâ-mâ
 DEM = one two HSAY count~PLAC DEM say NF while move.fast-get.lost-ITER
 khrâk]_{main.cl}
 EVI
 ‘(The tiger) kept counting “one, two” (but the fish) kept moving and disappearing.’
 [TA22_tiger 35 102]

Conversely, active verbs can have a willful or determined interpretation on top of a continuative/durative reading, as illustrated in Examples 28a and 28b from elicitation. The speaker employs stem reduplication to emphasize his/her intention and determination.

- (28) a. katêkatêya
 katê~katê-ya
 swim-PLAC-ALLO
 ‘I’ll definitely swim.’
 b. bôbôya
 bô~bô-ya
 go-PLAC-ALLO
 ‘I’ll definitely go.’ [both field notes 20102024]

A more fine-grained semantic classification of verbs is based on the similarities in their meaning, e.g. verbs of sitting or cutting. This is also of typological interest, as Tawrã has a wealth of verbs of movement, cutting and grabbing, among others. Such verbal groups not only share a similar meaning but often display a similar valency profile (cf. Post 2007, 251).

Tables 3.11 and 3.12 display several verbs of cutting and stabbing. It could be hypothesized that the high precision of these verbs bespeaks the hunter-gatherer heritage of the Tawrã, where specialized vocabulary would have been particularly helpful. However,

without other typological cues, this remains highly speculative (cf. Bickel and Nichols 2020).

Table 3.11: Verbs of hitting

Verb of hitting	Meaning
<i>kadyâng</i>	push, kick (lying down)
<i>nyâng</i>	push, step on
<i>pu</i>	punch downwards (e.g. a table)
<i>phôk</i>	slap (outer side of the hand)
<i>shyo</i>	kick
<i>tabí/dabí</i>	slap (inner side of hand)
<i>tẽ</i>	hit, crash, crack open
<i>lacrâp tẽ</i>	elbow strike
<i>pe tẽ</i>	flick (with fingers)
<i>sâpum tẽ</i>	hip strike
<i>tshâp</i>	punch (fist)
<i>urêi</i>	pinch

Table 3.12: Verbs of hitting (with an instrument)

Verb of hitting	Meaning
<i>hatúw</i>	cut (e.g. a branch from a tree)
<i>ì</i>	spear (fish)
<i>lô</i>	pierce, make a hole, thread
<i>nyê</i>	slash, cut, saw (also reap, harvest)
<i>nyûm</i>	pierce (e.g. with an arrow), pound (e.g. rice), wrestle (of horned animals)
<i>pi</i>	axe, peck
<i>plû</i>	stab (with a knife)
<i>tashyâ</i>	scar
<i>tyû</i>	hit (with an instrument), play (e.g. a gong)
<i>tyé</i>	cut (generic: used for slaughtering animals and reaping)
<i>tyúw</i>	slice (wood)

Table 3.13 presents a list of verbs of grabbing. Two of these verbs also occur as mensural classifiers (§5.1). Table 3.14 displays all verbs of carrying.

Table 3.13: Verbs of grabbing

Verb of grabbing	Meaning	Note
<i>grũ</i>	whole hand	also noun ‘leg, arm’ and mensural classifier for clenched fists
<i>katúw</i>	two-four fingers (pincer grasp)	
<i>tapûng</i>	with claws (as of an animal)	also classifier for half clenched fists
<i>wô</i>	five fingers	

Table 3.14: Verbs of carrying

Verb of carrying	Meaning	Note
<i>(a) ba</i>	carry child on the back, piggy-back	cf. Proto-Tani * <i>bak</i> ‘carry on back’ (Sun 1993, 214)
<i>glâi</i>	carry on the back	cf. Proto-Tani * <i>ga:</i> ‘carry on, be pregnant’ (Sun 1993, 97)
<i>habüw</i>	carry on shoulder (e.g. a log)	
<i>kará</i>	carry in front (in a cloth)	
<i>katyú</i>	carry on the head	
<i>pûw</i>	carry in a basket	
<i>rá</i>	carry in pocket	also ‘wear (a bag)’
<i>tyã</i>	carry on a stick (by two people)	

Table 3.15 displays several position verbs. More specific positions are described with specialized adverbs like *maputyak* ‘on all fours’ (cf. §3.4.1).

Table 3.15: Position verbs

Verb of position	Meaning
<i>agâw</i>	crouch
<i>dě</i>	stand
<i>dě-katã</i>	stand on tiptoes
<i>di</i>	sit
<i>hadi</i>	give space to sit (also ‘tell a story’)
<i>kha</i>	lie down (face up)
<i>khülâng-pum</i>	kneel down
<i>plã</i>	lie down (face down)

3.3.2.3 Humanness and animacy restrictions in verb selection

A small set of verbs have an in-built system that imposes a selection of referents based on agentive value. These form verbal pairs with a similar meaning (like *bite*, *be born* and *be there*) that differ in the assignment of semantic roles to their projected argument based on the animacy and agentivity of the referent⁹ (briefly mentioned by Sun 1993, 407; also for Lisu cf. Li and Thompson 1976, 473f. The case of Idu has been discussed by Aikhenvald 2000, 155). For instance, the verbs *thâk* ‘bite (of animals and insects)’ and *krâng* ‘bite (of humans)’ select a S/A argument based on humanness and animacy. Thus, a sentence like Example 29 could only mean that an *animal* might bite a person, on account of the verb *thâk* ‘bite (of animals and insects)’, and not the other way around.

- (29) mề thâkbinla (...)
 [mề]_{P.hum} thâk-binla
 person bite-POT.ALLO
 ‘It might bite people (...) (and not *the person might bite (something)).’
 [270823_STBT_eli1 371]

The use of the concept of *humanness* in the case of Tawrã is a simplification of the overall picture, as animals may also occur with a verb seemingly restricted to humans in certain circumstances (note the discussion on *differential object marking*, cf. Bossong 2021, 24f.). In traditional stories, one often encounters animals that share human-like skills, i.e. the ability to speak and (some) reasoning power. In this sense, the use of the predicate assigns a high agentive value to the referent instead of humanness *per se*, although there is a strong correlation between the two¹⁰. In Example 30, the man orders the tiger to bite him in the back of the neck if he were to lose in a competition against it. The verb employed *krâng* ‘bite’ is used for referents that have high agency as opposed to its low-agency counterpart *thâk* ‘bite’. Table 3.16 presents a list of verb pairs based on animacy restrictions.

⁹German has one equivalent verb pair *essen* ‘eat (as of humans)’ and *fressen* ‘eat (as of animals)’. The use of *fressen* ‘eat (as of animals)’ is also used for people to imply animal attributes like gluttony.

¹⁰Also cf. the discussion on the criteria that govern the choice of posture and position verbs in Galo. As discussed by Post (2007, 255f), the verb *dâk*- ‘stand’ commonly selects S arguments with high animacy and agentivity, i.e. with “*legs or leg-like features*, and in a more-or-less *vertical/upright posture*” (idem. 257, emphasis in the original).

(30) “adyângra gũ krângna” la la

[adyângra gũ]_{obl.loc} krâng-na la la

back.of.neck LOC bite-IMP say HSAY

‘He said (to the **tiger**): “Bite me on the back of the neck”, it is said.’ [270823_tiger_45 122]

Table 3.16: Verb pairs based on humanness

Verb	+ humanness	-humanness
bite	<i>krâng</i>	<i>thâk</i>
be born, give birth	<i>mẽ</i> (cf. <i>mề</i> ‘person’)	<i>brô</i>
be pregnant	<i>a-pu</i> (also ‘womb’)	<i>zã</i> (also ‘pack in bag’)
be there	<i>ì</i>	<i>ã</i>
grab	several verbs, cf. Table 3.13	<i>tapûng</i> (with claws; also a classifier for half clenched fists)
dig	<i>tyû</i>	<i>prâw</i> (also ‘make a borrow’)

3.3.3 Adjectives

This section will discuss the semantic sub-grouping of adjectives. The morphosyntactic arguments for a distinct class of adjectives have been discussed in §3.2.4. The syntactic features of adjectives in relation to the NP are discussed in §4.3.1.

Notionally, adjectives refer to qualities and attributes (Sasse 2015, 183; Schachter and Shopen 2007, 13). This function in Tawrã is also shared with stative verbs and several nouns. The class of adjectives is open and includes a large number of members. According to the typology presented by Dixon (1982, 16f., 34f), adjectives in Tawrã cover six of the seven core groups: *dimension*, *color*, *human propensity*, *age*, *value* and *physical property*. The second group *color* has only two members since other colors are expressed by verbs. The seventh group *speed* is covered only by verbs. Also within the core six groups, it should be noticed that many adjectives occur as other word classes (cf. §3.2.1). Table 3.17 illustrates the six core semantic groups of adjectives along with some

examples.

Table 3.17: Main semantic groups of adjectives

<u>Dimension</u> <i>daráng</i> ‘big’, <i>lê</i> ‘big (only for pig and rain)’, <i>a</i> ‘small (also ‘child’), <i>kaláng</i> ‘long’, <i>katyûqu</i> ‘short’	<u>Color</u> <i>lyó</i> ‘white’, <i>má</i> ‘black’
<u>Human propensity</u> <i>halyo</i> ‘happy’, <i>dra</i> ‘worthless’, <i>dyâ</i> ‘rich’	<u>Age</u> <i>mêqẽ</i> ‘new, first’, <i>bîya</i> ‘old (of abstract things)’, <i>ûm</i> ‘old (of concrete things)’
<u>Value</u> <i>pra</i> ‘good’, <i>jüm</i> ‘true, real’, <i>amũ</i> ‘false’, <i>chî</i> ‘pure (of rice beer)’	<u>Physical property</u> <i>asâng-ra</i> ‘raw, alive’, <i>bâqa</i> ‘shallow, thin’ <i>gãwã</i> ‘round’

As discussed above, many qualities and attributes are expressed by stative verbs, as is common in the language family (*adjectival verbs* cf. Matisoff 2003, 484; also Schachter and Shopen 2007, 18; *adjective-deficient* languages, cf. Dixon 1982, 15) but less common among Indo-European languages cf. Tawrã *shyaw* ‘be sweet’ vs. English *sweet* or Nepali *guliyo* (both are adjectives). Further examples of verbs that express qualities and attributes include *kâ* ‘be bitter’, *tamě* ‘be crazy’, *lyâng* ‘be evening, be dark’. On top of that, there are many cases of double class membership between stative verbs and adjectives, including *lâi* both ‘be half grown (v)’ and ‘young, first (e.g. of a crop) (adj)’ or *halyo* ‘be happy, make happy (v)’ and ‘happy (adj)’ (also discussed in §3.2.4).

It should be noted that several roots have dual membership with adjectives and adverbs (as discussed in 3.4.1). In Example 31a, a noun is modified by the adjective *pra* ‘good’. Conversely, in Example 31b the homophonous adverb *pra* ‘nicely’ modifies the action, i.e. a predicate phrase, which constitutes a prototypical adverbial function (in the sense of Sasse 2015, 186).

- (31) a. *tako pra*
 book good
 ‘The book is good.’
- b. *pra marô-na*
 good speak-IMP
 ‘Speak nicely!’

A few versatile nouns also display a functional overlap with adjectives, as they provide information on gender and age or stage of development¹¹ (specifically for animals) when used as compound elements. Differences in grammatical gender are lexicalized in nouns like *mûwa* ‘husband’ and *mâyã* ‘wife’¹². For animals, gender corresponds to sexual dimorphism¹³, and it may also encode a specification on the stage of development: *chê* ‘pullet, sexually immature female animal’, *krù* ‘post-reproductive female bird, mother bird’ and *kâri* ‘male animal (i.e. fully developed)’, as in *kwâg kâri* ‘male dog’. For humans, age is encoded in nouns such as *tapiiw* ‘elder, respected person’, e.g. in *taraw tapiiw* ‘elder sister-in-law’ or *shyâ-brê* ‘young man’.

3.4 Other open classes

This section includes a discussion on the morphosyntactic properties of adverbs (§3.4.1) and classifiers (§3.4.2).

3.4.1 Adverbs

The concept of adverb, although a cross-linguistically valid grammatical category, has been described as problematic since it often displays a high degree of morphosyntactic diversity (Givón 2001a, 87f; Sasse 2015, 186). For the present work, adverbs are defined

¹¹Such nouns could be interpreted as *noun classifiers* following Aikhenvald (2000, 17,81f). In the present analysis, these are considered nouns, although they are versatile compound elements.

¹²Note that there is a special *feminine* suffix *-yã* for human and human-like referents (cf. §4.5.2).

¹³Sastry (1991, 76,80) already identified this group of lexemes and mentioned their lexicalized gender system but he did not provide positive evidence for their identification.

as being adjoined syntactically to verbs and thus modifying the event semantically. Adverbs are an uninflected word class with many members sharing class membership with adjectives. They occupy a pre-verbal position, with fronting of the adverb perceived as an emphasis. In Example 32, the adverb *hayũ* ‘first’ modifies the verb *tsâi* ‘count’. Similarly, in Example 33 the adverb *ali* ‘like this’ modifies the verb. In Example 34, the adverb takes a marked position before the head noun, which emphasizes the manner expressed by the adverb.

- (32) we “nyu hayũ tsâidângdyibô ka byẽ hã we (...) adyângra gũ krângna” la la
 we [[nyu **hayũ** tsâi-dâng-dyibô ka byẽ]_{cond} hã we
 DEM 2S first count-COMPL.I-PFV.EGO COUNT.EXP COND 1S DEM
 adyângra gũ krâng-na la la]_{main.cl}
 back.of.neck LOC bite-IMP say HSAY
 “‘If you finish counting first, then, bite me in the back of the neck”, he said, it is
 said.’ [TA22_tiger 43-44 118]

- (33) ali adiga
ali adi-ga
 like.this tell-PL.I
 ‘Like this they narrated.’ [TA23_talyu_clan 73 258]

- (34) (...) bé.bé sokarê hanana
bé.bé sokarê hana-na
 slowly government(<IA) come-FACT
 ‘Slowly the government came.’ [TA22_clan 20 78]

The relation between adjectives and adverbs in Tawrã requires a brief digression, for it is central to the definition of this word class. Adverbs and adjectives are related classes in the face of the high number of members that exist in both of these categories. The following examples illustrate this situation through the use of *gâqa* ‘close’ as an

adjective, adverb and noun. In no case is special morphology required. In Example 35a, *gâqa* modifies the head noun as an adjective in a verbless clause. In Example 35b, *gâqa* ‘close’ modifies the predicate, displaying the syntactic function of an adverb. Finally, in Example 35c *gâqa* ‘close’ takes a postposition (cf. §9.5), suggesting that *gâqa* is employed as the head of a NP (a property of nouns, cf. §4).

(35) a. *nyu àng gâqa*

[*nyu àng gâqa*]_{VLC}

2S house close

‘Your house is close-by.’ [110823_STBT_voc 695]

b. *cya ma gâqa ìho gũ (...)*

[*cya ma gâqa ìho gũ*]_{cosub.cl}

3 CONTR.TOP close be.there-STAT NF

‘S/he being close-by (...)’ [TA22_kutyemba_phla 398 750]

c. *phlá krâ wrûm yüm gâqa kê ìho*

[*phlá krâ*]_{NP} [*wrûm*]_{NP} [*yüm*]_{VLC} [[*gâqa*]_{NP} *kê ìho*]_{main.cl}

stone.MOD hole depth NEG close LOC.INS be.there-STAT

‘He was not deep within the crack in the stone, he was nearby.’

[TA22_kutyemba_phla 407 766]

These examples suggest that the distinction between adjectives and adverbs relies on purely syntactic and semantic considerations: while adjectives modify the head of a NP, adverbs modify the verb, particularly the semantics of the whole event (cf. §3.2.1). Modification of the verb also correlates with a pre-position in the clause.

A further field where adverbs display a functional overlap is with predicate derivations (§6.2), as the latter has many members that do not affect the tense or aspect of the verb but rather add a specific manner or modality (akin to adverbs in Indo-European languages), such as *-plûw* ‘excessively, too much’. Furthermore, some predicate derivations have an adverbial rather than a verbal counterpart, cf. continuative CONT *-(ta)la* (§6.2.4.1), related to the adverb *(ta)la* ‘still’.

There is a particular phonological word structure present in several adverbs and verbs: (C $\hat{V}_1$).(V₁) - a trochaic syllable (cf. §2.2.6) composed of a CV mora with a high-falling tone and a syllable composed of one vowel that copies the preceding vowel. A glottal stop always separates the two heterosyllabic vowels which may also be nasalized. Table 3.18 displays all CVV adverbs and adjectives found in the language. The fact that this structure is shared by both adjectives and adverbs further indicates that these word classes are linked.

Table 3.18: CVV adjectives and adverbs

adjective/adverb	Meaning
<i>bâqa</i>	shallow, thin
<i>bêqe</i> (also <i>bé</i> or <i>bé-bé</i>)	slowly
<i>bûqü</i>	recently
<i>dêqe</i>	before
<i>gâqa</i>	nearby
<i>mêqẽ</i>	first, new
<i>tâqa</i>	light (weight)

Adverbs and adjectives can be modified by the adverb *tyâng* ‘very’, which takes a strict pre-adjective/adverb position. In Example 36 the adverb *tyâng* ‘very’ modifies another adverb. In Example 37 it modifies an adjective. Note that *tyâng* occurs before the adjective and often triggers a different word order within the NP: N Adj vs. *tyâng* Adj N (cf. §3.2.4).

- (36) *majarêi tyâng karo cék künam byúwkûbô*
 [*[majarêi **tyâng** karo cék künam]*_{sub.cl} *byúw-kû-bô*]_{main.cl}
 cat(<IA) very fast walk because run-COMPL.II-PFV.ALLO
 ‘The cat run away because it is very fast.’ [200622_JO_eli 600]

- (37) *tyâng daráng phlă*
***tyâng** daráng phlă*
very big stone
'a very big stone' [TA22_kutyemba_phla 8 81]

The remainder of this section is devoted to sketching the semantic fields of adverbs. These include *temporal adverbs*, *position adverbs*, *manner adverbs* and *location adverbs*. §3.4.1.2 describes ideophones.

3.4.1.1 Classification of adverbs

Adverbs can be classified into four groups based on semantic considerations: temporal adverbs, position adverbs, manner adverbs and location adverbs.

A) temporal adverbs include *tahîng* 'today' or *sûhîng* 'tomorrow'. Most members of this group include the formant *hîng* 'day' (a noun). There are, furthermore, specialized classifiers for counting days, months and years (§5). Table 3.19 includes all temporal adverbs found in the language.

Table 3.19: Temporal adverbs

Adverb	Meaning
<i>bûra</i>	earlier
<i>talyâng</i>	evening
<i>abe</i>	morning
<i>anaya</i>	early morning
<i>tahîng</i>	today
<i>sûhing</i>	tomorrow (when counting days)
<i>abena</i>	tomorrow
<i>anûhing</i>	overmorrow, one day after tomorrow
<i>kanehing</i>	two days after tomorrow
<i>kasahîng</i>	three days after tomorrow
<i>bûlehîng</i>	yesterday
<i>kambâ/kampâ</i>	already
<i>tya</i>	now

Some temporal adverbs have a prefixed root *bû-*, e.g. *bûlyâng* ‘last (only in fixed forms)’ and *abe-/ta-* ‘this (only in fixed forms)’. These forms are not productive in the modern language. Note that *abe-na* yields ‘tomorrow’ and not the expected *‘this morning’.

Table 3.20: Temporal adverbs: last and this

Adverb	Meaning
<i>bûlyâng na</i>	last morning
<i>bûlyâng lyâng</i>	last evening
<i>bûya</i>	last night
<i>abelyâng</i>	this evening
<i>abeya, taya</i>	tonight

B) position adverbs indicate corporal positions and occur with verbs. Table 3.21

displays two adverbs found in the language.

Table 3.21: Position adverbs

Adverb	Meaning
<i>maputyâk</i>	on all fours
<i>thrûga, putyâ</i>	on the back

C) Manner adverbs modify verbs by indicating how an action is to transpire. Manner adverbs are displayed in Table 3.22.

Table 3.22: Manner adverbs

Adverb	Meaning
<i>alâng-pâ, züm-pâ, ki-pâ</i>	together
<i>alí</i>	like this
<i>bupâ</i>	suddenly
<i>phati</i>	suddenly
<i>hapyê</i>	straight
<i>kadyêsyem</i>	hardly, with difficulty
<i>(ga)phlũ</i>	by himself
<i>tagrû</i>	quickly
<i>karô</i>	quickly

D) Location adverbs are employed to place a referent in a relative space. Demonstrative *we* (§4.3.3.1) and relator nouns (§3.3.1.1) are also used to indicate location.

The adverbs *(a)ci/ashi* ‘east, west’, *atyu* ‘north’ and *(a)mã/ama* ‘south’ express what has been termed *topographical deixis*, a set of references anchored on topographic features (cf. Post 2019). In this case, the flow of water from the Eastern Himalayan range to the Brahmaputra river. Rather than ‘north’ or ‘south’, these concepts indicate ‘uphill’ or ‘downhill’. Furthermore, *(a)ci/ashi* ‘horizontal axis (east/west)’ seems to indicate ‘across, same-level’. In this sense, the noun *mathûw* is worth noticing, as it refers to the place where the rivers originate and is also employed to mean ‘up in the mountains’ and ‘north’.

Table 3.23: Location adverbs

Adverb	Meaning
<i>bum-plâ</i>	here and there
<i>ewe-kyà</i>	here and there
<i>rôtyu</i>	yonder
<i>(a)ci/ashi</i>	horizontal axis (east/west)
<i>atyu</i>	north
<i>(a)mã/ama</i>	south

3.4.1.2 Ideophones

Ideophones impressionistically render a situation through an onomatopoeic root that is commonly reduplicated, rhymed or echoed (cf. Rai et al. 2005; Rai and Winter 1997). Ideophones are, however, different from all other adverbs, as the latter commonly cannot be reduplicated save for a few exceptions. Ideophones are included in the ranks of adverbs based on their syntactic features of modifying a verb and occurring in a strict pre-verbal position. However, it is common for ideophones to occur as their own clause, as the speaker wishes to render a given situation without the need for other clause constituents.

In Example 38, the ideophone modifies the verb, specifying a type of bodily reaction, in this case, shame. Table 3.24 displays a number of ideophones. The number of ideophones in the language is likely higher but further elicitation is required to achieve a more concrete picture.

- (38) *webô hwěga hwěga tili.tulu.tili tshângdâng gũ bô la*
 [webô hwě-ga hwě-ga [tili.tulu.tili tshâng-dâng gũ
 and.then see-PL.I see-PL.I manner.of.having.a.rash turn.pale-COMPL.II NF
*bô]*_{co/sub.cl} *la]*_{main.cl}
 TEMP.SUB HSAY
 ‘(When everybody) saw (him), he got a rash.’ [TA22_clan2 191 364]

Table 3.24: Ideophones

Adverb	Meaning
<i>shîm-shim</i>	manner of seething
<i>ma-ma</i>	close (distance)
<i>hari-hera</i>	zigzag motion
<i>tili-tulu</i>	manner of having a rash
<i>khika-khika</i>	in each
<i>doq-doq</i>	manner of ascending (e.g. smoke)
<i>baq-ba</i>	manner of being long (of hair)

3.4.2 Classifiers

Classifiers in Tawrã are an independent word class that often occur with numerals. Tawrã has a large number of classifiers, most of which are recruited from nouns. Classifiers *classify* nouns according to arbitrary semantic parameters such as shape or material. Classifiers are also known as *counting words* as they are often employed in counting and hence link nouns to numerals. Another crucial function of classifiers is to individualize referents (Bisang 1993). Classifiers are further described in §5. Appendix II (§13) includes a list of all classifiers found during my study of the language .

In Tawrã, classifiers are distinguished from other word classes such as nouns on morphological grounds: while regular nouns are counted with the full numeral one *khing-gê*, classifiers occur with the numeral's second element *gê* 'one'. In Example 39, the classifier *bra* 'CLF:generic' triggers a reduced form of the numeral 'one'.

(39) pahwã bra gê rôrĩ halyo gã bônârebô delé (...)

[pahwã bra gê rô-rĩ]_{main.cl} [[halyo gã]_{cosub.cl} bônâ-rẽ-bô

frog CLF:generic **one** catch-PL.II be.happy NF go-back-PL.II-PFV.ALLO

delé]_{main.cl}

it.seems

'They caught one frog and have returned happily (...)' [TA22_frog_135 719]

Employing the *reduced* numeral *gê* ‘one’ with a regular noun is ungrammatical, hence **àng gê* (lit. house one) intended: ‘one house’ is judged ungrammatical. Instead, ‘house’ should be counted with a full numeral (*àng khing.gê* (lit. house one)) or a classifier phrase (*àng dũ gê* (lit. house CLF:man.made.structures one) ‘one house’).

Classifiers are historically derived from nominal roots and many classifiers are homophonous with nouns, as described in §5.4. There are thus many nominal roots that exist both as nouns and classifiers. They are, however, distinguished morphosyntactically: In Example 40a, the lexeme *kahã* occurs as a classifier, for it triggers the reduced numeral *gê* ‘one’ (instead of *khing.gê*). Furthermore, in Example 40b the same root occurs as a noun, modified by an adjective.

- (40) a. *kahã* *gê*
 CLF:node one
 ‘one node’
- b. *kahã* *daráng*
 node big
 ‘a big node’

3.5 Closed classes

A word class is considered *closed* if its members are few in number and are basically the same for all speakers (Schachter and Shopen 2007, 3). In Tawrã, closed classes include pro-forms (§3.5.1), particles (§3.5.2), numerals (§3.5.3) and demonstratives (§3.5.4).

3.5.1 Pro-forms

Pro-forms have the basic function of replacing other words belonging to open classes or larger constituents (Schachter and Shopen 2007, 24). This section discusses personal pronouns (§3.5.1.1), interrogative pronouns (§3.5.1.2), indefinite pronouns of person *mê*

‘someone, somebody’ (§3.5.1.3), the reflexive pronoun *ga* (§3.5.1.4) and quantifier *mé* ‘all’ (§3.5.1.5). Finally, several indefinite pronouns meaning ‘something’ are discussed under the label ‘other indefinite pronouns’ (§3.5.1.6).

3.5.1.1 Personal pronouns

Personal pronouns form a closed class of independent lexemes that can be inflected for number and person. They commonly replace human referents and non-human referents with high agency. The morphological category of number sets pronouns apart from nouns; while the former have three values: singular, plural and dual, nouns do not have number (cf. §3.3.1.4). As it is common for related languages of the area, pronouns do not make any distinction for gender or politeness. Note that 3rd person singular *cya* is distinct from the demonstrative *we* (§4.3.3).

Table 3.25 displays all personal pronouns in the language. Note that *làng* (§3.3.1.4) is a collective suffix and it therefore occurs in parentheses.

Table 3.25: Personal pronouns

	singular	dual	plural
1 st person	<i>hã</i>	<i>ni-îng/ni-kaqîng</i>	<i>ning</i>
2 st person	<i>nyu</i>	<i>ané-îng</i>	<i>ane</i>
3 st person	<i>cya</i>	<i>cyá-îng</i>	<i>cya (làng)</i>

The plural forms are suppletive. In the case of the 3rd person, there is a dual but no plural form. A collective suffix *làng* (§3.3.1.4) is employed if the quantity of a referent greater than two is relevant. Dual forms are derived from a reduced form of the numeral *kaqîng* ‘two’, as is well-attested throughout the language family (DeLancey 2019, 123).

Personal pronouns replace any referent that is perceived as active or having agency. Therefore, in stories where animals play an important role, they may also take the pronoun *cya* ‘3’. Example 41 is from a story in which a tiger is the brother of the *first man*. Since the former has human attributes, like the ability to talk, it is replaced by *cya*:

(41) cya asáng thâthâ thĩ gũ (...)

[cya asáng thâ~thâ thĩ gũ]_{cosub.cl}

3 raw eat~PLAC CONC NF

‘By eating and eating raw (meat) (he the tiger) (...)’ [TA22_tiger_127 281]

Pronouns, without any modification, also indicate possession in Tawrã as in Examples 42, 43 and 44. Personal pronouns used as possessors constitute independent phonological words. This is not the case for the specialized (and irregular) *na*- possessive prefix found in several kinship terms (§4.5.3).

(42) ah nyu aye we we

ah nyu a=ye we we

ah 2S child=DEM DEM DEM

‘Oh that child of yours!’ [TA22_origin_052 153]

(43) kadényu hã a alí

kadényu hã a alí

how 1S child like.this

‘How come my child is like this!’ [TA22_origin_040 122]

(44) cya mlã nyu shìkũbô la

cya mlã nyu shì-kũ-bô la

3 village FOC take-COMPL.II-PFV.ALLO HSAY

‘(He) took (it) to his village, they say.’ [TA22_CLAN2_051 100]

3.5.1.2 Interrogative pronouns

Interrogative pronouns in Tawrã refer to an unknown referent or piece of information and commonly head a content question (i.e. a type of interrogative clause, cf. §7.3.1.2).

Table 3.26 displays the interrogative pronouns found in the language along with a note on compositionality and etymology (where available).

Table 3.26: Interrogative pronouns

Interrogative pronoun	Meaning	Notes
<i>shya</i>	who	PTB * <i>su</i> ‘who’ (Matisoff 2003, 180)
<i>ja</i>	what	
<i>jadi</i>	what, why, how	probably <i>ja</i> + *- <i>di</i>
<i>kadé</i> (often as <i>kadé.nyu</i>)	how	
<i>kadí</i> (often as <i>kadí.mu</i> , <i>kadí.gũ</i>)	when	
<i>kadé-ge</i>	how much	<i>kadé</i> + * <i>ge</i> (potentially <i>gê</i> ‘one’)
<i>hanûk</i>	where	
<i>hanû.we</i>	which one	probably <i>hanûk</i> + <i>we</i> DEM

As illustrated in Table 3.26, some question words seem to be morphologically complex while others are basic. As basic forms, we can distinguish *shya* ‘who’, *ja* ‘what’, *hanûk* ‘where’, *kadé* ‘how’ and *kadí* ‘when’. The pronouns *kadí* and *kadé* often occur with a further element: *kadí-mu*, *kadí-gũ* ‘when’ and *kadé-nyu*. The form *hanû-we* ‘which one’ is compositional.

The relation between *ja* and *jadi* is unclear. The pronoun *jadi* is rare in the language and it seems to conflate different interrogative pronouns (‘what, why, how’) as a sort of general question word.

The directive suffix -*zâ* (cf. §6.3.3.1) can be employed with interrogative pronouns to render a distributive sense, i.e. the marked pronoun refers to more than one option. In Examples 45 and 46, the directive suffix forces a distributive sense on the pronouns (here: many places, many people). A similar effect is achieved through the reduplication of the pronouns *ja* ‘what’ and *who* ‘who’, as will be discussed in §3.6.1.

- (45) *hanûkzâ khìge*
hanûk-zâ khì-ge
 where-DIR reach-and.back
 ‘Where have you been (referring to many places)?’ [271024_JO_eli1 1295]
- (46) *shyazâ hanabing*
shya-zâ hana-bing
 who-DIR come-POT.ALLO
 ‘Who might be coming (referring to many people)?’ [271024_JO_eli1 1354]

Interrogative pronouns can also function as indefinite pronouns with the sense of *something*, *anything*, *somewhere* or *anywhere*. In such cases, the pronoun is often reduplicated (cf. §3.6.1). In Example 47, the interrogative pronoun *ja* ‘what’ serves as an indefinite pronoun ‘whatever’. Similarly, in Example 48 the interrogative pronoun *hanûk* ‘where’ renders ‘wherever’. This also applies to *shya* ‘who’ (as ‘whoever’), *kadi* ‘where’ (as ‘wherever’), etc.

- (47) *ja (ja) hwě.tyâng we hã atyâbiya*
ja ja hwě.tyâng we hã atyâ-biya
 what what see DEM 1S tell-FUT.ALLO
 ‘Whatever he sees, he will tell me.’ [281124_BAB_eli1 742]
- (48) *hanûk trûga tatyâna*
hanûk trû-ga tatyâ-na
 where meet-PL.I tell-IMP
 ‘Wherever you see him, tell (me).’ [281124_BAB_eli1 820]

Interrogative pronouns can also be negated with the general negator *yûm* ‘NEG’ and the adverb *mu* ‘also’ in order to have an indefinite negative function - equivalent to English negative pronouns *nobody*, *nothing*, etc. Thus, in Example 49 English *no one*

corresponds to the indefinite pronoun *shya* ‘who’, the adverb *mu* ‘also’ and the negative *yüm* ‘NEG’. Similarly, in Example 50 the construction involves the pronoun *ja* ‘what’, rendering ‘anything, nothing’. In Example 51 the construction uses *kadí-gũ* ‘when’ to render ‘never’. Finally ‘anywhere’ is built with ‘where’ and the negative in Example 52.

- (49) *kwâg wúŋg we shya mu katyâng yüm*
kwâg wúŋg we shya mu katyâng yüm
 dog bark DEM who also listen NEG
 ‘No one heard the dog’s bark.’ [200622_JO_eli2 240]

- (50) *ja mu düm la*
ja mu dü=m la
 what also need = NEG say
 ‘S/he didn’t want anything, s/he said.’ [240122_AM_eli1 140]

- (51) *kadí.gũ mu hwê̂m*
kadí.gũ mu hwê̂=m
 when also see = NEG
 ‘I have never seen (it).’ [230122_JW_eli 1208]

- (52) *kadâm.kara hanûk nyu mu hana yüm*
kadâm.kara hanûk nyu mu hana yüm
 a.clan where FOC also come NEG
 ‘The Kadam Kara came from nowhere.’ [TA23_phlakam_clan 14 64]

As discussed by Post (2015b, 443), equivalent constructions are found in the language family and they are also widespread in South Asia (e.g. Palula (Indo-Aryan, Pakistan) cf. Liljegren 2016, 145; also in Kharia (Munda, India) cf. Peterson 2010, 187f).

The interrogative pronoun *shya* ‘who’ stands out regarding indefinite reference, as it can modify a noun to indicate an indefinite and total reference ‘every X, all X’. In

Example 53, *shya* ‘who’ in *shya bá-mě* (who clan) refers to all clans. This can be negated to render a negative indefinite reference ‘no X’, as in Example 54, where *shya* ‘who’ in *shya khũnyũm* (who god) plus *mu* ‘also, even’ and the general negator *yũm* renders ‘no god’.

- (53) (...) *wegũ shya bá.mě la dũ* (...)
wegũ shya bá.mě la dũ
 and.then who clan say CO.TEMP
 ‘(...) and then all clans saying (...)’ [TA23_newmyth_58 344]

- (54) *shya khũnyũm mu dyóna yũm magũ*
shya khũnyũm mu dyó-na yũm magũ
 who god also be.different-FACT NEG that’s.how.it.is
 ‘No god was different.’ [TA22_origin_102 297]

3.5.1.3 Indefinite pronouns of person *mě* ‘someone, somebody’

The noun *mě* ‘person’ also occurs as an indefinite pronoun of person ‘someone’. It is employed to indicate an unknown or otherwise generic human referent. The grammaticalization path from the noun meaning ‘person’ to an indefinite pronoun is typologically common, cf. German *Man* ‘man’ and the homophonous indefinite pronoun (subject only) *man* ‘one’.

In Examples 55 and 56 the pronoun *mě* stands for non-specific (generic) human referents.

- (55) *tya kabêna bô am mē ahwêna dyibô* (...)
[tʲa kabê-na bô]_{cond.cl} am [mē ahwê-na dyibô]_{cond.cl}
 now mediate(<IA)-FACT COND.ALLO FS somebody send-FACT COND.EGO
 ‘If we (try to) mediate now, if I send somebody (...)’ [TA24_land_97 167]

- (56) we mề laya gũ we ning khri ne hî (...)
 we [mề la-ya gũ]_{cosub.cl} we ning khri ne hî
 DEM somebody say-ALLO NF DEM 1P a.clan only EMPH
 ‘If someone says (anything), only we the Khri (...)’ [TA24_land_119 216]

3.5.1.4 Reflexive pronoun *ga*

The reflexive pronoun *ga* ‘self’ is co-referential with the preceding nominal referent and it may also occur anaphorically. The effect of the reflexive pronoun is pragmatic, as it reinforces the identity of an otherwise unlikely referent. This pronoun may occur as *ga-tyàng* (cf. *tyàng* ‘body’¹⁴) for certain speakers. Furthermore, it may often take the emphatic morpheme *hîq*.

In Example 57a, the referent’s son is highlighted, as both father and son are priests. Thereafter, in Example 57b the other interlocutor asks for clarification stating that the father was also a priest (‘oh, he himself was also a priest’). In Example 58 using the reflexive pronoun is pertinent as it makes it clear that the character in the story disemboweled himself.

- (57) a. cya ayûwaq ga mu gwók shyânabô
 cya ayûwaq ga mu gwók shyâ-na-bô
 3 son REFL also priest become-FACT-PFV.ALLO
 A: ‘His own son also became a priest.’
- b. oh cya ga mu gwók shyânabô
 oh cya ga mu gwók shyâ-na-bô
 oh 3 REFL also priest become-FACT-PFV.ALLO
 B: ‘Oh, so he himself also became a priest.’ [both TA23_talyu_84 511]

¹⁴The grammaticalization path *body* > *reflexive pronoun* is typologically common (cf. Kuteva et al. 2019, 78).

- (58) *khlâi we bēdângkũ (...)* *cya ga thĩ kwíkũna*
 [khlâi we bē-dâng-kũ]_{main.cl} [cya ga thĩ
 intestines DEM disembowel-COMPL.I-COMPL.II 3 REFL EMPH
kwí-kũ-na]_{main.cl}
 dig.a.hole-COMPL.II-FACT
 ‘He took out his intestines, he himself dug a hole.’ [TA22_CLAN2 200 385]

The reflexive pronoun may also occur anaphorically if the referent is unstated, as in Example 59. A similar effect is achieved in 60, as the speaker wishes to highlight that two people are brothers. This is a verbless clause (cf. §7.1) in which the complement is replaced by the reflexive pronoun.

- (59) *ga bôbô*
ga bô-bô
 REFL go-PERF.ALLO
 ‘He himself went.’ [281124_BAB_eli1 880]
- (60) *pû.mru ga na (...)*
pû.mru ga na
 siblings REFL INSIST
 ‘They themselves were brothers, right? (...)’ [TA24_land_270 500]

As mentioned above, some speakers employ a variant *ga-tyàng*, as in Example 61. It is not clear to me which dialect(s) employ this form. In this case, the speaker is from the Deb valley, Anjaw.

- (61) *ga.tyàng manûmtyûya*
ga.tyàng manûm-tyû-ya
 REFL wash-about.to-ALLO
 ‘I myself wash (the dishes).’ [201122_JA_eli1 2308]

The reflexive pronoun along with the emphatic particle *hîq* can also convey a sense of autonomy or independence of the referent, who might otherwise lack these abilities, cf. Example 62.

- (62) *garêi ga hîq keya*
garêi ga hîq ke-ya
 car REFL EMPH drive-ALLO
 ‘The car drives by itself.’ [171024_JA_eli 309]

Finally, a reduplication of *ga* ‘self’ adds a distributive sense ‘separately’ as in Example 63. Note the tonal change mid (˩) to high (˥) as discussed in §2.5.7.

- (63) *nyu hã gá.ga bôkêq*
nyu hã gá.ga bô-kêq
 2S 1S separately go-HORT.PL
 ‘Let us go separately!’ [281124_BAB_eli1 1100]

3.5.1.5 Quantifier *mé*

The quantifier *mé* with the meaning ‘all’ can also occur as a pronoun with the same meaning. It is likely related to the noun *mê* ‘person’, which goes back to PTB **mi* (Matisoff 2003, 661), a widespread etymon in the language family. The quantifier precedes the modified referent as in Example 64, where *mé* occurs before *kalyaw* ‘field’. As a pronoun, it means ‘all, everyone, everything’, cf. Example 65. When used pronominally, it commonly occurs in a reduplicated form, as in Example 66, which also displays the negation of the pronoun.

- (64) *cya karàng prûwbi mé kalyaw tshêng la*
[cya karàng prûw-bi]_{main.cl} [mé kalyaw tshêng la]_{main.cl}
 3 clear set.ablaze-PFV all field be.wet HSAY
 ‘He cleared (his field) (while) all (other) fields were wet, they say.’ [TA23_tayang_clan 63 395]

- (65) mé gru lyüwdângkũya ã yüm bô (...)
 [mé gru lyüw-dâng-kũ-ya]_{main.cl} [ã yüm bô]_{sub.cl}
 all burn come.up-COMPL.I-COMPL.II-ALLO be.there NEG TEMP.SUB
 ‘Everything was burned down. Nothing being there (...)’ [TA23_new_myth 27 162]

- (66) we mé.mé nyu bô yüm
 we mé.mé nyu bô yüm
 DEM all FOC go NEG
 ‘(But) not all went.’ [TA24_land 141 259]

3.5.1.6 Other indefinite pronouns

Tawrã has several indefinite pronouns to mean ‘something’. Most of these pronouns are morphologically complex, as they occur with the interrogative pronoun *ja* ‘what, something’ and *kadé* ‘how’, such as *ja-klu*, *ja-darâi* and *ja-palang/kadé-palang* all meaning ‘something’. The pronoun (*sim*)*thâ* ‘something’ also occurs as *sim-p(a)lang*. These forms are remarkably variable and speakers mold them creatively without altering their meaning as indefinite pronouns.

In Example 67, the complex indefinite pronoun *ja-klu* provides the sense of ‘something’. These pronouns can also be negated with the standard negator and the adverb *mu* ‘also, even’ to render ‘nothing’, as in Example 68.

- (67) natya làng kasâ tüpwi la âm ewe ja.klu gê düya tüpwi hanaho la
 na-tya làng kasâ tüpwi la âm [ewe ja.klu gê
 1.POSS-grandfather COLL know surely HSAY yes DEM something one
 dü-ya tüpwi hana-ho]_{compl.cl} la
 need-ALLO surely come-STAT HSAY
 ‘Our grandparents knew that he had come with something (i.e. against us), yes, they say.’ [TA22_CLAN2_151 280]

- (68) (...) ja.darâi mu lü blúm ãna yüm bô gũ (...)
ja.darâi mu lü blúm ã-na yüm bô gũ
 something also spare LOC.NMLZ be.there-FACT NEG TEMP.SUB NF
 ‘(...) being nothing else left to burn (...)’ [TA23_newmyth_21 127]

The indefinite pronouns also have other functions: they can occur as fillers (cf. §9.3) and they may convey the meaning that a list or enumeration continues in a similar manner (cf. English *and so on*). In Examples 69 and 70, the pronouns (*plang* and *simthâ*) act as fillers. In Example 71 the pronoun signals the continuation of a list (*and so on*). The indefinite pronoun *palang* also occurs as a tag for rhetorical questions (cf. §7.3.1.2.1).

- (69) krõ làng mlầ takyã kadé.nyu *plang*
*krõ làng mlầ takyã kadé.nyu *plang**
 a.clan COLL village LOC.ON how FILL
 ‘This Kro village up there, what to say.’ [TA22_clan_75 242]

- (70) mthâ we rôho ni
simthâ we rô-ho ni
 FILL DEM catch-STAT INSIST
 ‘This, what to say, is caught, right?’ [TA22_chicken_thief 54 211]

- (71) “shya tahù? ning tahù! shya tahù?” ja.palang
shya tahù ning tahù shya tahù ja.palang
 who group 1P group who group and.so.on
 ‘(Saying) “whose clan? Our clan! Whose clan?” and so on.’ [TA24_land_127 229]

3.5.2 Particles

Particles are understood in this grammar as usually small, non-inflectional functional morphemes that are syntactically dependent (i.e. they cannot head phrases), occur at the

right edge periphery of the clause and take multiple types of phrases under their scope. Based on their distribution, we can distinguish between five sub-types of particles: modal particles (§3.5.2.1), fillers (§3.5.2.2) and interjections (§3.5.2.3), postpositions (§3.5.2.4) and non-final markers (§3.5.2.5). This section provides an overview of the particles and their sub-types in terms of morphosyntax and semantics. §9 is devoted to the description of particles and offers a fuller discussion.

3.5.2.1 Modal particles

Modal particles have propositional scope over phrases and clauses. They are diverse in the types of meanings that they convey, modifying the epistemic, evidential and pragmatic reading of the affected clause/phrase. Some particles carve out very specific semantic niches and are less common than others. Table 3.27 displays all modal particles found in the language along with their references in the text.

Table 3.27: Overview of modal particles

Modal particle	Meaning	Reference
<i>matyô</i>	maybe	§9.2.1
<i>dě</i>	inferential evidential INFR	§9.2.2
<i>pà</i>	‘unfortunately’ ALAS	§9.2.3
<i>khrâk</i>	non-visual evidential EVI	§9.2.4
<i>thâ</i>	‘sadly/ for all I care’	§9.2.5
<i>delé</i>	‘it seems’	§9.2.6
<i>kho</i>	‘controversial’	§9.2.7
<i>ki/kale/kalai</i>	counter-expectational COUNT.EXP	§9.2.8
<i>ta</i>	advisative ADVS	§9.2.9
<i>la</i>	hearsay HSAY (also verb ‘say’)	§10.7.1
<i>shyüw</i>	observational negative	§6.4.2

3.5.2.2 Filler

Fillers mark a pause in speech, commonly as hesitation, giving the speaker more time to think about what s/he is saying. The most common filler is *mane* ‘FILL’, a loan from Assamese. Fillers are discussed in §9.3.

3.5.2.3 Interjections

The category of interjections presents a challenge for the classification of word classes since they do not represent a syntactically uniform category. Furthermore, as discussed by Sasse (2015, 209), they share properties both of function and content words (cf. §3.1). Morphologically, interjections cannot take any suffixes. Syntactically, they have the ability to form whole sentences depending on the semantic profile of the interjection, although they cannot form syntactic clause constituents. The main function of interjections is to express the attitude and intention of the speaker, e.g. a wish to do something (*chôq!* ‘let’s (e.g. go)’ or his/her consent (*âm* ‘yes!’). Interjections are further discussed in §9.4

3.5.2.4 Postpositions

Postpositions affect the whole NP (commonly adjuncts) and have a fixed, post-NP position (hence their name). A few postpositions also occur as non-final marker. Postpositions are fully described in §9.5. Table 3.28 displays all postpositions found in the language.

Table 3.28: Overview of postpositions

Postpositions	Reference	Notes
General locative/instrumental <i>gũ</i>	§9.5.1	cosubordinated marker
‘inside’ locative <i>kẽ</i>	§9.5.2	
‘atop’ locative <i>kyã</i>	§9.5.3	
‘with’ <i>dũ</i>	§9.5.4	co-temporality marker
‘for’ <i>(a)ba</i>	§9.5.5	
comparative ‘like’ <i>dahi</i>	§9.5.6	
NP quantifiers ‘each, every’ <i>(la) tsâi, (la) mang/müng</i> and <i>tyô</i>	§9.5.7	
additive <i>ma</i>	§4.4.1	

3.5.2.5 Non-final markers

Non-final markers are particles with scope over whole clauses. Their function is to turn clauses non-final, i.e. embed them in a multi-clause construction. §9.5 provides an overview of these particles but they are better described in functional terms in §10. Table 3.29 displays all non-final particles found so far in the language.

Table 3.29: Overview of non-final markers

Construction(s)	Device(s)	Reference	Notes
Cosubordination	<i>gũ</i>	§10.3	also locative/instrumental
Co-temporality	<i>dũ</i>	§10.3.4.1	also ‘with’
Temporal and conditional subordination	<i>bobyẽ</i>	§10.4.1	cf. <i>bô</i> ‘go’ and <i>byẽ</i> TOP
Unrealized alternative and contra-factual cond.	<i>(go)ji</i>	§10.4.2.3	<i>ji</i> ‘do, make’
- <i>thĩ</i> group			also emphatic marker
Temporal and hypothetical subordination	<i>bũ-thĩ</i>	§10.4.3.1	
‘as if’	<i>da-thĩ</i>	§10.4.3.2	
Concomitant action	<i>gũ-thĩ</i>	§10.3.4.2	

3.5.3 Numerals

Numerals are uninflected independent lexemes used to count and quantify referents. The numeral ‘one’ is a bimorphemic lexeme *khing-gê*. The *khing(-gê)* full form is used to count and modify nouns while the reduced, classifier form *gê* only surfaces in the classifier phrase (cf. §5). Similarly, numerals two *(kaq)îng* and three *(ka)sang* also have a shorter classifier form, although this feature seems to be falling out of use, as it is not maintained by all speakers. This special morphological characteristic of numerals sets them apart from other word classes.

Table 3.30 presents cardinal numerals 1-10.

Table 3.30: Numerals 1-10

numerals	gloss	numerals	gloss
<i>khing-gê</i>	1	<i>tahrô</i>	6
<i>kaqîng</i>	2	<i>wê</i>	7
<i>kasang</i>	3	<i>lyûm</i>	8
<i>kaprêi</i>	4	<i>kûnyâng</i>	9
<i>mânga</i>	5	<i>halâng</i>	10

Numerals 2-4 have an initial syllable *ka-* that is semi-productive for numerals 2-3 (cf. §3.7.3) and completely fused with the root for numeral ‘four’.

Numerals above ‘ten’ are formed in the following manner: tens (base ten) are formed by adding *halâng* ‘ten’, thus ‘thirty’ is *kasang halâng* and ‘seventy’ *wê halâng*. Ones (base one) are formed by adding a numeral after the ten: *kasang halâng kaqîng* is ‘thirty-two’ and *kasang halâng wê* ‘thirty-seven’. The numeral for ‘one hundred’ is *malum* and ‘one thousand’ *rüzang*. Both have a position to the left of the tens/hundreds.

3.5.4 Demonstrative *we* and complex demonstratives

Tawrã has one demonstrative: *we* (cliticized form *e=*). It serves the pragmatic function of “focus[ing] the hearer’s attention on objects or locations in the speech situation” (Diessel 1999, 2). This demonstrative is discussed to the full extent in §4.3.3. Three discourse operators are diachronically derived from demonstrative *we* and thus termed ‘complex demonstratives’. These are: *webo*, *weli* and *we cuwe*. They are described in §10.8.

3.6 Reduplication

Reduplication is a morphological process that affects different aspects of the grammar with the main effect of adding a distributive sense, although many forms seem to be lexicalized. Aspects of reduplication have already been discussed in relation to word classes

(§3.2), where it is used as an argument for class membership. Verbal stem reduplication is handled separately in §6.1.1.4, because it is related to number. Reduplication of classifiers is discussed in §5.3. This section presents an overview of the reduplication of question words, adverbs and adjectives and NP/nouns. This grammar distinguishes (true) reduplication from echo formations (§4.2.3), which involve the partial reduplication of a noun.

3.6.1 Reduplication of question words and pronouns

Reduplication of question words involves repeating the whole root to achieve a distributive sense. This process only affects two question words *ja* ‘what’ and *shya* ‘who’, as displayed in Table 3.31.

Table 3.31: Reduplicated question words

Form	Meaning	Reduplicated form	Meaning
<i>shya</i>	who	<i>shya shya</i>	who (among many)
<i>ja</i>	what, something	<i>ja ja</i>	whatever, anything

The reduplication of question words and pronouns is a feature found in other TB languages across the Himalayan range (in Galo, cf. Post 2007, 288; in Duhumbi (Kha-Bwo, North East India), cf. Bodt 2020, 116; in Kham (Magaric, Nepal), cf. Watters 2009, 172). An equivalent pattern is found in Indo-Aryan (henceforth IA), e.g. in Nepali *ko* ‘who’ and *ko ko* ‘who (out of many)’. However, in IA, it affects more question words, like *where* and *when* (also compare David 2015, 175). It cannot be ruled out that this feature has been borrowed from IA to Tawra.

Pronouns are not commonly reduplicated except for *mé* ‘all, everything’, that often surfaces as *mé.mé* (§3.5.1.5) and *ga* ‘self’ as *gá.ga* ‘individually’ (§3.5.1.4).

3.6.2 Reduplication of adjectives

Adjectives can be reduplicated to increase or intensify their meaning, as discussed in §3.2.4, but this is seldom employed in the language. Adjectives *gatyô* or *pütyô* meaning

‘different, other’ are often reduplicated to obtain a distributive effect (‘different from each other’)¹⁵. In Example 72, the narrator explains the origin of different religions in the world. He concludes by saying that everyone is different. The reduplication of the adjective is consistent with a distributive sense (‘everyone is different’). Similarly, in Example 73 the reduplication of the adjective renders a distributive sense: the body of the anti-hero is chopped to smithereens (lit. ‘small small’) and then eaten.

(72) *we gatyô gatyô*
DEM different different

‘(That is why) they are all different.’ [TA22_origin_112 326]

(73) (...) *katyê katyê ne thâya shyâ* (...)
katyô katyô ne thâ-ya shyâ
small small only eat-ALLO happen

‘(...) (into) small small pieces (and) eat it (...)’ [TA23_tayang_clan 37 222]

3.6.3 Reduplication of adverbs

Adverbs are not commonly reduplicated except for *ideophones*, as discussed in §3.4.1.2. Some adverbs, however, often occur in a reduplicated form like *bú-bú* ‘earlier’, which does not occur as **bú*. In other cases, the reduplication of an adverb does not yield a difference, cf. *rã* ‘late’ and its reduplicated form *rã-rã* (note the tonal change discussed in §2.5.7).

3.6.4 Reduplication of nouns and NPs

The partial reduplication of nouns or echo word formations is discussed in §4.2.3. The reduplication of classifiers is discussed in §5.3. The reduplication of nouns and NPs often yields a distributive effect, i.e. it includes several members. It should be noted

¹⁵The reduplication of words meaning ‘(be) different’ is common throughout South Asian languages regardless of their phylogenetic affiliation cf. Kurux (Dravidian) *p^huʔaː-p^huʔaː* ‘different’ (Kobayashi and Tirkey 2017, 710) or Pnar (AA) *dzar dzar* ‘different from another’ (Ring 2015, 75).

that, comparatively, the reduplication of nouns is widespread among IA languages (as in Hindko, cf. Bashir and Connors 2019, 577f) while it is more restricted among TB languages (Shaowu, cf. Ngai 2021, 109; or Karbi, cf. Konnerth 2014, 212). In the case of Tawrã, although the distributive effect is available, it is seldom exploited in speech. Nouns affected by reduplication commonly trigger the tonal change characteristic of nominal modification, discussed in §2.5.7, in which low and mid tones become high tone. In Example 74, the noun *mlă* ‘village’ is reduplicated to imply all villages.

- (74) *mlǎ́ mlǎ̀ kě*
mlǎ́ mlǎ̀ kě
 village.MOD village LOC.INS
 ‘in each village’ [field notes 21102024]

Relator nouns may also be reduplicated to yield a continuous sense. In Examples 75 and 76, it is understood that the action is distributed in time, i.e. it is supposed to continue for some time, e.g. during the whole walk (cf. Zemp 2018, 145).

- (75) *plũng plũng dyângya*
plũng plũng dyâng-ya
 behind behind chase-ALLO
 ‘S/he chases from behind.’ [291024_STBT_eli1 161]

- (76) *nya nya bôna*
nyá nya bô-na
 in.front.MOD in.front go-IMP
 ‘Go in front (e.g. in a narrow path)!’ [291124_BAB_eli1 1215]

A whole NP may be reduplicated to yield a distributive effect. In Example 77, a NP composed of a possessive pronoun and a head noun is repeated to indicate sets of discrete groups, such as families.

- (77) (...) *nyu khandan nyu khandan nyu khandan*
 2s family(<IA) 2s family(<IA) 2s family(<IA)
 ‘(He created) your family, your family, your family (i.e. different families).’
 [TA22_origin_104 303]

There are two cases of nouns that are commonly reduplicated: *mề* ‘person’, which is reduplicated as *mề mề* to mean ‘all men, everybody’ and nouns referring to time, like *hông* ‘day’, which becomes *hông hông* ‘daily’ (a feature also found in Sinitic languages, cf. Ngai 2021, 109f.). In both cases, a distributive sense led to an adverbial meaning ‘everyone’ and ‘every day’. In Example 78, the action is described as occurring daily (*hông hông*).

- (78) *bîya we tadâwe mề a dyềdâng la hông hông dyề la*
[bîya we tadâw=e mề a dyề-dâng la]_{main.cl} [hông hông
 past DEM eagle=DEM person.MOD child catch-away HSAY day REDUP
dyề la]_{main.cl}
 catch HSAY
 ‘In the past, this eagle caught the children of men, daily (lit. day day) did it caught (them).’ [TA22_origin_173 481]

3.7 Diachronic aspects of root morphology

Many lexical roots, particularly nouns and verbs, display a root structure that is synchronically derived from an etymological prefix + root collocation. Although in the modern language these lexemes are non-compositional, three characteristics suggest that these lexemes had complex roots that were reanalyzed as simplex roots: their syllable structure (strictly iambic foot cf. §2.2.6), their semantic distribution and comparative data. It thus seems that these root prefixes had a now-defunct semantic value.

This section will describe the prefixes *a-*, *ha-*, *ka-*, *me-*, *ma-* *ta-* and *we-*. It will comment on how widespread these forms are, hypothesize about their now-defunct semantic value and present a table with examples where each root occurs.

These roots do not share many features among them apart from being an (unproductive) initial non-stressed syllable. If we assume that all prefixes had a function at an earlier stage of the language, then we must imagine a language with an overwhelming number of monosyllabic roots - since prefixed roots would have turned monosyllabic roots disyllabic. Furthermore, the merger of the old Prefix + Root would have led to a change in the foot structure, now dominated by iambic words. In this sense, it is noteworthy that there are only two *true* prefixes in modern Tawrã: the kinship prefix *na-* (§4.5.3) and the prefix *ka-* (cf. §3.1).

Table 3.32 illustrates all defunct prefixes that will be described in this section. The list of possible semantic domains is by no means exhaustive, as it is based on the distribution of modern forms; it must be assumed that semantic and phonological shifts have altered the whole system after the old prefixes were reanalyzed as part of the root. Where relevant, I will mention connections to literature on related languages and the reconstruction of Proto Tibeto-Burman.

Table 3.32: Prefixed roots: Overview

Root prefix	Broad semantic group	Examples
a-	Body parts	<i>ahrûk</i> ribs
	Flora & Fauna	<i>anâtsi</i> snail
	Kinship terms	<i>anáng</i> brother-in-law
	Directions and deictic terms	<i>atyû</i> north
	Diverse verbs	<i>amé</i> follow
	Material culture	<i>alaîkã</i> bow
ha-	Flora & Fauna	<i>habra</i> foxtail millet
	Body parts	<i>hanyâgum</i> nose
	Diverse verbs	<i>habûm</i> stretch
ka-	Small animals and insects	<i>kame</i> type of squirrel
	Adverbs, classifiers and numerals	<i>kahîng</i> CLF:daytimes
me-	Diverse verbs (two attestations)	<i>mesê</i> kill <i>mese</i> inquire
ma-	Diverse animals	<i>macyû</i> sambar deer
ta-	Fauna & Fauna	<i>tadâw</i> eagle
	Bodies of water and related terms	<i>talû</i> Dibang river
	Material culture	<i>tahwê</i> strap
we-	Psychological verb (two attestations)	<i>wetyâng</i> remember

Prefixed roots are a pervasive feature in several branches of TB (Matisoff 2018; Wolfenden 1929, 12f.), which suggests that they could already have existed at an early stage of TB (LaPolla 2003). Particularly in Northeast India, many TB languages seem to have retained a similar system at the root level (for Bulu Puroik, cf. Lieberherr 2017, 90f; for Mongsen Ao, cf. Coupe 2008, 246f; for Karbi, cf. Konnerth 2014, 165f; for Galo, cf. Post 2007, 102f). In general terms, leaving certain particular cases aside, it can be stated that these prefixes have either been replaced or blended into the root and thus lost the function and productivity that they once had. An important factor that might have led to such a development is their unstressed nature (cf. Benedict 1972, 96, 103; Sun 1993,

300f.), at least in the TB languages of Northeast India and Nepal.

3.7.1 a- root prefix

The *a-* prefix is found across a wide array of semantic fields including fauna, body parts, demonstrative and deictic terms. Many of these semantic fields are shared with the prefix *ha-*. Some speakers drop the *a-* root prefix in casual speech (cf. §2.2.5). Dropping the initial *a-* leads to an increase in the homophony in the language, cf. (*a*)*shyâ* ‘mithun’ and *shyâ* ‘happen, become’.

A reconstructed **a-* prefix is described by Matisoff (2018) as a multi-functional prefix often associated with kinship terms, body parts, genitive constructions and 3rd person indexation. These semantic and functional fields have been subsumed under the idea of *inalienable possession*. The author, however, recognizes that the fossilization of this prefix may have allowed it to spread through analogy to multiple semantic fields. If we aim at connecting this old prefix with the Tawrã forms, we would have to account for all its diverse occurrences and/or assume that the modern *a-* prefix root has more than one origin and that these have converged into this polysemantic prefix.

An exploration of neighboring Kera’a reveals some information. The word list produced by Bufan and Qingxia (1992) of Chinese (Tibetan) Kera’a lists similar form for some body parts (*a*⁵⁵ *huŋ*⁵⁵ *koŋ*³⁵ ‘nail’ and *a*⁵⁵ *po*⁵⁵ ‘lungs’) and verbs (*a*⁵⁵ *hu*³⁵ ‘ask’, *a*³¹ *po*³⁵ ‘open’ and *a*⁵⁵ *ku*⁵³ ‘steal’), as illustrated in Table 3.33.

Table 3.33: Comparison of Kera’a and Tawrã. Data from Kera’a is from Bufan and Quinxia (1992)

(Tibetan) Kera’a	Tawrã	Example
<i>a</i> ⁵⁵ <i>huŋ</i> ⁵⁵ <i>koŋ</i> ³⁵	alhûng	nail
<i>a</i> ⁵⁵ <i>po</i> ⁵⁵	apó	lungs
<i>a</i> ⁵⁵ <i>hu</i> ³⁵	ahu	ask
<i>a</i> ⁵⁵ <i>ku</i> ⁵³	akâp	steal
<i>a</i> ³¹ <i>po</i> ³⁵	pó	open

In terms of material culture, we find one sound cognate *a*³¹ *liŋ*³⁵ ‘road’ and two

problematic ones: *i*⁵⁵*pe*⁵⁵ ‘arrow’ and *i*⁵⁵*li*⁵⁵ ‘bow’. An initial *i*- in Kera’a is either due to the loss of an onset consonant (cf. *ili* ‘pig’ in Midu dialect and *bili* in Mithu (Reinöhl 2022, 245), cf. Tawrã *bülik*) or an epenthetic *i/e*- prefix: *i*⁵⁵*gu*⁵⁵ ‘ash’ in Chinese (Tibetan) Kera’a (Bufan and Qingxia 1992) and *gû* ‘ash’ in Tawrã. For these two words (‘arrow’ and ‘bow’), the second option, i.e. the epenthetic insertion, is most likely. However, we don’t know whether there was an *a*- root prefix at all in Kera’a and whether the epenthetic *i/e*- prefix replaced it.

3.7.1.1 A brief note on *a-/*ha- prefixes

The distinction between the old root prefixes *a-/*ha- in the modern language is an unsettled matter: Both prefixes are found in similar semantic domains in the modern language (e.g. body parts and diverse verbs), there is an ongoing change in the language in which light syllables like *ha*- are sometimes reduced to *a*- or completely eroded (§2.2.5). Furthermore, word lists from other dialects suggest that this merge has been going on for a long time, as some *a*- lexemes occur with an initial *ha*- in the Tibetan dialect listed by Sun (1991), e.g. ‘lungs’ as *apó* in my data and *xa*³¹*po*⁵⁵ in the Tibetan word list¹⁶. It thus seems that these suffixes display a high level of intermixing and confusion, which is to be expected considering that they have merged with roots and lost their semantic meaning.

An interesting parallelism has been described for the Tani languages. Post and Sun (2017, 324) discuss a breathy/glottal onset distinction among vowel-initial words that are maintained in some languages, e.g. Northwestern Galo *ʰáa*- ‘come; set (sun)’ and *ʔáp*- ‘shoot’, while other Tani languages have merged this distinction. The analogy, however, is not complete, as the glottal onset goes back to an actual glottal stop while the breathy form corresponds to an old fricative segment, most often *h-, but sometimes *f-, *v- or *z- (Post and Sun in prep.). It remains to be established whether an equivalent distinction can be postulated for the *a-/*ha- forms in Tawrã.

3.7.1.2 a- root prefix body parts

In Table 3.34 we find several body parts with the *a*- prefix.

¹⁶Note that initial <x> in the Tibetan data from Sun (1991) corresponds to <h> in my data.

Table 3.34: a- root prefix: body parts

Tawrã	Meaning
<i>acîng</i>	navel
<i>adyângra</i>	back of neck
<i>ahrûk</i>	ribs
<i>alhûng</i>	finger nail
<i>apó</i>	lungs
<i>aprû</i>	arm
<i>asâ</i>	artery, vein
<i>atyâng</i>	thumb
<i>atyôka</i>	palm

3.7.1.3 a- root prefix flora & fauna

In Table 3.35 we find diverse terms for flora and fauna. The noun *ashyâ* ‘mithun’ (*Bos frontalis*) is most enigmatic, as it does not fit into the semantic group of *small animals*. Due to the cultural importance of the mithun for many peoples of Arunachal Pradesh (cf. Blench and Post 2014, 78), we should expect some degree of irregularity in this word.

Table 3.35: a- root prefix: flora & fauna

Tawrã	Meaning
<i>acûne</i>	firefly
<i>adâityang</i>	scorpion
<i>agrû</i>	freshwater prawn
<i>akhrûng</i>	type of wild banana
<i>alâwka</i>	type of bird
<i>alô</i>	garlic
<i>anâtsi</i>	snail
<i>apâwgra</i>	spider (generic)
<i>arêqto</i>	type of eel (probably <i>spotted snakehead</i>)
<i>ashyâ</i>	mithun

3.7.1.4 a- root prefix: kinship terms

Several kinship terms in Tawrã have an initial *a-* prefix. This seems to be an old form, as it is found in many languages from Northeast India and beyond (cf. Wolfenden 1929, 101f). It is noteworthy that the word for ‘child’ is *a*. Perhaps it was reanalyzed from the said prefix and further expanded to kinship terms.

Some Tani languages also have a vocative/intimate prefix that could potentially be related, cf. Galo *áa* and *áa~âa* (cf. Post 2007, 672). Table 3.36 displays a number of kinship terms with *a-* root prefix. Kinship terms are described in §3.3.1.2.

Table 3.36: a- root prefix: kinship terms

Tawrã	Meaning
<i>anyu</i>	family, relatives
<i>anyuwa</i>	nephew
<i>atyô</i>	younger daughter
<i>ayĩ</i>	granddaughter
<i>a-yûwak</i> / <i>a</i>	son

3.7.1.5 a- root prefix: directions and deictic terms

A small number of words for directions and deictic terms have a *a-* root prefix.

Table 3.37: a- root prefix: directions and deictic terms

Tawrã	Meaning
<i>(a)ci/ashi</i>	horizontal axis (east/west)
<i>amã</i>	south
<i>atyu</i>	north
<i>alú</i>	up there

3.7.1.6 a- root prefix: diverse verbs

A large number of verbs in Tawrã have an initial *a-* root prefix. Some of these verbs are displayed in Table 3.38:

Table 3.38: a- root prefix: diverse verbs (excerpt)

Tawrã	Meaning
<i>abrêi</i>	be bifurcated, be double
<i>agã</i>	set up (as of a trap)
<i>ahrú</i>	shock, be shocked
<i>ahu</i>	ask
<i>akâp</i>	steal
<i>amé</i>	follow
<i>apüü</i>	dry (e.g. cloths)
<i>asáng</i>	grow
<i>atí</i>	leave, be lost
<i>athûk</i>	notch

3.7.1.7 a- root prefix: material culture

The root prefix *a-* is found in several terms in Tawrã including tools and man-made structures, here subsumed under the ample term *material culture*. Some of these terms are displayed in Table 3.39:

Table 3.39: a- root prefix: material culture

Tawrã	Meaning
<i>akâ</i>	granary
<i>alaĩkã</i>	bow
<i>alô</i>	hanging bridge
<i>alũ</i>	big mortar
<i>alyüm</i>	road, way
<i>alhũng</i>	nail
<i>apũ</i>	arrow
<i>atyũm</i>	rack (over <i>côri</i>)

3.7.2 ha- root prefix

The *ha-* prefix is widespread in the language. As discussed in §3.7.1.1, speakers commonly drop the initial *h-*, thus *harèi* ‘root (for a plant; generic)’ is also pronounced *arèi*. Certain lexemes, e.g. *hatũm* ‘sky’ lose the entire prefix in casual speech (*tũm*). In rare cases, there is a free variation of initial *ha-* and *ma-*, e.g. *hamyú/mamyú* ‘after, later’ (cf. §2.2.5). The *ha-* prefix bears many similarities with the *a-* prefix described above in terms of the broad semantic field that it encompasses.

3.7.2.1 ha- root prefix: flora

Table 3.40 displays some terms for flora with an initial *ha-* root prefix.

Table 3.40: ha- root prefix: flora and related terms

Tawrã	Meaning
<i>hablâng</i>	type of banana
<i>habra</i>	buckwheat
<i>hajîng</i>	ginger
<i>hamûng</i>	be ripe
<i>hare</i>	banyan tree
<i>harèi</i>	root
<i>haro</i>	branch
<i>hazúm</i>	classifier: bunches (of lychee, grapes)

3.7.2.2 ha- root prefix: body parts

Table 3.41 displays some body parts with the initial *ha-* root prefix.

Table 3.41: ha- root prefix: body parts

Tawrã	Meaning
<i>habrâng</i>	finger
<i>hanyâgum</i>	nose
<i>haprû</i>	hand
<i>harû</i>	backbone, spine
<i>hawrêi</i>	blood
<i>halêi</i>	scale (of fish)

3.7.2.3 ha- root prefix: diverse verbs

Table 3.42 displays some diverse (i.e. uncategorizable) verbs that share an initial *ha-* root prefix.

Table 3.42: ha- root prefix: various verbs (excerpt)

Tawrã	Meaning
<i>habai</i>	be able
<i>habûm</i>	stretch, pull
<i>hadâ</i>	borrow, loan
<i>hadû</i>	fight
<i>hagû</i>	guard, watch
<i>hana</i>	come
<i>hara</i>	be crooked
<i>hatû</i>	help
<i>hatyâ</i>	be lazy
<i>hawa</i>	waste time, delay

3.7.3 ka- root prefix

The *ka-* prefix is found in some small animals and insects, classifiers and adverbs. These semantic groups will be discussed separately.

3.7.3.1 ka- root prefix: small animals and insects

The root prefix *ka-* is found in a handful of small animals and insects. This prefix is also found in the adjective *katyêq* ‘small’ and also in the adverb *katyêne* ‘less’. Table 3.43 displays several examples.

Table 3.43: ka- root prefix: small animals and insects

Tawrã	Meaning
<i>kacêk</i>	type of mole
<i>kame</i>	type of squirrel
<i>tâ kapá</i>	type of fish (probably <i>banded gourami</i>)
<i>kapê</i>	(aquatic) leech
<i>kaprêi</i>	type of bird
<i>kaphrûlyu</i>	butterfly (generic)
<i>kasô</i>	type of (jungle) rat
<i>katài</i>	type of grasshopper
<i>katyâw</i>	crab
<i>katyê</i>	tick
<i>kazû</i>	type of squirrel

This root prefix has at least two possible cognates among the Tani languages: A prefix **ka-* for rodents reconstructed for Proto-Tani (Post and Sun in prep.), cf. Milang *karok* ‘rodent (variety)’ (Modi 2017, 84) and a *ta-* prefix, reserved for insects and small animals, among others, e.g. Galo *taká* ‘squirrel’ (Post 2007, 84).

Further afield, Matisoff (2003, 134-139) has discussed a **g-* ‘animal prefix’. Attempts to reconstruct a semantic value for this velar prefix face the challenge that the prefix is found in a “number of semantically disparate nouns” (idem, 134). However, it does include some animals¹⁷ and numerals ‘two’ and ‘three’ (cf. §3.7.3.2). In the case of the animal prefix **k-* in Lolo-Burmese, Matisoff (2003, 138-139) suggests that it could have been an AA loan, probably related to Proto AA **kuun* ‘child’ (#1127a-b Shorto 2006). In Proto Palaungic (AA), we find some animals with an initial *kə-*, cf. **kənee?* ‘rat, mouse’ **kətəl~kətɿl* ‘hive (animal)’ or **kəthaam* ‘crab’ (adapted from Sidwell 2015). Although modern Palaungic is spoken 500 kilometers away from the Mishmi Hills (cf. §1.1.3), it is noteworthy that this prefix has a similar semantic value in Tawrã. The diachronic

¹⁷Cf. in Japhug the derivational prefix *qa-*, e.g. on the noun *qapyɿmtuɿmtu* ‘hoopoe’ (from *pya* ‘bird’ and the (semi) reduplicated form of the noun *u-mtu* ‘crest’) (Jacques 2021, 167).

development of this prefix should be further investigated.

3.7.3.2 *ka-* root prefix: adverbs, classifiers and numerals

A small number of adverbs, temporal classifiers and numerals have an initial *ka-* syllable. We may assume that this form used to have a function as a derivational marker, e.g. an adverbializer, crafter of classifiers and/or numerals. Syntactically, the root prefix *ka-* retains some syntactic independence and functions, e.g. as a prefix for certain classifiers (e.g. years *kanáng*) and numerals ('two' and 'three'). For classifiers, the prefix disappears when modified: in Example 79a, the classifier occurs with the prefix while in Example 79b, the prefix is dropped, for the classifier is modified by an adjective. Numerals lose the prefix when modifying a classifier and the 3rd person pronoun *cya*, as has been discussed in §3.4.2.

- (79) a. *kanáng gê*
CLF:year one
'one year'
- b. *kapyù náng*
next CLF:year
'next year'

Table 3.44 displays all adverbs, classifiers and numerals with the *ka-* root prefix.

Table 3.44: ka- root prefix: adverbs and temporal classifiers

Tawrã	Meaning	Comment
<i>kanehîng</i>	two days after tomorrow	cf. <i>hîng</i> ‘day’
<i>kasahîng</i>	three days after tomorrow	idem. + <i>kasang</i> ‘three’
<i>kalyâng</i>	classifier of evenings	cf. <i>lyâng</i> ‘evening, half’
<i>kapyéi</i>	classifier for <i>handfuls</i> (e.g. of mustard seeds)	*pyéi does not exist
<i>kahîng</i>	classifier for counting day-times	cf. <i>hîng</i> ‘day’
<i>kanáng</i>	classifier for counting year	cf. <i>náng</i> ‘year’ (only when modified)
<i>kapyù</i>	next	
<i>kaqîng</i>	two	semi-productive prefix, cf. §3.5.3
<i>kasang</i>	three	idem.
<i>kaprêi</i>	four	

3.7.4 me- root prefix

The *me-* prefix is attested with only two verbs: *mesê* ‘kill’ and *mese* ‘inquire’ and it is often dropped. This prefix likely shares the etymon with the noun *mê* ‘person’, perhaps through a reanalyzed nominal incorporation (§6.5). It should be noted that these verbs can be used with non-human referents as well, as illustrated in Example 80.

- (80) *tyû mesêya majárei we ning hwě.tyângya*
tyû mesê-ya majárei we ning hwě.tyâng-ya
 chicken kill-ALLO cat DEM 1P find-ALLO
 ‘We found the cat(<IA) that killed the chicken.’ [030622_JO_eli1 654]

A clue to the origin of this prefix could be found in neighboring Kera'a as some noun roots carry an initial *mi-/me-*, that may also go back to the PTB etymon for 'person' (Matisoff 2003, 661), cf. Tawrã *mě* 'person'. As described by Reinöhl (2022, 250f.), this root prefix in the Kera'a words *miku* 'dog' and *meto* 'chicken' could stand for a mark of domestication. However, this could also have been an independent development, since these nouns do not have the said prefix in Tawrã.

3.7.5 ma- root prefix

Several names of animals and related verbs in Tawrã have an initial *ma-* prefix. The animals are neither all large nor all domesticated (in order to create an opposition with *wild* animals with *ta-* cf. §3.7.6.1). One possibility is that this prefix was initially related to sources of food and was later expanded to mean 'big animals' and related activities (like hunting and cooking). This could be related to the root for 'paddy, rice plant' *²a-ma in Proto-Tani (Post and Sun in prep.).

Table 3.45 lists all animals with a *ma-* root prefix (and related terms) so far found in the language.

Table 3.45: ma- root prefix: large animals

Tawrã	Meaning
<i>mabaw</i>	fox
<i>macyû</i>	sambar deer
<i>majéi</i>	buffalo
<i>malyû</i>	rhinoceros
<i>marû</i>	horse, mule
<i>matsâp</i>	cow, ox
<i>mazúm</i>	deer
<i>malûng</i>	fish (vt)
<i>manyu</i>	cook, prepare food
<i>mapyâ</i>	graze
<i>matyâ</i>	set free, release

3.7.6 ta- root prefix

The *ta-* prefix is firmly entrenched in the broad groups of flora & fauna, bodies of water (and related terms) and material culture. Since the presence of this prefix in these three diverse groups can hardly have one single explanation, we will discuss each semantic group in detail.

3.7.6.1 ta- root prefix: flora & fauna

The prefix *ta-* is common among animals, insects and related terms. The latter refers to parts of animals (e.g. *talũ* ‘feather’) or classifiers used in relation to animals, like *tarum* ‘CLF (employed for areas of grazing for animals)’. Among animals, it seems to occur mostly with nondomesticated animals. In this line, the lexeme *takwâg* ‘wolf’ deserves attention as it is composed of the prefix *ta-* + *kwâg* ‘dog’, seemingly a *wild dog*¹⁸. The word for ‘fish’ *tangâ* or *tâ* is noteworthy as it is a reflex of a widespread etymon of PTB **ŋa/ŋya* ‘fish’ (Matisoff 2003, 650) and it occurs as a class term (§3.3.1.3) for fish cf. *tâ kapá* ‘a fish (probably *banded gourami*)’ (cf. §3.3.1.3).

A similar root prefix has been described for Galo (Post 2007, 265f.), where it occurs, among other semantic groups, with lower animals, e.g. *taká* ‘squirrel’ or *taci* ‘crab’ and insects, e.g. *takà* ‘body louse’ or *taik* ‘head louse’ and in some flora terms e.g. *ta-gú* ‘wild mango’. Post goes on to reconstruct the value of proto-*ta-* as (*masculine*) *diminutive*, based on the presence of the prefix in certain masculine terms e.g. *taìi* ‘youngest son’¹⁹ and on the diminutive prefix for male names *ta-* which is most likely cognate with the said prefix. In Milang (Macro-Tani, North East India) there are two similar root prefixes: *ta-* and *taa-*. While the former occurs in varied sets of words (and thus seems to be the result of Tani-internal borrowing), the latter occurs in three semantic groups: names of birds, words related to the harvesting of the culturally important aconite (arrow poison) and certain terms related to flora (Modi 2017, 110-112).

Table 3.46 displays some terms for flora/fauna with initial *ta-* in Tawrà. All members of this group make up approximately fifty elements.

¹⁸This term seems to refer to the Himalayan wolf and not to the dhole (*Cuon alpinus*), which is called *taprũww*. Also Cf. Milang *ta-pun*² ‘wild dog’ (Post and Sun in prep.).

¹⁹Cf. the paradigmatic opposition ‘youngest daughter’ *ja-ii* Post (2007, 857).

Table 3.46: ta- root prefix: wild animals, related terms and flora

Tawrã	Meaning
<i>tabâw</i>	snake
<i>tabri-tagra</i>	big hunt
<i>tabrû-tsû^a</i>	hunter
<i>tabû/nabû^b</i>	maize (cf. Galo <i>təpə</i> ‘maize’) (Post 2007, 265).
<i>tabyê</i>	honey
<i>tachà</i>	type of banana
<i>tadâw</i>	eagle
<i>tadyá</i>	type of squirrel (probably <i>ratufa bicolor</i>)
<i>tahum</i>	bear
<i>takang</i>	takin, gnu goat
<i>takwîkã</i>	type of fish
<i>takwing</i>	mushroom (generic)
<i>takwâg</i>	wolf
<i>takwûi</i>	monitor lizard
<i>talà</i>	male bird
<i>talâ</i>	type of deer (probably <i>moschus leucogaster</i>)
<i>talõ</i>	wing
<i>talyâng</i>	cardamom
<i>tamê</i>	animal (generic) cf. Milang <i>tamon</i> ‘wild animals’
<i>tamẽ</i>	boar
<i>taméqĩng</i>	elephant
<i>tamû</i>	type of banana
<i>tamyá</i>	tiger
<i>tanũng^c</i>	crocodile
<i>tangâ/tã</i>	fish (generic)

Notes

^a The formatives of *tabrû-tsû* ‘hunter’ are not known. Perhaps *tabrû-* is related to *tabrâ* ‘hunt’ or *tabrê* meat.

^b Alternation *ta-/na-* is discussed in §2.4.3.2.

^c The difference between *tanûng* and *kà* ‘crocodile, gharial’ is not clear, for there is only one species found in the area, namely the *gharial* (*Gavialis gangeticus*). Perhaps the existence of two terms can be explained by misidentifying of the *Ganges river dolphin* (*platanista gangetica*), which might have reached Lohit district in the past (Chintan Seth, p.c.). Alternatively, one of these lexemes could be a loanword but this has not been proven so far.

3.7.6.2 ta- root prefix: bodies of water and related terms

This group includes names for bodies of water, elements that are semantically related to water bodies and some names of rivers. All terms are displayed in Table 3.47.

Table 3.47: ta- root prefix: bodies of water and related terms

Tawrã	Meaning
<i>talâ</i>	a mythical rain
<i>talû</i>	Dibang river
<i>tapyêk</i>	sand
<i>taphra</i>	river, stream
<i>taphrûsã</i>	bridge
<i>tatyû</i>	Digaru river
<i>tawì</i>	lake, pond
<i>tawrî</i>	mud

3.7.6.3 ta- root prefix: material culture

Many terms related to utensils and other man-made objects, subsumed under *material culture*, have a root prefix *ta-*. Table 3.48 illustrates some of these nouns. The total number of terms in my data ranges is approximately thirty.

Table 3.48: ta- root prefix: material culture

Tawrã	Meaning
<i>tacông</i>	furniture
<i>tachêi-giüüm</i>	bamboo figure used in <i>puja</i>
<i>taglú</i>	provisions
<i>tago/nago</i>	needle
<i>tahwẽ</i>	strap (general)
<i>taji</i>	bait (when fishing)
<i>takã</i>	type of pot
<i>tako</i>	book
<i>taku</i>	cylindrical trap for fishing (<i>digre</i> in Asm.)
<i>takhi</i>	medium size sack (for rice)
<i>takhrâi</i>	snare trap
<i>takhrẽ</i>	(walking) stick
<i>talû</i>	type of jewelry (a belt worn from the shoulder to the hips)
<i>talyâ</i>	ammunition
<i>tama</i>	medicine
<i>tanyâng-tazô</i>	weaving pattern
<i>tapâ</i>	spear, punja
<i>taplá</i>	(wooden) spoon
<i>tâplă</i>	hammer
<i>tapôm</i>	fishing rod
<i>taprôtyu</i>	stairs
<i>tapûm-darũ</i>	altar (where vigil is kept over a deceased person)
<i>tawra^a</i>	sword (generic)

Notes

^a Some terms take *tawra* ‘sword’ as the head of the compound (class terms, cf. §3.3.1.3), e.g. *tawrá-bro*

‘type of sword’, *tawrá-hablûm* ‘large sword’, *tawrá-jû* ‘type of sword’ and *tawrá-za* ‘knife’.

3.7.7 we- root prefix

The prefix *we-* is found in two verbs with a clear *psychological* sense (cf. *psycho-collocations* in Karbi in Konnerth 2014, 377). These verbs are *wesâ* ‘forget’ and *wetyâng* ‘remember’. The origin of this prefix is probably related to the noun *tawê* (or *wê*), ‘(psychological) mind, heart, opinion’ or the verb *wê* ‘be tired’. In other verbs, this formative still retains its initial stress (and thus would not qualify for this list), cf: *wêlang* ‘believe’ and *wêmara* ‘forgive’. The difference in stress pattern probably reflects different levels of reanalysis.

4 Noun phrase

This chapter is devoted to the structure of the noun phrase (henceforth *NP*). §4.1 discusses the structure of the NP, including its slot architecture and head. Nominal compounding and echo word formation are discussed in §4.2. §4.3 discusses NP-head modifiers, including adjectives, the classifier phrase, numerals and the demonstrative. Note that the relative clause is discussed in §10.5.2. The topic of §4.4 is the conjunction of NPs, both marked and unmarked. Finally, §4.5 describes nominal suffixes.

4.1 NP structure

This section addresses the ordering of constituents within the NP, i.e. its slot structure. §4.1.1 presents the maximal shape of a NP. The attributes of the head are discussed in §4.1.1.1. §4.1.2 presents two marked subtypes: NPs headed by a personal name (§4.1.2.1) and by a pronoun (§4.1.2.2). These can be considered marked or defective subtypes within the maximal structure. Headless NPs are discussed in §4.1.3.

4.1.1 Maximal structure

Table 4.1 illustrates the maximal shape of a NP in the language organized in slots, taking the **head noun** as the starting point. Other NP constituents may also take the head position, as will be discussed in §4.1.1.1. Note that no constituent occurs obligatory within the NP, not even the head noun. Crucially, the word order illustrated in Table 4.1 reflects a pragmatically neutral setting and is, therefore, the most common order.

The pre-head domain has only two slots. The leftmost constituent (slot nr. -3) corresponds to the emphatic demonstrative marker (§4.3.3). It is in parenthesis since it is only triggered in particular pragmatic settings. It occurs in tandem with the demonstrative

proper (slot nr. 3), forming a double demonstrative construction. No other constituent occurs *double* at NP level. Slot nr. -2 is occupied by the relative clause (§10.5.2), which commonly takes a pre-nominal position. Slot -1 corresponds to nominal modification of the type N-N, further discussed in §4.2.

Slot zero is reserved for the head of the NP. This position may be occupied by any noun (including proper names) or, otherwise, adjectives, the demonstrative and the ClP (including stand-alone numeral and bare classifiers). Other properties of the NP head are discussed in §4.1.1.1.

The post-domain includes four constituents: Adjectives (§4.3.1) take slot 1 and may also occur pre-nominally. The adverb *tyâng* ‘very’ forces a pre-nominal position when modifying an adjective. Slot 2 is reserved for the demonstrative *we* (§4.3.3), which may occur in a double demonstrative construction. Finally, slot 3 includes numerals and the classifier phrase (§4.3.2).

Table 4.1: NP headed by a noun

	[(emph. DEM)	[Rel.Cl]	N.mod-	Head	ADJ	DEM	Num [ClP]	]NP
slot nr.	-3	-2	-1	0	1	2	3	

4.1.1.1 NP head

The nominal head of a NP may be simple or complex. A simple nominal NP head may consist of a noun (including personal names), a pronoun, an adjective, a demonstrative or a classifier phrase. A NP headed by an adjective (§4.3.1), a demonstrative (§3.5.4) or a classifier phrase (§4.3.2) does not have considerable differences from a NP headed by a common noun (i.e. excluding personal names), i.e. it may take most of the NP-internal modifiers. The structure of a NP headed by either of these constituents (adjectives, demonstrative and classifier phrase) is discussed in each dedicated section. A NP headed by a personal noun is discussed in §4.1.2.1 and by a pronoun in §4.1.2.2.

NP complexity has two layers: An internally complex nominal head of a NP consists of two nominal lexemes that stand in one of two possible configurations: the preceding

noun may either modify the head noun thus forming a *hierarchical* or *asymmetrical* nominal compound (treated in §4.2.1) or two nouns may be juxtaposed without a clear head, forming a *symmetrical* compound (cf. §4.2.2). Furthermore, two or more NPs can also be conjoined either with the additive postposition *ma* (§4.4.1) or through unmarked apposition. Conjoined NPs are further discussed in §4.4. Table 4.2 illustrates NP complexity with *nominal head* complexity on one side and structural complexity on top.

Table 4.2: NP complexity

	simple NP	complex NP
simple nominal head	[N] _{NP}	[[N] _{NP} (ma) [N] _{NP} (ma)] _{NP}
complex nominal head	[N-N] _{NP}	[[N-N] _{NP} (ma) [N-N] _{NP} (ma)] _{NP}

4.1.2 Restricted subtypes

This section will go over two subtypes within the maximal NP structure defined in §4.1.1. These are NPs headed by a personal name (§4.1.2.1) and headed by a pronoun (§4.1.2.2).

4.1.2.1 NP headed by a personal name

A NP headed by a personal name has only one restriction compared to common nouns: the use of adjectives. Thus, Example 81 is not accepted in elicitation. There is, however, one exception, although the adjective must be pre-nominal: *jüm Câne* (true a.name), the ‘first man’. Further details of this name are discussed in §4.3.1.

(81) **[kutyemba daráng]_{NP}*

PN Adj

a.name big

Intended: ‘the big Kutyemba’

Personal names can take all other NP modifiers and markers, including the demonstrative (cf. Example 82), the collective marker (cf. §3.3.1.4) and the feminine suffix *-yã* (cf. Example 83).

- (82) we kutyemba cya lâlhak tapwi
 [we kutyemba]_{NP} [cya]_{NP} lâlhak tapwi
 DEM a.name 3 trick HAB
 ‘This Kutyemba used to cheat.’ [TA22_JAstory_51 167]

- (83) kabrũyà tüyà
 [[kabrũ-yà]_{NP} [tüyà]_{NP}]_{VLC}
 a.name-FEM lord
 ‘(Goddess) Kabru is a lady.’ [TA22_origin_19 67]

4.1.2.2 NP headed by a pronoun

A NP can be headed by a pronoun (including personal pronouns (§3.5.1.1) and interrogative pronouns (§3.5.1.2)). The only restriction that applies to pronouns is that they cannot be modified by an adjective (same as personal names, cf. §4.1.2.1). They can be modified by the demonstrative (as in Example 84) and the feminine marker (as in Example 85). The use of the collective marker (§3.3.1.4) is also allowed, although it is part of the paradigm of personal pronouns discussed in §3.25. In Example 86, the collective marker suffixed to the question word *shya* ‘who’ implies that the speaker has enough information to assume that the referent is non-singular.

- (84) (...) rô trûwa gũ cya we (...)
 rô trû-ga gũ [cya we]_{NP}
 catch meet-PL.I NF 3 DEM
 ‘(...) catching him, this (...)’ [TA22_SOstory_9 46]

- (85) cya-yà
 3-FEM
 ‘she’

(86) *shya làng?*

who COLL

‘who? (assuming it is more than one)’ [011024_STBT_eli2 773]

4.1.3 Headless NP

A NP may occur without an overt head. This is often the case when the head is recoverable from context or when it has been mentioned earlier in discourse. Omitted arguments in multi-clause constructions are discussed in §8.2.5.

The following examples are from a story in which tiger and man compete for who can finish a task first: the tiger must count a fleeing school of fish in a pond while the man must count the knots on a hanging bridge. In Example 87a, the tiger is trying to count the fish. In Example 87b, we learn that it cannot complete the task. There is in the first sentence no overt NP head (marked by a $\emptyset$) but we may assume the continuity of the tiger as the topic. In the same example ‘the man’ is overtly marked as he breaks topic continuity (cf. §8.2.5).

(87) a. *tamyá weli tangâ we taprâw we tsâimâ*

[tamyá weli]_{NP} [tangâ we]_{NP} [taprâw we]_{NP} tsâi-mâ]_{main.cl}

tiger DISC.DEIC.DEM fish DEM school DEM count-ITER

‘This tiger kept counting the school of fish.’ [TA22_tiger_62 152]

b. ‘*ekhün gâqing*’ *la tsâiya bai yüm mu (...)* *mê we tawrê lô tsâidângge matyô*

[$\emptyset$ [e = khün kaqîng la] tsâi-ya habai yüm mu]_{main.cl} [[mê

DEM = one two HSAY count-ALLO be.able NEG also person

we]_{NP} [tawrê lô]_{NP} tsâi-dâng-ge matyô]_{main.cl}

DEM a.cane hanging.bridge count-COMPL.I-and.back maybe

‘(The tiger) could not count ‘one, two’ (i.e. beyond two). (Meanwhile) the man finished counting (the knots on the) hanging bridge and came back, it seems.’

[TA22_tiger_65 156]

Similarly, the following examples tell the story of a great catastrophe. In Example 88a, the god *Shûm-takã-dün* is burning down debris that is accumulating on a river due to an ongoing unweather. In succeeding Example 88b, we learn that the fire produced by *Shûm-takã-dün* has started to spread; now the unweather has turned into a wildfire. Here the topic is the catastrophe itself. In Example 88c, the topic shifts to people who have no place to go as a result of the event. Finally, Example 88d is a headless clause; the topic (the people) was mentioned in the preceeding clause.

- (88) a. *shûm.takã.dün prûwga we labû prûwmayûya wegû (...)*
 [[*shûm.takã.dün*]_{NP} *prûw-ga*]_{main.cl} [[*we*]_{anaph.NP} [*labû*]_{NP} *prûw-mayû-ya*
 a.name burn-PL.I DEM debris burn-shrink-ALLO
wegû]_{main.cl}
 and
 ‘*Shûm.takã.dün* burnt (them), he burnt (lit. burn-shrink) the debris and then (...)’ [TA23_Newmyth_18 106]
- b. *we grukûnatyû byẽ (...)* *talâ lê karâp grukûnaya (...)*
 [[*we*]_{NP} *gru-kû-na-tyû* *byẽ*]_{main.cl} [[*talâ* *lê*]_{NP} *karâp*
 DEM burn-COMPL.II-FACT-DIR TOP catastrophe big burn
gru-kû-na-ya]_{main.cl}
 burn-COMPL.II-FACT-ALLO
 ‘It burned down, the big catastrophe burned down.’ [TA23_Newmyth_19 115]
- c. (...) *mê we lîna blûm ãna yûm bô (...)*
 [[*mê* *we*]_{NP} [*lî-na* *blûm*]_{NP} *ã-na* *yûm bô*]_{sub.cl}
 person DEM escape-FACT NMLZ.LOC be.there-FACT NEG TEMP.SUB
 ‘These people having no place to escape (...)’ [TA23_Newmyth_20 121]

d. àng bâya gũ ì la (...)

$[[[\emptyset_i]_{NP} \text{ àng} \quad \text{bâ-ya} \quad \text{gũ}]_{\text{cosub.cl}} [\emptyset_i]_{NP} \quad \text{la}]_{\text{main.cl}}$
house build-ALLO NF be.there HSAY

‘Building a house, (they) were there, it is said (...)’ [TA23_Newmyth_23 143]

4.2 Nominal compounds and echo formations

This section discussed nominal compound and echo formations. A nominal compound in Tawrã is a juxtaposition of two nominal lexemes without any ancillary device. The relation between these nouns may be one of *modification*, i.e. the former modifies the latter, or of *coordination*, i.e. there is no clear head between the two nouns. Both compound types exhibit a tonal alternation discussed in §2.5.7, in which an initial low or mid tone in the compound becomes high tone.

Echo formations are the partial reduplication of a noun. This process often renders a collective sense or the paraphernalia associated with the reduplicated noun. Echo formations are discussed in §4.2.3.

There is a special type of compound in the language formed by a noun and a classifier. Such compounds have different degrees of semantic integration. *Classifier compounds* are discussed in §5.5.

4.2.1 Hierarchical nominal compounds

Hierarchical noun + noun compounds are head-final compounds in which the initial noun modifies or adjusts the meaning of the latter, e.g. via possession or modification. This is a productive feature of the language, cf. *majéi ráw* ‘horn of buffalo’ (from *majéi* ‘buffalo’ and *ráw* ‘horn’). There is no genitive marker as in related and neighboring languages. Table 4.3 illustrates the structure of hierarchical compounds in the language.

Table 4.3: Structure of hierarchical compounds

[X] + [Y]	>	[[X] + [Y]]

Even though the compounded elements form one syntactic unit, they are internally layered (cf. Matthews 1991, 85f) and thus nominal constituents may be modified within the compound. In Example 89a, the compound *khí mlằ* ('Chinese place', i.e. China) modifies the noun *boder* 'border'. In Example 89b, the noun *khí* 'Chinese' is modified by the demonstrative *we*, which further proves the layered structure of the whole compound.

(89) a. $[[khí \quad mlằ]_{NP} \quad boder]_{NP}$

Chinese place border(<E)

'the border with China'

b. $[[[khí \quad we]_{NP} \quad mlằ]_{NP} \quad boder]_{NP}$

Chinese DEM place border(<E)

'the border with the country of these Chinese' [both 011024_STBT_eli2 780]

Compounds form a continuum between transparent and idiosyncratic meanings, as is typologically common (cf. Matthews 1991, 86f). In transparent compounds, the first constituent modifies the second, cf. *kwâg tshwîng* ('dog + flea') 'dog's flea'. In idiosyncratic or obscure compounds, one (or both) element is either unknown or has an unexpected meaning. Many reasons may prompt the opacity of a compound. One is creative association. Thus, the word for 'plane' is a compound made up of 'sky + car' *hatûm-garêi*. In this case, the speakers made use of a language-internal, productive strategy to name a new concept. This is also common for body parts: *blüm-klâi* 'rheum' (transparently composed of *blüm* 'eye' and *klâi* 'excreta') or *thrûm-ûm* 'mustache' (*thrûm* 'mouth' and *ûm* 'hair'). A further possibility is semantic shift or the use of metaphors which often make compounds obscure. One example is the compound *rîng-kwâg* 'rabbit', where the first element *rîng* could either be a) a modifier form of 'sun' and hence the whole compound would be 'sun + dog', perhaps through a figure of speech that is not known (at least to my informants) or b) there could have been a semantic shift or a sound change in which the first element became homophonous with 'sun'. Compounds that have undergone some sort of semantic change are considered *true* compounds; a *combined lexeme* is more than the mere sum of its parts (Matthews 1991, 82f; Aikhenvald 2007, 24). This lexical process enriches the lexicon by creating new, unpredictable forms.

As discussed in §3.2.1, a typological feature of Tawrā is multiple class membership, i.e. many lexical roots occur underived as verbs, nouns, adjectives etc. This may create structural ambiguity between N-N compounds and homophonous N V clauses. If the first element (in both cases the noun) is eligible for the tonal change discussed in §2.5.7, i.e. if it either a) a monosyllabic noun with low or mid tone or b) a polysyllabic word that ends on a low or mid tone, then these form minimal pairs. The following examples illustrate this point: Example 90a is a complement clause; the noun does not change its low tone since the second element is a verb. However, in Example 90b the change of tone discloses the nominal nature of the second element and the two nouns form a nominal compound ‘popcorn’. Conversely, in Example 91 the personal name is not eligible for tonal change and thus there is a structural ambiguity between a nominal compound and a (verbless) clause interpretation that context alone has to resolve.

- (90) a. *nabũ* *bó*
 maize burst
 N V
 ‘The maize bursts.’

- b. [*nabũ* *bó*]_{NP}
 maize.MOD pop
 N N
 ‘popcorn’ [both: field notes 211024]

- (91) *kutyemba hadi*
 a.name story/tell
 ‘kutyemba’s story’ also ‘Kutyemba tells (something).’
 [based on TA22_kutyemba_phla 514 934]

§3.3.1.3 introduced a type of noun labeled *class term*. These are super-ordinate elements in a nominal compound that express *meronyms* like biological and functional clades, as in ‘things related to water and water bodies’. These compounds have a class

term as the first element and follow a [generic-specific] pattern and thus form hierarchical compounds.

Spatial relations are also commonly expressed through compounding, with *relator nouns* (cf. §3.3.1.1) as the second element in a nominal compound. Apart from their spatial semantics, relator nouns are no different from ordinary nouns.

4.2.2 Symmetrical nominal compounds

Symmetrical noun + noun compounds are two juxtaposed nouns without a clear head that often stand for a mass or collective group (cf. *copulative compounds*, cf. Booij 2005, 80). These compounds are lexicalized, i.e. may not be productively derived in the language and the order of their constituents may not be altered. Some of these compounds correspond to what is called *natural conjunction* in the literature (cf. Haspelmath 2007, 23f), meaning that they form a conceptual unit, e.g. *lông apó* (from *lông* ‘liver’ and *apó* ‘lungs’), which symbolizes life, e.g. a hunter might shoot an animal through its *lông apó* to kill it.

Symmetrical compounds usually consist of two noun roots but there is one example of three roots: *pû-pramu-pathì* ‘siblings’ (composed of *pû*, ‘elder brother’, *pramu* ‘younger brother’ and *pathì* ‘younger sister’). Kinship terms are particularly prone to build symmetrical compounds (e.g. mother-father i.e. parents). In some cases, the meaning of one element is unknown, usually the second element in the compound. Table 4.4 displays a list of symmetrical compounds. Further lexical study is bound to find many more forms. A special type of symmetrical compounds, termed *echo formations* includes the partial reduplication of a disyllabic noun. These forms are discussed in §4.2.3.

Table 4.4: Symmetrical noun + noun compounds (excerpt)

Compound	Meaning	Term 1	Meaning	Term 2	Meaning
yú-tapè	food (as in a feast)	yu	wine	tapè	rice, food
harwêi-pünyûm	blood (usually from the head, e.g. in an offering)	harwêi	blood	pünyûm	brains
khalâi-phlă	fields	khalâi	land, ground, soil	phlă	stone
kalyáw-karà	fields	kalyaw	field	karà	stone
hadí-atyà	stories	hadi	story	atyà	unknown
mế-jüm	mankind	mề	person	jüm	truth
má-bà	parents	ma	mother	bà	father
yâw-bra	in-laws	yâw	mother-in-law	bra	father-in-law
tîng-kazüm	cloths	tîng	shirt	kazüm	cloths
matyâng-haqûm	villagers, neighbors	matyâng	neighbor	haqûm	unknown
ashyâ-matsâw	(big) animals	ashyâ	mithun	matsâw	cow
lêng-apó	heart (rather a sensitive organ, a symbol of life)	lêng	liver	apó	lungs
klûm-dyâ	wealth	klûm	unknown, cf. klûm v. 'greet'	dyâ	unknown
mlă-matyâng	villages	mlă	village	matyâng	neighbor
pawüing-masyă	money	pawüing	money	masyă	unknown
hatûm-khalai	everywhere	hatûm	sky	khalâi	land, ground
tapé-tanà	food (of various kinds)	tapè	food	tanà	vegetable
tabrê-tathâ	food	tabrê	meat	tathâ	grain
byêk-talya	grasses and such	byêk	a grass	talya	unknown

4.2.3 Echo formations: partial reduplications of nouns

In Tawrã, an echo word or echo formation is the partial reduplication of a disyllabic noun (no monosyllabic nouns have been attested), commonly in (imperfect) rhymes. The reduplication of POS other than nouns is discussed in §3.6.

The partially reduplicated element (the *echo*) does not have a meaning on its own but rather the resulting collocation stands often for a) paraphernalia associated with the reduplicated noun or b) a collective mass, usually including different semantic subgroups of the noun, which is consistent with equivalent formations across the area. In plain words, echo words often render the phrase *and things like that* or *etc* (also note *General Extender Constructions* in Karbi, Konnerth 2014, 574f). Syntactically, echo formations are not a form of N-N nominal compounding, since the echo does not trigger a tonal change on the first noun. Semantically, however, echo formations can be considered to be a subtype of symmetrical compounds (§4.2.2) as the elements involved do not have a hierarchical order. Echoes commonly follow the original noun.

In Example 92, the echo *kahlâi* after *agrâi* ‘type of squirrel’ yields ‘squirrels and such’. Similarly, in Example 93 the echo *tacî* after *tathâ* ‘grain’ means ‘grains and the-like’ and implies that nothing will grow.

- (92) *agrâi kahlâi (...) tanî elàng gwók la*
 [*agrâi* ***kahlâi***]_{NP} [*tanî*]_{NP} [[*e = làng*]_{NP} [*gwók*]_{NP} *la*]_{main.cl}
 type.of.squirrel ECHO type.of.rodent DEM = COLL priest HSAY
 ‘Squirrels and such animals (used to be) priests.’ [TA23_NewMyth_35 215]

- (93) *ja tathâ tacî ja mu shyâ yüm*
ja [*tathâ* ***tacî***]_{NP} *ja* *mu* *shyâ* *yüm*
 what crop ECHO what also become NEG
 ‘No crop will grow.’ [TA22_Kutyemba_phla 48 147]

Table 4.5 depicts all echo formations found so far. Further lexicographic research is bound to find many more of these forms.

Table 4.5: Echo words in Tawrã

Single noun	Meaning	echo formation	Meaning
<i>agrâi</i>	type of squirrel	<i>agrâi kahlâi</i>	squirrels and such animals
<i>alyüm</i>	road, way	<i>alyüm aklüm</i>	different types of ways
<i>tathâ</i>	grain	<i>tathâ tachâ</i> (also <i>tachî</i>)	different types of grain
<i>kalyaw</i>	field	<i>kalyaw salyaw</i>	fields and such
<i>tako</i>	book	<i>tako tazo</i>	related to books (e.g. studied people)
<i>khrai</i>	paralytic	<i>khrai barei</i>	paralytic

Indo-Aryan (henceforth IA) loanwords may be incorporated along with their own echoes. In Example 94, the Assamese noun *hâki* is followed by its own echo *bâji*¹.

(94) we las *hâki* *bâji* hanabiya (...)

we las *hâki* *bâji* hana-biya

DEM at.the.end(<E) witness(<Asm) ECHO(<IA) come-POT.ALLO

‘At last, the witnesses will come (...)’ [TA24_land 48 147]

Echo words are an areal feature of South Asia, ranging across all major linguistic families (for IA, cf. Masica 1993, 80; for Tamil, cf. Schiffman and Harold 1999, 172; for Kharia, cf. Peterson 2010, 129f; for an overview, cf. Abbi 2018). Even though the strategies for forming echoes vary greatly, the semantic function remains roughly the same: either to add ‘etc.’ or exaggeration. In some languages like Hindi, echo formation is so productive that it can even be applied to loan words, cf. *seṭiṇ peṭiṇ* ‘all the things (e.g. arranged in a room)’ (< Eng. *setting*; *peṭiṇ* does not have a meaning on its own).

As noted by Hyslop (2017, 329f.), echo formations are not a prevalent feature of Tibeto-Burman, where we mostly find common reduplication, e.g. in ideophones, ono-

¹Although echoes generally lack a meaning of their own, nouns with a related meaning can also be recruited for echoes in IA languages, thus blurring the line between symmetrical compounds and echo formations. In this case, *bâji* is a noun meaning ‘bet’. Its function as an echo can be explained by the fact that witnesses are unpredictable and thus a *bet*.

matopoetic forms or binomials (Rai et al. 2009, 26). Furthermore, echo formations among Tibeto-Burman languages seem to correlate with intense language contact with IA (for Rabha cf. Joseph 2006, 287, 328; for Dhimal cf. King 2009, 58f), suggesting that the partial reduplication of nouns in Tibeto-Burman may have been adopted or at least reinforced through contact.

4.3 NP-head modifiers

NP-head modifiers are constituents of the NP that may occur pre- and post-nominally. The NP-head modifiers discussed in this section are: adjectives (§4.3.1), the classifier phrase and numeral (§4.3.2) and the demonstrative (§4.3.3). These constituents share the ability to head a NP. Note that the relative clause is discussed in §10.5.2 and nominal complements in §10.5.3, as part of event nominalization. Nominal compounding has been discussed separately in §4.2.

4.3.1 Adjectives

Adjectives as a word class have been discussed in §3.3.3 and §3.2.4. The section aims to discuss the syntactic features of adjectives in relation to the noun phrase. There is a small note on the adjective *a* ‘small’ in §4.3.1.1.

Adjectives, as a device of nominal modification, have two main functions: they may modify a noun (*attributive* function) as in ‘the **big** rat’ or they may form a (verbless) clause predicate (§7.1) in which the S argument is equated to the adjective (*predicative* function), as in ‘the rat is **big**’. Example 95 illustrates the attributive function of adjectives as *daráng* ‘big’ modifies the NP head *phlá pû* ‘rock’. Conversely, Example 96 corresponds to the predicative function of adjectives: the adjective *pû* forms a verbless clause with the NP headed by *a* ‘child’.

- (95) mlă kēq gũ phlá pû daráng khing ã la
 mlă kēq gũ phlá pû daráng khing ã la
 village LOC.INS LOC stone.MOD rock big one be.there HSAY
 ‘In a village there was a big rock, they say.’ [TA22_kutyemba_phla 4 73]

- (96) (...) a we nyu lyûm nang gũ mu we pũ
 [a we nyu [lyûm nang gũ]_{cosub.cl} mu we pũ]_{VLC}
 child DEM TOP eight year NF also DEM unchanged
 ‘Even being eight years old this child was unchanged (i.e. it did not grow).’
 [TA23_newmyth 49 292]

Adjectives modify whole NPs and do not require any morphological device to occur. They have, however, two suffixes available that they share with verbs: stative *-ho* (§6.3.4.5) and iterative *mâ* (§6.3.4.4), as has been discussed in §3.2.4. Concerning their position in the NP, they commonly occur *after* the head noun but this position is not fixed. In Example 97, the adjective follows the noun head. Similarly, in Example 98 a speaker recounts a story from the past and mentions *bá-mê daráng* ‘big clan(s)’ and *bá-mê aqa* ‘small clan(s)’ with the pragmatically neutral N-Adj order.

- (97) hazâ lyó làng wegũ (...)
 hazâ lyó làng wegũ
 officer white COLL and.then
 ‘These white officers (i.e. the Britishers) (...)’ [TA22_CLAN2 8 16]
- (98) bá.mê daráng byě bá.mê aqa pô mu kwâg mu bâ la
 [[bá.mê daráng byě]_{NP} [bá.mê aqa]_{NP} [[pô]_{NP} mu [kwâg]_{NP} mu]_{co-ord NP} bâ
 clan big TOP clan small slave also dog also do
 la]_{main.cl}
 HSAY
 ‘Big clans would enslave small clans (lit. do like slave like dog), it is said.’
 [TA22_clan 12 56]

The language seem to allow a degree of functional ambiguity between *predicative* and *attributive* adjectives since there are no morphological means to distinguish these two functions and both normally take the post-nominal slot (cf. Hengeveld and Siewierska 2004, 555f). Example 99 illustrates the ambiguity between both functions of adjectives.

(99) *bá-mề daráng*

clan big

‘The clan is big’ or ‘the big clan.’

Conversely, pre-nominal order of adjectives is only attested from elicitation. Speakers perceive the pre- and post-nominal orders as interchangeable. If there is a pragmatic difference (as there is with the classifier phrase in relation to the NP head, cf. §5.1) I have not detected it. Example 100a is the standard word order, as found in the corpus. The order found in 100b has only been attested in elicitation.

(100) a. *àng daráng*

house big

‘a big house’

b. *daráng àng*

big house

‘a big house’ [both 011024_STBT_eli2 804]

One exception regarding the standard N + Adj word order is the name of a mythical figure *jüm Câne* (‘true a.name’), lit. ‘the true Chane’, where the adjective precedes the noun. An opposite order **Câne jüm* is not accepted by speakers. It is unlikely that this particular name would bear an ancient word order (as in the case of Turung, cf. Morey 2012, 495). It cannot be ruled out that this be a calque from a different language, since e.g. IA loan adjectives retain their pre-nominal position (see below).

When the adverb *tyâng* ‘very’ modifies an adjective, it prompts a change in word order, placing the adjective before the head noun. In Examples 101 and 102, the modified adjective precedes the head noun (instead of the common N Adj order).

(101) *tyâng daráng phlă*

tyâng daráng phlă

very big stone

‘A very big stone’ [TA22_kutyemba_phla 8 81]

- (102) *tyâng pra mîyaqa hadyû gũ (...)*
tyâng pra mîyaqa hadyû gũ
 very good girl resemble NF

‘(...) (the fish) resembling such a beautiful girl (...)’ [TA23_Fishstory 84 157]

Adjectives borrowed from an IA language may retain the word order of the source language. In Example 103, the IA adjective *sarkari* ‘governmental, of the government’ is placed before the noun.

- (103) (...) ehm *sarkari mề we (...)*
ehm sarkari mề we
 ehm government(<IA) person DEM

‘(...) ehm, this person from the government (...)’ [TA22_CLAN2 31 55]

The slot position of adjectives is free relative to the head noun (i.e. it can occur before or after it) but not relative to other constituents of the NP. As such, adjectives placed in a different slot are interpreted as ungrammatical or as a verbless clause (§7.1). As an example, the slot of the demonstrative *we* occurs further away from the head noun than the slot of the adjective, as was discussed in §4.1 and illustrated in Example 104a. Conversely, placing the adjective *after* the demonstrative triggers a predicative reading, as in Example 104b.

- (104) a. [*àng daráng we*]_{NP}
 house big DEM
 ‘this big house’

- b. [*àng we*]_{NP} *daráng*]_{verbless.cl}
 house DEM big
 ‘This house is big.’ [both 011024_STBT_eli2 815]

Adjectives with an attributive function in Tawrã can be *stacked*, i.e. up to two adjectives can co-occur modifying the head of the noun phrase without an intervening

constituent. In elicitation, the preferred order is post-nominal, as in Example 105a. If adjectives are separated by the head noun, the sentence must be read as a verbless clause (cf. §7.1) with a modified nominal head, as in Example 105b (cf. Morey 2012, 298). Pre-nominal stacked adjectives are perceived as strange, as in Example 105c.

- (105) a. *kwâg má daráng*
 dog black big
 ‘a big, black dog’
- b. *má kwâg daráng*
 black dog big
 ‘The black dog is big’. It does not yield: ‘a big, black dog’
- c. ?*má daráng kwâg*
 black big dog
 (?) ‘a big, black dog’ [all three 011024_STBT_eli2 854]

Adjectives can head a noun phrase without any special derivation if there is no overt noun, as mentioned in §3.2.4. The ellipsed noun is usually understood from the context. Examples 106a, 106b and 106c are from a story in which a tiger eats raw meat and becomes wild. In Example 106a, the speaker uses the adjective *asâng-ra* ‘raw’ without a nominal head. A noun is added as post-posed noun (§8.2.3) for clarification. Once the referent has been established, the noun *tabrê* ‘meat’ does not have to be repeated. In Examples 106b and 106c, the noun ‘meat’ is omitted and both NPs are headed by adjectives.

- (106) a. *asâng.ra thâbô (...) tabrê we*
 [*asâng.ra*]_{NP} *thâ-bô* [*tabrê we*]_{postposed.NP}
 raw eat-PFV.ALLO meat DEM
 ‘(The tiger) ate raw (...) meat.’ [TA22_tiger_117_265]

- b. cya asâng thâthâ thĩ gũ (...)

cya [*asâng*]_{NP} *thâ~thâ* *thĩ* *gũ*

3 raw eat~PLAC CONC NF

‘By eating and eating raw (meat) (...)’ [TA22_tiger_127_281]

- c. kâca kâca thâ gũ (...)

[*kâca*]_{NP} [*kâca*]_{NP} *thâ* *gũ*

raw(<IA) REDUP eat NF

‘By eating raw (meat) (...)’ [TA22_tiger_129_286]

A NP headed by an adjective does not have any type of restriction regarding NP head modifiers. In Example 107a, the NP still takes a demonstrative. In Example 107b, the adjective is modified by a further adjective. Finally, in Example 107c the feminine marker docks on the adjective.

- (107) a. [*daráng we*]_{NP} *bô-bô*

big DEM go-PFV.ALLO

‘A big one has gone.’

- b. [*muhwa daráng*]_{NP} *bô-bô*

rich big go-PFV.ALLO

‘A big rich (person) has gone.’

- c. [*muhwa-yă*]_{NP}

rich-FEM

‘a rich woman’ [all three 011024_STBT_eli2 879]

4.3.1.1 Small/diminutive *a(qa)*

The adjective *a(qa)* ‘small’ deserves a separate note for it displays a few features not found with other adjectives. This adjective is related to the noun *a* ‘child’, which is a

common source for the diminutive suffix (Kuteva et al. 2019, 88; also for Tibeto-Burman languages, cf. Matisoff 2003, 485-487). The sense of ‘child’ is still present when used with animals, cf. *pahwã a* ‘small frog, frogling’ from *pahwã* ‘frog’. However, one distinctive feature of this adjective is that it is commonly reduplicated, cf. *bá-mě a-qa* ‘small clans’ (i.e. *a a*) without yielding a clear difference from its non-reduplicated form. This is noteworthy since reduplication of adjectives is generally associated with exaggeration (cf. §3.6.2). Furthermore, the noun *a* ‘child’ is never reduplicated.

The noun/formative *a* further occurs in four compositional nouns: *mâqa* ‘father’s younger sister’, *bâqa* ‘father’s younger brother’, *mîyaqa* ‘girl’² and *mûwaqa* ‘man’. The first two terms are derived from ‘mother’ and ‘father’ respectively (in the sense of ‘younger’) while the last two from the word for ‘wife’ and ‘husband’, respectively.

Lastly, *a* cannot be attached to a whole clause as other NP modifiers do. In Example 108, the clause yields ‘I know the small one’; the adjective operates as an autonomous NP meaning ‘the small one’. In other words, it has not undergone decategorization and it retains its adjectival features.

(108) *kasâde aqa we*

kasâ-de aqa we
know-HAB.EGO small DEM

Intended: ‘I know the small one’ (it does not yield ‘The little I know’).

[011024_STBT_eli2 924; Modeled after Lehmann (1984: 61)]

4.3.2 Classifier phrase and numerals

Classifier phrases and numerals occupy the same slot in the NP structure and therefore they are treated together in this section. The classifier phrase is composed of a classifier and the numeral (in that strict order) and is hierarchically below the noun phrase, as it will be described in §5.2. Numerals have been described in §3.5.3.

Both the classifier phrase and the numeral may occur pre- or post-nominally, the latter being the unmarked position, as illustrated in Example 109. The inverse order ClP-N/ Num-N implies a specific referent, as discussed in §5.2.

²This term is further employed in other kinship terms, cf. *búm-mîyaqa* ‘maternal uncle’s daughter’.

- (109) pahwã bra gê (...) rôĩ halyo gũ bônarêbô delé (...)
 [[pahwã [bra gê]_{CLP}]_{NP} rô-*rĩ*]_{main.cl} [[halyo gũ]_{cosub.cl} bô-na-rê-bô
 frog CLF:generic one catch-PL.II be.happy NF go-FACT-PL.II-PFV.ALLO
delé]_{main.cl}
 it.seems
 ‘They caught (...) one frog (and) have gone happily (...)’ [TA22_frog1 135 719]

Certain nouns do not trigger a classifier when counted, as discussed in §5.1, and are modified by a numeral alone. There is considerable variation between speakers in this regard, which is likely due to an ongoing semantic change in which the classifier *bra* ‘CLF:small.round’ is being re-analyzed as a generic classifier. In Example 110, the head noun is modified by a numeral (a reduced form of the full numeral ‘one’, cf. §3.5.3) without a classifier.

- (110) ble lê khing münbîna
 [ble lê khing]_{NP} mün-bî-na
 pig big one roast-EMPH-IMP
 ‘Roast the biggest pig!’ [TA22_kutyemba_phla 33 120]

Both the classifier phrase and numerals may replace a NP, often anaphorically, as mentioned in §4.1.1.1. The only restriction is that they cannot be modified by a further classifier phrase/numeral. In Example 111, a classifier phrase heads the NP, which is further modified by an adjective and a demonstrative.

- (111) pum gê darang we
 [pum gê]_{CLP} darang we]_{NP}
 CLF:3D.round one big DEM
 ‘a big, round one’ [011024_STBT_eli2 939]

4.3.3 Demonstrative *we*

This section describes the demonstrative *we*. §4.3.3.1 discusses the main features of the demonstrative, while its pragmatic function is discussed in §4.3.3.2. There are two

complex discourse operators termed *complex demonstratives* since they are derived from the demonstrative *we*. These are *weli* and *webo* and they are discussed in §10.8.

The only demonstrative found in the language is *we* (sometimes *ewe*) with a cliticized form *e=* and an emphatic double demonstrative form *we X we*. This demonstrative is likely related to Rawang distal demonstrative *wē* (LaPolla 2008b, 50f) and similar forms in related languages (indicative particle *ve*, cf. Matisoff 1973a, 407f; *wa* in Turung, cf. Morey 2012, 346), although it should be noted that in Tawrã the demonstrative is not involved in nominalization/relative clause constructions. The demonstrative is different from the 3rd person pronoun *cya* (cf. §3.5.1.1).

4.3.3.1 Basic features

The main function of the demonstrative is to mark a prominent referent. Crucially, it does not convey spatial orientation. This comes as a surprise since demonstratives in related languages spoken in mountainous settings often have a grammaticalized system of topographical deixis (Bickel 2000a, 227; Post 2011c, 145f). Tawrã thus belongs to a small set of languages without deictically contrastive demonstratives (cf. Diessel 1999, 37f).

The demonstrative may occur as *we* or cliticized as *=e* (and *e=*). Post-nominally, the demonstrative may also be reduplicated as a form of emphasis: (*X we we*). Example 112 illustrates the demonstrative in a post-head position, which can be considered the least marked form. Example 113 displays both the cliticized form and the emphatic reduplication of the demonstrative. Example 114 displays the *ewe* form.

(112) *khalâi khanyùm we byě* (...)

khalâi khanyùm we byě

ground earthquake DEM TOP

‘as of the earthquake (...)’ [TA23_tayang_clan 6 26]

(113) ah nyu aye we we

ah nyu a=e we we

ah 2S child=DEM DEM DEM

‘Ah this child of yours!’ [TA22_origin 52 153]

(114) bîya ma ning ewe (...)

bîya ma ning ewe

past ADD 1P DEM

‘In the past we (...)’ [TA22_clan 140 421]

The demonstrative may be also employed in an emphatic manner, which is a double demonstrative construction (cf. Dryer 2007b, 163; also in Milang, cf. Modi 2017, 232f, 305f) around the modified NP head: *we X we*. Example 115 has two double demonstratives, one around the ‘village people’ and one around ‘the stone’.

(115) we tyû mlá mẽ we “tya hîng we phlâ we glâikũvidyô”

[*we tyû mlá mẽ we*]_{NP} [*tya hîng [we phlâ we]*]_{NP}

DEM north village.MOD people DEM now day DEM stone DEM

*glâi-kũ-vidyô*_{direct.sp}

carry-COMPL.II-CERT.FUT

‘(He said) to the people from the northern village “today I’ll carry this stone away.”’ [TA22_kutyemba_phla 25 103]

A demonstrative may head a noun phrase, often with an anaphoric use (cf. §4.3.3.2). As head of a NP, the demonstrative can be modified by most constituents of the NP. In Example 116a, the demonstrative takes the feminine suffix, although this form is perceived as strange. In Example 116b, the demonstrative is modified by a classifier phrase (ClP). In Example 116c, the demonstrative is modified by an adjective.

- (116) a. *we-yǎ̀*
 DEM-FEM
 ?‘she, this (female)’ [011024_STBT_eli2 972]
- b. *we [bra gê]_{CLP}*
 DEM CLF:generic one
 ‘this one’
- c. *we daráng bô-bô*
 DEM big go-PFV.ALLO
 ‘This big one is gone.’ [all three 011024_STBT_eli2 972]

4.3.3.2 Pragmatic functions

Based on the work of Diessel (2015) and Himmelmann (1996) we may discuss the main pragmatic functions of the demonstrative in Tawrã: *exophoric* use and *reference tracking*. This subsection will also go over the general features of demonstratives as pragmatically marked devices. Finally, based on the work of Becker (2021), the incompatibility between the demonstrative *we* and *contextually unique referents* will become evident.

The *exophoric* use of demonstratives refers to their ability to focus the “hearer’s attention on entities in the situation surrounding the interlocutors” (Diessel 1999, 94). This involves taking the speaker as the deictic center and is often accompanied by gesturing (e.g. pointing). In Example 117, the speaker tells where he is before telling a story. The use of the demonstrative with *talyâng* ‘evening’ works to anchor it in the speech situation (*discourse deictic*, cf. Himmelmann 1996, 221f).

- (117) *hã etalyâng pasigat kě*
hã e = talyâng pasigat kě
 1S DEM = evening a.place LOC.INS
 ‘I am at Pasigat this evening.’ [TA23_pineapple 3 13]

Examples 118a, 118b and 118c are the beginning of an exercise in which the speaker was asked to narrate a story based on images from a storybook (Rodriguez 2008). The speaker points at the animals, tries to guess what they are and asks for confirmation. The speaker employs the demonstrative *ewe* and pointing gesture to refer to visible referents from the book.

(118) a. *ewe* *tahum weligũ tyû thazâ* (...)

ewe tahum weligũ tyû tha-zâ

DEM bear and.then chicken what-DIR

‘This is a bear and then a chicken and what else ...?’ [TA22_chicken.thief 2 21]

b. *ewe pyâ!*

ewe pyâ

DEM bird

‘This is a chicken!’ [TA22_chicken.thief 4 33]

c. *ewe da na? khorguj na kya?*

ewe da na khorguj na kya

DEM what EMPH rabbit(<IA) INSIST Q

‘What is this? a rabbit or what?’ [TA22_chicken.thief 6 41]

A further use of demonstratives is *reference tracking*, i.e. they help track discourse participants either as an *anaphora* (i.e. the discourse deictic use from a preceding section in discourse) or as coreferential within a noun or noun phrase in the same clause (Himmelmann 1996, 226f; Diessel 1999, 94f). Demonstratives are the main device for anaphora but not the only ones; the 3rd person pronoun (§3.5.1.1) and classifiers (§5.3) may also occur with this function, although not nearly as often as demonstratives.

The following examples are from a story in which the son of a criminal stayed at the speaker’s grandfather’s village with the hidden intention of killing the speaker’s grandfather. Example 119a describes how the son of a criminal (ellipsed) stayed at the village. In Example 119b, we find the demonstrative tracking that same referent.

- (119) a. *kanang lyang gê ìya la âm*
kanang lyang gê ì-ya la âm
 CLF:year half one be.there-ALLO HSAY AFF
 ‘For half a year he stayed, they say, yes.’ [TA22_CLAN2 148 275]
- b. *we ìì webogũ (...)*
we ì~ì~ì webogũ
 DEM be.there~PLAC~PLAC and.then
 ‘He kept staying and then ...’ [TA22_CLAN2 150 278]

Similarly, in Example 120a, a referent (Pashya) is introduced into the discourse as a central character. Thereafter, in Example 120b, the same Pashya is retrieved as an anaphora with a demonstrative *we*.

- (120) a. *pâshyã we gũ hana la (...)* *pâshyã we mlã we tîya we pâshyã tîya we*
[pâshyã we gũ hana la]_{main.cl} [pâshyã we [mlã we tîya]_{NP}
 a.name DEM LOC come HSAY a.name DEM village DEM lord
we]_{main.cl-VCS} [pâshyã tîya we]_{main.cl-VCS}
 DEM a.name lord DEM
 ‘(They) came at the time of Pashya, it is said. Pashya was the lord of that village. Pashya was the lord.’ [TA23_Tayang_clan 17 96]
- b. *we macimâ hî ìna la*
[we]_{NP} maci-mâ hî ì-na la]_{main.cl}
 DEM harm-ITER EMPH be.there-FACT HSAY
 ‘He did a lot of harm, they say.’ [TA23_Tayang_clan 19 108]

As discussed by Himmelmann (1996, 229f) and Diessel (1999, 98f), a cross-linguistically common strategy is to employ the demonstrative right after the introduction of a new, major discourse participant. Once this referent has been established (and becomes the topic), the demonstrative is not the preferred device for anaphora and the topic may be

referred to by other devices, such as a 3rd person pronoun or left unmarked. The following examples correspond to the beginning of a story. In Example 121a, the main figure (*Kutemba* a trickster) is introduced. This first mention is through a classifier (§5.3). In the following sentence, Example 121b, the name of the main character is introduced with an emphatic form of the demonstrative (i.e. *we X we*). Thereafter, the speaker retrieves the referent with a NP that includes a demonstrative (albeit not as an anaphora) in Example 121c. Finally, in Example 121d the main referent is replaced by a 3rd person pronoun.

- (121) a. hayú bîya mề bra gê lâlhakya mề ìya
hayú bîya [mề bra gê]_{NP} [[lâlhak-ya]_{rel.cl} mề]_{apposed.NP}
 past old person CLF:generic one cheat-ALLO person
ì-ya
 be.there-ALLO
 ‘Long time ago there was a trickster.’ [TA22_SOstory_1 0]
- b. we amang we kumtemba
we amang we kumtemba
 DEM name DEM a.name
 ‘His name was Kutemba.’ [TA22_SOstory_2 26]
- c. we mề lâlhakmâ
we mề lâlhak-mâ
 DEM person cheat-ITER
 ‘He used to cheat (i.e. was a trickster).’ [TA22_SOstory_3 31]
- d. kehîng ge cya mề lâhlakya
kehîng ge cya mề lâhlak-ya
 CLF:daytime one 3 person cheat-ALLO
 ‘One day he cheated a man.’ [TA22_SOstory_5 35]

The use of demonstratives as anaphora is often *marked*, i.e. it implies an unexpected referent in the context (Himmelfmann 1996, 227; also discussed by Diessel 1999, 99), and

it often stands in functional contrast to other, less marked devices employed for anaphora (a point also made in §8.2.5 on omitted arguments in multi-clause constructions).

The pragmatically marked use of the demonstratives *we* in Tawrã is illustrated in the following series of examples, although a brief note on the context of the story is required: After the proverbial trickster *Kutyemba* convinced the village people that a big rock was about to destroy the village, everyone got together to help *Kutyemba* push the rock away. The trickster asked the villagers to close their eyes and push the rock with their backs. In Example 122a, *Kutyemba*, the trickster, commands the village people to push with their eyes closed. In the following sentence, Example 122b, the A argument is the village people and not *Kutyemba*, which is the main character and the most topical referent. The use of an anaphoric demonstrative highlights a relevant change in reference, although it is also specified by a post-posed NP.

- (122) a. “bülüm màho tayãrëna” la la
 [bülüm mà-ho tayã-rẽ-na]_{direct.speech} la la]_{main.cl}
 eye close-STAT shout-PL.II-IMP say HSAY
 “Shout with your eyes closed!” he said, it is said.’ [TA22_kutyemba_phla_105
 235]
- b. we tayãga khrâk mlá mẽ we
 [we tayã-ga khrâk [mlá mẽ we]_{postposed.NP}]_{main.cl}
 DEM shout-PL.I EVI village.MOD person DEM
 ‘They shouted, it seems, these village people.’ [TA22_kutyemba_phla_109
 241]

A final note on the pragmatic use of the demonstrative in Tawrã and *contextually unique referents* is in order. Contextually unique referents, in the parlance of Becker (2021, 79f), are referents constructed as “as unique in a larger context based on (general) knowledge shared by the speaker and the hearer” (idem.), regardless of whether they are “familiar or visible to the discourse participants” (idem.). Crucially for this discussion, demonstratives are expected to be *infelicitous* to mark contextually unique referents (idem. 115f.). This is perhaps the case because *contextually unique referents* rely

on general, world knowledge, whereas demonstratives are mostly associated with *in-the-world*, physical objects that are present in the discourse situation (i.e. deictic referents) (idem. 60).

Be it as it may, the demonstrative *we* is not allowed with contextually unique referents. In Example 123, the speaker refers to the market in his village. Since there is only one market in that village, it qualifies as a unique referent. Employing a demonstrative would not be acceptable here.

(123) *tĩng làng bréya künüm bojar (*we) khìna*

*tĩng làng bréi-ya künüm bojar (*we) khì-na*

cloths COLL buy-ALLO PRP marker(<IA) DEM reach-FACT

‘I went to the (*this) market to buy clothes.’ [based on TA23_Pineapple_07 35]

4.4 Conjoined NPs

Conjoined NPs can be divided into two groups based on the device (or lack thereof) employed to indicate coordination: the use of the additive marker and asyndetic coordination (i.e. no special device). These are treated in subsections §4.4.1 and §4.4.2 respectively. Nominal compounding has been discussed in §4.2.

4.4.1 Additive *ma*

The additive *ma* is a postposition that follows all NPs involved in a complex noun phrase. This device corresponds to the preposition *with* and the adverb *also* in English. In Example 124, both the tiger and the man take this suffix as both are involved in the story. The use of this postposition with *Cane* ‘a name’ could have been primed by the earlier additive construction. The additive postposition is likely related to the homophonous contrastive topic marker (cf. §8.3.1.2).

- (124) tamyá ma mề ma kapâq maroga ki la bîya be byẽ cane ma tayîm gu
 [[tamyá]_{NP} **ma** [mề]_{NP} **ma** kapâq maro-ga ki la]_{main.cl} [bîya bebyẽ
 tiger ADD person ADD alone talk-PL.I COUNT.EXP HSAY past TOP
 [[cane]_{NP} **ma** [tayîm]_{NP} gũ]_{NP}]_{main.cl}
 a.name ADD time(<Eng) LOC
 ‘In the past, in Chane’s time, only the tiger and the man talked (i.e. could talk).’
 [TA22_tiger 15 47]

4.4.2 Asyndetic coordination

Asyndetic coordination or mere juxtaposition of nouns or pronouns without intervening morphology is a common strategy for indicating coordination (Haspelmath 2007, 7f). Depending on the context, it may imply conjunction (*and*) or disjunction (*or*). In Example 125, two NPs (headed by a personal pronoun, cf. §4.1.2.2) form a complex NP with an internal conjunctive (*and*) relation. Similarly, Example 126 includes three referents with an internal conjunctive (*and*) relation.

- (125) nyu hã durum gatyô shìnade bawan we gatyô rônade
 [[nyu]_{NP} [hã]_{NP}]_{asyndetic.rel} durum gatyô shì-na-de]_{main.cl} [bawan
 2s 1s religion(<IA) different take-FACT-HAB.EGO god(<IA)
 we gatyô rô-na-de]_{main.cl}
 DEM different catch-FACT-HAB.EGO
 ‘You **and** I we take different religions, we follow (lit. catch) a different god.’
 [TA22_origin 111 323]

- (126) natya sângdom natya tapôdo natya jakrê we kasang we a lûmkũ kyẽ
 [[*na-tya* *sângdom*]_{NP} [*na-tya* *tapôdo*]_{NP} [*na-tya*
 1.POSS-grandfather a.name 1.POSS-grandfather a.name 1.POSS-grandfather
jakrê]_{NP} *we*]_{NP} [*kasang we*]_{NP} *a lûmkũ kyẽ*]_{main.cl}
 a.name DEM three DEM son inside LOC.INS
 ‘My grandfathers Sangdom, Tapodo and Jakre, these three are from the same son
 (lit. inside).’ [TA23_Karap_clan 94 503]

Asyndetic coordination is also commonly employed for making lists. Example 127 includes a list of NPs that are part of a list.

- (127) ning we shyâ matsâw tyû ble hawrêi pünyûm sú.sü hamà nyâ.blâng nyâ.po
ning we [*shyâ*]_{NP} [*matsâw*]_{NP} [*tyû*]_{NP} [*ble*]_{NP} [*hawrêi pünyûm*]_{NP}
 1P DEM mithun caw chicken pig blood brain
*[sú.sü hamà]*_{NP} [*nyâ.blâng*]_{NP} [*nyâ.po*]_{NP}
 burned.bones smoky.smell a.seed sesame.seed
 ‘(When we worship) we (offer) mithuns, cows, chicken, blood (lit. blood brain)
 the smoky fragrance of burned bones and seeds.’ [TA22_origin 127 369]

Juxtaposed NP may also imply a disjunctive relation (X or Y). In Example 128, the speaker couples two NPs, one with an overt head *tükü aqa* ‘small matter’ and one with an ellipsed NP head (marked by a $\emptyset$). These are understood to be in a disjunctive relation.

- (128) we já.sû tashi ewe we mu ali түкү ақа дарáng byẽ (...)
we já.sû tashi ewe we mu ali [[*tükü aqa*]_{NP} [$\emptyset$]
 DEM whatever fight DEM DEM also like.this matter small
daráng]_{NP}]_{disjunctive.rel} *byẽ*
 big COND
 ‘Whatever fight (would occur), whether small **or** big matter (...)’ [TA22_clan_41
 144]

4.5 Nominal affixes

Nominal affixes modify a NP. This group includes two suffixes: *-pa* ‘place of many’ (§4.5.1) and the feminine suffix *-yã* (§4.5.2), and one prefix: *na-* prefix for kinship terms (§4.5.3). While suffixes dock on the right edge of the NP, the prefix is attached directly to the NP head, which must be a kinship term.

Nominal affixes are attached to NP heads. If there is no NP, e.g. in a headless relative clause, they are not allowed. Example 129a is ungrammatical because there is no NP head. Suffixes *-pa* ‘place of many’ (§4.5.1) and the feminine suffix *-yã* (§4.5.2) share the property of docking on the right edge of a NP. They dock directly on nouns or pronouns (cf. Example 129b) unless they are modified by an adjective, in which case they dock on the adjective, as discussed in §4.1 and shown in Example 129c.

- (129) a. **tabrẽ thâyayã*
 [*tabrẽ thây*]_{rel.cl} *∅*]_{NP}-*yã*
 meat eat-ALLO -FEM
 Intended: ‘A female who eats meat.’

- b. *cya-yã*
 3-FEM
 ‘she’

- c. (*mẽ*) *lyó-yã*
 person white-FEM
 ‘a white female’

4.5.1 ‘place of many’ *-pa*

The nominal suffix ‘place of many’ *-pa* occurs with nouns and implies ‘high numbers’, cf. *phlã-pa* (stone-place.of.many) ‘stony place, place of many stones’ or *abã-pa*

(type.of.bamboo-place.of.many) ‘place of many *abâ* bamboos, forest of *abâ* bamboo’. The origin of this suffix is not clear.

As a nominal suffix, this morpheme occurs at the right edge of the NP and is thus ungrammatical in Example 130a. It is uncommon to include adjectives on a NP that hosts this suffix, as depicted in Example 130b with the adjective *daráng* ‘big’. Nominal suffixes can, however, be attached to a NP headed by an adjective, as in Example 130c.

- (130) a. **masáng-pa daráng*
 tree-place.of.many big
 Intended: ‘place of many big trees’
- b. *?masáng daráng-pa*
 tree big-place.of.many
 ‘place of many big trees’ [both field notes 19102024]
- c. *má-pa*
 black-place.of.many
 ‘place of blacks (i.e. where black people live)’ [field notes 19102024]

4.5.2 Feminine suffix *-yã̀*

The feminine nominal suffix *-yã̀* indicates that the referent is female³. This suffix seems to be related to the *ya-* (written <ja->) prefix found among Tani languages, which, among other things, occurs with certain female nouns, cf. Galo *yai* (written <jai>) ‘last daughter’ *yapóm* (written <japóm>) ‘fairy, demon’ (note that it stands in paradigmatic opposition to the **ta-* prefix found on some terms relating to males, cf. Post 2006, 49f;

³In many languages, the assignment of gender to inanimate objects follows attributes like shape and size (Kemmerer 2017, 408). As an example, many IA languages have masculine-feminine pairs of inanimate objects that often mirror differences in size, cf. Panjabi *churī* ‘knife’, the feminine (unmarked form) as opposed to *churā*, the masculine form which refers to a much bigger knife (Bashir and Connors 2019, 93f). In Tibeto-Burman languages grammatical gender commonly corresponds to a sex distinction male/female (cf. the case of Jero Opgenort 2005, 79f), although there are some exceptions, cf. the augmentative and female suffix *pī* in Karbi (and the Galo example cited therein, cf. Konnerth 2014, 206).

Post 2007, 266f). It is, however, noteworthy that they differ in their position relative to the noun: while the Tawrã form is a suffix, the Tani form is a prefix.

In Example 131a, the suffix *-yã* indicates a female member of the Bõ clan. Similarly, in 131b it indicates a female member of the Tawrã tribe. Any word class that can head a noun phrase may be marked by this suffix. In Example 131c (repeated from Example 129b for the sake of clarity), the feminine suffix is attached to a personal pronoun.

- (131) a. *bõ-yã*
 a.clan-FEM
 ‘a woman of the Bõ clan’
- b. *tawrã-yã*
 a.tribe-FEM
 ‘a Tawrã woman’
- c. *cya-yã*
 3-FEM
 ‘she’

The feminine suffix *-yã* generally refers to female *human* referents but this is dependent on context and what one means to express. Example 132a is from a story in which a *takwĩkã* fish transforms into a woman. The use of the feminine marker is allowed since the referent has human-like features in the story. The use of this suffix with non-human nouns and adjectives may also imply a derogatory use, as in Example 132b. Lastly, in Example 132c, the adjective *tabró* ‘colorful’ is only used for animals and, therefore, the use of *-yã* is understood as a joke, implying colorful or mottled females.

- (132) a. *takwĩkã-yã*
 type.of.fish-FEM
 ‘a female *takwika* fish’ [TA23_FishStory_63 126]

b. *kwâk-yâ*

dog-FEM

‘a dog-like woman’ (derogatory)

c. *?tabró-yâ*

colorful-FEM

(?)‘a colorful female’ [last two from field notes 161024]

4.5.3 Prefix for kinship terms *na-*

There are no specialized possessive pronouns in Tawrã; an apposed pronoun + noun is enough to express possession: *nyu àng* (2S house) ‘your house’. However, a group of close kinship terms (§3.3.1.2) in Tawrã have an irregular 1st person singular possessive prefix *na-* that is otherwise not found among nouns. This is also the only functional prefix in the language, cf. the discussion on *ka-* in §3.7.3.

Having more than one strategy of possession depending on certain parameters (be it stylistic, semantic or discursive, among others) has been termed *split possession*. North and South America and Micro- and Polynesia are hot spots for split possession, especially regarding kinship terms and body parts (i.e. inalienable terms)⁴ (cf. Stolz et al. 2008, 30f. Aikhenvald 2012b, 3f.). As an example, Dene Sųłiné (Athabaskan, Canada) has an irregular 1st person singular possessive prefix for only one close kinship term and therefore resembles the Tawrã model: *ɛnɛ* ‘my mother’ but *-q* ‘mother’, e.g. *bq* ‘his/her mother’ cf. the regular 1st person singular possessive prefix *s(ɛ)-/s(i)-* e.g. *sɛliɛ* ‘my daughter’ and 3rd person possessive prefix *b(ɛ)-/b(i)-* e.g. *bɛliɛ* ‘his/her daughter’ (adapted from both Elford and Elford 1998, 17;43;356;357; and Holden 2013, 42).

Closer to the Himalayas, it is noteworthy that neighboring Idu and Kaman do not have different types of possessors (although Meyor is reported to, cf. Jacquesson 2015,

⁴It is often the case that nouns like body parts and kinship terms belong to a group of inalienable, possessed nouns, i.e. nouns that cannot occur without a possessor (Aikhenvald 2012b, 4). This is not the case for Tawrã, as there are no nouns that must obligatorily be possessed. Even those kinship terms that occur with the *na-* prefix may occur without it, e.g. when referring to a general statement, e.g. ‘fathers should ...’ would be uttered as *bà (lâng)* (father COLL) i.e. it would not be possessed.

102f.). A larger overview of possessive strategies across the languages of the Himalayan region is a desideratum.

Turning to Tawrã, the possessive prefix *na-* (instead of the expected *hã* 1st person singular cf. §3.5.1.1 on personal pronouns) is only attached to a few close kinship terms and not to regular nouns, thus **na-àng* (intended: ‘my house’). For those kinship terms that do take the *na-* 1st person singular possessive prefix, that is also their citation form. Furthermore, the *na-* kinship prefix is sometimes extended to include a 1st person plural, as in Example 133.

(133) (...) *weligũ ning nayã natya ìna*

weligũ ning na-yã na-tya ì-na

and.then 1P 1.POSS-grandmother 1.POSS-grandfather be.there-FACT

‘(...) after that our grandmother and grandfather stayed there.’ [TA23_Malo_clan 57 335]

Table 4.6 displays some kinship terms that take the prefix *na-*. Most close kin terms take the *na-* prefix but there are some remarkable exceptions, like *pamru* ‘younger brother’ and *pathi* ‘younger sister’⁵. Furthermore, in-laws and generations below EGO do not take the *na-* prefix either. A description of kinship terms in Tawrã is available in §3.3.1.2.

Table 4.6: Kinship terms with initial *na-* (excerpt)

Kinship term	Meaning
<i>na-bâpak</i>	my uncle (father’s younger brother)
<i>na-bà</i>	my father
<i>na-bi</i>	my elder sister
<i>na-kâw</i>	my maternal uncle
<i>na-pû</i>	my elder brother
<i>na-tya</i>	my grandfather (also ‘ancestor’)
<i>na-ma</i>	my mother
<i>na-yã</i>	my grandmother

⁵The initial syllable *pa-* could have had a similar possessive function but, in the modern language, it is fused to the root and has neither syntactic independence nor any traceable semantic value left.

5 Classifiers

Classification lies at the core of human cognition. Although there is a neurobiological base for the prominence of shapes in classification systems (cf. Kemmerer 2017, 406f), the literature on classification has recognized the role that environment and metaphorical thinking play in influencing systems of nominal classification (cf. Ellen 2020, 162-164; Seifart 2005, 96; Adams 1989, 16f; Adams and Conklin 1973). Therefore, the study of classifiers lies at the intersection between morphosyntactic, semantic and cognitive questions, making it a promising field for interdisciplinary research.

Nominal classifiers (henceforth *classifiers*, for the sake of simplicity) occur along with numerals in numeral NPs and expressions of quantity (cf. Aikhenvald 2000, 98-100; Craig 1992, 277-289). The scholarly literature has also pointed out that classifiers are engaged in other grammatical functions, such as agreement, possession, definiteness, nominal modification and discursive uses among others (cf. Aikhenvald 2000, 318; Rose and Van linden 2023; Bisang 2017, 199f.). Tawrã is a classifier language and, although rare in discourse, classifiers are associated with anaphora and referentiality (individuation) (cf. Hotz forthcoming 2025).

This chapter is organized in the following manner: §5.1 defines the word class of classifiers by their occurrence with the numeral *one*. It also describes its main properties such as obligatoriness and basic semantic distinction. §5.2 introduces the classifier phrase (henceforth CLP) and its basic word order. §5.3 describes the main functions of the CLP: anaphora and referentiality, i.e. the individuation of referents. §5.4 discusses diachronic aspects including the lack of a nominal counterpart for many classifiers in the language. Finally §5.5 introduces *classifier compounds*, i.e. nominal compounds of the form N + CLF.

A list of Tawrã classifiers is found in Appendix II (§13). It includes all classifiers found so far in the language as well as a note on usage and etymology.

5.1 Basic properties of classifiers

As described in (Hotz forthcoming 2025), classifiers in Tawrã constitute independent grammatical and phonological words (§3.4.2). Phonologically, they possess their own stress and tone and do not fuse phonologically with either nouns or numerals (cf. §2.2.7).

Classifiers in Tawrã trigger a truncated form of the numeral ‘one’ *gê* (as opposed to *khing(.gê)*)¹. Numbers two and three also exhibit an equivalent morphological alternation when modifying classifiers but this is only known to conservative speakers of Anjaw district. Table 5.1 displays all three numerals and their alternation, i.e. normal and classifier forms. Forms for ‘two’ and ‘three’ are given in parentheses as most speakers either do not use them or employ them as normal numeral forms, i.e. without a classifier. Numerals are further discussed in §3.5.3.

Table 5.1: Allomorphs of numerals 1, 2 and 3

numeral	normal form	classifier form
1	<i>khing(.gê)</i>	<i>gê</i>
2	<i>kaqîng</i>	<i>(îng)</i>
3	<i>kasang</i>	<i>(sang)</i>

Example 134a displays a classifier in connection with a (truncated) numeral ‘one’. Conversely, in Example 134b the noun *àng* ‘house’ does not trigger the truncated numeral and the example is deemed incorrect; a noun can be modified by a numeral (without a classifier) in its normal form, as illustrated in Example 134c.

- (134) a. *ka gê*
 ka *gê*
 CLF:mouthful one
 ‘one mouthful (of food)’

¹Kera’a has equivalent morphological processes up to number six. Kaman has a similar process but it is restricted to the numeral *one* (*kîmu* vs. *mu*). Also cf. Chinese numeral ‘two’: *èr* when counting numbers and *liǎng* when used with classifiers and counting e.g. days. For an overview of different allomorphs of numerals ‘one’, cf. Greenberg (1974).

b. *àng gê

àng gê

house one

Intended: ‘one house’

c. àng khing.gê

àng khing.gê

house one

‘one house’ [all three: field notes 22102024]

These examples showcase a relevant property of classifiers in Tawrã: numerals may modify nouns without a classifier, a feature shared by some classifier languages of NEI².

On top of the non-mandatory nature of classifiers in Tawrã, animals and human nouns only seldom take a classifier. In Example 135, a big rock is modified by the numeral, without a classifier. Similarly, in Example 136 the frogs do not have a classifier.

(135) mlă kē gũ phlă pû daráng khün ã la

mlă kē gũ phlă pû daráng **khing** ã la

village LOC.INS LOC stone.MOD rock big one be.there HSAY

‘In a village there was a big rock.’ [TA22_kutyemba_phla 4 73]

(136) pahwă kaqîng masáng hela múhorẽ

pahwă **kaqîng** masáng hela mú-hô-rẽ

frog two tree under hide-STAT-PL.II

‘Two frogs were hiding under the tree.’ [TA22_frog1 132 680]

Human nouns rarely take classifiers in Tawrã³, a feature shared with Tani languages (cf. Doley and Post 2012, 258; also cf. Post 2022). In Example 137, a ‘child’ is modified

²This contrasts with well-known classifier languages such as Thai or Mandarin where classifiers are obligatory when a numeral modifies a noun (Enfield 2021, 290; cf. Iwasaki and Horie 2005, 74f).

³It should be noted that some classifiers, that are generally employed for plants, have an extended meaning that refers to people. One example is *hadûm* ‘CLF:groups of plants and thickets’, which also applies to same-gender siblings. The metaphorical extension of some classifiers from plants to humans should be further explored.

by the numeral alone. Similarly, in Example 138 the noun *m̃ya* ‘wife’ does not take a classifier.

- (137) *webo a khing m̃ya khrâk*
webo a khing m̃-ya khrâk
 and.then child one give.birth-ALLO EVI
 ‘And then (they) had a child.’ [TA23_tayang_clan 52 322]
- (138) *m̃ya îng k̃ gũ khing m̃ya we mu shêi na*
 [*m̃ya îng*]_{NP} *k̃ gũ [khing m̃ya we mu]*_{NP} *shêi na*
 wife two LOC.INS LOC one wife DEM also die INSIST
 ‘From the two wives, one died, right?’ [TA24_land 256 473]

Semantically, classifiers are generally subdivided into two groups: *sortal* and *mensural* classifiers (cf. Aikhenvald 2000, 114f; also Adams 1989, 3; cf. Becker 1975, 114). *Sortal* classifiers refer to individuals based in terms of kind while *mensural* classifiers refer to a particular quantity (Lyons 1977, 463; Aikhenvald 2000, 114f). Thus, the classifier *na* ‘CLF:2-D.foldables’ distinguishes individuals based on their materiality and is sortal while *asû-grũ* refers to the measurement from the chest to the fingertips (half an arm span), may be employed for quantities e.g. wood, and is mensural. In terms of morphosyntax, however, there are no differences between mensural and sortal classifiers.

The syntactic characteristics of classifiers are discussed in §5.2.

5.2 Syntactic constituency

The classifier and the numeral form a syntactic constituent within the NP: the *classifier phrase* (CLP) (Hotz forthcoming 2025). Evidence for this constituency comes from a) its fixed Clf-Num order, in which no other element can intervene, and b) its alternative position before the head noun, as will be discussed below. Table 5.2 illustrates the structure of the CLP. All elements in brackets can be ellipsed. This word order can be considered to be pragmatically the most neutral.

Table 5.2: Structure of the classifier phrase

(Dem)	(N)	[Clf	(Num)] _{CLP}

The word order CLF-NUM is typologically uncommon, although found in neighboring Tibeto-Burman languages (cf. Evans 2022; Jiang et al. 2013, 90). An alternative order NUM-CLF has only been attested with classifiers borrowed from Indo-Aryan (henceforth IA) languages (the same phenomenon has been described for Boro, cf. Aikhenvald 2000, 105; citing Greenberg 2013[1972], 326).

The structure of the CLP illustrated in Table 5.2 has a further complication: if the numeral is higher than ten, the classifier forms a unit with the lower enumerator (from right to left) while all other enumerators occur to the left of that unit. Table 5.3 illustrates this structure based on three numbers: thirty, thirty-two and one hundred thirty-two. In the case of thirty, there are two variations: in the first variation the classifier bonds with *kasang halâng* ‘thirty’ (lit. three ten). An alternative is to have a proxy ‘one’ (cf. English universon *one hundred*, also IA, cf. Hindi *ek sau* ‘one hundred’). Crucially, in both cases the classifier bonds with the lower enumerator, even if it is a multiplier of ten. This same structure is found in thirty-two and one hundred thirty-two: the classifier bonds with the lower enumerator from right to left. It should be noted that classifiers rarely occur with numbers higher than ten with the exception of the classifier *pla*, which is used for money (as well as blades, planks, etc.).

Table 5.3: Structure of the CLP higher numerals

Expression	Tawrã	Structure
30 rupees	<i>pla kasang halâng</i>	[CLF 3 10]
variation	<i>kasang halâng pla gê</i>	3 10 [CLF 1]
32 rupees	<i>kasang halâng pla kaqîng</i>	3 10 [CLF 2]
132 rupees	<i>malûm gê kasang halâng pla kaqîng</i>	100 1 3 10 [CLF 2]

The CLP may occur before or after the head noun. This apposition is triggered by referential and topical factors: the CLP after the NP head is the most neutral configuration,

while the opposite (CLP occurring before the NP head) is marked and indicates specificity. Table 5.4 illustrates both word orders. Note that all elements in parentheses can be omitted.

Table 5.4: Possible structure of the CLP

Possible structures	Constituents
A) post-noun	(N) [CLF (NUM)] _{CLP}
B) pre-noun	[CLF (NUM)] _{CLP} (N)

Examples 139a and 139b illustrate the difference in word order. In Example 139a, the CLP occurs after the head noun. Conversely, in Example 139b, the CLP occurs before the head noun, which prompts a specific reading of the NP.

(139) a. *tyû pum kasang malâde*

tyû [pum kasang]_{CLP} malâ-de
 chicken CLF:3D.round three search-EGO.HAB
 ‘I am looking for (any) three chickens.’

b. *pum kasang tyû malâde*

[pum kasang]_{CLP} tyû malâ-de
 CLF:3D.round three chicken search-EGO.HAB
 ‘I am looking for three (specific) chickens (e.g. the ones I lent to somebody).’
 [both 011024_STBT_eli2 175]

5.3 Functions

Classifiers are often employed as devices for anaphora⁴ and as markers of referentiality. It should be noted that referential properties, such as identifiability, specificity or individuation, are intertwined and can hardly be taken apart (e.g. *uniqueness* for Löbner

⁴In some languages, classifiers may introduce new participants (cf. Hopper 1986, 312f). This use is not available in Tawrã.

2011, 283f.). In many cases, the referential value of classifiers in Tawrã corresponds to individuation (Lowe 2012, 103), especially the use of the bare classifier construction, as will become clear below.

Classifiers can be employed to substitute a referent, i.e. anaphorically (cf. Pichetpan and Post 2021, 464; Evans 2022, 8; Bisang 1999, 147f). In Example 140, the classifier *pí* ‘CLF:2-D folds’ stands for the noun ‘knee’.

(140) grũ mu pí lyũm la

grũ mu pí lyũm la

leg also CLF:2D.folds eight HSAY

‘(His) legs were also eight knees long (lit. eight foldings), it is said.’

[TA23_newmyth 14 84]

Examples 141a and 141b are from the same story, in which a big rain and a wildfire destroy the world. In Example 141a, the protagonists try to save themselves by making an improvised hut of eight layers of banana leaves. In Example 141b, we learn that four layers remained after the fire came closer and started to burn down the hut. The use of the classifier for layers *halá* stands here for the house.

(141) a. tatyâng we gũ halá lyũm àng bâya gũ ì la

tatyâng we gũ [[halá lyũm]_{CLP} àng]_{NP} bâ-ya gũ ì

a.leaf DEM INS CLF:2D.layers eight house build-ALLO NF be.there

la

HSAY

‘This (banana) leaf, having made the house of eight layers, (they) lived (there), it is said.’ [TA23_newmyth 22 136]

b. halá kaprêi we lüya la

[[halá kaprêi]_{CLP} we]_{NP} lü-ya la

CLF:2D.layers four DEM remain-ALLO HSAY

‘Four layers (i.e. the house made of banana leaves) remained, it is said.’

[TA23_newmyth 24 152]

The rich semantics of classifiers, when employed anaphorically, can be exploited to reduce the number of potential referents. In Example 142, the use of *na* ‘CLF:2-D foldables’ can only refer to things like leaves or shirts, i.e. the head noun can be omitted if it is clear from the context.

(142) *na gê hàngna*

na *gê hàng-na*
CLF:2D.foldables one give-IMP

‘Give me one (foldable, e.g. a shirt)!’ [field notes 20102024]

Commonly, however, the topic is maintained in a conversation and there is no need to narrow down potential referents. Hence the generic classifier *bra* is commonly employed, as in Example 143. The generic classifier may readily replace other more specific classifiers (cf. Vittrant and Allasonnière-Tang 2021, 748).

(143) *we bra tuluqa ìya (...) tuluqa ìya tahũ bônasâ we*

we bra tuluqa ì-ya tuluqa ì-ya tahũ
DEM CLF:generic a.place be.there-ALLO a.place be.there-ALLO group
bô-na-sâ we
go-FACT-POT DEM

‘This one living (at) Tuluqa (...) the group living (at) Tuluqa, where might they have gone?’ [TA23_talyu_clan 43 264b]

Classifiers are associated with a range of semantico-pragmatic features that revolve around referentiality (Aikhenvald 2000, 93; Bisang 1999, 114f; Bisang 2017, 216f). This requires some exemplification, as other grammatical devices may also single out a referent, e.g. demonstratives (Hanks 2018[1996], 181) or adverbs like ‘alone’ or ‘only’. In concrete terms, the pragmatic effect of classifiers often surfaces as a referent being individualized, i.e. singled out as one identifiable argument. In Example 144, a nephew plans to kill the speaker’s grandfather. After an unsuccessful attempt by the nephew to ambush the speaker’s grandfather with a group of clansmen he is forgiven and allowed to come and live in the speaker’s grandfather’s village. The speaker employs the CLP to refer to the nephew alone, as opposed to his clansmen.

- (144) we jîng bô ja mu di hanéna yüm gũ mane dâi bra gê adyû bra gê mane (...) hanatabô mane mlâ kě

[we jîng bô]_{sub.cl} [ja mu di hané-na yüm gũ]_{cosub.cl} mane
 DEM end TEMP.SUB what also do be.able-FACT NEG NF FILL(<Asm)
 [dâi [bra gê]_{CLP}]_{NP} [adyû [bra gê]_{CLP}]_{NP} mane
 a.tribe CLF:generic one nephew CLF:generic one FILL(<Asm)
 hana-ta-bô mane mlâ kě
 come-be.allowed-PFV.ALLO FILL(<Asm) village LOC.INS

‘When it ended and he could not do anything, I mean, this one Idu, this one nephew, I mean, was allowed to come, I mean, to the village.’ [TA22_CLAN2 140 261-2]

Classifier phrases in Tawrã may occur *bare*, i.e. without a numeral. Following Pichetpan and Post (2021, 495), *bare classifiers* reveal the distilled function of classifiers in Asian languages as markers of the referential function of individuation (also cf. Bisang 1999, 147,178; Li and Bisang 2012, 337). In Example 145a, the head of the NF (‘bee’) occurs without any modifier (demonstrative or classifier) and, without context, it may either refer to a unit, a generic bee or a mass. Conversely, the bare classifier in Example 145b forces an individualizing reading; a single bee. This form may not take the collective marker, as illustrated in Example 145c.

- (145) a. tawà katyângdyibô
 tawà katyâng-dyibô
 bee hear-EGO.PFV
 ‘I heard a bee/a swarm of bees.’
 b. tawà bra katyângdyibô
 tawà bra katyâng-dyibô
 bee CLF:generic hear-EGO.PFV
 ‘I heard one (single) bee.’

c. *tawà bra làng

tawà bra làng

bee CLF:generic COLL

Intended: ‘a group of bees.’ [all three 011024_STBT_eli2 263]

Bare classifiers can be reduplicated (cf. §3.6) to indicate a particular way of performing a task: one by one or one after the other (Hotz forthcoming 2025; also cf. Matisoff 1973a, 93). In Examples 146 and 147, the reduplication of the classifier implies ‘one after the other’.

(146) mề bra bra bôga

mề bra bra bô-ga

person CLF:generic CLF:generic go-PL.I

‘The men go one after the other.’

(147) tîng na na shîna

tîng na na shî-na

shirt CLF:2D.foldables CLF:2D.foldables take-IMP

‘Take the shirts one by one!’ [both field notes 18102024]

The bare classifier construction seems to be the origin of the participation of classifiers in nominal compounding, termed here *classifier compounds*. These will be further discussed in §5.5.

5.4 Origin of classifiers

Tawrã has a large number of classifiers, with more than one hundred members (cf. §13) and further research is bound to find many more. Nouns are cross-linguistically the main source of classifiers (Aikhenvald 2000, 100, 120, 442; cf. DeLancey 1986b, 438; Dixon 1986, 110-111). In Tawrã, a large number of classifiers have a clear nominal counterpart (cf. §3.2.1), e.g. *taplá* CLF:spoonfuls (from the noun *taplá* ‘spoon’) or

tyàng ‘CLF:bodies’ of large animals (including people) (from *tyàng* ‘body’). Mensural classifiers are commonly derived from nouns, including body parts, e.g. *khalyâng-pà-hapyê* (< *khalyâng-pà* ‘shoulder’) ‘CLF: water up to the shoulder’ and tools and objects, cf. *wing* both ‘cup’ and ‘CLF:cups’. Furthermore, nouns with a potential for being used as measurements are particularly prone to be adapted as classifiers, cf. *garêi* ‘CLF:carfuls’ (from *garêi* ‘car’, an adapted IA loanword, cf. Hindi *gāḍī* ‘car’) or *pis* CLF:pieces (ultimately from English *piece*). In such cases, the classifier can be distinguished from the noun (the physical object) by the form of the numeral *one* that it triggers. Thus, in Example 148a, *taplá* stands for the noun ‘spoon’ whereas in Example 148b *taplá* stands for the classifier for spoons (spoonfuls) and triggers the numeral in its classifier form (cf. §5.1).

(148) a. *taplá khing*

spoon one

‘one spoon’

b. *taplá gê*

spoon one

‘one spoonful (e.g. of food)’ [both field notes 18102024]

IA words for time are also readily borrowed as classifiers, cf. *gânta* CLF:hours (cf. IA *ghaṇṭā* ‘hour’) or *hâpta* ‘CLF:hours’ (cf. IA *haphtā* ‘week’, ultimately a Persian loanword).

Three classifiers have an initial syllable *ka-*, these are *kahîng* ‘CLF:day(times)’ (cf. *hîng* ‘day’), *kanáng* ‘CLF:years’ (cf. *náng* ‘years’) and *kalyâng* ‘CLF:evenings’ (cf. *lyang* ‘half’). This syllable could be related to the prefix in numeral two *kaqing*, three *kasang* and four *kaprei*. Furthermore, it occurs as an old prefixed root in names of small animals and insects, e.g. *kacêk* ‘mole sp.’ or *katài* ‘grasshopper sp.’, as discussed in §3.7.3.

For an important number of classifiers in Tawrã, however, there is no nominal equivalent. This suggests that the system of classifiers in the language is considerably old and the relation to nouns has been severed by a) natural language change (cf. Hock 1991, 627f; Grinevald 2000, 84) and potentially b) contact phenomena. The latter point will not be expanded here but it suffices to say that several classifiers in Tawrã have a corre-

spondence in Kaman, cf. §13. The effects of language contact between Kaman and Tawrã remain still to be thoroughly studied.

5.5 Classifier compounds

Classifiers may occur as the second element of a nominal compound (cf. §4.2), as illustrated in Table 5.5.

Table 5.5: Structure of classifier compounds

first element	second element
N	Class

Such compounds, here termed *classifier compounds* (cf. Hotz forthcoming 2025), can be considered to be nouns in syntactic and semantic terms, since they a) can be counted with a (generic) classifier or a normal numeral (i.e. not a numeral in its truncated, classifier form cf. §5.1), b) can take the collective marker *làng* (§3.3.1.4) and c) the meaning of the compound often exceeds the mere sum of its constituent parts, a basic definition of hierarchical (nominal) compounds (cf. Matthews 1991, 86f). In Example 149a, the classifier compound *krú-na* (< *k(u)rù* ‘head’ and *na* ‘CLF:2-D.foldables’) ‘ear’ is modified by a CLP. The classifier *pla* ‘CLF:2-D.flat’ can be replaced by the generic *bra* and the whole phrase could be replaced by a sole numeral (in this case, *khing(.ge)*). Conversely, in Example 149b the classifier compound cannot be modified by a truncated form of the numeral ‘one’; either a CLP or a numeral in its normal form would be appropriate. In Example 150a, a classifier compound *tükü-bra* (< *tükü* ‘matter, subject, language’ and *bra* ‘CLF:3-D.small.round/generic’) ‘letter, story’ takes the collective marker *làng*. Note that this is otherwise not available for bare classifiers, as discussed in §5.3 and illustrated in Example 145c (here repeated for clarity as Example 150b)

- (149) a. *krú.na pla gê*
 ear CLF:2D.flat one
 ‘one ear’

b. **krú.na gê*

ear one

‘Intended: one ear’ [both 011024_STBT_eli2 395]

(150) a. *tükü.bra làng*

letter COLL

‘letters’ [field notes 21102024]

b. **tawà bra làng*

tawà bra làng

bee CLF:generic COLL

Intended: ‘a group of bees.’ [011024_STBT_eli2 263]

The relation between the meaning of the whole compound and its constitutive parts is not always straightforward, as it might reflect a) an emic perspective of classification that might not be accessible or transparent to an outsider (Denny 1979, 117) and b) a complex grammaticalization process. Regarding the former point, exploring the driving forces behind classification would exceed the goals of this grammar. Regarding the latter point, these compounds display a certain level of variation in their compositionality, cf. the case of *sâi-pum* (< *sâi* ‘iron’ and *pum* ‘CLF:3-D.round’), which stands for any round thing made of metal but also, specifically, for a weighing stone. It thus seems that the compounds transit from a transparent X + Y meaning to a lexicalized, new meaning Z that exceeds the sum of the constituent parts - this is a central aspect of grammaticalization (cf. Givón 1986, 92).

Table 5.6 displays all classifier compounds found so far in the language. For some forms, the constituent parts are not transparent anymore. The classifier compound for ‘finger’ has two variants: one with the classifier *brü* ‘CLF:3-D.bamboo-shaped things’ and one with the classifier *bra* ‘CLF:generic’ *habrang-brü/habrang-bra*.

Table 5.6: List of classifier compounds in Tawrä

Classifier compound	First element (N)	Second element (Class)
<i>apó-pum</i> ‘heart’	<i>apó</i> ‘lung’	<i>pum</i> ‘CLF:3-D.round.mid-size’
<i>habra</i> ‘millet sp. (<i>Setaria italica</i>)’	*ha- ???	<i>bra</i> ‘CLF:3-D.round.small’
<i>krú-na</i> ‘ear’	<i>k(u)rù</i> ‘head’	<i>na</i> ‘CLF:2-D.foldables’
<i>sâi-pum</i> ‘weighing stone’, also any round thing made of metal	<i>sâi</i> ‘iron’	<i>pum</i> ‘CLF:3-D.round.mid-size’
<i>shâ-brã</i> ‘testicles’	*shâ- ???	<i>brã</i> ‘CLF:1-D.rope.like’
<i>tükü-bra</i> ‘word, letter’	<i>tükü</i> ‘matter, thing, language’	<i>bra</i> ‘CLF:generic’
<i>thalyâng-na</i> ‘tongue’	*thalyâng- ???	<i>na</i> ‘CLF:2-D.foldables’
<i>takhrê-brü</i> ‘stick, lathi’	<i>takhrê</i> ‘stick, lathi’	<i>brü</i> ‘CLF:3-D.bamboo-shaped things’
<i>habrang-brü/habrang-bra</i> ‘finger’	<i>habrang</i> ‘finger’	<i>brü</i> ‘CLF:3-D.bamboo-shaped things’, <i>bra</i> ‘CLF:generic’

6 Predicate complex

The predicate complex is the amalgamation of elements that constitute the head of a predicate clause (cf. §7.2). The head of a predicate clause projects both the clausal argument structure and the TAM structure of the predicative clause.

This chapter is organized thus: §6.1 discusses the predicate structure, including stem types (§6.1.1). §6.2 discusses predicate derivations and includes an introduction to theoretical considerations (§6.2.1). Predicate derivations are divided into three groups: valency operators (§6.2.2), manner/modality (§6.2.3) and aspect/Aktionsart (§6.2.4). §6.3 describes TAM and Number suffixes including a discussion on relevant concepts (§6.3.1.1). TAMN (tense, aspect, mode and number) suffixes are divided into five groups: number (§6.3.2), associated motion (§6.3.3), Aktionsart (§6.3.4), command moods (§6.3.5) and other suffixes (§6.3.6). Negation markers are discussed in §6.4. Finally, nominal incorporation is described in §6.5.

6.1 Predicate structure

A predicate in Tawrã consists minimally of a verbal stem. The predicate may be optionally expanded into a *complex predicate* by one or more predicate derivations that occur between the stem and the final suffixal chain. The predicate can also take *tense, aspect, mode and number* (henceforth TAMN) suffixes. TAMN suffixes and predicate derivations are two position classes that do not overlap syntactically, even though some morphemes may have meanings that are related (§6.2.1). The slot for negation comes after TAMN suffixes (cf. §6.4.1). Table 6.1 displays an example of a predicate complex with a predicate derivation, a TAMN suffix and the general negator *yüm*.

Table 6.1: Predicate structure: Overview

Structure:	Verbal stem	Predicate derivation	TAMN suffix	Negation
Example:	<i>yúm</i>	<i>-dâng</i>	<i>-q</i>	<i>yüm</i>
	fly	away	CONT	NEG
	‘S/he does not fly away.’			

A predicate head can occur without a TAMN suffix, i.e. bare verbal stems are allowed and their temporal interpretation is dependent on the context (cf. Timberlake 2007, 305f). In Example 151a, the predicate lacks a TAMN suffix. This example is grammatical, although it requires some context to be interpreted as either past or non-past. Example 151b portrays a TAMN suffix on the verbal stem. In Example 151c, the verbal stem is expanded by a predicate derivation and it also takes a TAMN suffix.

- (151) a. *cya* *thâ*
 3 eat
 pronoun V.stem
 ‘s/he eats/ate.’ [021024_STBT_eli2 41]

- b. *cya* *thâ-q*
 3 eat-CONT
 pronoun V.stem-TAMN.suffix
 ‘S/he is eating.’ [021024_STBT_eli2 25]

- c. *hã* *khô-jî-dyibô*
 1S cry-CAUS-PFV.EGO
 pronoun V.stem-predicate.derivation-TAMN.suffix
 ‘I made him/you cry.’ [071023_STBT_eli3 646]

In Example 152, the predicate is negated. Negation occurs after the predicate by the negator, either *yüm* or *shyüw*. Negation is discussed in §6.4.1.

- (152) pra mu hwě.tyâng yüm
 pra mu hwě.tyâng yüm
 good also see NEG
 ‘I do not see well.’ [TA22_chicken_thief 66 241]

6.1.1 Stem types

The stem constitutes the core of the predicate. It can be simple, compound, discontinuous or reduplicated. This section will discuss all stem types. Table 6.2 summarizes the main stem types along with a scheme.

Table 6.2: Stem types

Type of stem	Scheme	Example	Reference
simple	V	<i>mará</i> ‘laugh’	§6.1.1.1
compound	V-V	<i>ì-kha</i> (lit. be there-lie down) ‘sustain oneself’	§6.1.1.2
discontinuous	V V	<i>câng-ya ru-ya</i> (get-ALLO get.together-ALLO) ‘get rich’	§6.1.1.3
reduplication	$V_1 \sim V_1$	<i>tsâi~tsâi</i> (count~PLAC) ‘keep counting’	§6.1.1.4

6.1.1.1 Simple stem

A simple verbal stem consists of one single verb root, as in *îng* ‘sleep’. This is the most common type of stem in the corpus. The stem alone does not hold an *a priori* temporal interpretation but is dependent on the context.

6.1.1.2 Compound stem

A compound stem consists of two lexicalized verbal stems, neither of which can be considered to be the head; they are in a symmetrical relation. Compound stems have one single, non-compositional meaning, i.e. the *denotatum* is not a mere addition of its constitutive parts (Matthews 1991, 82f; also Aikhenvald 2007, 24; note *elaborate expressions*, cf. Matisoff 1973a, 81-82; Konnerth 2014, 578). An example would be *chêi-dě* ‘be alive’, which is composed of *chêi* ‘walk’ and *dě* ‘stand’. These compound stems also operate as one stem in terms of how they are referenced by the larger clausal: the final suffixal chain agrees the compound stem. No element can intervene between the two stems in a compound. In Example 153, the compound stem occurs with the verb *shyâ* ‘become’. Similarly, in Example 154 the verbs *la* ‘say’ and *khû* ‘fight’ render ‘have an argument’. Table 6.3 displays all known compound stems.

(153) mé ning (...) brútshù shyâdi

[mé ning [brú.tshù shyâ-di]_{VP}]_{main.cl}

all 1P increase become-IPFV.EGO

‘We all, we expanded.’ [TA22_origin 2 16]

(154) (...) ja shyâna? webo gũ lakhûganabi gũ (...)

[ja shyâ-na]_{main.cl} [webo gũ la-khû-ga-na-bi gũ]_{cosub.cl}

what happen-FACT and.then INS speak-retaliate-PL.I-FACT-PFV NF

‘(And then) what happened? having had an argument (...)’ [TA23_talyu_clan_15 83]

Table 6.3: Complex stems

Compound stem	Meaning	Term 1	Meaning	Term 2	Meaning
<i>chêi-dế</i>	be alive	<i>chêi</i>	walk	<i>dế</i>	stand
<i>ì-kha</i>	sustain oneself	<i>ì</i>	be there	<i>kha</i>	lie down
<i>zû-lô</i>	write	<i>zû</i>	write	<i>lô</i>	pierce
<i>brú-tshù</i>	increase	<i>brú</i>	be enough	<i>tshù</i>	increase, multiply
<i>nyûm-ûm</i>	be soft	<i>nyûm</i>	pierce	<i>ûm</i>	be old
<i>kasâ-shì</i>	consider	<i>kasâ</i>	know	<i>shì</i>	take
<i>thâ-akâp</i>	steal	<i>thâ</i>	eat	<i>akâp</i>	steal
<i>la-khû</i>	have an argument	<i>la</i>	talk, say	<i>khû</i>	fight

6.1.1.3 Discontinuous stems

Discontinuous stems are paired verbal stems which share the same suffixal chain. They look like two adjoined finite predicates but they form a syntactic unity with varying degrees of lexicalization. This structure has so far only been found with one gifted storyteller; it is not clear how common it is in naturalistic speech.

In Example 155, the non-finite verbs *tyé* ‘cut’ and *apyék* ‘spread’ form a transparent coordination of actions. Conversely, Example 156 has two finite verbs with one non-compositional (i.e. lexicalized) meaning that is similar to the compound stems described in §6.1.1.2. Similarly, in Example 157 the two finite verbal stems have a non-compositional meaning and share the same suffix.

(155) *tyékũ apyékkũ gũ yang túm khalay pwôktè*

[[tyé]_{V.root-kũ} [apyék]_{V.root-kũ}]_{disc.VP} gũ]_{cosub.cl} yang [hatûm khalai]_{NP}

cut-COMPL.II spread-COMPL.II NF xxx sky ground

pwôk-tè

blow-IMP.NARR

‘By cutting (and) spreading it, blow it everywhere (lit. (through) sky and land)!’

[TA22_origin 54 157]

(156) cângya ruya tahwîbô (...)

[[[câng]_{v.root-ya} [ru]_{v.root-ya}]_{disc.VP/compl.cl} tahwî-bô]_{main.cl}

get-ALLO get.together-ALLO teach-PFV.ALLO

‘(God) taught us how to get rich (lit. get and get together) (...)’ [TA22_origin 84 245]

(157) ah alí gũ kûmna bwîna(...)

ah alí gũ [[kûm]_{v.root-na} [bwî]_{v.root-na}]_{disc.VP}

ah like-this INS bow.down-IMP dance-IMP

‘Worship (lit. bow down and dance) like this!’ [TA22_origin 106 308]

6.1.1.4 Stem reduplication

Reduplication of verbs is not a common strategy in Tawrã; there are few attestations in the corpus and there are grammatical morphemes that compete in having a similar semantic effect as reduplication (cf. §6.3.3.1). Stem reduplication is here understood as a subtype of reduplication as discussed in §3.6 that differs in its morphosyntax.

A simple stem may be reduplicated right after the predicate head. I follow Gil (2005, 31) in distinguishing reduplication from repetition: repetition is driven by discourse and syntax whereas reduplication is a morphological process that affects the grammatical word; a reduplicated stem is still *one* predicate whereas repetition renders two or more predicates. A repeated predicate can be identified by having TAMN suffixes §6.3 on both stems, something that is not possible in verbal reduplication. However, since a verbal stem does not require a TAMN suffix or a derivation to occur as a finite predicate, the distinction between reduplication and repetition is convoluted, at best. In Example 158, the predicate *hanâ-mâ* is repeated along with an iterative marker. They form conjoined clauses (§10.2). Conversely, on a reduplicated stem there can be no intervening morphemes between the stem and the reduplicant. Example 159 portrays a case of reduplication; the TAMN suffixes occur after the reduplicated stem.

- (158) hanâmâ hanâmâ hrû yüm bô khümyûa mamýú bô (...)
 [hanâ-mâ hanâ-mâ hrû yüm bô]_{sub.cl} khümyû-ya]_{main.cl} mamýú bô
 press-ITER press-ITER burn NEG TEMP.SUB get.angry-ALLO after TEMP.SUB
 ‘Having squeezed and squeezed, (the fire) did not burn and then he got angry.
 Later (...)’ [TA22_tiger_116 262]

- (159) bôdângkûnabô mane e lyâlyâkûnabô
 bô-dâng-kû-na-bô mane we
 go-away-COMPL.II-BACK-PFV.ALLO FILL(<Asm) DEM
 lyâ~lyâ-kû-na-bô
 return~PLAC-COMPL.II-BACK-PFV.ALLO
 ‘(The Britishers) were already gone, they had already gone back.’ [TA22_CLAN2
 10 17]

The semantic effect of verbal stem reduplication can be understood within the broad frame of *pluriactionality* (PLAC). Pluriactionality has been defined as the plurality or multiplicity of states and events and, as such, it is related to number (Mattiola 2019, 20-28; also cf. *verbal number*, Mithun 2001, 83-86). The main function of reduplication in Tawrã is the distribution of the single occurrences of an event through time, something akin to iteration (cf. Mattiola 2019, 22f). In Example 160, the reduplication of the stem renders an iterative event: (the tiger) kept eating and eating. Crucially, the effect of reduplication can apply to longer periods of time, although always as part of the single overarching event or action. In Example 161, the reduplicated verb ì ‘be there, stay’ refers to the event (‘staying’) over a long period of time.

- (160) cya asáng thâthâ hĩ gũ (...)
 cya asáng thâ~thâ hĩ gũ
 3 raw eat~PLAC CONC NF
 ‘By eating and eating raw (meat) (...)’ [TA22_tiger 127 281]

- (161) kanáng lyâng gê ìya la âm (...) we ìì (...)
 [kanáng lyâng gê ì-ya la]_{main.cl} âm [we ì~ì~ì]_{main.cl}
 CLF:year half one be.there-ALLO HSAY yes DEM be.there~PLAC~PLAC
 ‘Half of the year he stayed (with us) (...) he kept staying (...)’ [TA22_CLAN2 148 275]

A reduplicated action naturally gives way to a plurality of referents, especially if the reduplicated verb is telic and/or the event cannot be performed twice on the same person (like ‘kill’) (Mattiola 2019, 26). In Example 162, the tiger is set to count the fish. The reduplication of the stem suggests an iterative action and, therefore, a high number of fish (assuming that each can be counted only once). A clearer case is Example 163 (repeated from Example 159), where the reduplication of the stem does not imply iteration but rather a large or otherwise indeterminate number of Britishers.

- (162) “ekhün gâqing” la tsâitsâi we la gũ thĩ tshâwhamâmâ khrâk
 [e=khing kâqing la tsâi~tsâi we la gũ thĩ]_{sub.cl} [tshâw-amâ-mâ
 DEM=one two say count~PLAC DEM say NF CONC move.fast-get.lost-ITER
 khrâk]_{main.cl}
 EVI
 ‘(The tiger) kept counting “one, two”, (but the fish) kept moving and disappearing.’ [TA22_tiger 35 102]

- (163) bôdângkũnabô mane e lyâlyâkũnabô
 bô-dâng-kũ-na-bô mane we
 go-away-COMPL.II-BACK-PFV.ALLO FILL(<Asm) DEM
 lyâ~lyâ-kũ-na-bô
 return~PLAC-COMPL.II-BACK-PFV.ALLO
 ‘(The Britishers) were already gone, they had already gone back.’ [TA22_CLAN2 10 17]

On top of the effect on number of reduplication, Tawrã makes a distinction between stative and active verbs, in which the latter have an effect on agency apart from pluralization interpretation and aspect, i.e. the S/A argument is understood to perform the act willfully. This is further discussed in §3.3.2.2.

6.2 Predicate derivations

A verbal stem may be optionally expanded by a derivational morpheme. These are morphologically bound and occur between the verbal stem (head of the predicate) and the TAMN suffix(es)¹. Semantically, predicate derivations are diverse: some predicate derivations have an abstract, grammatical meaning (e.g. benefactive, causative, etc.) while other predicates have a meaning that is close to their verbal counterpart. Predicate derivations are also diverse regarding their relation to other POS: although most predicate derivations have a clear origin in verbs, some of these morphemes are also cognate with adverbs/adjectives as well as verbs. This diversity of recruited material reflects different degrees of grammaticalization that co-exist in this construction (cf. Hopper 1991, 22).

This section is organized in the following manner: Some theoretical considerations on the label *predicate derivations* are discussed in §6.2.1. For the sake of clarity, predicate derivations are organized into three rough groups: valency operators (§6.2.2), manner/modality (§6.2.3) and aspect/Aktionsart (§6.2.4). Table 6.4 displays the predicate derivations covered in this section organized into groups.

¹This distribution is in line with the observation of Bybee (1985, 96) that derivational affixes (in this case predicate derivations) occur closer to the root than inflectional affixes (in this case, TAMN suffixes).

Table 6.4: Predicate derivations organized in groups

Groups	Predicate derivation
Valency operators	Associative <i>-né/nê</i> (§6.2.2.1)
	Causatives (§6.2.2.2)
	Bene/malefactive <i>-hàng</i> (§6.2.2.3)
	Permissive <i>-shyüw</i> (§6.2.2.4)
Manner/modality	Inopinative <i>-ta</i> (§6.2.3.1)
	Frustrative <i>-hã</i> (§6.2.3.2)
	‘by force’ <i>-mazâi</i> (§6.2.3.3)
	‘at last’ <i>-ha</i> (§6.2.3.4)
	‘in exchange’ <i>-dá</i> (§6.2.3.5)
	‘thoroughly’ <i>-makáw</i> (§6.2.3.6)
Aspect/Aktionsart	Permansive <i>-(ta)lá</i> (§6.2.4.1)
	Ambulative <i>-wi</i> (§6.2.4.2)
	Completive <i>-dâng</i> (§6.2.4.3)

6.2.1 Theoretical considerations: inflection, derivation and predicate derivations

6.2.1.1 Inflection, derivation

The distinction between inflection and derivation is a central issue in the history of linguistics. According to Bybee (1985, 81f.), discussing Greenberg (1954), inflectional morphology is obligatory in certain expressions while derivational morphology is not. Thus the progressive suffix *-ing* is obligatory in *the duckling was swimming* while the *-ling* diminutive suffix in the same sentence (*duckling*) is optional. In other words, inflectional morphemes are thought to be required by the syntax of the sentence. Furthermore, derivation is considered to craft new lexemes with their own meaning and sometimes even a different word class while inflection creates new forms of the same lexeme (Booij 2000). Even though these principles are widespread, comparative work has pointed out the huge amount of variation found in the world’s languages; finding a

common definition is a difficult undertaking (cf. Bickel and Nichols 2007; also cf. Booij 2005, 85f).

The work on complex predicates from South Asian languages has provided interesting examples that point to the wide spectrum between the parameters of inflection and derivation and to the complexity that lies therein. Furthermore, much of the study of grammaticalization has focused on the recurrent diachronic development in which independent forms are reanalyzed as bound elements and eventually absorbed by the stem (Givón 1971; Givón 1985; Hopper and Traugott 2003).

Closer to the Tawrã area, Coupe (2008, 297f.) describes the difficulty of categorizing some of the verbal suffixes in terms of inflection or derivation in his grammar of Mongsen Ao (Ao, North East India). This is particularly the case for those suffixes that are cognate with independent verbs, as these share lexical content and only occur with small semantic classes of verbs; they are therefore not part of the verbal paradigm (i.e. they are optional) and do not change the word class of the lexemes to which they attach. The author borrows a model of *layered grammaticalization* from Hopper (1991, 22f) to describe the situation: independent verbs form compounds and, after some time, the second verb in the complex predicate is recruited as a suffix. Gradually, the suffix relinquishes its lexical, additive, i.e. original meaning and acquires a new, grammaticalized meaning. The process of grammaticalization entails various changes: from limited productivity (only in a particular compound) the second element in the complex predicate may expand its scope of verb hosts with which it may occur and is commonly described as undergoing phonological erosion².

This discussion is central to making sense of the verbal complex in Tawrã, although it should be noted that I will not employ the term *inflection*, but rather *TAMN* (tense, aspect, mode and number) for the functional equivalent to inflection in other languages. This is because these verbal suffixes do not behave syntactically as inflection, as will become clear in §6.3. The crucial point is that there are two position classes (predicate derivations and TAMN) that do not overlap morphosyntactically, as presented in §6.1. Semantically, the picture is more complex. Although we may make a generalization in which

²The double change of phonological erosion and semantic bleaching is central to grammaticalization theory (Givón 1971; Hopper and Traugott 2003, 1f; Lehmann 1995[1982]; also cf. Campbell 2000).

i) derivation is associated with event semantics (e.g. valency changing derivation) and
 ii) TAMN suffixes with the speaker's attitude (e.g. egophoricity), many morphemes defy this simplification and share characteristics of distribution, productivity and semantic complexity of both. One example is the pair of completive morphemes: the derivational *-dâng* COMPL.I and the TAMN suffix *-kũ* COMPL.II. Although they occupy non-overlapping morphosyntactic positions, their meanings are often indistinguishable. We thus observe complex layers of (often ongoing) grammaticalization. Rather than pointing to a gray area between inflection and derivation, these cases remind us that diachronic processes are gradual and they imply a co-habitation of features of both source and target categories since the recruitment of material for these constructions is never simultaneous and thus different morphemes exhibit different degrees of semantic abstraction and phonological bleaching.

For the present description, the label *predicate derivation* has been chosen to refer to verb-selecting and predicate-expanding morphemes. The term *predicate derivation* was proposed as an analytical category to provisionally reference these constructions in a cross-linguistic sense without disregarding the ample (and largely unexplored) variations found in individual languages until a framework is established that allows both a synchronic and diachronic analysis (Post 2010, 175-177; also Konnerth 2014, 245f; Modi 2017, 138f).

Predicate derivations commonly display some of these features: they select a verbal host, they may display phonological dependency, they are optional (non-inflectional), they are semantically rich/complex, they often occur as independent verbs and are often translated by independent lexemes in other languages (Post 2010, 177). These points largely coincide with other labels found in the literature such as *compound verbs* (Doornenbal 2009, 473f.), *versatile verbs* (Matisoff 1973a, 220f) or *multiverb-constructions* (Aikhenvald 2012a, 304f.).

Crucially, however, the term *predicate derivations* highlights the morphological aspect of this construction as a type of word-formation in which a morphological application creates a new base from a simpler one (cf. Post 2010, 176).

In Tawrã, most predicate derivations have their origin in independent verbs and thus, as in other agglutinating languages, the question of morpheme identity is raised: are in-

dependent verbs and homophonous verb-selecting morphemes two different morphemes or just one with an abstract sense that is pinned down with a concrete meaning when used alone (cf. Bickel and Zúñiga 2017, 163; Widmer et al. 2021, 713). For example, the verb *dâng* ‘finish’ also occurs as a predicate derivation with both a concrete meaning (‘to finish doing X’; cf. Example 164a) and an aspectual (completive), more grammaticalized meaning where there is no *finishing* part (cf. Example 164b).

- (164) a. *hã thâdângli*
hã thâ-dâng-li
 1S eat-COMPL.I-REC.PST
 ‘I just finished eating.’ [100823_STBT_voc 1011]
- b. (...) *tili.tulu.tili tshângdâng gũ* (...)
tili.tulu.tili tshâng-dâng gũ
 manner.of.having.a.rash turn.pale-COMPL.I NF
 ‘(...) turning completely pale (...)’ [TA22_CLAN2_191 364]

Furthermore, the coexistence of cognate lexical and grammaticalized material gives way to a certain degree of ambiguity that is highly dependent on the context: In Example 165, the predicate derivation *-hàng* does not yield bene/malefaction (cf. §6.2.2.3) but a lexical sense ‘X and give’. This is probably because it would be implausible for a bird to have an intention that could be framed as inherently beneficent or maleficent (as opposed to a dog, cf. Example 175, also §6.2.2.3) but it helps to illustrate problematic aspects of morpheme identity (cf. Bickel and Zúñiga 2017, 163f).

- (165) *pyâ hã pihàngya*
pyâ hã pi-hàng-ya
 bird 1S peck-give-ALLO
 ‘?The bird pecks and gives me (something) (it does not yield: ‘the bird pecks (at) me’).’ [250322_JO_eli2 814]

In this sense, the choice of the term *predicate derivation* is orthogonal to the question of morpheme identity as this construction is a unified morphological category that recruits verb-selecting morphemes that are related not only to verbs but also to adverbs (cf. Post 2007, 460,564f.). Conversely, a label that manifests the relation between these morphemes and their verbal counterparts, like *V2*, would therefore only address but a subset of all intended morphemes.

Lastly, I will employ the term *versatile verb* for those morphemes that occur as part of either the derivation or TAMN paradigm of the verb complex and bear both the form and function of an independent verb (cf. Matisoff 1991, 402; Daudey 2014, 305). This term comes as a hermeneutic approach to the dichotomy inflection/derivation, since it takes a pre-theoretical stance regarding the morphosyntax of grammaticalized verbs: As an example from Tawrã, the verb *bô* ‘go’ occurs as a TAMN verbal suffix (akin to inflectional suffixes in other languages, cf. §6.3) meaning ‘do V and go’. Thus, *bô* ‘go’ is a versatile verb in this account.

6.2.2 Valency operating predicate derivations

This group of predicate derivations includes five morphemes: the associative *-né/nê* (§6.2.2.1), two causative markers (§6.2.2.2), the benefactive/malefactive *-hàng* (§6.2.2.3) and the permissive *-shyüw* (§6.2.2.4).

6.2.2.1 Associative *-né/nê*

The associative derivation adds a further, unspecified participant (S with intransitive verbs and A with transitive verbs) which mirrors the thematic role of the S/A argument of the base clause (i.e. it does not replace it). Semantically, the base S/A argument leads the action while the new participant is thought to ‘do along’³, i.e. accompanies the base S/A. The nature of the tone variation (*-né/nê*) is not clear.

Example 166 is from a story in which a chicken and a fox escape together. The suffix *-né/nê* allows the inclusion of a new S argument that mirrors the S of the base clause.

³It may be easily translated through the particle *mit* in German separable verbs such as *mitkommen* ‘come along’ and equivalent constructions in other Germanic languages (cf. sub-standard English ‘come with’) or with a prepositional phrase containing ‘along’ in English.

Example 167 also introduces a new S argument (in this case, the speaker him/herself). Example 168 is a transitive verb and the predicate derivation *-né/nê* adds a new, unstated A argument to the clause.

- (166) lyó, lyóya byûnéq nyu tyû tâzi nyu byûné, tyû nyu
 lyó [[lyó-ya]_{rel.cl} byû-né-q nyu]_{main.cl} [tyû tâzi nyu byû-né
 FS be.white-ALLO run-ASSO-CONT FOC chicken upper.part FOC run-ASSO
 tyû nyu]_{main.cl}
 chicken FOC
 ‘The white, the white one is running (with the fox). The chicken runs along (the fox) to the upper part.’ [TA22_chicken_thief 73 255]

- (167) cya îngnê yûm!
 cya îng-nê yûm
 3 sleep-ASSO NEG
 ‘I won’t sleep with him/her!’ [071023_STBT_eli3 616]

- (168) cya àng bânnêli
 cya àng bâ-nê-li
 3 house build-ASSO-REC.PST
 ‘S/he built a house (together with someone).’ [071023_STBT_eli3 545]

6.2.2.2 Causative predicate derivations

Tawrã has two predicate derivations with a causative meaning⁴. Their distribution is partly determined by semantic considerations: *-ji* (§6.2.2.2.1) is a general causative, i.e.

⁴Sastry (1984, 154f) also lists two causative morphemes: *bo* (cf. <-bô>) and *syig* (here <-shyüw>). The former, based on the examples provided, is likely a confusion between egophoric habitual *-de* and egophoric non-past *-dyibô*, i.e. *-dyibô* was probably analyzed as **-de-bô*, which is perhaps the historical origin of this suffix (cf. §6.3.4.8.1) but, crucially, neither *-dyibô*, *-bô* nor *-de* have a causative meaning. The latter suffix (*syig* (here <-shyüw>)) is a causative but the semantic distinction provided (i.e. the former “used when the executor is the subject” and the latter “when the agent is the subject” (idem.)) does not match the present analysis. Sastry’s faulty analysis has been reproduced by Dixon (2000, 47) as the *unusual*

it is neutral regarding the undergoer's /causee's willingness to do something. Conversely, *-shyüw* (§6.2.2.2.2) is a volitional causative: the intentions of the undergoer/causee and the actor/causer are thought to agree (to a certain extent). They do not form a volitionality-pair.

In both cases, the causative increases the argument structure of the host verb by adding a volitional A argument (causer/instigator). If the base clause is intransitive, it is transitivized and the original S argument becomes P (upon whom/what the causativized action is carried out). In a transitive clause, the causer becomes A while both the original A (the causee) and the original P retain their status (Dixon 2000, 30, 47f; Zúñiga and Kittilä 2019, 15f.). Both causative suffixes can imply direct or indirect causation.

Before proceeding to the causative predicate derivations, a brief etymological note is in order. The verb pair *shêi* 'die' and *mesê*⁵ 'kill' seem to go back to an old PTB causative alternation based on the suffix **s-*. As in other Tibeto-Burman languages, this device is not productive in Tawrã and this is the only verb pair found so far to have a trace of this old suffix alternation (cf. LaPolla and Huang 2003, 127f; Matisoff 1973a, 243).

6.2.2.2.1 Causative *-ji*

The general causative *-ji* is likely derived from the versatile verb *ji* 'make, be about to' and expresses causation without providing information regarding the undergoes/causee's will (as opposed to the volitional causative/permissive *-shyüw*, cf. §6.2.2.2.2). In Example 170, the intransitive verb *khro* 'cry' is made mono-transitive and the A argument is understood to be the causer of the action. It should be noted that the causative *-ji* does not imply unwillingness but is rather a product of the verb ('cry'). Conversely, Example

Mishmi causative that reportedly does not augment the valency of transitive verbs by one. The faulty **-de-bô* concatenation is displayed in Example 169. A better translation would be 'I ate rice/food', without a causative.

- (169) *hă tapě thá-de-bo*
 1S + NOM rice eat-TENS-CAUS
 'I (causee) was made to eat rice (by someone – unstated but implied causer)' [Sastry 1984, 154, also Dixon 2000, 47]

⁵Note that initial *me-* is a later development which probably mirrors the noun *mê* 'person', cf. §3.7.4 (also an equivalent development in Kera'a, cf. Reinöhl 2022, 250).

171 corresponds to the causativization of a two-argument clause. In this case, we only know that the undergoer/causee has been commanded or exhorted to burn the books; causative *-ji* is neutral regarding the willingness of the undergoer/causee.

- (170) *hã cya khròjidyibô*
hã cya khrò-jì-dyibô
 1S 3 cry-CAUS-PFV.EGO
 ‘I made him cry.’ [150122_AM_WA_eli 1304]

- (171) *cya tako prûjidyibô*
cya tako prû-jì-dyibô
 3 book burn-CAUS-PFV.EGO
 ‘I had him burn the book.’ [021024_STBT_eli2 203]

6.2.2.2.2 Volitional causative/permissive *-shyüw*

The volitional causative *-shyüw* is related to the versatile verb *shyüw* ‘allow, appoint, let’ and the observational verbal negator *shyüw* (§6.4.2). Its use implies that the causee agrees to some degree to the action. Commonly, the causer perceives some benefit from the action. This predicate derivation is also a permissive, i.e. it allows or gives permission to an action, which mirrors its origin as an independent verb.

In Example 172, the dog sits atop the child while he swims; the use of this suffix implies that the child willingly swims the dog over. Similarly, in Example 173 one person had another inquire about a plot of land. We may assume that this person was hired or willingly engaged.

- (172) *tya kwâg we macîk kē gũ katêshyühoya*
tya kwâg we macîk kē gũ katê-shyüw-ho-ya
 now dog DEM water LOC.INS LOC swim-CAUS.V-STATE-ALLO
 ‘Now this dog made him swim in the lake.’ [TA22_frog1_126 635]

- (173) men.rol we mane Tayang Chaitom we seshyüw
men.rol we mane a.clan a.clan we se-shyüw
 main.roll(<E) DEM FILL xxx xxx DEM inquire-CAUS.V
 ‘This Tayang person had Chaitom inquire (about the land)’ [TA24_land_292 539]

6.2.2.3 Bene/malefactive *-hàng*

As defined by Kittilä and Zúñiga (2010, 2f.), the benefactive marks “a participant that is advantageously affected by an event without being its obligatory participant (either A argument or primary target, i.e. patient). Since normally only animate participants are capable of making use of the benefit bestowed upon them, beneficiaries are typically animate” (also cf. Lehmann et al. 2004, 68; Van Valin and LaPolla 1997, 382). There are different ways in which the semantic effect of the benefactive can be expressed syntactically. Commonly, the stem takes a benefactive suffix and the introduced beneficiary takes both the slot in the clause and the morphology of a direct object (Kittilä and Zúñiga 2010, 9f).

In Tawrá, the benefactive and malefactive *-hàng* is transparently derived from the versatile verb *hàng* ‘give’, which is a typologically common development (cf. Kittilä and Zúñiga 2010, 8f). In syntactic terms, the benefactive is an applicative derivation in the sense that it introduces an (optional) non-A argument that could not be expressed by the base verb (i.e. without the applicative derivation) (Pacchiarotti and Zúñiga 2022, 1f; Zúñiga and Creissels 2024, 1, 18f). This is illustrated in the following examples: In Example 174a, the verb *zã* ‘pack’ takes the applicative benefactive derivation which licenses the P argument (the applied patient), *nyu* ‘you’. Conversely, Example 174b has the same verb but without the benefactive applicative; an applied P is here not allowed. Note that there is no marking of core argument NPs in the language (cf. §8.2.4). The beneficiary does not have to occur overtly in the clause (here in parentheses).

- (174) a. hã (nyu) tabrẽ zãhàngdyibô
 hã (nyu) tabrẽ zã-hàng-dyibô
 1s 2s food pack-BEN-PFV.EGO
 ‘I packed meat for you.’
- b. hã (*nyu) tabrẽ zãdyibô
 hã nyu tabrẽ zã-dyibô
 1s 2s food pack-PFV.EGO
 ‘I packed meat.’ [both 021024_STBT_eli2 275]

An A argument promotion, however, is not always the case, as in Example 175, where the P of the mono-transitive verb ‘bite’ may only be recast as the semantic bene/maleficiary and there is no addition of a new direct P (also cf. Modi and Post 2022). This example also illustrates that the new applied object can be adversely affected (thus *malefactive*). In Example 176, we see that the bene/maleficient participant can be omitted and it is understood from the context (cf. §8.2.1).

- (175) kwôg hã thâkhàngya
 kwôg hã thâk-hàng-ya
 dog 1s bite-MAL-ALLO
 ‘The dog bit me.’ [250322_JO_eli2 747]
- (176) manyuhàngya
 manyu-hàng-ya
 cook-BEN-ALLO
 ‘S/he cooked (for me).’ [250322_JO_eli2 520]

The postposition (*a*)*ba* ‘for’ (§9.5.5) can be employed to make a distinction between a *recipient* argument and a *beneficiary* (i.e. writing **for** somebody who might not be able to write for him/herself) (cf. Zúñiga and Kittilä 2019, 16f). In Example 177a, the standard interpretation is that of a recipient. In Example 177b, however, the use of the postposition *-ba* forces a beneficiary reading (akin to English *for*).

(177) a. *hã cya cithi zûhàngne*

hã cya cithi zû-hàng-ne

1s 3 letter(<IA) write-BEN-INT

‘I will write him/her a letter.’

b. *hã cya ba chiti zûhàngne*

*hã cya **ba** chiti zû-hàng-ne*

1s 3 for letter(<IA) write-BEN-INT

‘I will write a letter **for** him/her.’ [both 021024_STBT_eli2 165]

In Example 178 an intransitive verb *dế* ‘stand’ is rendered mono-transitive through the benefactive. The S argument is promoted to an A and the new P is the beneficiary.

(178) *lain gũ dếhàngne*

lain gũ dế-hàng-ne

line(<Eng) LOC stand-BEN-INT

‘I will stay in line (for you).’ [250322_JO_eli2 930; modeled after Schikowski (unpublished)]

6.2.2.4 Permissive -*shyüw*

Permissive -*shyüw* is related to the versatile verb *shyüw* ‘allow, appoint, let’, the homophonous observational verbal negator (§6.4.2) and volitional causative (§6.2.2.2.2). The permissive (‘allow to V’) is closely related to the independent verb. In Example 179, the permissive can be paraphrased with the verb *allow*.

(179) *ning gâmbra lûmkû he dângkûdùya we sokarei bôshyü yüm*

ning gâmbra lûmkû he dâng-kû-dù-ya we

1P gaonburha(<IA) inside EMPH finish-COMPL.II-need-ALLO DEM

sokarei bô-shyüw yüm

government(<IA) go-PERM NEG

‘We had to settle (lit. finish) (our disputes) under the gaonburha (leader of the village), we didn’t allow (them) to go to the government.’ [240122_clan 42 150]

6.2.3 Manner/modality predicate derivations

This group includes an eclectic assortment of predicate derivations that affect the manner in which an action takes place, i.e. how an action is executed as well as modality, here understood as a semantic category that reflects the speaker's attitude (e.g. ability, necessity or probability) (Givón 2001a, 151; Noonan 2007, 102; Talmy 2007, 150). The manner/modality predicate derivations are: inopinative *-ta* (§6.2.3.1), frustrative *-hã* (§6.2.3.2), 'by force' *-mazâi* (§6.2.3.3), 'at last' *-ha* (§6.2.3.4), 'in exchange' *-da* (§6.2.3.5) and 'thoroughly' *-makáw* (§6.2.3.6).

6.2.3.1 Inopinative *-ta*

The inopinative *-ta* expresses the speaker's surprise in reaction to an action. As such, it is appropriate for events that occur unexpectedly. In Example 180, the predicate derivation *-ta* frames the action as unexpected.

- (180) lyâtakûbô
lyâ-ta-kû-bô
throw-INOP-COMPL.II-PFV.ALLO
'He just threw it away (unexpectedly).' [291124_BAB_eli1 2209]

This predicate derivation is related to the hortative marker *-ta* (§6.3.5.0.3.1) as well as to the advisative modal particle *ta* (§9.2.9).

6.2.3.2 Frustrative *-hã*

The predicate derivation *-hã* can be paraphrased as either 'in vain' or 'unknowingly/mistakenly'; it thus resembles the use of *frustrative* as found in some Amazonian languages (cf. Overall 2017). In Example 181, the speaker tells a story of how his tribe was generally backwards compared to other tribes. The use of *-hã* 'in vain' expresses the purposelessness of their existence. Similarly, in Example 182 the main character waited the whole day for his horse to be brought back. His waiting was in vain, as the horse had been stolen.

- (181) *tawê yüm ihădyibô*
tawê yüm ì-hă-dyibô
 mind NEG be.there-FRUS-PFV.EGO
 ‘We lived in vain without a brain (lit. mind).’ [TA22_origin 157 447]

- (182) *kalyûmâhăbô la*
kalyû-mâ-hă-bô la
 wait-ITER-FRUS-PFV.ALLO HSAY
 ‘He kept on waiting in vain, they say.’ [TA22_JAstory 45 144]

A further sense of this predicate derivation is to express a mistake, often the product of a lack of information. Example 183 is from a storybook (Mayer 1969): the dog managed to put his head (lit. mouth) in a jar and get stuck. The predicate derivation frames *-hă* ‘in vain’ it as an obvious mistake.

- (183) *kwâg thrûm.bum we bôtlo kě halúhăho*
kwâg thrûm.bum we bôtlo kě halú-hă-ho
 dog mouth DEM bottle(<E) LOC.INS put.in-FRUS-STAT
 ‘The dog’s mouth got stuck in the bottle by mistake.’ [TA22_frog1 51 220]

6.2.3.3 ‘by force’ *-mazâi*

The predicate derivation *-mazâi* indicates that an action transpires forcefully or through coercion. It is not a causative as it does not affect the valency of the verb and it also does not include an inherent telic component. In Example 184, the villagers compelled the wife to speak. The use of the verb *ahu* ‘ask’ with the predicate derivation *-mazâi* renders a sense of ‘asking coercively’. Similarly, in Example 185 the verb *tî* ‘wrestle’ with *-mazâi* leads to ‘being subjugated’.

- (184) shêina yüm webo gũ mîya we ahumazâigabô
 [shêi-na yüm webo gũ]_{cosub.cl} mîya we ahu-mazâi-ga-bô
 die-FACT NEG and.then NF wife DEM ask-force-PL.I-PFV.ALLO
 ‘Since he would not die, they compelled (his) wife to speak.’ [TA23_tayang 33 198]

- (185) tîmazâikũho
 tî-mazâi-kũ-ho
 wrestle-force-COMPL.II-STAT
 ‘S/he has been subjugated by force.’ [161024_JOJA_eli 1462]

6.2.3.4 ‘at last’ -ha

The predicate derivation *-ha* ‘at last’ expresses the speaker’s view that an has action occurred at an inopportune time. In Example 186, the doctor came later than what would have been desirable (but not too late). In Example 187, the speaker expresses his wish to finally settle somewhere after roaming for a long time.

- (186) daktar hanahabô
 daktar hana-ha-bô
 doctor(<E) come-at.last-PFV.ALLO
 ‘Finally, the doctor came.’ [281024_JO_eli2 670]
- (187) webo gũ we “ekê ihakéq” la gũ (...)
 webo gũ we [e=kê ì-ha-kéq]_{dir.speech} la gũ
 and.then INS DEM DEM=LOC.INS be.there-at.last-HORT say NF
 ‘And then saying “let’s finally settle here” (...)’ [TA24_land 215 386]

6.2.3.5 ‘in exchange’ -*dá*

The predicate derivation -*dá* is related to the versatile verb *dá* ‘return (vt), finish paying’. As a predicate derivation, it implies that an action is reciprocated, e.g. in an exchange of gifts or marriage partners among clans. This is illustrated in Example 188a: a trickster is hiding in a cave and, after the village people learn of his whereabouts, they take shifts as guards in order to catch him crawling out of the cave. During one of these shifts, the trickster convinces a guard to put his hand into the entrance of the cave. Once the fingers of the guard touch the hands of the trickster, he pulls him down, causing the guard to activate a tramp intended for the trickster. The narrator casts the action as an *exchange*, as both parties were trying to reach each other. A rough translation would be that the trickster ‘caught him in turn’. Example 188b is from the same story and the narrator goes back to the same episode, describing the luck of the trickster, who became free ‘in exchange’ (for the guard).

- (188) a. *agûya mề we nyu sêdá khrâk*
agû-ya mề we nyu sê-dá khrâk
 guard-ALLO person DEM FOC catch-in.exchange EVI
 ‘The guard was caught (in turn).’ [TA22_kutyemba_phla 471 860]
- b. *cya hwă.wră shyâdána matyô*
cya hwă.wră shyâ-dá-na matyô
 3 free become-in.exchange-FACT maybe
 ‘Maybe he became free (in turn).’ [TA22_kutyemba_phla 483 883]

6.2.3.6 ‘thoroughly’ -*makáw*

The predicate derivation -*makáw* is related to a homophonous adverb meaning ‘thoroughly, enough’. It implies that an action was carried out to the best of one’s capabilities. In Example 189, the priest is said to have done a complete inquiry of an area before settling there.

- (189) (...) gwók we semakáw gũ “emacêiye tyûmna” na (...)
 [gwók we se-makáw gũ]_{cosub.cl} e = macêi = ye tyûm-na na
 priest DEM inquire-thoroughly NF DEM = water = DEM drink-FACT INSIST
 ‘(...) after the priest inquired thoroughly (he said) “this water (can) be drunk”,
 right?’ (...)’ [TA24_land_297 551]

6.2.4 Aspect/Aktionsart predicate derivations

The group aspect/Aktionsart includes three members: permansive *-(ta)lá* (§6.2.4.1), ambulative *-wi* (§6.2.4.2) and completive *-dâng* (§6.2.4.3). Tense and aspect/Aktionsart are further discussed in §6.3.1.2.

6.2.4.1 Permansive *-(ta)lá*

The marker for permansive Aktionsart is *-talá*, or *-lá*, and is related to the adverb *talá* ‘still, only’. The permansive aspect is used for actions that began in the past and continue into the present or the future (cf. Jacques 2021, 977f). The start of the event in the past and its often tiresome or weary continuation into the present sets the permansive predicate derivation apart from the continuative suffix *-q* (§6.3.4.2). Usually, actions that started in the past and have a repercussion in the present are marked with the perfect marker *-bô* (§6.3.4.8.1). However, the permansive *-talá* is used for actions that are ongoing; it does not emphasize a prior change of state.

Examples 190a and 190b are from a storybook (Mayer 1969). While the boy and the dog are looking for their lost frog, they encounter different animals in the forest. In Example 190a, the speaker digresses into the identification of a small rodent. Thereafter, in Example 190b, the speaker returns to the feat of the boy and the dog - here the use of the permansive indicates an ongoing process.

- (190) a. kacî nyu lêzâ delé
 kacî nyu lê-zâ delé
 mouse FOC exit-DIR it.seems
 ‘It’s a mouse that came out, it seems.’ [TA22_frog1 76 350]

- b. kwâg we cuwe samplang tabrâtaláq
 kwâg we cuwe samplang tabrâ-talá-q
 dog DEM DISC.DEIC.DEM FILL hunt-PERMA-CONT
 ‘This dog, eh, is still hunting.’ [TA22_frog1 77 354]

In Example 191, the speaker describes a story that explains the present state of its tribe. The use of the permansive emphasizes that this particular, pointless state (-hâ ‘in vain’) persists to this day.

- (191) we ni tya khréq mu alí ìhâládi
 we ning tya khréq mu alí ì-hâ-talá-di
 DEM 1P now up.to also like.this be-FRUS-PERMA-HAB.EGO
 ‘Even today we remain like this.’ [TA22_origin 169 470]

As mentioned above, the form *talá* ‘still, only’ also occurs as an adverb. In Example 192, *lá* ‘still, only’ modifies the predicate and occurs before it, corresponding to the function and position of adverbs (cf. §3.4.1).

- (192) talyũ hîq lá zûgaq
 talyũ hîq talá zû-ga-q
 a.name EMPH still write-PL.I-CONT
 ‘They still write *Talyu*.’ [TA23_talyu 31 192]

6.2.4.2 Ambulative -wi

The ambulative -wi implies that an action is performed repetitively, often in more than one place, cf. *to drink around*. In Example 193, the use of the ambulative suggests that the person is homeless, as s/he eats ‘here and there’. Similarly, in Example 194 the speaker is talking about the inheritance of land and the point is that the grandfather in question never lived there, as he was always away doing business, i.e. the action was carried out in different places or a repetitive manner.

(193) *thâ-wi*

eat-AMB

‘S/he eats here and there (i.e. like a homeless person).’ [271024_JO_eli1 682]

(194) (...) *natya we ji asâng nyu îng dukan sukan bîsnes jiwina hî weli kamazâna*

[*na-tya* *we* *ji* *asâng* *nyu* *îng* *dukan* *sukan*

1S.POSS-grandfather DEM EMPH alive FOC a.bit store(<IA) ECHO

bîsnes *ji-wi-na* *hî*]_{main.cl} [*weli* *kama-zâ-na*]_{main.cl}

business(<E) do-AMB-FACT EMPH DISC.DEIC.DEM be.lost-DIR-FACT

‘(...) my grandfather, while he was alive, he did some business here and there.

And like that, he passed away.’ [TA24_land 149 273]

The predicate derivation *-wi* is likely related to the verb *guwi* ‘roam’. The grammaticalization path verb of motion > ambulative is attested in the language family (cf. Jenny 2000, 166-167; the ambulative in Kiranti, cf. Schackow 2015, 235-236; Bickel 1996, 163-165). However, the reduction of the recruited verb is uncommon in Tawrä; there is neither a verb **wi* nor a predicate derivation **-guwi*. While root prefixes are commonly deleted (§3.7), there is no other case where a syllable *gu-* is erased.

6.2.4.3 Completive *-dâng*

The completive aspect marker *-dâng* COMPL.I is transparently derived from the versatile verb *dâng* ‘finish’ and it emphasizes a telic reading or the fulfillment of an action, either as a past event or as a prospect. As such, it resembles the common completive TAMN suffix *-kũ* (§6.3.4.3, glossed as COMPL.II). There are, however, certain differences: completive *-dâng* often implies a state of affairs that can be translated by the English phrasal adverb ‘already’, i.e. it depicts a relevant change of state or a crucial point that has been reached (cf. similar constructions in languages from MSEA, cf. Jenny 2021, 847; cf. *iamitive* Olsson 2013; Kuteva et al. 2019, 174; completive *-ròk* in Karbi, cf. Konnerth 2014, 266). Furthermore, completive *-dâng* often involves a large number of referents, i.e. counting *all* the fish, or *everybody* leaving a place. In Example 195, the completive

marker emphasizes that the action was carried out to the limit of exhaustion by several villagers. In Example 196, the completive indicates a relevant change of state (the withdrawal of the Britishers) that corresponds to English *already*. Here again, the referent is plural. The plural reference, however, is not essential to this predicate derivation but merely a common feature. In Example 197, the referent having a rash is one person.

(195) tayàkadyângholi, tayâdângli gũ (...)

[tayâ-kadyâng-ho-li]_{main.cl} [tayâ-dâng-li gũ]_{cosub.cl}

shout-push.with.legs-STAT-REC.PST shout-COMPL.I-REC.PST NF

‘(They) shouted and pushed. They shouted (completely) (...)’

[TA22_kutyemba_phla 128 271]

(196) bôdângkũnabô mane e lyâyâkũnabô

bô-dâng-kũ-na-bô mane we

go-COMPL.I-COMPL.II-BACK-PFV.ALLO FILL(<Asm) DEM

lyâ~lyâ-kũ-na-bô

return~PLAC-COMPL.II-BACK-PFV.ALLO

‘(The Britishers) were already gone, they had already gone back.’ [TA22_CLAN2 10 17]

(197) tili.tulu.tili tshângdâng gũ (...)

tili.tulu.tili tshâng-dâng gũ

manner.of.having.a.rash have.a.rash-COMPL.I NF

‘Having a rash (...)’ [TA22_CLAN2_191 364]

Completive *-dâng* can occur along with TAMN completive *-kũ* (§6.3.4.3). It is associated with particular (often telic) verbs (e.g. die, disappear, fall, etc.) and verbs that affect large numbers of people while TAMN completive *-kũ* emphasizes the terminal point of the action (to *completely* die). In Example 198, *-dâng* is linked to the verb *shêi* ‘die’ and it further implies a large number of people while *-kũ* stays closer to its aspectual/Aktionsart

function. Similarly, in Example 199, the *-dâng-kũ* collocation refers to a (large) number of referents who are irremediably affected.

(198) shîdângkũya

shêi-dâng-kũ-ya

die-COMPL.I-COMPL.II-ALLO

‘(They all) died.’ [141222_ST_BT_eli2 356]

(199) kadé labũ yu lüho mé grudângkũnabô

[kadé labũ yu lü-ho]_{main.cl} [mé gru-dâng-kũ-na-bô]_{main.cl}

how.much debris FS remain-STAT all burn-COMPL.I-COMPL.II-FACT-PFV.ALLO

‘How much debris remained! Everything was burned away.’ [TA23_newmyth 21 127]

6.3 Tense, aspect, mode and number (TAMN) suffixes

This section offers a description of TAMN (i.e. tense, aspect, mode and number) suffixes, which come at the last slot of the verbal complex described in §6.1. This slot position is termed *TAMN suffixes* and it agglomerates grammatical categories that are called *inflection* in other languages. It is relevant to note that there is no obligatory agreement/argument cross-reference on the predicate as known, e.g. from Indo-European languages (sometimes termed *canonical agreement*, cf. Corbett 2006, 4f) or from other Tibeto-Burman languages, such as the Kiranti group or rGyalrong (DeLancey 2011; also cf. van Driem 2001, 566; Jacques 2016). In Tawrã, TAMN suffixes exhibit “person-based sensitivities” (Post 2007, 689); they do support the identification of a grammatical relation that we might define as S/A alignment but none of these suffixes can be straightforwardly described as “subject agreement suffixes”, as each of them is also sensitive to the semantico-pragmatic factors of animacy, engagement and egophoricity (§6.3.1.1). Predicate marking is discussed in §8.2.1.

TAMN suffixes in Tawrã occur after the end of the verbal predicate. These suffixes add a semantic entailment as well as TAMN. They do not license a finite predicate, as

verb stems can occur without suffixing. More than one of these suffixes can occur on the verb complex. TAMN suffixes follow a slot order, although suffixes *-kũ* ‘COMPL.II’ and *-na* ‘FACT’ may occur before the predicate derivation if their scope is only on the first verb stem. Their distribution is described in the following subsection (§6.3.1.3). TAMN suffixes have a local, word-level scope as opposed to particles, which have a clause-level scope, as discussed in §3.5.2.4.

TAMN suffixes cover a wide spectrum of grammatical categories: TAM, number, egophoricity and direction. In this section, the suffixes have been grouped in semantic/grammatical functional groups for the sake of clarity. Table 6.5 presents an overview of all TAMN suffixes discussed in this section according to groups. §6.3.1 will introduce three concepts that are relevant for TAMN suffixes: Egophoricity (§6.3.1.1), tense and aspect (§6.3.1.2) and suffixal chain (§6.3.1.3).

Table 6.5: TAMN suffixes organized in groups

Groups	TAMN suffixes
Number	Plural I <i>-ga</i> (§6.3.2.0.1) Plural II <i>-ring</i> (§6.3.2.0.1) Dual marker <i>-nega</i> §6.3.2.0.2
Associated motion	Directive <i>-zâ</i> (6.3.3.1) Go (and come back) <i>-ge</i> (§6.3.3.2)
Aspect and Aktionsart	‘ahead’ <i>-gadyô</i> (§6.3.4.1) Continuative <i>-q</i> (§6.3.4.2) Completive <i>-kũ</i> (§6.3.4.3) Iterative <i>-mâ</i> (§6.3.4.4) Stative <i>-ho</i> (§6.3.4.5) Intentional (egophoric) <i>-ne</i> §6.3.4.6) The <i>bô</i> group (§6.3.4.8) Recent past <i>-li/lü</i> (§6.3.4.9) Certain future <i>-(vi)dyô</i> (§6.3.4.10) ‘permanently’ <i>-tá</i> (§6.3.4.11)
Command mood	Imperatives <i>-na</i> (and <i>-tè</i>) (§6.3.5.0.1) Prohibitive <i>-yâq</i> (§6.3.5.0.2) Hortatives <i>-ta</i> and <i>-kéq</i> (§6.3.5.0.3) Jussive <i>-tyêyu</i> (§6.3.5.0.4)
Other suffixes	Allophoric <i>-ya</i> (§6.3.6.1) Factitive <i>-na</i> (§6.3.6.2)

6.3.1 Relevant concepts

6.3.1.1 Egophoricity (and allophoricity)

Tawrã has several suffixes that are sensitive to *egophoricity*. Egophoricity is commonly understood as “personal knowledge, embodied experience, or involvement of a conscious self” (San Roque et al. 2018, 2). It flags privileged access to personal information or the

involvement of the speaker. In morphological terms, it refers to a (semi-)paradigmatic, person-sensitive system in which an egophoric marker, which is canonically found in declarative sentences with first person singular S/A arguments and interrogative sentences with second person singular S/A arguments, contrasts with an allophoric (inaccessible knowledge) form. Egophoric systems are commonly sensitive to parameters such as volition, agency, epistemic authority and shared knowledge. Languages with egophoricity have diverse strategies to express this category, such as bound morphemes or copulas. Egophoric morphemes are often portmanteaus, also flagging tense or aspect and they are commonly obligatory (Hale 1980; Hargreaves 2005, 31; San Roque 2018, 406; Widmer and Zemp 2017).

The egophoric system overlaps with person marking in the broad frame of Bickel and Nichols (2007, 220), in which person is the “grammaticalization of conceptual distinctions between participants involved in speech activities” (also cf. Widmer 2020, 266f.). However, this overlap with person can also be thought of as the outcome of the speaker’s interpretation of participant knowledge-states (or its access) in relation to event semantics (semantics in the sense of DeLancey 1990; also cf. Post 2013) and not as a form of agreement inflection (cf. Bergqvist and Kittilä 2019, 1f.).

In Tawrã, egophoricity intersects with tense/aspect in four portmanteau suffixes while two other markers do not have inherent tense/aspect values. All six egophoricity-sensitive suffixes are displayed in Table 6.6.

Table 6.6: TAMN suffixes sensitive to egophoricity

Suffix	
- <i>de/di</i> Egophoric habitual	§6.3.4.7
- <i>ya</i> Allophoric	§6.3.6.1
- <i>dyibô</i> Egophoric perfective	§6.3.4.8.1
- <i>bô</i> Allophoric perfective	§6.3.4.8.1
-(<i>vi</i>) <i>dyô</i> Egophoric certain future	§6.3.4.10
- <i>ne</i> Egophoric intentional	§6.3.4.6

Egophoric markers are loosely associated with first person (the category of num-

ber will be discussed below) in statements and second person questions (excluding non-volitional forms, see below). Conversely, *allophoric* markers are associated with all other contexts.

Semantically, egophoric markers stand for privileged access to personal information (experiential-involvement (cf. Post 2013, 115)). The *assertor*, i.e. the semantico-syntactical role from which the epistemic status is defined (cf. Widmer and Zúñiga 2017, 420), has inside knowledge over him/herself and may request such knowledge from a second person. In Example 200a, the assertor is involved in the action. In Example 200b, the assertor requests access to this form of information. Conversely, Example 200c illustrates the epistemic impossibility of requesting inside knowledge from a third person. An allophoric suffix would be expected in this scenario.

(200) a. *juta hrê-dyibô*
 shoe(<IA) remove-PFV.EGO
 ‘I removed (my) shoes.’

b. *juta hrê-dyibô kya?*
 shoe(<IA) remove-PFV.EGO Q
 ‘Did you remove (your) shoes?’

c. **cya juta hrê-dyibô kya?*
 3 shoe(<IA) remove-PFV.EGO Q

Intended: ‘Did he remove (his) shoes?’ [all three 071023_STBT_eli3 653]

The assertor does not normally have scope over non-volitional events, such as falling or psychological or physiological states (also cf. San Roque et al. 2018, 14; DeLancey 1986a): In Example 201, the act of resting is described with an allophoric form. Similarly, in Example 202 a question indexing a second person takes a non-egophoric form. In these two examples, sleeping and being hungry are conceived as lacking volition and receive allophoric marking. Note that this pattern mirrors the marking of verbs of experience in

South Asian languages that do not have an egophoric system⁶.

- (201) (...) *aram jinaya*
aram ji-na-ya
rest(<IA) do-FACT-ALLO
'(...) I rested.' [TA22_pineapple 18 88]

- (202) *nyu nâshik tyûng ke?*
nyu nâshik tyûng ke
2s appetite be.hungry Q
'Are you hungry?' [AJ 14_02_23]

Allophoric markers are associated with third person. This is not a matter of volition but of the epistemic impossibility of accessing a third person's personal knowledge. This is illustrated in Examples 203 and 204, where the verb takes the allophoric marker *-ya*. An egophoric marker would not be allowed.

- (203) *cya ja dúa la*
cya ja dú-ya la
3 what need-ALLO say
'What did s/he (say s/he) want?' [240122_AM_eli1 86]

- (204) *oh khirki gã dûkûbô kwâg we*
[*oh khirki gã dû-kû-bô* [*kwâg we*]_{NP}]_{main.cl}
oh window(<IA) LOC jump-COMPL.II-PFV.ALLO dog DEM
'Oh, it has jumped out of the window, this dog.' [TA22_frog1 53 239]

⁶A non-conventional or oblique marking on the S/A argument of a verb associated with perception and states of mind is a typologically widespread phenomenon (cf. Belletti and Rizzi 1988, for a discussion about this construction in Italian), particularly in South Asia, where it is sometimes termed *oblique experiencer* (Hock 2016d, 544), *transimpersonal* pattern (Malchukov 2015, 80) *experiencer-as-goal* or *experiencer/experiential* frame (particularly in the literature on Kiranti cf. Schackow 2019, 482f; Schikowski et al. 2015, 677f; but also beyond, cf. the case of Kham, cf. Watters 2009, 223f). The usual markers associated with the experiencer S are the dative, genitive and nominative/absolutive (unmarked), which often correspond to Goal-marking cases (Bickel 2000b, 600f; Hock 2016d, 544).

A further point concerns (self-attributed) epistemic authority associated with egophoricity (cf. Hargreaves 2005, 34). As hinted above, a first person **plural** may also be an assertor and be flagged with an egophoric marker, something that might seem to run counter the idea of *personal* knowledge (cf San Roque et al. 2018, 2).⁷ As discussed by San Roque et al. (2018, 17f.), the use of egophoricity with a plural assertor is problematic for how epistemic authority is understood, although it is “universally acceptable to extend an assertion of personal knowledge over a group (or certain kinds of groups) of which one is part”. In this line, consider the following cases: Example 205 is from a clan story of the *Tayang* clan. The speaker himself was not present at the time when the events took place. Similarly, Example 206 is from a mythical story of the origin of the Tawra. In both cases, the verbs take an expected allophoric *-ya* suffix.

- (205) ning ali (...) mlá mẽ mlá kě gâmbra bâya (...) bârēna na?
 [ning ali mlá mẽ mlá kě gâmbra
 1P like.this village.MOD person village.MOD LOC.INS gaonburha(<IA)
 bâ-ya]_{main.cl} [bâ-rē-na na]_{main.cl}
 make-ALLO make-PL.II-FACT INSIST
 ‘We, like this, appointed (lit. made) gaonburha (from) village people, right?’
 [TA22_clan 37 123]

- (206) ning byē macêi mathûw nyu ìkûna-ya
 ning byē macêi mathûw nyu ì-kû-na-ya
 1P TOP water place.of.origin FOC be.there-COMPL.II-FACT-ALLO
 ‘As for us, we lived towards the origin of the water (i.e. the north).’ [TA22_origin
 164 460]

Conversely, Examples 207 and 208, which are from the same text, display a different pattern. Here the assertor (a 1st person plural) uses an egophoric marker. In Example 207, the assertor employs an egophoric form to embody a direct link to the group (the

⁷Some languages with an egophoric system have two separate sets of markers for singular and plural, as described e.g. for Tororó (Barbacoan, Colombia) (Castaño 2019, 180f.) or Bunan (Tibeto-Burman, India) (Widmer 2014, 571f; also cf. Widmer 2015).

Tawrã tribe), which also includes the hearer. In Example 208, the assertor speaks for the gods, who directly address the men.

(207) ning tawrã bá.mě kade ìdi!

ning tawrã bá.mě kade ì-di

1P a.tribe clan how.many be.there-HAB.EGO

‘Our Tawrã were plenty (lit. how many we were!).’ [TA22_clan 10 51]

(208) welügũ “ning hwéne kathùnegaya”

[weligũ ning hwé-ne]_{main.cl} [kathù-nega-ya]_{main.cl}

and.then 1P look.after-INT take.care-DU-ALLO

‘And then (they said): we will look after (you), we (will) take care (of you).’

[TA22_origin 143 410]

One way to approach this phenomenon is to follow the analysis proposed by Modi (2017, 212f.) concerning the involvement of a coherent group in the system of personal pronouns (termed *ideoclusivity*, as opposed to mere clusivity). Egophoricity markers thus seem also to be sensitive to ‘coherent group’ constituency (cf. verbal plural suffixes -*ring* in §6.3.2.0.1). In Example 207, a tribe constitutes a prime example of a coherent group and certain members are *de facto* authorized to attribute to themselves the role of epistemic authority. Similarly, in Example 208 the gods display enough internal cohesion to constitute a coherent group. The epistemic source of egophoricity in Tawrã, it seems, also draws from ‘(coherent) group knowledge’ (rather than *personal* knowledge).

Further research should explore the cognitive aspect of group membership, particularly in the case of small-scale communities, where there often are adaptative social strategies that privilege small and homogeneous groups (Modi 2015, 248f).

A final note is required for the egophoric suffixes -*ne* ‘intentional’ INT and -(*vi*)*dyô* ‘certain future’ CERT.FUT. Although they are sensitive to egophoricity, they also indicate intention and determination and, in particular contexts, this can be extended to scenarios where egophoric markers do not commonly occur, e.g. with a 3rd person singular S/A argument (also cf. San Roque et al. 2018, 3f). In Examples 209a and 209b, the narrator

tells of a nephew who came to the village with the intention to kill the narrator's grandfather. The use of the intentional suffix in Example 209b for a third person refers to the intentions of the nephew. The use of this egophoric suffix is allowed in a story since the intentions of the characters are always known to the narrator. It should also be noted that Example 209b is not an instance of direct speech, which would render an utterance *verbatim* (cf. §10.6.2), but rather the verb 'say' has the function of linking discourse (cf. §10.7.3).

- (209) a. *webo byẽ abe.yâ jûzâ matyô*
webo byẽ abe.yâ jû-zâ matyô
 and.then TOP midnight crawl-DIR maybe
 'And then (the nephew) crawled in at midnight.' [TA22_CLAN2_161 304]
- b. *sêne la*
sê-ne la
 kill-INT say
 'In order to kill.' [TA22_CLAN2_163 306]

The same situation occurs with *-(vi)dyô* 'certain future' CERT.FUT. In Examples 210a and 210a, a priest wants to blow down a group of trees. The 'certain future' suffix occurs with a 3rd person singular since the speaker is in a position to presume the intention of a character in his/her story.

- (210) a. *hîng gũ we masáng we piathûwshyüwa la wegũ pûtsîn.thana.grõ gwók we*
[hîng gũ we masáng we pi-athûw-shyüw-ya la wegũ
 day INS DEM tree DEM ax-notch-CAUS.V-ALLO HSAY and.then
[pûtsîn.thana.grõ gwók we]_{NP}main.cl
 a.name priest DEM
 'One day, he wanted to fell these trees, this Putsin Thana Grõ priest.'
 [TA23_newmyth_39 244]

b. haráng jizâdyô magû

haráng ji-zâ-dyô magû

wind do-DIR-CERT.FUT that's.how.it.was

‘He would blow them down, that is how it was.’ [TA23_newmyth_40 250]

6.3.1.2 Tense and aspect

Tense and aspect are categories used to describe the temporal structure of a predicate. Verbs in Tawrã may take suffixes that affect their aspectual and temporal interpretation, although neither of these is obligatory on the verb. **Tense** refers to points in time based on three anchored categories: past, present and future. In Tawrã, the temporal interpretation of the non-marked verb stem depends entirely on the context. Furthermore, only a few suffixes indicate tense, e.g. certain future suffix *-(vi)dyô* (§6.3.4.10) or the recent past suffix *-li* (§6.3.4.9).

Aspect can be thought of as “different ways of viewing the internal temporal constituency of a situation” (Comrie 1976, 3; also defined minimally as “situations, and changes of situations” by Timberlake 2007, 284). This means that a verbal action can be described as complete, continuous, finished or repetitive. **Aktionsart** or lexical aspect depends on the semantic characteristics of a verb or morpheme and refers to the inherent temporal configuration, which can often be pictured in binary sets, such as dynamic vs static or durative vs. punctual, telic vs. atelic, etc. (Bickel 1996, 16f. Brinton 1988, 3; De Swart 2012, 174). The definition of aspect is often associated with inflectional morphology while that of Aktionsart with the event structure inherently encoded in a lexeme. However, this is bound to be problematic if the distinction between these two is not straightforward in the language (§6.2.1.1). In many grammars of South Asian languages, we find that there is no clear relation between these temporal categories and morphological marking (be it suffixal or derivational), especially concerning the second element in a complex predicate (here termed *predicate derivation*, cf. §6.2), as it often adds a wide array of aspectual notions, tense and Aktionsart as well as other grammatical categories (cf. Bickel 1996, 207f; Joseph 2006, 181f. Liljegren 2016, 244).

In Tawrã there is no isomorphism between aspect/Aktionsart and inflection/derivation. This is partly because the distinction between inflection and derivation lies on morphological considerations (cf. §6.2). Thus, for example, both the predicate derivation -*dâng* ‘completive’ and the TAMN suffix -*mâ* ‘iterative’, which correspond to two different syntactic positions, can be categorized as Aktionsart. It should also be noted that aspectual/Aktionsart and temporal categories are not coupled in Tawrã, i.e. there are no portmanteau morphemes that combine reference to aspect/Aktionsart and to tense (cf. egophoric habitual -*de* §6.3.4.7, which is neutral regarding tense).

For the remainder of this chapter, the term Aktionsart and aspect may be employed jointly to categorize those TAMN suffixes that influence the aspectual/Aktionsart interpretation of the predicate.

6.3.1.3 Suffixal chain

The precise order of suffixes is organized in syntactic *slots*: the numbered position in a chain that determines the occurrence of one morpheme in relation to others. This is illustrated in Table 6.7. This systematization is based on suffix collocation: if -*bô* ‘perfective (allophoric)’ and -*ga* ‘plural’ co-occur in a verb complex, the one closer to the stem will be assigned a number that is smaller relative to the one further away from the stem. This exercise is repeated until all suffixes are numbered.

The predicate head, the reduplicated stem (cf §6.1.1.4) and the predicate derivation(s) (cf. §6.2) occupy the slots 1, 2 and 3 respectively. The rest of the suffixes commonly follow in the order given by the table.

Table 6.7: Slots in the verbal complex

#	Suffixes	rough category
1	Verb stem	
2	reduplication	
3	predicate derivation	stem boundary
4	<i>-mâ</i> ITER	
5	<i>-nega</i> DU <i>-kũ</i> COMPL.II	
6	<i>-ho</i> STAT	
7	<i>-na</i> FACT <i>-li</i> REC.PST	
8	<i>-ga</i> PL.I <i>-ring</i> PL.II	plural
9	<i>-zâ</i> DIR <i>-gè</i> and.back <i>-bô</i> and.go	directional
10	<i>-ya</i> ALLO <i>-dyibô</i> PFV.EGO <i>-bô</i> PFV.ALLO <i>-ne</i> INT <i>-de</i> HAB.EGO	TAM/egophoricity
11	<i>-na</i> IMP <i>-ta</i> HORT.SING <i>-kéq</i> HORT.PL <i>-yâq</i> PROH	command mood

An inspection of Table 6.7 reveals that there is a rough distribution of aspect, i.e. event semantics, being closer to the stem than TAMN suffixes, which stand for the attitude of the speaker towards the action.

It should be noted that suffixes that have scope only over the predicate head (and not

the predicate derivation) occur between slots 1 and 2. This is illustrated in Example 211, where the completive only extends over *carry* and not the volitional causative⁸.

- (211) glâikûshyürêtakéq labô
 glâi-kû-shyü-rê-ta-kéq la-bô
 carry-COMPL.II-CAUS.V-PL.II-permanently-HORT.PL say-PFV.ALLO
 ‘(They) said: “let’s have (him) carry (the stone) away for good!”’
 [TA22_kutyemba_phla 60 163]

6.3.2 Number

Tawrã has three suffixes that indicate that a referent is non-singular. These are the plural suffixes *-ga* and *-ring* (both under §6.3.2.0.1) and dual suffix *-nega* (§6.3.2.0.2). These suffixes differ from agreement suffixes as known e.g. from European languages, as they are not obligatory and are routinely omitted once the number of the referent has been established. As discussed in §8.2.1, non-singular markers normally overlap with S/A arguments although they may instead agree with a P argument if its non-singularity is somehow prominent in the context. The label *number*, however, captures the association between these suffixes and the composition of the referent(s).

The effect of stem reduplication has been described as affecting number, as understood within the broad frame of *pluriactionality* discussed in §6.1.1.4.

6.3.2.0.1 Plural suffixes *-ga* and *-ring/-rê*

Plural suffixes *-ga* and *-ring/-rê* indicate a plural referent. As mentioned in §3.3.1.4 on the collective marker, verbal plural suffixes are often optional and are driven by the animacy of the referent (also cf. §8.2.1) and the relevance of stating plurality for the context. Crucially, the difference between these two suffixes (*-ga* and *-ring/-rê*) is related to the internal cohesion of the arguments; while *-ring/-rê* presupposes high cohesion, i.e.

⁸Directive *-zâ* would constitute an exception to the slot rule when following *karà* ‘rain’, as it precedes all other slots. However, the explanation lies in that this suffix is lexicalized to that verbal stem (cf. the discussion in §6.3.3.1).

it often refers to a coherent group, *-ga* is neutral in this respect. It does not necessarily presuppose group coherence. The exemplification of this difference requires a brief digression, since it is only observable in certain examples from the corpus and speakers perceive no difference between the two and even exchange them freely in elicitation. For the sake of clarity, *-ga* will be glossed as PL.I and *-ring/-rẽ* as PL.II. The sister language Kera'a also has a *-ga* plural suffix while a cognate of *-ring/-rẽ* has, so far, not been found.

The suffix *-ga* PL.I has a broader and more general meaning and occurs more often than *-ring/-rẽ* PL.II in the corpus. Its use does not imply group coherence or coordinated actions. It therefore occurs as a neutral plural suffix, as in Example 212. It may also occur with the hearsay particle (grammaticalized from the verb 'say', cf. §10.7.1), as in Example 213.

(212) *bîya ali dyângga tyéga we lana tyé (...)*

[*bîya ali dyâng-ga tyé-ga*]_{main.cl} *we la-na tyé*

past like.this chase-PL.I cut-PL.I DEM say-FACT cut

'In the past, like this, they chased and cut (each other), they say. (They) cut (...)'

[TA22_clan 6 36]

(213) *hwèshyûbô laga mlá mẽ we làng we*

[*hwè-shyû-bô la-ga [mlá mẽ we làng we]*]_{NP}]_{main.cl}

tighten-CAUS.V-PFV.ALLO HSAY-PL.I village.MOD person DEM COLL DEM

'They tighten (the rock), they say, these village people.' [TA22_kutyemba_phla 90 211]

The suffix *-ring/-rẽ* PL.II is associated with coordinated actions and coherent groups. Examples 214a and 214b are from the same story in which Kutyemba tricks the village people into hiring him to get rid of a deadly stone. In Example 214a, Kutyemba orders the village people to push the stone while lying on their backs. In Example 214b, he further orders them to push with their eyes closed. In these three examples, the villagers constitute a coherent group which makes decisions and acts together. They are therefore indexed with *-ring/-rẽ* PL.II.

- (214) a. “thrûga lya gũ kadyângrêna” la
 [thrûga lya gũ]_{cosub.cl} kadyâng-rê-na la
 on.the.back lie NF push-PL.II-IMP say
 “‘Push while lying on your back!’ he said.’ [TA22_kutyemba_phla_99 226]
- b. “bülüm màho tayârêna” la la
 [bülüm mà-ho tayârê-na]_{direct.speech} la la]_{main.cl}
 eye close-STATE shout-PL.II-IMP say HSAY
 “‘Shout with your eyes closed!’ he said, it is said.’ [TA22_kutyemba_phla_105 235]

Once the referent has been identified as a coherent group, the speaker switches to the neutral suffix *-ga* PL.I. Example 215a is a follow-up from Example 214b; the villagers are shouting after Kutyemba’s command. Similarly, in Example 215b (which occurs much later in the story) the villagers discover that Kutyemba has tricked them and they shout in anger. These examples take the neutral plural suffix *-ga*.

- (215) a. we tayâga khrâk mlă mẽ we
 [we tayâ-ga khrâk [mlă mẽ we]_{postposed.NP}]_{main.cl}
 DEM shout-PL.I EVI village.MOD person DEM
 ‘They shouted, these village people.’ [TA22_kutyemba_phla 109 241]
- b. “büle lê ne thâ wridyâ” atyângmâga
 [büle lê ne thâ wri-dyâ]_{direct.speech} atyâng-mâ-ga
 pig big only eat cheat-and.leave shout-ITER-PL.I
 “‘He only ate our biggest pig and left!’ they kept shouting.’
 [TA22_kutyemba_phla_147 308]

Certain actions are prototypically undertaken by coherent groups, such as migrating and building houses, and have a higher chance of being indexed with *-ring/-rê* PL.II (instead of *-ga* PL.I). In Example 216a, the speaker tells of a group of his clansmen who settled in Roing. In Example 216b, other clansmen are said to have left a place as part

of their ancient migration route. In both cases, clans seem to constitute coherent groups *par excellence*. Similarly, Example 217 takes the *-ring/-rẽ* plural suffix since building a house is prototypically done by a coherent group, e.g. a clan.

(216) a. Roing gũ bojar kě ìhoring

roing gũ bojar kě ì-ho-ring

a.place LOC market(<IA) LOC.INS be.there-STATE-PL.II

‘They live at Roing bazaar.’ [TA23_talyu_clan 29 175]

b. lêatîkũ lêtakũhoring

lê-atî-kũ lê-ta-kũ-ho-ring

exit-leave-COMPL.II exit-permanently-COMPL.II-STAT-PL.II

‘(They) left, (they) left forever.’ [TA23_talyu_clan 45 281]

(217) tûw gê we la we la K.-T. ladi we làng àng tâwrühoring

[*tûw gê we la*]_{main.cl} *we la* [*K.-T. la-di*]_{main.cl} [*we làng àng*

CLF:half one DEM HSAY DEM HSAY a.name say-HAB.EGO DEM COLL house

tâwrü-ho-ring]_{main.cl}

build.house-STAT-PL.II

‘One party, this, this, I call them K. T., they have built a house.’ [TA24_land 25 43]

Finally, the more general, neutral plural suffix *-ga* PL.I occurs with actions that lack inherent coordination and where referents are not deemed as having the ability to form coherent groups. In Example 218, the brothers fought each other. Such cases could not take the plural suffix *-ring/-rẽ* PL.II, for they are understood as chaotic (i.e. not coherent). In Example 219, the omitted A argument (‘the bees’) also triggers the neutral plural suffix *-ga* PL.I. We may presume that this is due to the fact that non-human referents often lack the perceived agency to undertake coordinated actions.

- (218) komjuri jina gũ webo hwûwga khûga weli gũ pûmru hî abli sêgakûbî khrâk
 [komjuri ji-na gũ webo hwûw-ga khû-ga]_{main.cl} [weli gũ
 quarrel(<IA) do-FACT NF and.then argue-PL.I fight-PL.I DISC.DEIC.DEM INS
 pû-mru hî abli sê-ga-kû-bî khrâk]_{main.cl}
 elder.brother-younger.brother EMPH some kill-PL.I-COMPL.II-PFV EVI
 ‘They had a fight and they argued and fought and some siblings killed (each other).’ [TA23_talyu_clan 40 241]

- (219) tya kwâg (...) we dyânggabô
 tya kwâg we dyâng-ga-bô
 now dog DEM chase-PL.I-PFV.ALLO
 ‘Now (the bees) have chased (...) the dog.’ [261122_frog1 93 429]

In sum, while the two plural suffixes have a similar meaning, *-ring/-rẽ* PL.II has carved out a specific semantic niche related to coordinated actions and coherent groups, while *-ga* PL.I is neutral. This morphosyntactic sensitivity for group coherence is in line with the use of egophoricity markers for plural referents (§6.3.1.1) (i.e. ‘we, a group’) and further resembles the *ideocclusive* pronoun set in Milang, which marks construed groups (cf. Modi 2017, 212) (also §6.3.1.1 on egophoricity). The relation between special morphological marking for construed groups and the socio-environmental conditions that have shaped the indigenous populations of the Eastern Himalayas offers a promising field for further research.

6.3.2.0.2 Dual marker *-nega*

The dual suffix *-nega* is likely composed of **-ne* + *-ga* (plural marker; §6.3.2.0.1). It signals the referent as being composed of two. In Example 220, a brother and a sister survive both floods and a big fire and then consider how to increase the population of the human race. The first verb is marked by the dual *-nega*; once their dual reference has been established, it need not be further marked.

(220) lünega bobyẽ (...) mẽ we (...) kadé.nyu tshûnaya îng la? (...) tshûblúm ãna yüm
bô (...)

[lü-nega bobyẽ]_{sub.cl} mẽ we kadé.nyu tshû-na-ya îng la
survive-DU TEMP.SUB person DEM how expand-FACT-ALLO a.bit HSAY

[tshû blúm ã-na yüm bô]_{sub.cl}
expand LOC.NMLZ be.there-FACT NEG TEMP.SUB

‘Having survived, these people, how (would they) thrive? Having no place to thrive (...)’ [TA23_newmyth 33 198]

6.3.3 Associated motion suffixes

Associated motion suffixes are verbal suffixes that associate translational motion to a verbal event (e.g. ‘eat and come back’) (cf. Guillaume and Koch 2021, 1; Genetti et al. 2021, 348). These suffixes often imply a telic reading for the verb they are attached to (Jacques et al. 2021, 841-842). Suffix *-bô* ‘and go’ is also a motion suffix but it will be discussed along with the ‘bô group’ (§6.3.4.8).

6.3.3.1 Directive *-zâ*

Directive *-zâ* can be paraphrased as ‘do X and bring/come (often towards the speaker)’ but it can also simply mean an action that is cast in a particular direction, not necessarily towards the speaker. This suffix is commonly used to describe events that occur to someone’s detriment. The link between direction markers and bene/malefactive is well-attested in the world’s languages (cf. Kuteva et al. 2019, 53,76, also cf. Jing and Liu 2023). Just like other motion suffixes, *-zâ* can imply completeness, emphasizing the telic nature of a verb.

In Example 221, the directive *-zâ* emphasizes the transgressive aspect of the action since the goods were stolen and taken away. Similarly, Example 222 emphasizes that the stone will bring harm to the village.

- (221) kya we (...) tsapakwa nyu haza simplang glâiringza mane na
kya we tsapakwa nyu haza simplang glâi-ring-zâ mane na
 Q DEM a.place FOC FS FILL carry-PL.II-DIR FILL(<Asm) INST
 ‘This (...) they carried it to Chapakhowa, right?’ [TA22_CLAN2 43 84]

- (222) maminzâya wêq adi ephlâe wegũ
[mamin-zâ-ya]_{sub.cl} wê-q ali [e=phlâ=e]_{top.NP}main.cl
 bring.misfortune-DIR-ALLO think-CONT like.this DEM = stone = DEM
wegũ
 and.then
 ‘(Kutyemba said:) “This stone is thinking about bringing misfortune.”’
 [TA22_kutyemba_phla 18 93]

The negative implication of *-zâ* is evident in some lexicalized forms. In the elicited Example 223, the use of the directive implies slanderous or otherwise defamatory statements (and not simply ‘say’ or ‘say to me’). Example 224 displays this suffix with a telic and possibly also malefactive reading.

- (223) cya ja lazâq
cya ja la-zâ-q
 3 what say-DIR-CONT
 ‘What is he saying about/against me.’ [261122_JO_eli 1221]

- (224) shêkũnazâ mane
shêi-kũ-na-zâ mane
 die-COMPL.II-FACT-DIR FILL(<Asm)
 ‘He died.’ [TA22_CLAN2 107 210]

The directive suffix *-zâ* is found as a lexicalized component with two verbs: *shì* ‘take’ and *karà* ‘rain’. In both cases, the lexicalization entails a fixed post-verbal root position

of the directive, i.e. *skipping* the slot order displayed in Table 6.7. Phonologically, however, there are no signs of these collocations becoming a single word, as they keep their tone/stress. In Example 225, the directive *-zâ* adds a direction to the verb *shì* ‘take’, rendering ‘bring’.

- (225) *shìzâna!*
shì-zâ-na
 take-DIR-IMP
 ‘Bring it!’ [021024_STBT_eli2 395]

In Example 226, *-zâ* is used with the verb *karà* ‘rain’ to imply rainfall. The directive does not imply a telic reading which is commonly associated with the directive. Example 227 illustrates the grammaticalized character of the directive *-zâ*, as it occurs before the predicate derivation *-talá* CONT (§6.2.4.1), *skipping* the suffixal chain presented in §6.3.1.3

- (226) *hã karàzâ künüm abang bô yüm*
[hã [karà-zâ künüm]_{sub.cl} habang bô yüm]_{main.cl}
 1S rain-DIR PRP outside go NEG
 ‘I won’t go out because of the rain.’ [240122_AM_eli2 19518 93]

- (227) *karàzâlabiya*
karà-zâ-talá-biya
 rain-DIR-PERMA-FUT.ALLO
 ‘It will keep raining.’

Finally, *-zâ* can be employed with interrogative pronouns to render a distributive sense, i.e. the marked pronoun refers to more than one referent (e.g. ‘where (referring to many places)’). This function is described in §3.5.1.2.

6.3.3.2 Go (and come back) -ge

The suffix *-ge* implies that the S/A argument of the verb has gone somewhere to do the action and came back (after finishing). In Example 228, the tiger and the human are competing; the tiger is to count small fish in an elusive school while the human is to count nodes in a hanging bridge. The tiger could not count beyond the number ‘two’ before the human came back after finishing counting all nodes (*tsâi-dâng-ge* count-COMPL.I-go.and.back).

- (228) “ekhing kâqing” ma tsâiya ba yüm mu mẽ we tawrê lô tsâidângge matyô
e = khing kâqing ma tsâi-ya ba yüm mu mẽ we
 DEM = one two CONTR.TOP count-ALLO be.able NEG also person DEM
tawrê lô tsâi-dâng-ge matyô
 a.grass hanging.bridge count-COMPL.I-and.come.back maybe
 “‘One two”, (the tiger) could not count (beyond that) (and) the person finished counting the (nodes) of the *tawrê* hanging bridge.’ [TA22_tiger 65 156]

The telic tense of this suffix is highlighted in Example 229, where the *chambri* bird ‘shot and came back’, i.e. killed the eagle - the fact that it came back is not relevant for the story but rather indicates that it succeeded in killing the eagle (i.e. shoot and came back) (cf. Jacques et al. 2021, 841-842).

- (229) câmbriyâ ne ôgeya
câmbri-yâ ne ô-ge-ya
 a.bird-FEM only shoot-and.come.back-ALLO
 ‘Only the chambri bird could kill it (lit. shoot it down and come back).’ [TA22_origin 190 521]

Although the suffix *-ge* adds a new motion to the main verb, these actions are integrated to the point that a negation marker affects both the action of the verb and the motion. This is illustrated in Example 230, where the negator applies to both the verb and the associated motion suffix (which here has a telic function) (cf. Jacques et al. 2021, 842).

- (230) *shya mu ônagem*
shya mu ô-na-ge = m
 who also shoot-FACT-and.come.back = NEG
 ‘No one could kill it.’ [TA22_origin 189 519]

6.3.4 Aspect/Aktionsart

The suffix group ‘aspect/Aktionsart’ includes suffixes that seem to affect aspect/Aktionsart, as discussed in §6.3.1.2.

6.3.4.1 ‘ahead’ *gadyô*

The suffix *gadyô* ‘ahead’ implies that an action is carried out earlier than expected. In Example 231, the speaker uses this suffix as s/he decides to eat *ahead*, even though s/he is expected to eat with a group of people. In Example 232, the speaker decides to part earlier than expected.

- (231) *hã thâgadyô*
hã thâ-gadyô
 1s eat-ahead
 ‘I’ll eat ahead.’ [221024_JO_eli2 1162]

- (232) “îng halo di ìnagadyô” *la*
îng halo di ì-na-gadyô la
 some month at.least stay-FACT-ahead say
 ‘He said: “I’ll go ahead, for some months at least.”’ [TA22_Clan2 180 345]

6.3.4.2 Continuative -*q*

The continuative -*q* suffix expresses that an action is ongoing. In Example 233, the continuative -*q* indicates that the action has not yet finished and is ongoing.

(233) “hî àng khě mẽ” wêq

[hî àng khě mẽ]_{compl.cl} wê-q

EMPH house LOC.INS person think-CONT

‘He is thinking: “there is a person at (my) house!” [TA23_FishStory 28 69]

6.3.4.3 Completive *-kũ*

The completive *-kũ* COMPL.II is a common suffix in the language that marks the completive aspect. This is probably related to completive suffixes in the Tani language family (*-kú* in Galo, cf. Post 2007, 585; *-ku* in Tangam, cf. Post 2017, 62). Also cf. perfective suffix *-khal* in Pangkhua (Kuki-Chin) (Akter 2022, 238).

This suffix is used to express that an action has been carried out in a way that cannot be continued any longer (i.e. *completely* done) and may indicate that the changes brought about by an action cannot be easily undone. In Example 234, the deer throws the child and his dog down a slope. The completive emphasizes the irrevocable nature of the action. Similarly, in Example 235 the stone has to be carried *away*, i.e. to get rid of the stone. In Example 236, the completive occurs with the verb ‘say’, suggesting that the wife said everything there was to say (regarding her husband’s weak spot) and that her statements could not be taken back.

(234) tya azâkũge delé

tya hazâ-kũ-ge delé

now drop-COMPL.II-and.back it.seems

‘(The deer) went to drop them (and came back) it seems.’ [TA22_frog1 78 364]

(235) we mu, ephlâe glâikũya knüm gũ (...)

we mu e=phlâ=e glâi-kũ-ya knüm gũ

DEM also DEM=stone=DEM carry-COMPL.II-ALLO PRP INS

‘This, also, in order to carry this stone away (...)’ [TA22_kutyemba_phla 27 111]

- (236) “tapî dũ atyûya ganei shi” la la atyâkûbi (...) uh mîya we jûm la jûm la jûm la gũ
 [tapî dũ atyû-ya ganei shi la la]_{main.cl} [atyâ-kû-bi]_{main.cl} [uh mîya we
 sand with fry-ALLO SEQ die say HSAY tell-COMPL.II-PFV oh wife DEM
 jûm la jûm la [jûm la gũ]_{cosub.cl}]_{main.cl}
 true say true say true say NF
 ‘She said “fry (him) with sand and (he) will die”, they say. Oh, this wife was say-
 ing the truth, she was saying the truth, she was saying the truth.’ [TA23_tayang_clan
 35 210]

The irrevocable implications of *-kû* are most evident with stative verbs. In Example 237, man and tiger became different from each other in a way that cannot be amended. The completive emphasizes this definitive change.

- (237) mề ma tamyá ma pütyô shyâkû
 mề ma tamyá ma pütyô shyâ-kû
 person ADD tiger ADD different become-COMPL.II
 ‘Human and tiger became separated/different.’ [TA22_kutyemba_phla 144 318]

In rare instances, the completive marker can be reduplicated to render a distributive sense on top of the completive Aktionsart. In Example 238, it is important that the body of the protagonist was cut into small pieces. The reduplication of both the verbal head (cf. §3.6.2) and the completive suffix help render a distributive effect.

- (238) (...) tyêkûkûya gũ (...) katyô katyô ne thâyâ shyâ (...) harûmga gũ tyêkû gũ byẽ
 [tyê-kû-kû-ya gũ]_{cosub.cl} katyô katyô ne thâyâ]_{main.cl} shyâ
 cut-COMPL.II-REDUP-ALLO NF small REDUP only eat-ALLO happen
 [harûm-ga gũ]_{cosub.cl} [tyê-kû gũ]_{cosub.cl} byẽ
 get.together-PL.I NF cut-COMPL.II NF TOP
 ‘(...) cutting it to smithereens (...) (into) small small pieces (and) eat it (...) people
 gathered, to cut (him).’ [TA23_tayang_clan 37 222]

6.3.4.4 Iterative *-mâ*

The iterative suffix is an aspectual marker that indicates that an action is repeated, often in an exaggerated or otherwise undesirable manner. In Example 239, we are introduced to a trickster who habitually cheats on people. In Example 240, the villagers continuously pull and tie ropes. In Example 241, the officer kept on waiting (in vain) for his stolen horse to return.

(239) *we mẽ lâhlakmâ*

we mẽ lâhlak-mâ

DEM person cheat-ITER

‘This person was a cheater.’ [TA22_SOstory 3 31]

(240) (...) *we tarâw tatyế hrêrêmâkũya*

we tarâw tatyế hrê-rê-mâ-kũ-ya

DEM type.of.creeper rope pull-tie.up-ITER-COMPL.II-ALLO

‘(They) kept pulling and tying the creeper ropes.’ [TA22_kutyemba_phla 81 168]

(241) *kalà lyó kalyũmâ (...) güwri hanalyâm*

[*kalà lyó kalyũ-mâ*]_{main.cl} [*güwri hana-lyâ = m*]_{main.cl}

officer white wait-ITER horse come-return = NEG

‘The white officer kept on waiting (but) (his) horse did not come back.’

[TA22_JAstory 53 179]

The iterative marker *-mâ* shares with the stative *-ho* (§6.3.4.5) the trait of also occurring with adjectives, as has been discussed in §3.2.4.

The canonical way that the iterative indicates repetition gives way to a tedious, wearisome or otherwise negative, if not malignant, interpretation of a verb. In Example 242a, the iterative frames the action of drinking as inherently undesirable (a drunkard). Similarly, in Example 242b the verb *marô* ‘talk’ marked with the iterative suffix does not

render a neutral ‘talk continuously’ but it refers to an *annoying* way of talking. Furthermore, this form underwent an additional semantic development, as it (also) specifically refers to a mentally ill person. Note that verbs do not require a special morpheme to be nominalized and to head a NP; the iterative marker *-mâ* is not a nominalizer, but rather the nominal interpretation (‘drunker’, ‘crazy person’) springs from the semantics of the verb, as discussed in §3.3.2. Some combinations of verbs and the iterative suffix, however, cannot stand as heads of a NP or are perceived as strange. This is the case of Example 242c, where the verb is interpreted as tedious or wearisome but an interpretation of ‘a flyer’ is not unanimously accepted.

- (242) a. *yu-mâ*
 drink-ITER
 ‘(S/he) drinks too much, a drunkard’
- b. *marô-mâ*
 talk-ITER
 ‘(S/he) talks too much, talkative person, a talker, also crazy person.’
- c. *yüm-mâ*
 fly-ITER
 ‘It kept flying (annoyingly)’ (?‘flyer (e.g. flying squirrel)’ is perceived as strange). [all three 021024_STBT_eli2 469]

Finally, a semantically negative interpretation of the iterative marker *-mâ* seems not to be available for adjectives. In Example 243, the adjective has a *neutral* iterative interpretation.

- (243) *kanú-mâ*
 dark-ITER
 ‘It keeps getting dark (It does not yield *‘it gets annoyingly dark’ or the like). [021024_STBT_eli2 469]

6.3.4.5 Stative *-ho*

The stative marker *-ho* implies a state that is stable in time or where change is not foreseeable (or somehow not relevant). Like the iterative marker *-mâ* (§6.3.4.4), the stative marker can also occur with adjectives, as has been discussed in §3.3.3.

With non-motion verbs, the stative commonly expresses a timeless state corresponding to adjectives in other languages like English (cf. §4.3.1). It may also imply a resultative reading, i.e. it refers to states that have *become* like that. For active verbs, it describes ongoing events.

In Example 244, the speaker uses *-ho* to express a state of affairs upon seeing a shining moon in the image he is describing. Similarly, in Example 245 from the same storybook (Mayer 1969), the dog appears with its tail between its legs, which prompts the use of *-ho*. In this example, a resultative reading ('it has become X') is also possible.

- (244) atûm kê halo gûngho
hatûm kêq halo gûng-ho
sky LOC.INS moon shine-STAT
'In the sky the moon is shining.' [TA22_frog1 9 38]

- (245) kwâg we, lûmûng pâho dlé
kwâg we lûmûng pâ-ho dlé
dog DEM tail be.curved-STAT it.seems
'This dog, its tail is curved (also: has become curved), it seems.' [TA22_frog1 103 475]

In Examples 246 and 247 (both active verbs), the stative marker *-ho* indicates a state of affairs that is true and ongoing at the time of the story. In Example 248, an active verbs also has a resultative reading due to the context: the villagers are in shock upon discovering that the trickster has left them while they had their eyes closed. In this sense, it is not relevant for the story whether the action is ongoing but the fact that the trickster is suddenly gone.

- (246) *tyû nyu kya rôhonéq?*
tyû nyu kya rô-ho-né-q
 chicken TOP Q catch-STAT-ASSO-CONT
 ‘Is it the chicken that is being caught (lit. taken with)?’ [TA22_chicken_thief 71 255]
- (247) *tayâkadyângholi (...) tayâdângli gũ (...)*
[tayâ-kadyâng-ho-li]_{main.cl} [tayâ-dâng-li gũ]_{cosub.cl}
 shout-push.with.legs-STAT-REC.PST shout-COMPL.II-REC.PST NF
 ‘(they) shouted and pushed. After having shouted (completely) (...)’.
 [TA22_kutyemba_phla 128 271]
- (248) *byûwatîhona gũ (...)*
byûw-atî-ho-na gũ
 run-leave-STAT-FACT NF
 ‘By running away (...)’ [TA22_kuyemba_phla 120 259]

6.3.4.6 Intentional (egophoric) *-ne*

The intentional *-ne* implies a purpose to act (cf. *-(vi)dyô* certain future §6.3.4.10). Although it is an egophoric marker, its role as a marker of intention allows it to occur with 3rd persons S/A in particular contexts, e.g. in stories. This feature is not common for other egophoric suffixes (further discussed in §6.3.1.1). The intentional suffix may only refer to non-past actions; a past interpretation is not allowed.

In Example 249, the hero of the story tells the British officer that he intends to go and fetch his magical bag. In Example 250, the speaker employs this form to guess the addressee’s intention.

(249) hã karô hananane hã tya ge.ne bô gû tya shìnagene
 [hã karô hana-na-ne]_{main.cl} [hã tya ge.ne bô gû tya shì-na-ge-ne]_{main.cl}
 1S quickly come-FACT-INT 1S now once go NF now take-FACT-go.and.back-INT
 ‘(He said:) “I’ll come quickly, I’ll go at once, I’ll go now and bring it now.”’
 [TA22_JAstory 37 126]

(250) ning dũ thâne kya?
 ning dũ thâ-ne kya
 1P with eat-INT Q
 ‘Do you want to eat with us?’ [AM_WA_2022_07_28]

6.3.4.7 Egophoric habitual *-de/di/d*

The egophoric habitual suffix *-de/di* refers to actions that either do not have a clear endpoint or are carried out on a regular basis⁹. Crucially, this marker is not specified for tense and a past or non-past interpretation depends on the context. It is regularly translated into (Arunachali) Hindi with the habitual aspect which coincides with a habitual, not punctual reading. This suffix does not have an allophoric counterpart.

The variation in vowel quality (*-de/di*) seems to reflect dialectal differences: the *-de* form is more extended in the plains (Lohit) while the *-di* form is found in the mountains (Anjaw). The *-d* form occurs when the negative marker is cliticized, i.e. *thâdyüm* (*thâ-d=yüm*, eat-HAB.EGO=NEG) ‘I won’t eat’ (cf. §2.2.7.1).

Example 251 corresponds to an imprecise time that extends into the present. In Example 252, the speaker recalls an indefinite time, as opposed to a punctual event. Similarly, Example 253 refers to a common event in the past (that is not true anymore). Crucially, these examples show that the habitual suffix is not anchored to a particular temporal interpretation but is a purely aspectual marker.

⁹Having a continuous and habitual reading coupled in one verbal category (imperfective) is a common aspectual configuration in the world’s languages (cf. Comrie 1976, 26).

(251) *we ni tya khré mu alí ìhaládi*

we ning tya khré mu alí ì-ha-talá-de

DEM 1P now up.to also like.this be-FRUS-still-HAB.EGO

‘And we, even today remain like this.’ [TA22_origin 169 470]

(252) *ning tawrã bá.mě kade ìdi!*

ning tawrã bá.mě kade ì-de

1P a.tribe clan how.many be.there-HAB.EGO

‘Our Tawrã were plenty (lit. how many we were!).’ [TA22_clan 10 51]

(253) *hayũ ké câng yũm gũ eyô gũ yu kwêde*

[hayũ ké câng yũm gũ]_{cosub.cl} eyô gũ yu kwê-de

past rice be.available NEG NF millet INS beer distill-HAB.EGO

‘In the past, rice was not available; we used to make beer from millet’ [talk 050123_JA]

This suffix is likely related to the verb *di* ‘sit’. A verb meaning ‘sit’ becoming a marker of habitual aspect is a well-attested grammaticalization path (Kuteva et al. 2019, 405; also noted by Grierson 1909, 617).

6.3.4.8 The *bô* group

Several heterogeneous morphemes are historically derived from the versatile verb *bô* ‘go’. These are: a) the perfective aspectual suffixes (§6.3.4.8.1), b) a TAMN suffix with the meaning ‘V and go’/‘go and V’ (§6.3.4.8.2), c) a conditional and temporal subordinator (§10.4.1) and d) a sentence connector (§10.8.2). This section will only discuss the first two suffixes, as the rest of these morphemes are described in other chapters. Section 6.3.4.8.3 will discuss some synchronic aspects of the *bô* group that are of typological interest.

6.3.4.8.1 Perfective *-dyibô/-bô*

The egophoric perfective suffix *-dyibô* and the allophoric perfective suffix *-bô* are discussed together as they only differ in terms of egophoricity (cf. §6.3.1.1). Both markers cast a situation as *completed*, without an “explicit reference to the internal temporal constituency of a situation” (Comrie 1976, 21f) and are often opposed to an ongoing or otherwise incomplete situation. In Example 254, the allophoric perfective marker casts the situation as a complete whole, disregarding its internal complexity; a *taló puja* has several stages and lasts up to four days but this internal complexity is not relevant in this case. The perfective can also herald a new state of affairs, especially if it occurs in contrast with a previous state or if it is perceived as an abrupt transition (i.e. not gradual). In Example 255, the use of the perfective with the stative verb ‘forget’ is consistent with the idea of a change of state.

(254) we talóbô we mane atyu

we taló-bô we mane atyu

DEM perform.funerary.rite-PFV.ALLO DEM FILL(<Asm) up.there

‘(They) performed the *taló* , I mean, up there.’ [TA24_land 151 279]

(255) hã lâlhak phlê we àng nyu wesâdyibô

hã lâlhak phlê we àng nyu wesâ-dyibô

1S magical bag DEM house FOC forget-PFV.EGO

‘I forgot my magical bag at home.’ [TA22_JAstory 23 90]

Punctual and non-motion verbs commonly occur with the perfective, such as *shêi* ‘die’, *shyâ* ‘happen, become’, *lû* ‘survive’ or *tamê* ‘be crazy’, as these verbs commonly do not have an internally complex composition.

The perfective aspect offers a broad reference to a completed event, one which can be complemented with other aspectual markers that highlight aspects of its internal complexity (Comrie 1976, 21f). For example, the completive *-kû* highlights the completion of an event. This is illustrated in Example 256, as the thief takes the horse away for good.

The stative *-ho* points to the static or ingressive nature of an event, as in Example 257. The permanent marker *-tá* also highlights the permanent state of an event, as in Example 258.

- (256) shikũbô khrâk kadé.mu hanalyam
 [shì-kũ-bô khrâk]_{main.cl} [kadé.mu hana-lyá = m]_{main.cl}
 take-COMPL.II-PFV.ALLO EVI when come-back = NEG
 ‘It seems he took it (the horse) away (and) will never come back.’ [TA22_JAstory 59 197]

- (257) àng mu bâhobô alyüm ah gorã alyüm we (...)
 àng mu bâ-ho-bô alyüm ah gorã alyüm we
 house also build-STAT-PFV.ALLO road ehm horse(<IA) road and.then
 ‘They have also built houses, roads, ehm, horse trails and then (...)’ [TA22_clan 145 436]

- (258) (...) yúw mlâ we makángtábô mane
 yúw mlâ we makáng-tá-bô mane
 region village DEM take.control-permanently-PFV.ALLO FILL(<Asm)
 ‘(...) he took permanent control of the whole place, I mean.’ [TA22_clan 145 436]

As hinted above, the perfective suffixes *-dyibô* and *-bô* are portmanteau morphemes since they also encode egophoricity and constitute an egophoricity pair, i.e. *-dyibô* is sensitive to egophoricity while *-bô* is to allophoricity further discussed in §6.3.1.1). It is noteworthy that this is the only egophoricity pair in the language; no other egophoricity-sensitive marker has a functional counterpart, e.g. the habitual egophoric marker *-de/di* (§6.3.4.7) lacks a habitual *allophoric* counterpart.

It is not clear what morphological process led to the creation of two distinct perfective suffixes *-dyibô* and *-bô* since there is no segment **dyi-* in the language. One possibility is that the egophoric habitual suffix *-de/di* could have merged with *-bô* rendering **-dibô* (as

seemingly implied by Sastry (1984, 154f.)). The unstressed first syllable could have been palatalized at some point, rendering the modern form. More historical work is required to confirm this hypothesis.

6.3.4.8.2 ‘and go/go to’ suffix

The root *bô* occurs as a TAMN suffix meaning ‘do X and go’. This suffix may either imply a sequence of actions, i.e. movement after the completion of the action (*X and go*), as in Example 259, or a coordination (almost purposive) of actions: ‘go to do X’, as in Example 260.

- (259) *hã thâbône*
hã thâ-bô-ne
 1S eat-go-INT
 ‘I will eat and go.’ [021024_STBT_eli2 580]

- (260) *cya matsâw matyâbô*
cya matsâw matyâ-bô
 3 cow release-go
 ‘He went to release the cows.’ [261122_JO_eli 951]

§6.3.4.8.3 will discuss typological and diachronic aspects that are relevant to mapping the development of the *bô* group.

6.3.4.8.3 Typological and diachronic overview

The versatile verb *bô* ‘go’ has been prolific in rendering grammatical morphemes with diverse meanings and syntactic positions in Tawrä. Figure 6.1 illustrates four grammatical morphemes (and their pairs when available) derived from the verb *bô* ‘go’.

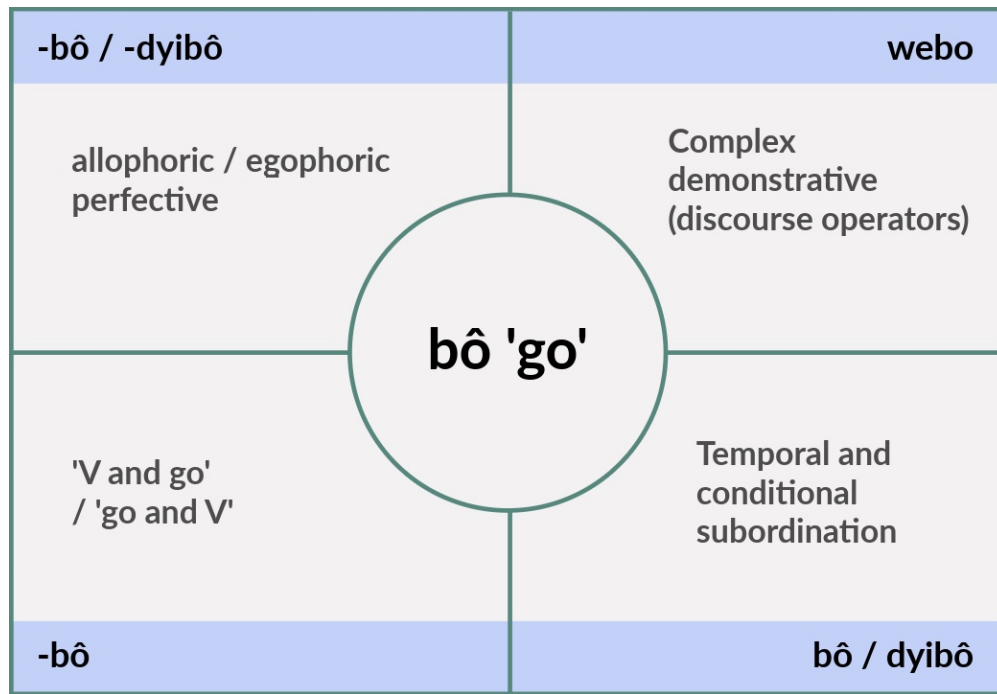


Figure 6.1: Grammatical morphemes derived from the verb *bô* 'go'

Concerning the grammaticalization path of these morphemes, the predicate derivation meaning 'V and go' is the closest to the source form: the verb 'go' > suffix 'V and go/go and V' and closest to the slot position of the independent verb. We may therefore presume that this was the first grammatical morpheme that was derived from the independent verb. Furthermore, the TAMN perfective suffix pair *dyibô*/-*bô* would have been a further derivation, further away from the stem (TAMN slot, cf. §6.3.1.3) and with new morphological material (**dyi-*, cf. §6.3.4.8.1). These two derived morphemes cannot co-occur in the same predicate, cf. Example 261, and occupy two different slot positions that do not overlap. Grammaticalization and the inflection/derivation divide have been discussed in §6.2.1.1.

(261) **cya thâbôbô*

cya thâ-bô-bô

3 eat-and.go-PFV.ALLO

Intended: 'S/he ate and went.' [021024_STBT_eli2 595]

The conditional and temporal subordinator *dyibô*/-*bô* is likely a further development

of the grammaticalized predicate derivations and TAMN suffix pair. It occupies a post-predicate slot and has clause-level scope (cf. §10.4). In Example 262, the temporal subordinator *bô* has scope over whole clauses.

- (262) tamûn hrû yüm hanâmâ hanâmâ hrû yüm bô khümyûa mamýú bô (...)
 [tamûn hrû yüm]_{main.cl} [hanâmâ hanâmâ [hrû yüm bô]_{sub.cl}]_{main.cl}
 fire burn NEG press-ITER press-ITER burn NEG TEMP.SUB
 [khümyû-ya mamýú bô]_{sub.cl}
 get.angry-ALLO after TEMP.SUB

‘The fire did not burn. He squeezed and squeezed (his testicles) because it (the fire) did not burn. He got angry and later (...)’ [TA22_tiger_116 262]

The complex demonstrative *webo* is probably composed of the demonstrative *we* DEM and morpheme **-bo*, related to the versatile verb ‘bô’, which lost its tone/stress as part of the grammaticalization process. This complex demonstrative functions as a discourse connector and often occurs in a clause-initial position (§10.8.2).

Typologically, the verb ‘go’ commonly grammaticalizes into a progressive or a future marker although it also became a marker of *change of state* in many languages (Kuteva et al. 2019, 204f; 212f; for a cognitive explanation, cf. Langacker 2002, 149). The grammaticalization path of ‘go’ to a perfective marker is, however, marginally found in some languages of MSEA (e.g. Thai, cf. Jenny 2000, 117f) but prominently found in Indo-Aryan languages, as well as in Bodic (DeLancey 1991, 6-11). Concerning this development in Indo-Aryan, bilingual speakers of Tawrã are well aware of the semantic similarity between the perfective marker *-bô* and the perfective vector verb based on ‘go’ in languages like Hindi and Nepali (cf. Masica 1993, 276; cf. Slade 2016). As an example, the perfective marker in the elicited Example 263a ‘s/he died’ was spontaneously translated to Nepali as Example 263b. I do not mean to imply that Indo-Aryan might have influenced the recruitment of the verb ‘go’ as a marker of the perfective aspect but it is worth noting that these developments went parallel ways in both languages and that speakers are aware of it.

(263) a. *shêi-bô*
 die-PFV.ALLO
 ‘S/he died.’

b. Tawrã Nepali
mor-ga-yo
 die-go[.PFV]-PST.3S
 ‘S/he died.’ [both 021024_STBT_eli2 650]

The grammaticalization of the verb ‘go’ as a discourse operator is a much rarer development typologically. One noteworthy case is found in Kxoe (Khoe, mostly Namibia), where Heine (2000) describes the role of the grammatical element *cîl* in the periphrasis ‘watch out, now something new is going to happen that is relevant to what follows’, i.e. a *new event*. Crucially, *cîl* also exists as a verb meaning ‘proceed, reach’ (cf Kuteva et al. 2019, 94f. also Bourdin 2008). The grammaticalization of the verb ‘go’ as a conditional and subordination marker is also a rare phenomenon.

Finally, the etymology of *bô* ‘go’ is not entirely clear and requires further historical work. Kera’a has verb *ba* ‘go’. There is a potential cognate *-boŋ* suffix found in some Eastern Tani languages where it is a marker of ‘change of state’. The velar nasal coda *-ŋ* could be an innovation of Eastern Tani (Post 2024, p.c.).

6.3.4.9 Recent past *-li/lü*

The recent past marker *-li/lü* is not a common marker in the corpus and it often occurs along with other temporal and aspect markers. When elicited alone, speakers depict the event as having ‘just occurred’. Therefore *matyâ-li* (release-REC.PST) yields ‘just released/already released’ and *tsâng-li* (drown-REC.PST) ‘just drowned’. In Example 264, an Idu prisoner was taken to jail where he died shortly thereafter. The use of the recent past marker emphasizes the close succession of these actions. Similarly, in Example 265 a trickster tells the village people that he will perform the desired action (dispose of a deadly rock) after eating at their village. He assures the people that he will carry out the duty immediately after eating.

(264) webobyẽ dâi we shêli labô jîngbô

[webobyẽ dâi we shê-li la-bô]_{main.cl} [jîng-bô]_{main.cl}

and.then a.tribe DEM die-REC.PST HSAY-PFV.ALLO end-PFV.ALLO

‘And then this Idu died, they said. It was finished.’ [TA22_CLAN2 110 214]

(265) thâli ge.ne ane mlă kě (...)

thâ-li ge.ne ane mlă kě

eat-REC.PST once 2P village LOC.INS

‘(Kutyemba said:) (after) eating once at your village (...)’ [TA22_kutyemba_phla 40 134]

6.3.4.10 Certain future (egophoric) *-(vi)dyô*

The certain future *-(vi)dyô* expresses an impression that an act will surely be carried out. Although it is an egophoricity-sensitive marker (cf. §6.3.1.1) it also implies intention. This leads this suffix to occur with 3rd person S/A arguments, something that is uncommon for egophoric suffixes (a similar case occurs with ‘intentional’ *-ne*, cf. §6.3.4.6). For some speakers, this suffix has a voiced labiodental fricative [v] < v > on the first syllable that is otherwise not found in the language; it is a defective phoneme (cf. §2.4.9). It is not clear to me whether this phoneme is bound to a particular dialectal zone. In Example 266, the main character of the story announces to the village people that he will carry the deadly stone away. He uses *-vidyô* to reinforce his certainty.

(266) we tyû mlă mẽ we “tya hîng we phlă we glâikũvidyô”

we atyû mlă mẽ we [tya hîng we phlă we

DEM north village.MOD people DEM now day DEM stone DEM

glâi-kũ-vidyô]_{dir.speech}

carry-COMPL.II-CERT.FUT

‘(Kutyemba said) to the northern village people: “today I’ll (surely) carry this stone (away).”’ [TA22_kutyemba_phla 25 103]

6.3.4.11 ‘permanently’ -*tá*

The TAMN suffix *-tá* expresses events that occur permanently. As such, it includes a telic component. In Example 267, the village people hope that an appointed person will carry a stone away for good. The use of *-tá* entails a permanent development. Similarly, in Example 268 the speaker employs the permanent suffix to indicate that everybody left his clan’s ancestral place; all clan members settled somewhere else. These actions are permanent and cannot be undone.

(267) “glâikûshyürêtakeq” labô

glâi-kû-shyü-rê-tá-keq

la-bô

carry-COMPL.II-CAUS.V-PL.II-permanently-HORT.PL say-PFV.ALLO

‘(They said:) “let’s have (him) carry (the stone) away for good”!’

[TA22_kutyemba_phla 60 163]

(268) *tuluqa we kê ning we shyâ mu ina yüm lêatîkû lêtakûhoring*

[tuluqa we kê ning we shyâ mu i-na yüm]_{main.cl}

a.place DEM LOC.INS 1P DEM happen also be.there-FACT NEG

[lê-atî-kû]_{main.cl}

[lê-tá-kûhoring]_{main.cl}

migrate-leave-COMPL.II migrate-permanently-COMPL.II-STAT-PL.II

‘At Tuluqa we are no more. All left, they left for good.’ [TA23_talyu_clan 45 281]

6.3.5 Command mood suffixes

The command mood suffixes include the imperatives *-na* and *-tè* (§6.3.5.0.1), the prohibitive *-yâq* (§6.3.5.0.2), the hortatives *-ta* and *-keq* (§6.3.5.0.3) and the jussive *-tyêyu* (§6.3.5.0.4). Command moods do not commonly occur with other TAMN suffixes. The jussive differs morphologically from the other command moods as it includes two suffixes (the second is the hortative). In terms of the organization of slots, however, this bi-morphemic marker occupies only one slot (nr. 11) as no other morpheme may occur in between.

6.3.5.0.1 Imperatives *-na* and *-tè*

An imperative sentence is employed for direct orders and requests, as well as invitations and advice (König and Siemund 2007, 303; Davies 1986, 30). Tawrà has two dedicated suffixes *-na* and *-tè* for expressing the positive imperative (termed *true imperative*, cf. König and Siemund 2007, 304). The negation of the imperatives is achieved through a specialized prohibitive suffix *-yâq* (§6.3.5.0.2). The difference between the two suffixes seems to lie in register: while *-na* is the most common suffix, one speaker consistently employs *-tè* in stories (hence *narrative imperative* IMP.NARR). These suffixes do not seem to differ in their syntax and meaning.

In Example 269, the officer commands Kutymba to try and trick him. Similarly, in Example 270 the village people are commanded to roast a pig. The imperative suffix used in narratives *-tè* is illustrated in Example 271. It should be noted that, in imperative clauses, S/A arguments are commonly ellipsed (cf. §8.2.2).

(269) nyu la lâlhakmâ laga. tya hã ing lâlhak hanekûna!

[nyu la lâlhak-mâ la-ga]_{decl.cl} [tya hã ing lâlhak hane-kû-na]_{imp.cl}

2S FS cheat-ITE HSAY-PL.I now 1S also cheat try-COMPL.II-IMP

‘They say you are a trickster. Try to trick me!’ [TA22_JAstory 15 62]

(270) ble lê ewe münbina

bülei lê we mün-bî-na

pig big DEM roast-EMPH-IMP

‘Roast this big pig!’ [TA22_kutyemba_pla 38 128]

(271) tyékû apyékkû gũ yang túm khalai pwôktè

[tyé-kû apyé-kû gũ]_{cosub.cl} yang [[hatûm khalai]_{NP} pwôk-tè]_{imp.cl}

cut-COMPL.II spread-COMPL.II NF xxx sky ground blow-IMP.NARR

‘After cutting (and) spreading it, blow it everywhere (lit. (through) sky and land)!’ [TA22_origin 54 157]

6.3.5.0.2 Prohibitive -yâq

The prohibitive mood is expressed by a specialized marker -yâq. This suffix is not related to the negative markers -yüm and -ma (cf. §6.4.1). As mentioned in §1.1.5, Sun (1993, 406) indicated a link with Proto Tani prohibitive *-yo (<-jo> in Sun's orthography).

The prohibitive can be thought of as a negative imperative; to command *not* to do something. In Example 272, the prohibitive expresses a command *not* to spit. Similarly, in Example 273 a trickster tells the village people not to stay near him.

- (272) khalai tyôyâq
khalai tyô-yâq
dirt spit-PROH
'Don't spit!' [241122_RI_voc 530]

- (273) hã nú nyu byě ìrěyâq
hã nú nyu byě ì-ring-yâq
1S near FOC TOP be.there-PL.II-PROH
'Don't stay near me!' [TA22_kutyemba_phla 92 114]

A prohibitive clause can be formed indirectly, without the prohibitive suffix but merely with the allophoric suffix -ya and the general negative suffix (cf. §6.4.1). This has the effect of lessening the directive force that the prohibitive carries (König and Siemund 2007, 311). In Example 274, a younger brother tells his elder brothers not to make traps lest the dholes (*Cuon alpinus*) die. The use of this indirect strategy is triggered by the hierarchical relation between brothers; an imperative could be felt as transgressing that relation.

(274) (...) pûmru we (...) cyang (...) tamêga shyâ mu thûde jiya yüm bum.kamh jiya yüm (...)

[pûmru we]_{stranded.NP} [cya làng tamê-ga]_{main.cl} [shyâ mu
 younger.brother DEM 3 COLL refuse-PL.I happen also
 thû-de ji-ya yüm]_{main.cl} [bum.kamh ji-ya yüm]_{main.cl}
 trap-HAB.EGO do-ALLO NEG trap.in.a.pit do-ALLO NEG

‘(...) this younger brother, they disagreed. He (said) not to trap (dholes) (and) not to make trapping pits (...)’ [TA23_talyu_clan 11 60]

6.3.5.0.3 Hortatives

There are two hortative suffixes in the language: the singular hortative *-ta* (§6.3.5.0.3.1) and the plural hortative *-keq* (§6.3.5.0.3.2).

6.3.5.0.3.1 Singular hortative *-ta*

The singular hortative suffix *-ta* is related to the homophonous advisative modal particle *ta* (§9.2.9). The singular hortative occurs in an imperative clause and renders it a (singular) hortative, a sort of advice or instruction for the hearer (*you better do X*). It occurs in a slot position right before the imperative suffix. In Example 275, the predicate derivation *-ta* occurs with an imperative clause.

(275) *di-ta-na*

sit-ADVS-IMP

‘It is better you sit!’ [221024_JO_eli1 45]

6.3.5.0.3.2 Hortative *-kéq*

The plural hortative mood is marked by the suffix *-kéq*. It is used to encourage a group (1st person plural inclusive) to act, often as a polite invitation ‘let us drink!’. In Example 276, the villagers agree to go look for somebody. In Example 277, the hortative has the effect of politely suggesting the hearer to settle somewhere.

(276) “chôq! hwénaringbôkéq” la
chôq hwé-na-ring-bô-kéq la
 let’s! look-FACT-PL.II-and.go-HORT.PL say
 ‘(They) said: “Alright! Let’s go look (for him)!”’ [240122_kutyemba_phla 494
 902]

(277) (...) “ekē ìhakéq” la gũ (...)

e = kē ì-ha-kéq la gũ

 DEM = LOC be.there-at.last-HORT.PL say NF

 ‘(...) saying “let’s finally settle here” (...)’ [TA24_land 215 386]

6.3.5.0.4 Jussive -tyêyu

The jussive mood expresses a desire or wish for a third person from the side of the speaker(s). It is similar to the English ‘may/let s/he ...’. The jussive is marked by the compound suffix *-tyêyu-keq*, where the *-keq* segment goes back to the hortative suffix (cf. 6.3.5.0.3). In Examples 279 and 278, the hero of the story expresses his wish for the tiger to eat raw meat and wilden.

(278) (...) abang bônatyêyukeq püttyô shyânaya künam
[habang bô-na-tyêyu-keq [püttyô shyâ-na-ya künam]_{sub.cl}]_{jus.cl}
 forest go-back-JUS-HORT.PL different become-FACT-ALLO PRP
 ‘Let it (the tiger) go (to) the forest, in order to become different!’ [240122_tiger
 131 288]

(279) “cya asángra thâtyêyukeq” we lana matyô
cya asángra thâ-tyêyu-keq we la-na matyô
 3 raw eat-JUS-HORT.PL DEM say-FACT maybe
 ‘Maybe he said: “let it (the tiger) eat raw (meat)!”’ [240122_tiger 134 298]

6.3.6 Other TAMN suffixes

The group ‘other suffixes’ includes two suffixes that cannot be assigned to any other group. These are the allophoric suffix *ya* (§6.3.6.1) and the factitive suffix *-na* (§6.3.6.2).

6.3.6.1 Allophoric suffix *-ya*

The allophoric suffix *-ya* is not sensitive to tense or aspect; it merely indicates an allophoric controller of the verb (cf. §8.2.1). In Example 280, the speaker asks for confirmation about an event of the past. We only know from context that this refers to the past. In Example 281, the bird went and shot the evil eagle. A past tense interpretation springs from the context and the directional *-ge*, which has a telic reading (cf. §6.3.3.2).

(280) *tayâng kalà làng bârěya, na?*

tayâng kalà làng bâ-ring-ya na

a.clan officer COLL make-PL.II-ALLO INSIST

‘They (the Britishers) made (i.e. appointed) the Tayang (as) officers, right?’

[TA22_clan 34 115]

(281) *câmbriyǎ́ ne ôgeya*

câmbri-yǎ́ ne ô-ge-ya

a.bird-FEM only shoot-and.come.back-ALLO

‘Only the *chambri* bird could kill it (lit. shoot it down and come back).’ [TA22_origin 190 521]

In Example 282, the allophoric suffix *-ya* stands for an ongoing, present event, based on the context. Similarly, in Example 283 the allophoric is employed in direct speech.

(282) *rûtyu welibogũ thîya kû shyânaya*

rûtyu welibo gũ thîya kû shyâ-na-ya

up.there and.then NF mountain plateau happen-FACT-ALLO

‘Up there, I mean, (at) the high plateau it’s going on.’ [TA22_chicken_thief 68 244]

- (283) *ewe ble e meya ja shyâya matyô*
we ble we me-ya ja shyâ-ya matyô
 DEM pig DEM rear-*allo* what happen-ALLO maybe
 ‘(We) might not rear any pigs.’ [TA22_kutyemba_phla 51 153]

The allophoric suffix is associated with a 3rd person S/A as well as with a 1st and 2nd person in non-volitional events (§6.3.1.1). In Example 284, the verb takes an allophoric suffix for the verb *khamyû* ‘get angry’ is non-volitional.

- (284) *khamyûya kya?*
khamyû-ya kya
 get.angry-ALLO Q
 ‘Are you angry?’ [150923_STBT_eli1 487]

In Example 285, the purposive construction with the particle *künüm* requires the allophoric *-ya* suffix of the main verb, as discussed in §10.4.2.2.

- (285) *asan.repho bü asang kaprêi khìge rôya künüm la*
asan.repho bü kasang kaprêi khì-ge [rô-ya
 Assam.rifles(<E) CLF:times three four reach-and.back catch-ALLO
*künüm]*_{sub.cl} *la*
 PRP HSAY
 ‘The Assam Rifles went to arrest him three-four times, they say.’ [TA22_CLAN2 75 151]

6.3.6.2 Factitive *-na*

The term *factitive* is understood here as a grammatical category that indicates the speaker’s confidence that the predicate is true, knowledgeable and relevant. As such, it is the speaker’s best guess about his/her statement (Post 2007, 636) or a *fact* (cf. Lustig 2010, 79; Breugel 2014, 427f) - hence the term *factitive*. The pragmatic effect of the factitive can be expanded to include requests for confirmation (*is it a fact that X?*) as

well as assumptions about the other speech act participant's missing information. In an extended sense, interrogative clauses with factitive suffixes are often understood as referring to a relevant and known implication, e.g. *at what time will you be here (since I am cooking for you)*. This category is expressed in Tawrã by the verbal suffix *-na*, which is common in natural speech while rare in stories or other forms of monologue. This suffix is likely related to Macro-Tani declarative clause-final particle *na*, e.g. Galo (Post 2007, 636), Tangam (Post 2017, 116f) or Milang *ni* (Modi 2017, 383). The rest of this section will analyze examples where this suffix is employed. The choice of this label is not unproblematic and there is a discussion at the end of this section that delves into this matter (§6.3.6.2.1).

As hinted above, the main function of factitive *-na* is to assert the speaker's confidence in his/her assertion. This would include unquestioned facts or something that stands beyond the speaker's own biased appreciation - or at least the speaker might want to portray it in this way. In Example 286, a speaker explains the genealogy of his clan; the use of the factitive presents those events as truthful, unquestionable or self-evident. Similarly, in Example 287 a speaker narrates an episode when he asked his clansmen why their numbers hadn't increased. The use of the factitive in a reported question indicates the attitude of the speaker towards the information (their meager thrive), i.e. this information is *given* or unquestionable.

(286) dyâze shyâna tashîpu shyâna

[dyâze shyâ-na]_{main.cl} [tashîpu shyâ-na]_{main.cl}

a.clan become-FACT a.clan become-FACT

'They became the Dyaze and Tashipu clans.' [TA23_dimha_clan 11 65]

(287) tshûnem nyu kadé.nyu shyâna hâ mu ning ahuge

[tshû-na = m nyu kadé.nyu shyâ-na]_{compl.cl} hâ mu ning ahu-ge

thrive-FACT=NEG FOC why happen-FACT AFF also 1P ask-go.and.back

"How come you haven't expanded (in numbers)?" we went to ask (them).'

[TA23_talyu_clan 35 216]

(288) (...) tya ning we abâpâ ìnadi gũ (...)
 tya ning we abâ-pâ ì-na-di gũ
 ancestor 1P DEM type.of.bamboo-place.of be.there-FACT-HAB.EGO NF
 ‘(...) while our ancestors stayed in a place of bamboos (...)’ [TA23 CLAN2 13 24]

- There are two morphosyntactic settings where the factitive occurs particularly often in my corpus of Tawrã: in a negated clause and along the stative marker. One way to think about this is that the confidence that a speaker places upon a statement by employing the factitive can be taken as a sign of its relevance in a given context. As such, the coupling of factitive and negation is only natural, because an important reason to utter a negative clause is to evoke “the positive conception of what is being denied” (Langacker 2008, 59). As such, a negative clause is pragmatically marked as a relevant fact in relation to a hypothetical, positive development. In Example 290, the speaker employs the factitive on a negated predicate, as he can state with confidence that the agreement was not found. This is relevant in the context, for a written document would have helped the speaker in his case. Similarly, in Example 291a we learn that a mythical figure could not be killed. Thereafter, in Example 291b, the speaker employs a factitive on the negated first predicate to logically connect it with the preceding clause: they tried

to kill him eight times but he still refused to die. Apart from coding a type of knowledge upon which the speaker may rely (i.e. clan genealogies), it highlights the truthfulness of an otherwise unlikely development.

(290) we agrimen nyu ãna yüm (...)

we agrimen nyu ã-na yüm

DEM agreement(<E) FOC have-FACT NEG

‘As of the agreement, we don’t have it (...)’ [TA24_land 43 69]

(291) a. “lêlyâ yüm tapwi” la gũ bü lyûm sênâ mu

[[lê-lyâ yüm tapwi]_{compl.cl} la gũ]_{cosub.cl} bü lyûm sê-nâ

resuscitate-BACK NEG lest say NF CLF:times eight kill-AGAIN

mu]_{main.cl}

also

‘Saying “lest he would resuscitate” they killed him 8 times.’

b. shêina yüm webo gũ mîya we ahumazâigabô khrâk (...)

[shêi-na yüm]_{main.cl} [webo gũ mîya we ahu-mazâi-ga-bô

die-FACT NEG and.then INS wife DEM ask-force-PL.I-PFV.ALLO

khrâk]_{main.cl}

EVI

‘He would not die and then (they) compelled (her) wife to speak (...)’ [both TA23_tayang_clan 32 193]

A further morphosyntactic instance where the factitive is also common is coupled with the stative marker *-ho* (§6.3.4.5). This can be explained in that the stative is often employed to describe situations that have come to be in a certain way. Example 292 is from a storybook where speakers were asked to describe what they saw: the state in which an elk got away. The use of the factitive here indicates the speakers’ attitude towards what they are describing: it is their best guess given what they have seen. A further case is Example 293, where the speaker recounts a story that happened to his

grandfather. He employs the *-ho-na* (-STAT-FACT) combination when describing a burial monument that was in place at the time of the story. This example is noteworthy, as the speaker was not alive at the time of the narrated events and he could not have seen the burial monument, as it is typically burnt after 3-4 days; the confidence in the truthfulness of the statement originates from his clan membership, a point that is further discussed in §6.3.1.1.

- (292) *kadyêsyem lîhona*
kadyêsyem lî-ho-na
 hardly escape-STAT-FACT
 ‘The elk hardly got away.’ [TA22_frog1 117 553]

- (293) *brũ bâhona âm*
brũ bâ-ho-na âm
 burial.monument build-STAT-FACT AFF
 ‘There was a burial monument in place, right?’ [TA22_CLAN2 120 228]

As hinted above, the semantics of the factitive can be exploited for other pragmatical ends. One use of the factitive is to supplement the statement of another speech act participant; a speaker might use the factitive when intervening or completing a sentence or when asking for clarification. Therefore, by framing a statement as a fact, one (tacitly) asks for the approval of the supplemented information. In Example 294a, speaker A recounts the fate of a man (here termed Z.) whose wife died after labor. Speaker B assumes that Z. remarried after the passing of his wife and thus completes the utterance in Example 294b. The use of the factitive attempts to insert that new piece of information as part of the logical narrative, which is informed both by the context and shared knowledge about the people involved and customs (in this case, that men may remarry after the passing of their spouse).

- (294) a. we me mēhōli gũ hîq nyu shêiatîbô xxx nyâra
 [we me mē-ho-li gũ hîq nyu]_{sub.cl} shêi-atî-bô]_{main.cl}
 DEM FS give.birth-STATE-REC.PST NF EMPH TOP die-be.lost-PFV.ALLO
 [xxx nyâra]_{postposed.main.cl}
 (inaudible) sickness
 A: ‘So she died after giving birth (...), (from) a sickness.’

- b. Z. harûmna
 Z. harûm-na
 a.name remarry-FACT
 B: ‘So Z. remarried.’ [both TA24_land 172 315]

Similarly, in the following examples one speaker infers that the other is not aware of some information: the fact that person X. is dead. In Example 295a, speaker A asks about X. and speaker B provides information on his identity in Example 295b. Furthermore, in Example 295d and after two repeated questions by speaker A, speaker B uses the factitive suffix to inform A that the person in question is dead. By framing it as a fact, speaker B is assuming that speaker A is unaware of this relevant information.

- (295) a. X. shyā?
 X. shyā
 a.name who
 A: ‘X. who?’
- b. K.
 K.
 a.clan
 B: ‘K.’ (i.e. X. K., a name)

c. shya mu? ekhalâi lade we á?

shya mu? e = khalâi la-de we á?
 who also DEM = field say-HAB.EGO DEM Q

A: ‘Who? Are we talking about this plot?’

d. ewe tya ehm (...) shêina we X. we (...)

ewe tya ehm [[shêi-na]_{rel.cl} we X. we]_{NP}
 DEM now ehm die-FACT DEM a.name DEM

B: ‘This now ehm (...) this X., who died, he (...)’ [TA24_land 26 46]

A further pragmatic extension of the factitive conveys the marked predicate as relevant or pertinent in a (shared) context, particularly in interrogative clauses. The factitive thus relates the predicate to an unstated but contingent issue: *why/when/etc. are you doing X (because of Y)*. In Example 296, two people discuss a land issue. One asks the other about the ownership of the land. The factitive frames the question in the larger context: *(since we are talking about this) under whose name is it?*

(296) e nyu tyá amang we shya (...) we shya klem jina? we cuwe jaga tya shy shyâ amang tshôna?

e nyu tyá amang we shya we shya klem ji-na
 DEM 2S grandfather.MOD name DEM who DEM who claim(<E) do-FACT
we cuwe jaga tya shy shy amang tshô-na
 DEM DEM place(<H) now FS who name be.kept-FACT

A: ‘(Under) your father’s name who, who is making the claim? (And) this place, (under) whose name is it now?’ [TA24_land 19 30]

In certain cases, after a speaker frames a statement or a question with the factitive, the response from the interpellated person can be subtle or at least less determined, since s/he might feel interpellated. In Example 297a, speaker A utters a statement based on the previous discussion, assuming that speaker B did not comply with sending the due notification. Although it is not constructed as a question (cf. §7.3.1), the illocutionary force of the factitive is that of a question by requesting a confirmation (*is it true that you*

didn't X?). The lack of a response from B confirms A's assumption. Thereafter, A tells B what he should do next (Example 297c).

(297) a. A: *we bra nôtis hànglyânam*

we bra nôtis hàng-lyâ-na = m

DEM CLF:generic notice(<E) give-BACK-FACT = NEG

A: 'So you didn't give (them) a notice again.'

b. B: ...

B: ...

c. A: *elüm hàngya shyâbô matyô*

e = lüm hàng-ya shyâ-bô matyô

DEM = these.days give-ALLO have.to-PFV.ALLO maybe

A: 'You'll have to give (them) (a notice) these days.' [both TA24_land 48 79]

Finally, the factitive occurs in lexicalized, formulaic expressions like *pra bô-na-na* (good go-FACT-IMP) 'farewell (for someone departing)' or *pra ì-na-na* (good stay-FACT-IMP) 'farewell (for someone staying)'.

In sum, the factitive is employed to express the speaker's confidence in a statement; it makes it a fact, or the speaker's best guess. This semantic feature is also exploited to intervene in a statement and indirectly ask for confirmation regarding its truthfulness (*Peter went to ... - The park, right?*). Furthermore, in a question, it may also relate the predicate to a larger context. The confirmation is, however, often tacit, as the new information is seamlessly integrated into the conversation. Finally, it is also employed in formulaic expressions of farewell.

6.3.6.2.1 The term *factitive* and its use in related languages

The term *factitive* has been employed for related languages to indicate eclectic categories that share similar semantics, as discussed for Tawrã in §6.3.6.2. This section will

address both the epistemic value of the factitive as well as discuss further grammaticalization processes found in related languages.

In his description of Atong, Breugel (cf. 2014, 427f) describes the function of the *factitive* suffix *-wa* on a main clause as a marker of reification (*Verdinglichung*), i.e. as a *fact*, something that *is the case* (idem.). He further describes it as a form of *epistemic modality*, a term he borrows from Lyons (1977, 797), which refers to a speaker's explicit commitment to the truthfulness of his/her proposition, as opposed to an *unqualified assertion* (idem.), which lays claim to truthfulness merely to the extent of being an illocutionary act of assertion. In a slightly different approach, Post (2007, 636f) discusses what he calls the *simple declarative* clause-final particle *na* in Galo, which has the value of 'I'm telling you this' (idem. 636). It thus seems to both communicate the speaker's own commitment to the truthfulness of his/her assertion, as well as to 'enhance the addressee's confidence in the speaker's knowledge accordingly' (idem.). Crucial for the discussion of Tawra is that both authors recognize that the locus of knowledge that defines this category is the speaker's assertion to the hearer, as opposed to being a fact in the world. In this line, and for the sake of terminological clearness, I believe it is important to circumscribe the term *factitive* to that definition because a plethora of vaguely related definitions for the same term will lead to *factitive* being devoided of meaning (cf. Zeisler 2024, 57). However, these categories are undergoing different phases of grammaticalization and this is bound to blur a clear picture of what the *factitive* is.

In several related languages, the marker of factitive is isomorphic with a subordinator/nominalizer. Such is the case in Lepcha (cf. Plaisier 2006, 212f), Zaiwa (cf. Lustig 2010, 68, 79) and Atong (cf. Breugel 2014, 427). A further, potentially related category is the article *-n(a)* found in some Kiranti languages (cf. Watters 2008, 2f). As described by Bickel (1999a), the Athpare article *-n* (cf. Belhare *-na*) is the only nominalizer found in the language and it also marks specificity and definiteness. Crucially, a function of nominalization (and this also applies to many Kiranti languages with different nominalizing devices) is to reinstate the proposition as a given fact, as something "that is the case" (idem., 287). There is thus a structural opposition between nominalized *this-is-the-case* clauses and non-nominalized clauses with a weaker epistemic stance. This meaning is similar to how the factitive has been described in the literature and how it is employed

in this grammar. Lastly, the Kiranti *-n(a)* article could be related to the Tawrã factitive *-na* as well as the Tani clause-final particle *na*. However, exploring this connection exceeds the limits of this grammar.

6.4 Negation

The slot for negation comes after TAMN suffixes, as discussed in §6.3.1.3, and it can be filled by one of two negators: *yüm* and *shyüw*. While the former is more common and can also negate verbless clauses, the latter has a meaning circumscribed to personal inferences and observation. §6.4.3 also discusses the obsolesced negator *-ma*. Although not productive suffix, it is of historical interest. The prohibitive has been discussed in §6.3.5.0.2.

Tawrã shares the feature of having a post-head negator, rather than pre-head element, with the Tani languages and Boro-Garo (cf. Post 2015b, 431). Conversely, a pre-head negation is widespread among many other Tibeto-Burman languages, including the neighboring language Kaman.

6.4.1 General negator *yüm*

The marker *yüm*, and its cliticized form *=(e)m*, marks negative polarity for verbs and verbless clauses (§7.1). The negator occurs with interrogative pronouns to create negative pronouns like ‘nothing’, ‘never’, etc. These forms are discussed in §3.5.1.2. Negation in multi-clause constructions is discussed in §10.3.3.

The origin of the general negative marker in Tawrã possibly goes back to a collocation of a defunct copula **yu* and a negative suffix **-m(a)* (for an equivalent analysis for Kuki-Chin languages, cf. DeLancey 2015). While this copula does not have an obvious descendant in the modern language, the negator is likely related to the widespread PTB **ma* etymon (cf. Matisoff 2003, 488), cf. §6.4.3. The negator *yüm* does not seem to be related to the prohibitive suffix *-yâq* (cf. §6.3.5.0.2).

The cliticized form of the negator *=(e)m* alternates freely with the full negator. The cliticized form also occurs in the lexicalized interjection *shyâm* ‘no!’ (< *shyâ* ‘happen, become’ plus negative *=m*).

In Example 298, both verbs are negated. As with the bare stem, a temporal interpretation is dependent on the context. In Example 299, the speaker is talking about the place where a thief in a story used to take stolen goods. The interviewer suggests a place (Anini) that is promptly discarded by the speaker (*it is not X*). Existential verbs are commonly overridden by the negative marker, as will be mentioned below.

- (298) ama byẽ khì-câng yüm ama hana câng yüm, âm
 ama byẽ khì-câng yüm ama hana-câng yüm âm
 here TOP reach-be.able NEG here come-be.able NEG yes
 ‘(They) couldn’t reach here, (they) couldn’t come here, yes.’ [TA22_clan 26 93]

- (299) Anini yüm (...) ewe (...)
 Anini yüm ewe
 a.place NEG DEM
 ‘Not to Anini, this (...)’ [TA22_CLAN2 63 118]

Speakers disprefer the co-occurrence of the negator and certain suffixes and regularly drop the latter if the verb is negated: Example 300a is a positive relative clause formed with the allophoric suffix *-ya*; its negative counterpart in Example 300b simply takes the negator *yüm*, although *-ya yüm* would not be considered ungrammatical.

- (300) a. [tabrẽ thâ-ya]_{rel.cl} mề
 meat eat-ALLO person
 ‘A person who eats meat.’
 b. [tabrẽ thâ yüm]_{rel.cl} mề
 meat eat NEG person
 ‘A person who does not eat meat.’ [both 021024_STBT_eli2 290]

When TAMN suffixes do occur on a negated verb, it is pragmatically marked. In Example 301, a retelling of a storybook, the negative occurs with the stative and factitive

suffixes. This can be explained because the speaker tries to be as informative and clear as possible while retelling the story.

- (301) *cya (...) pahwã we ìhona yüm*
cya pahwã we ì-ho-na yüm
 3 frog DEM be.there-STAT-FACT NEG
 ‘He (...) the frog was not there.’ [TA22_frog1 26 129]

Example 302 illustrates the cliticized form of the negative marker =*m*. Example 303 further shows the assimilation process that regularly occurs after the habitual suffix: *de* + *yüm* > *dyüm*.

- (302) *we dâi we mu (...) maroge dyángnam mane*
we dâi we mu maro-ge dyáng-na = m mane
 DEM a.tribe DEM also talk-and.back comply-FACT = NEG FILL(<Asm)
 ‘Now this Idu (...) (they) went to talk (to him) (but) he wouldn’t agree (to come).’
 [TA22_CLAN2 87 167]

- (303) *kasâdyüm*
kasâ-d = yüm
 know-HAB.EGO = NEG
 ‘I don’t know.’ [021024_STBT_eli2 315]

As noted in §7.1, the negator also occurs in verbless clauses to negate equational clauses (X is not Y). The negator commonly occurs at the end of the clause, as in Example 304a, although it can also be placed between the topic and the focus to emphasize the negative relation, as in Example 304b. The negator is not allowed before the topic and the focus, as illustrated in Example 304c.

(304) a. [cya]_{Top} [tawrã]_{Foc} [yüm]_{neg}

3 a.tribe NEG

‘S/he is not Tawrã.’

b. [cya]_{Top} [yüm]_{neg} [tawrã]_{Foc}

3 NEG a.tribe

‘S/he is NOT Tawrã.’

c. * [yüm]_{neg} [cya]_{Top} [tawrã]_{Foc}

NEG 3 a.tribe

Intended: ‘S/he is not Tawrã.’ [all three 310823_STBT_eli1 170]

6.4.2 Observational verbal negator *shyüw*

The negator *shyüw* is related to the versatile verb *shyüw* ‘allow, appoint, let’ and the volitional causative predicate derivation *-shyüw* (cf. §6.2.2.2.2). Compared to the main negator *yüm* (§6.4.1), *shyüw* is employed to establish observations, inferences or personal opinions.

The observational negator *shyüw* is a modal particle (§9.2) based on two syntactic reasons: First, it is associated with 3rd person S/A arguments, akin to the allophoric marker *-ya* (§6.3.6.1). In Example 305, the observational negator is not allowed with a first person A argument¹⁰ (although it yields a causative interpretation). This is in line with the conflict between an epistemic inference and the *assertor* (§6.3.1.1), a feature of some modal particles (discussed in §9.2). Furthermore, the negator *shyüw* cannot co-occur with a modal particle, suggesting that they share the same epistemico-pragmatic function and syntactic position. In Example 306, *shyüw* cannot be analyzed as a negator; instead, the only interpretation is that of the homophonous causative predicate derivation.

¹⁰It remains to be studied whether the negator *shyüw* can occur with a non-assertor 1st person S/A argument, e.g. with a non-volitional verb or in a situation where someone is unaware (cf. Zeisler 2024, 22-27).

(305) *hã tapê thâ-shyüw*

1s food eat-CAUS.V

‘I allow someone to eat’ (it does not yield *‘I won’t have food’) [field notes 251024]

(306) *thâ-shyüw delé*

eat-CAUS.V it.seems

‘It seems like someone is allowing someone else to eat.’ (It cannot yield ‘It seems like s/he won’t eat’ [field notes 251024])

Even though both negators differ in their syntactic position, they cannot co-occur. In Example 307a, the observational negator occurs after its homophonous predicate derivation. This same position can be taken by the general negator *yüm*, as illustrated in Example 307b. This example also shows that a co-occurrence of both negators is not allowed. As will be discussed presently, the negator *shyüw* has carved out a specific pragmatic/modal niche within negation.

(307) a. *lâng-shyüw shyüw*

enter-CAUS.V OBS.NEG

‘S/he is not allowed to enter.’

b. *tapê thâ-shyüw yüm*

food eat-CAUS.V NEG

‘S/he does not allow him to eat food’ (but not: *‘S/he does not take food.’)
[both 241024_JO_eli3 418]

The main function of the negator *shyüw* is to establish an observation. In Example 308a the speaker employs the negator *shyüw* to establish an inference based on observation (‘s/he won’t drink beer’). Conversely, the same sentence with the general negator *yüm* can be taken as a generalization or established fact: ‘he never drinks’, e.g. a teetotaler. Similarly, in Example 309a the cook might say that the food is not as good as s/he

him/herself prepared it, or s/he might presume that the commensals might not be fond of it for some reason. The use of the general negator *yüm*, in Example 309b, would make it a general statement, not (only) based on observation (e.g. ‘the food from there is not good’).

(308) a. *yu tyûm shyüw*

beer drink OBS.NEG

‘S/he won’t drink beer (an observation).’

b. *yu tyûm yüm*

beer drink NEG

‘S/he doesn’t drink beer (i.e. never).’ [both 271024_JO_eli1 536]

(309) a. *tapê thâ pra shyüw*

food eat be.good OBS.NEG

‘The food might not be good!’

b. *tapê thâ pra yüm*

food eat be.good NEG

‘The food is not good.’ [both field notes 251024]

Finally, a further effect of the negator *shyüw* is to lessen the pragmatic implications of negation (cf. Post 2015b, 448) in a manner akin to a modal verb in English (e.g. *might not*); the speaker can draw away responsibility and blame the circumstances instead of his/her decision. In Example 310, the news that a planned trip will not take place is expressed by the use of *shyüw*. This makes it clear that the rationale behind is based on perceivable circumstances, e.g. schedule problems. Conversely, the use of the general negator could be interpreted as impolite, a decision based on one’s own thoughts.

(310) *düing bô-câng shyüw*
 a.place go-be.able OBS.NEG

‘We won’t be able to go to Tidding.’ [field notes 221024]

6.4.3 Obsolesced negator *-ma*

There is an obsolesced negator *-ma* that has been found in two fossilized forms: *bülüm-ma* (lit. eye NEG) ‘blind’¹¹ and *tamê-ma* (lit. crazy-NEG) ‘crazy’. This old marker is likely related to the widespread PTB **ma* etymon (cf. Matisoff 2003, 488) and also occurs on the general negator *yüm* (§6.4.1).

6.5 Nominal incorporation

Tawrä has a number of conventionalized Noun + Verb collocations that are here discussed as nominal incorporations. The syntactic status of these N + V combinations is, however, a problematic issue, as the language does not have nominal case marking (cf. §8.2.4) and they could either be analyzed as phrases or as a form of lexical incorporation (Type I, according to Mithun 1984, 853-584). Semantically, some of these combinations resemble compounds in the sense that they often have idiosyncratic meanings that are not a mere sum of their constituents (cf. Matthews 1991, 82f; also cf. Aikhenvald 2007, 12f). The tension between incorporation and a normal argument of the predicate is illustrated in Examples 311a and 311b. In Example 311a, both interpretations are possible: The noun forms a conceptual or otherwise conventionalized bond with the predicate for it does not necessarily mean ‘writing (in) a book’ - but that interpretation would also be there, provided that the right context is available. In Example 311b, the modification of the NP head with a demonstrative *we* breaks the conventionalized (unmarked) reading and forces a specific interpretation (i.e. this particular book). This would suggest that the N + V form a form of syntactic phrase that does not allow other constituents; only a modal particle can intervene between the N + V group. This is illustrated in Example 312,

¹¹Note the equivalent form in Kera’a (Idu) *elo-ma* (eye-NEG) ‘blind’ and Proto-Tani **mik mang* (eye NEG/lack) ‘blind’, cf. Sun (1993, 256).

where the advisative modal particle *ta* (§9.2.9) occurs between an incorporated noun and the predicate.

- (311) a. [*tako zû-na*]_{nom.incorp}
 book write-IMP
 ‘Write! (not necessarily *a* book)’
- b. *tako we zû-na*
 book DEM write-IMP
 ‘Write this book!’ [both 071023_JOTHA_eli1 1380]

- (312) *darêk ta tako ta bâna*
*darêk ta [tako ta bâ-na]*_{nom.incorp}
 directly(<E) ADVS book ADVS make-FACT
 ‘(It is better to) go directly (and) file a complaint (lit. make a book).’ [TA24_land 46 77]

Some evidence for the incorporation of a noun into the predicate phrase comes from the adverb *tyâng* ‘very’ (cf. §3.4.1), which has a pre-verbal position. In Example 313a, the adverb *tyâng* ‘very’ modifies the whole predicate phrase and it may not be placed between the incorporated noun and the predicate head, as in Example 313b.

- (313) a. *hã tyâng [nasîk tyû-ya]*_{nom.incorp}
 1S very appetite be.hungry-ALLO
 ‘I am very hungry.’
- b. **hã nasîk tyâng tyû-ya*
 1S appetite very be.hungry-ALLO
 Intended: ‘I am very hungry.’ [both 291024_STBT_eli2 842]

In some cases, N-V collocation can be thought of as being “(...) limiters of the unspecified generality of the naked verb” (Matisoff 1973a, 309; echoed by Mithun 1984, 853). This means that the noun in the N-V collocation restricts the potentially multiple interpretations of semantically broad verbs to a particular meaning that fits the context. An example is *tapè thâ* (food/rice eat) ‘have a meal’. A further case is *mîya bréi* (wife buy) meaning ‘marry (male persp.)’. In other cases, however, these forms seem to form collocations without narrowing down meaning, as some verbs only occur with an adjoined noun, e.g. *bülüm patsüng* ‘blink’ (cf. *bülüm* ‘eye’ and *patsüng* ‘blink’). Table 6.8 displays some nominal incorporations. Further lexicographic work is bound to find many more forms.

Table 6.8: Nominal incorporation

V-N form	N	V
<i>àng tâwriüw</i> ‘make a house’	<i>àng</i> ‘house’	<i>tâwriüw</i> ‘make a house’
<i>bülüm patsüng</i> ‘blink’	<i>bülüm</i> ‘eye’	<i>patsüng</i> ‘blink’
<i>habûm ko</i> ‘make a roof’	<i>habûm</i> ‘roof’	<i>ko</i> ‘make a roof’
<i>karà manyâ</i> ‘downpour’	<i>karà</i> ‘rain’	<i>manyâ</i> ‘downpour’
<i>karà tshâng</i> ‘get wet’	<i>karà</i> ‘rain’	<i>tshâng</i> ‘get wet’
<i>khalâi léi</i> ‘shake (as in an earthquake)’	<i>khalâi</i> ‘ground’	<i>léi</i> ‘shake (as in an earthquake)’
<i>khalâi tyô</i> ‘spit’	<i>khalâi</i> ‘spit (n), dirt’	<i>tyô</i> ‘spit’
<i>mîya bréi</i> ‘marry (male persp.)’	<i>mîya</i> ‘wife’	<i>bréi</i> ‘buy, marry’
<i>nâshik tyûng</i> ‘be hungry’	<i>nâshik</i> ‘appetite’	<i>tyûng</i> ‘be hungry’
<i>pachà tyé</i> ‘make fire’	<i>pachà</i> ‘flint stone’	<i>tyé</i> ‘hit, cut’
<i>tako puri</i> ‘read, study’	<i>tako</i> ‘book’	<i>puri</i> ‘study’(<IA)
<i>tako zû</i> ‘write’	<i>tako</i> ‘book’	<i>zû</i> ‘write’
<i>tapè thâ</i> ‘have a meal’	<i>tapè</i> ‘food, rice’	<i>thâ</i> ‘eat’

As discussed above, some of these N-V forms have a specific meaning and therefore restricted usage. We may assume that these N + N combinations are a hotbed for grammaticalization and reanalysis and perhaps some disyllabic verbs in the language go back

to a N-V form, e.g. *shîwtsüm* ‘sneeze (v)’, potentially from **shîw* + **tsüm* (cf. *büliüm patsüng* ‘blink’). A similar case could have brought about the *me*- root prefix, e.g. *(me)sê* ‘kill’, perhaps from **me* ‘person’ and *sê* ‘kill’ (cf. §3.7.4).

7 Clause types

This chapter explores the different clause types available in the language. Verbless clauses are discussed in §7.1. Predicate clauses are discussed in §7.2. Within predicate clauses, §7.2.1 describes their constituents (arguments and adjuncts) and §7.2.2 discusses their differences in terms of valency. Non-declarative clauses are discussed in §7.3. It includes interrogative clauses (§7.3.1) and command moods (§7.3.2), although the verbal suffixes involved in command moods have been thoroughly described in §6.3.5.

The terms *final* and *non-final* clause (also employed in §10.3) refer to the structure of co-subordination in the context of multi-clause constructions. A final clause may take one modal particle (§9.5) and does not take a non-final marker, while a *non-final* clause may not have a modal particle and must take a non-final marker (e.g. *gũ* NF).

7.1 Verbless clause

This section describes the structure of verbless clauses. The locative construction is discussed in §7.1.1.

A verbless or appositive clause commonly serves to describe, identify or indicate the attribution of a noun phrase (henceforth NP), as opposed to verbal (transitive and intransitive) clauses which refer to actions¹. Verbless clauses do not have a predicate in their basic structure². A pragmatically-oriented way to describe the verbless clause is in terms of topic and focus, as discussed in §8.3. In this case, the *topic* corresponds to the

¹Verbless clauses functionally resemble copular clauses in related languages (for Rawang cf. LaPolla 2010; on Dhimal cf. King 2009, 106f).

²Only verbless clauses lack a predicate. Otherwise, the omission of the predicate is not widespread in the language. Neighboring Hindi and Nepali may omit the verbal predicate in routinely occurring questions, as in Example 314. Equivalent pragmatically driven omissions of the predicate in Tawra have not been attested so far.

NP that is being described, since it is definite, topic and given (the “center of attention”, cf. Li and Thompson 1976, 464). Conversely, the *focus* is the missing information, the attribute of the topic. The basic constituents of a verbless clause are illustrated in Table 7.1, including negation (in parenthesis). Structurally, a topic NP is *apposed* to a focus NP (complement), which may consist of another noun, a proper name or an adjective. Note that adjectives have structural properties that make them resemble nouns, as discussed in §3.3.3. The topic has several properties that make it prominent in discourse (e.g. known, given, topical) and therefore susceptible to being ellipsed, as it is understood from the context. As mentioned above, and differently from all other clause types, there is no overt verbal predicate between these two elements (hence *verbless*) (cf. Dixon 2022, 347f; cf. Post 2008, 131f.) like *copulas* in Tawrã. Topicalization within a verbless clause is further discussed in §8.3.1.

Table 7.1: Structure of a verbless clause

(Topic)	Focus	(negation)
(NP)	NP	(NEG)

A note on terminology is in order before we proceed to the examples. The use of the term *topic* for verbless clauses corresponds to ‘*verbless clause* subject’ in the description of other languages (cf. Dixon 2022, 347f. also cf. ‘copular subject’, which seems to be equivalent in syntactical terms, cf. Dixon 2006b, 8; Post 2007, 420f; Bentley 2016). This term is not employed here, as *Subject* is restricted to its semantic and grammatical use: A Subject as a semantic role is defined based on a verb’s argument structure (cf Andrews 2007a, 135). Therefore, there can hardly be a semantic Subject without a verb. Furthermore, there is no evidence of a strong grammaticalization of S/A arguments that would justify the use of the term Subject, as discussed in §8. The term Subject, however, may

-
- (314) Nepali (personal knowledge)
kahã (*jāne*)?
 kahã jā-ne
 where go-OPT
 ‘Where (are you going)?’

still capture the semantic roles motivated by the construction itself and the constituents of a verbless clause that “evoke semantic frames within which participant roles are construed, along with semantic roles” (Västi and Kittila 2014, 540f.). However, the topic in a verbless clause and the S/A argument differ in syntax and semantics (cf. Bentley 2016): the topic in a verbless clause cannot be omitted nor can it be relativized (through unmarked nominalization, cf. §10.5.2), as opposed to the S/A argument in other clause types.

Example 315 is an attributive clause in which the identity of the referent is expressed. As mentioned above, one of the main uses of a verbless clause is to identify the topic and establish an equation between the topic and the focus: X is Y. In this sense, it can also be employed in questions as interrogative appositive clauses (cf. Modi 2017, 334) (also cf. §7.3.1), as in Example 316. The constituents of a verbless clause can still be modified by NP-modifiers (§9.5), as in Example 317.

(315) *ewe pyâ*

[ewe]_{top} [pyâ]_{foc}

DEM bird

‘This is a bird.’ [161222_ST_BT_txt_004]

(316) *shya cya?*

[shya]_{top} [cya]_{foc}

who 3

‘Who is he?’ [230823_STBT_eli1]

(317) *khûnyüm dahi mề (...) dahi mề, na?*

[khûnyüm]_{top} dahi [mề]_{foc} dahi [mề]_{foc} na

god like person like person INSIST

‘Gods (were) like people, like people, right?’ [TA22_clan 20 78]

The act of naming is a common function of verbless clauses (or copula clauses in other languages, cf. LaPolla 2010, 1067). In Example 318, the focus (here a verbless complement) is a proper name.

- (318) we amang we kutyemba
 [we amang we]_{top} [kutyemba]_{foc}
 DEM name DEM a.name
 ‘His name was Kutyemba.’ [TA22_SOstory_02 26]

In Example 319, the topic is a NP consisting of a noun while the focus is an adjective. Adjectives are attributes of the topic and do not take any form of derivation (cf. §3.3.3).

- (319) tabrẽ hwâng.ra
 [tabrẽ]_{top} [hwâng.ra]_{foc}
 meat raw
 ‘The meat is raw.’ [field notes 25102024]

Verbless clauses cannot specify their tense and aspect through verbal suffixes and, therefore, the temporal interpretation of equational relations is bound to contextual information (also cf. §6.1.1.1). Thus, Example 320 can be ambiguous, as it may refer to both a non-past and past state of affairs. Further temporal precision can be achieved, if required, through temporal adverbs (e.g. ‘now’, ‘later’, etc.)

- (320) cya aphisî tîya
 [cya]_{top} [aphisî tîya]_{foc}
 3 office(<E) boss
 ‘S/he is/was the head of the office.’ [field notes 25102024]

As mentioned above, the topic is known and given and therefore commonly omitted in naturalistic speech. This is illustrated in the following examples: In Example 321a, the speaker’s grandfather is introduced. In Example 321b, which is a comment on Example 321a, the topic is omitted, as it is reconstructable from the context.

- (321) a. *webo ning natya we mane gôbmen cya mane ahwê mane ne*
webo ning na-tya we mane gôbmen cya
 and.then 1P 1S.POSS-grandfather DEM FILL(<Asm) government(<E) 3
mane ahwê mane ne
 FILL(<Asm) send FILL(<Asm) INSIST
 ‘Our grandfather ehm, the government sent him, I mean, right?’ [TA22_CLAN2 81 162]
- b. *âm, sokarêi mề nyu ein (...)*
âm [∅]_{top} [sokarêi mề nyu]_{foc} ein
 yes governmental(<IA>) person FOC FS
 ‘Yes, he was a man from the government (...)’ [TA22_CLAN2 83 166]

Verbless clauses allow pragmatically-driven apposition, e.g. to highlight one particular referent over other options. In Example 322, we learn that the topic (here *cya* ‘3’) is Tawrã and not *pâlyu* ‘non-tribal’. In this case, a topical marker is not employed (cf. §8.3).

- (322) *tawrã cya, pâlyu yüm*
[tawrã]_{foc} [cya]_{top} VLC [pâlyu yüm]_{VLC}
 a.tribe 3 non.tribal NEG
 ‘S/he is Tawrã, not *pâlyu*.’

The general negator *yüm* is employed for both verbal and verbless clauses (cf. §6.4.1). In both cases, the slot for negation comes after TAM and Number suffixes, as has been discussed in §6.3.1.3. §6.4.1 further discusses a pragmatically-driven apposition in which the negator occurs between the topic and the predicate. Example 323 illustrates a negated verbless clause.

- (323) *hã kapa yüm*
[hã]_{top} [kapa]_{foc} yüm
 1S fool NEG
 ‘I am not a fool.’ [180122_AM_eli 919]

7.1.1 Locative construction

A locative clause includes two NPs. The first NP corresponds to the topic of the clause and the second NP is understood to indicate a location, which may or may not take a locative marker. The structure of this construction is described in Figure 7.2. The NP in the focus slot is semantically specified, as it iconically refers to locational arguments. These arguments are generally unmarked in the language and treated as extended core arguments (E), cf. §7.2.1.1). A predicate is not required in this construction but, differently from the verbless clause described in §7.1, an existential verb may be added for a pragmatic effect, as will be discussed below.

Table 7.2: Locative clause

Topic	Focus	(V)	(Neg)
NP	location NP	(existential V)	(yüm)

Example 324 corresponds to a locative construction. No predicate is required in this case. The NP in the focus slot is a canonical locative NP, in this case, a place. Similarly, in Example 325 an interrogative clause follows the same structure. Note that there are no locative markers on the locative NPs (cf. §9.5).

(324) *hã àng Sunpura*

[hã àng]_{top} [Sunpura]_{foc}

1s house a.place

‘My house is in Sunpura.’ [230122_JW_eli 82]

(325) *thîya nyu bôya we tya hanûk mlă?*

[[thîya nyu bô-ya]_{rel.cl} we]_{top} tya hanûk [mlă]_{foc}]_{main.cl}

mountain FOC go-ALLO DEM now where village

‘The one who went to the mountain, where about is he now (lit. where place/village)?’

[TA23_Talyu_clan 18 106]

As hinted above, the locative clause can take a predicate when it is pragmatically marked or in non-canonical situations (cf. Boro 2017, 424). In Example 326, there is an existential predicate in the locative clause. This is employed as an emphasis.

- (326) *tuluqa we kē ning we shyā mu ìna yüm lê atî-kũ lêtakũhoring*
 [[*tuluqa we kē*]_{foc} [[*ning we*]_{top} [*shyā mu*]_{top} [*ì-na* *yüm*]_{pred}]_{main.cl}
 a.place DEM LOC.INS 1P DEM who also be.there-FACT NEG
 [[*lê atî-kũ*]_{main.cl} [[*lê-ta-kũ-ho-ring*]_{main.cl}
 exit leave-COMPL.II exit-permanently-COMPL.II-STATE-PL.II
 ‘At Tuluqa there is one. All left, they all left for good.’ [TA23_talyu_clan_045
 281]

Uncommon or odd scenarios also seem to trigger a locative verb in locative constructions: Example 327 describes an image from a storybook (Rodriguez 2008) in which a set of animals, arranged one atop the other, chase a fox that stole a hen. Probably the odd formation of the protagonists triggered the use of an otherwise uncommon verb in a locative construction (note the change in constituent order: Focus-Topic-V). Similarly, in Example 328 we learn about the fate of four brothers (only the last two have been included for the sake of clarity), the last of which settled (locative verb) in Tuluqa. The speaker might have uttered this verb, which is commonly ellipsed in locative constructions, in analogy to the other sentences (i.e. ‘one went, one stayed’).

- (327) *tamh we kyã mambaw ìho weli kyã tyû ìho*
 [[*tahum we kyã*]_{foc} [[*mambaw*]_{top} [*ì-ho*]_v]_{main.cl} [[*weli* *kyã*]_{foc}
 bear DEM LOC.ON fox be.there-STAT DISC.DEIC.DEM LOC.ON
 [*tyû*]_{top} [*ì-ho*]_v]_{main.cl}
 chicken be.there-STAT
 ‘Atop the bear is the fox and atop it the chicken.’ [TA22_chicken_thief_30 114]

- (328) (...) bra gê thîya bô gũ bô bra gê we tuluqa kē ìya la
 [[bra gê]_{top} [thîya]_{foc} [bô gũ bô]_{co/sub.cl} [bra gê we]_{top} [tuluqa
 CLF:generic one mountain go NF TEMP.SUB CLF:generic one DEM a.place
 kē]_{foc} ì-ya la]_{main.cl}
 LOC.INS be.there-ALLO HSAY
 ‘(...) one went (to) the mountains, one stayed in Tuluqa.’ [TA23_talyu_clan_017
 96]

The locative construction is negated with the standard negator *yüm*, as do all clauses (§6.4.1). Example 329 is a negated locative clause without a locative predicate.

- (329) cya Sunpura kē yüm
 [cya]_{top} [Sunpura kē]_{foc} yüm
 3 a.place LOC.INS NEG
 ‘He is not in Sunpura.’ [310823_STBT_eli1 133]

Note, however, that when the negated predicate is fronted to the left of the clause, the locative predicate cannot be ellipsed. Example 330 is a negated locative construction with a pragmatically marked configuration: the negated predicate is fronted to the left of the clause. The predicate cannot be omitted in that position.

- (330) *(i) yüm àng kē
 *(i) yüm [∅]_{top} [àng kē]_{foc}
 be.there NEG house LOC.INS
 ‘He is NOT at home.’ [060923_STBT_eli1 185]

7.2 Predicate clauses

A predicate clause in Tawrã is headed by an obligatory verbal predicate, as opposed to a verbless clause, which does not take a verbal complement (cf. §7.1). Arguments commonly precede the predicate, although word order is not fixed in Tawrã but is rather

subject to pragmatic considerations (cf. §8.3). Apart from the obligatory predicate, a predicate clause includes one or two core arguments (§7.2.1.1), the number depending on the argument structure of the predicate, and optional oblique adjuncts (§7.2.1.2). §7.2.3 discusses possession, a non-prototypical function of predicate clauses.

7.2.1 Constituents of a predicate clause

The predicate complex has been described in §6. This section is merely instrumental to understanding clause types in the language.

7.2.1.1 Arguments

Arguments are the main syntactic entities to carry out the action described by the head of a predicate clause. Arguments and their semantic roles are thus determined by the argument structure of the predicate, i.e. they fill up the roles determined by the predicate (e.g. *killer* and *killed* by the verb *kill*) (Andrews 2007a, 135). In Tawrã, there is no evidence for a syntactical integration of arguments into the clause in the form of grammatical relations (for Subject or Object), as discussed in §8.2.

Arguments are often omitted if they are recoverable from the context. This concerns most often S/A (cf. Lambrecht 1994, 118). In Example 331, a tiger competes with a man on counting fish. Both the A argument of the cosubordinate clause (a tiger) as well as the A argument of the main clause (the fish) are omitted as they are understood from the course of the story. Even if omitted, they are expressed in the semantic interpretation of the clause.

(331) “ekhing gaqîng” la tsâitsâi we la gũthĩ tshâw hamâmâ matyô

[*e = khing kaqîng la tsâi~tsâi we la gũ-thĩ*]_{cosub.cl} *tshâk*

DEM = one two say count~PLAC DEM say NF-CONC move.fast

hamâ-mâ matyô]_{main.cl}

get.lost-ITER maybe

‘(The tiger) counted and counted, saying “one two” (i.e. not more), (but the fish) kept moving and disappearing, maybe.’ [TA22_tiger 35 102]

It should be noted that elicited sentences commonly have a higher percentage of overt argument NPs and this is, in many cases, an artifact of a semi-controlled environment without large context, as informants attempt to be fully explicit in order to avoid ambiguity. In natural speech, this is only the case at the beginning of a narrative text or when there is a high referential competition and thus a high chance of ambiguity. Otherwise, naturalistic speech in Tawrã makes little use of overt NPs (also termed low *referential density*, cf. Bickel 2003), particularly with S and A roles (cf. transitive clauses §3.3.2.1).

Both transitive and intransitive predicates may take an extra, unmarked goal/source argument termed *extended core argument* or E. The language does not distinguish consistently between goal and source arguments. In Example 332, *hã* ‘1s’ is the core S argument of the intransitive predicate *bô* ‘go’, *gari gũ* ‘by car’ is an optional adjunct oblique and *tajap* ‘a city’ is an unmarked extended goal argument (E). In Example §333, a transitive predicate does not have any overt core arguments save for the unmarked extended core argument *tajap*, which is the semantic source. Note that *nyu* is a focus morpheme (§8.3.2) and does not encode source or goal.

- (332) *hã tajap gari gũ bôtapwide*
 [hã]_S [tajap]_E [gari gũ]_{Adju} *bô-tapwi-de*
 1s a.city car(<IA) INS go-HAB-HAB.EGO
 ‘I usually go to Tezu by car.’ [field notes 20102024]

- (333) *tsâpakwa nyu ja simplang glâiringzâ mane na*
 [tsâpakwa]_E *nyu ja simplang glâi-ring-zâ mane na*
 a.place FOC what FILL carry-PL.II-DIR FILL(<Asm) INSIST
 ‘(They) carried, what to say, (things) from Chapakhowa, I mean, right.’ [TA22_CLAN2 43 84]

7.2.1.2 Adjuncts

Differently from arguments, Adjunct NPs are optional, although they depend on the semantics of the verb and the general sense of the clause (cf. Andrews 2007a, 157f). As

optional elements of the clause, they are not retrievable if not overtly expressed. Obliques are marked by several particles depending on the semantic relation and purpose of the oblique. Table 7.3 displays all particles that occur with adjuncts. Most of these particles are postpositions (cf. §9.5).

Table 7.3: Particles that occur with adjuncts

Postpositions	Reference	Notes
General locative/instrumental <i>gũ</i>	§9.5.1	also cosubordination marker also cotemporality marker
‘inside’ locative <i>kẽ</i>	§9.5.2	
‘atop’ locative <i>kyã</i>	§9.5.3	
‘with’ <i>dũ</i>	§9.5.4	
‘for’ <i>(a)ba</i>	§9.5.5	
Purposive <i>kanüm</i>	§10.4.2.2	also a noun ‘reason’

In Example 334, the postposition *ba* expresses a beneficiary of the action, i.e. someone for whom the action is being carried out (cf. §6.2.2.3). In Example 335, the particle *kanüm* expresses the reason for the main action. In Example 336, the postposition *gũ* marks the instrument with which the action is carried out.

(334) *hã nyu ba tashêne*

hã [nyu ba]_{adjunct} tashêne

1s 2s for sing-INT

‘I’ll sing on your account/for you.’ [110823_STBT_eli2 29]

(335) *ning mîyaqa kanüm tîga*

[ning mîyaqa kanüm]_{adjunct} tî-ga

1P woman reason fight-PL.I

‘We fought for our women.’ [110823_STBT_eli2 89]

(336) drâpkap tâplã gã khalai tyûzâ

drâpkap [*tâplã* *gã*]_{adjunct} *khalai* *tyû-zâ*

a.name hammer INS land hit-DIR

‘Drapkap hit (i.e. flattened) the land with a hammer.’ [201122_JA_eli1 1322]

7.2.2 Predicate clause based on argument structure

This section will present an overview of different sub-types of predicate clauses available based on the argument structure of the predicate. Verbless clauses (§7.1) will not be discussed in this section as they do not have predicates and thus cannot have argument structure.

As discussed by Dryer (2007a, 250), the basic distinction in terms of the argument structure of the predicate is that between transitive and intransitive clauses, the former having two or three arguments and the latter only one (also cf. §3.3.2.1). A brief overview of the language may suffice to notice that there are some difficulties in the identification of core arguments, as they can be omitted if retrievable from the context (cf. §7.2.1.1). Thus, in Example 337 the transitive verb *dyang* ‘chase’ does not have any overt arguments (here included as empty sets). Through attestations from the corpus and targeted elicitation, however, we may observe the number of arguments that a verb can maximally take, as in the elicited Example 338, where the predicate *dyang* ‘chase’ takes maximally two arguments.

(337) dyangga (...) we cya tshâktyû tshâktyûmâ la

[$\emptyset_A$ $\emptyset_{Pi}$ *dyang-ga*]_{main.cl} [*we* *cya*]_{Si} *tshâk-tyû* *tshâk-tyû-mâ* *la*]_{main.cl}

chase-PL.I DEM 3 move.fast-DIR move.fast-DIR-ITER HSAY

‘(They) chased (him) (and/while) he kept running and running.’ [TA22_kutyemba_phla 165 339]

- (338) majerei kacîk dyang
 [majerei]_A [kacîk]_P [dyang]_{pred}
 cat mouse chase
 ‘The cat chased the mouse.’ [210823_STBT_eli1 56]

7.2.2.1 Atransitive clause

An atransitive clause (also termed *zero-intransitive (or ambient) clauses*, cf. Dryer 2007a, 267f) does not have a semantic argument, i.e. no argument can occur with an atransitive verb. Atransitive verbs are commonly found in weather and parts-of-the-day terminology. Example 339 is ungrammatical because *lyâng* ‘be evening’ is an atransitive verb and there is no argument projected by the predicate and thus no capacity for an argument NP to appear within the clause.

- (339) (...) lyângtyübô, kanútyübô
lyâng-tyü-bô *kanú-tyü-bô*
 be.evening-about.to-PFV.ALLO be.dark-about.to-PFV.ALLO
 ‘(...) it was about to become evening, it was about to become dark.’
 [TA22_JAstory_047 152]

- (340) **hing/lyâng* *lyâng-ho*
 day/evening be.evening-STATE
 Intended: ‘the days/evening has become evening.’ [140823_JOJA_eli2 1046]

7.2.2.2 Intransitive clauses

An intransitive clause is headed by an intransitive verb. Intransitive verbs can have only one argument (S). In Example 341, the verb only allows one argument. Similarly, in Example 342 adding a further argument to the verb *yúm* ‘fly’ would make the clause ungrammatical.

- (341) kacîk gûzâ
 kacîk gû-zâ
 mouse stick.out-DIR
 ‘A mouse stuck out.’ [TA22_frog1_75 347]
- (342) oh tayîmbum we làng mu yúmdângga delé
 oh tayîmbum we làng mu yúm-dâng-ga delé
 oh type.of.bee DEM COLL also fly-COMPL.I-PL.I it.seems
 ‘Oh these bees are also flying away!’ [TA22_frog1_091 418]

As discussed in §7.2.1.1, transitive and intransitive verbs may take an extra goal/source extended core argument (E). In Example 343, the goal of the motion verb (*Sulyang* village) is unmarked, which resembles a core argument although it is semantically the goal of the intransitive motion verb. Similarly, in Example 344 there is a further unmarked argument that corresponds to the location (Abapa village).

- (343) tûm mlâ gũ sulyang bôya
 [tûm mlâ gũ]_{obliq.source} [sulyang]_E bô-ya
 a.village village LOC a.village go-ALLO
 ‘From Tom village they went to Sulyang.’ [TA23_chaitom_clan 19 110]
- (344) bîya júm we gũ tya ning we abapâ ìnadi gũ
 bîya júm we gũ tya ning we [abapâ]_E ì-na-di gũ
 past true DEM LOC now 1P DEM a.place be.there-FACT-HAB.EGO NF
 ‘In the (true) past, we stayed in Abapa village.’ [TA23_CLAN2 13 24]

Such cases have been subsumed under an ‘extended intransitive’ label and the further argument has been termed E (*extended to core*) (Dixon 2006b, 7f). As discussed by Post (2007, 417f, 694) in the case of Galo, both intransitive verbs of motion and transitive verbs of locution and cognition have an extended argument E either as Goal or as Addressee, which has diverse markings, sometimes similar to a P and sometimes similar to

an Oblique. Post further points out that the marking of the argument is related to its *predictability*: recurrent and expected actions are less prone to be marked. In the case of Tawrã, further unsupervised data (e.g. a conversation) would be required to confirm an equivalent phenomenon.

7.2.2.3 Transitive clauses

A transitive clause is headed by a transitive verb, which takes two arguments: A and P. In Example 345, both A and P arguments are overt, but this is uncommon in speech, as discussed at the end of §7.2.1.1. Rather, one argument (if not both) is commonly ellipsed. Since A is more likely to be topical, it is also prone to be ellipsed, as opposed to P (cf. §8.3.1). Cf. Example 346, where the A argument (the *carrier*) is omitted.

- (345) cya khumpha thâ
 [cya]_A [khumpha]_P thâ
 3 bribe eat
 ‘He took/accepted a bribe.’ [261122_JO_eli 1689]
- (346) “phlǎ́ pû we glâikũvidyô” la la
 Ø_A [phlǎ́ pû we]_P glâi-kũ-vidyô la la
 stone.MOD rock DEM carry-COMPL.II-CERT.FUT.EGO say HSAY
 ‘He said: “(I)’ll surely carry this stone (away)”, they say.’ [TA22_kutyemba_phla_40
 134]

P arguments may also be omitted when they are high in referential continuity (or else highly predictable, such as *food* in *I ate food*, cf. nominal incorporation §6.8). Examples 347a and 347b are from the same story in which the *first man* challenges a tiger. In Example 347a, we learn that the tiger is counting the fish in a pond. The verb *tsâi* ‘count’ takes the fish as its object. Example 347b is from the same story, albeit in a later part. As the P argument of *tsâi* ‘count’ has been established, it is not repeated.

- (347) a. (...) tangâ a abrabra wa tshâwa “we we tsâinagana” la la
 [tangâ a abrabra wa tshâw-ya]_{main.cl} [we we
 fish small manner.of.moving xxx move-ALLO DEM DEM
 tsâi-na-ga-na la la]_{main.cl}
 count-FACT-PL.I-IMP say HSAY
 “‘(...) count those moving fish!’ he said, it is said.’ [TA22_tiger 28 84]
- b. (...) nyu hayû tsâidângdyibô ka byê (...)
 [nyu_A ∅_P hayû tsâi-dâng-dyibô ka byê]_{cond}
 2S first count-COMPL.I-PFV.EGO COUNT.EXP COND
 “‘If you finish counting (the fish) first’ (...)’ [TA22_tiger 43-44 118]

7.2.3 Predicative possessive construction

This section describes the predicative possessive construction. Nominal possession has been described in §6.1.1.2. Note also that postposition *dũ* also expresses possession (§9.5.4). The predicative possessive construction is built with the intransitive verbs *ì* ‘be there (animate)’ and *ã* ‘be there (inanimate)’, depending on the semantic nature of the referent (§3.3.2.3). There is otherwise no syntactic difference with a standard predicative clause. In Example 348, the verb *ì* ‘be there (animate)’ selects an animate argument. The negation simply takes the negative marker *yüm* (cf. §6.4.1) as in Example 349.

- (348) khing pâ.mru ne ìya
 khing pâ.mru ne ì-ya
 one younger.brother only be.there-ALLO
 ‘I only have one (younger) brother (lit. there is only one younger brother).’
 [180122_AM_eli 1910]

(349) jâp budi ãnam, jâ dimag ãnam

[jâp budi ã-na = m]_{main.cl} [ja dimag ã-na = m]_{main.cl}

a.tribe brains(<IA) be.there-FACT=NEG what brains(<IA) be.there-FACT=NEG

‘We didn’t have the Kaman’s brains (i.e. intelligence), we didn’t have any brains (lit. there is no Kaman brain. There is no brain).’ [TA22_origins_167 466]

7.3 Non-declarative clauses

This section provides an overview of the illocutionary force of clauses, i.e. the intention of the speaker, focusing on *non-declarative clauses*. A *declarative* clause is employed for “assertions, reports, acts of complaining and bragging, but also acts of predicting and promising” (cf. König and Siemund 2007, 285). As per this schematization, declarative sentences can be thought of as unmarked in relation to all other types (interrogative and imperative) because a) they are the most common type of clause and b) they are not restricted in the expression of tense and aspect (idem. 285). Conversely, *non-declarative* clauses include interrogative clauses (§7.3.1) and command moods (§7.3.2). Note that the suffixes of the command mood are described in §6.3.5.

7.3.1 Interrogative clauses

Interrogative clauses pursue the speech act of requesting information, be it rhetorical or non-rhetorical. As it is common in the world’s languages (König and Siemund 2007, 290), Tawrã has two sub-types of questions: polar tag questions (§7.3.1.1) and content questions (§7.3.1.2). Interrogative clauses are commonly subordinated to verbs of cognition, cf. §10.6.

Interrogative clauses are based on declarative clauses in their syntax; commonly the only deviation from a declarative clause is the use of an interrogative element: Polar tag questions have a clause-final question tag and content questions place the question word on the same position as the questioned constituent.

A rising intonation often accompanies interrogative clauses, especially those with question tags (less so a question pronoun). An interrogative clause may not, however,

only rely on intonation: an interrogative element, be it a pronoun or a question tag, must be part of the clause. A rising intonation on a question clause that only contains a sole NP is not accepted in elicitation and has not been attested in the corpus. This is illustrated in Examples 350a and 350b, which are from elicitation. In neither case is an interrogative interpretation available; both may only yield a declarative interpretation³.

(350) a. *mêqe*

new

‘It is new.’ (It does not yield ‘*is it new?’)

b. *Tajaw Lohit kē*

a.place a.district LOC.INS

‘Tezu is in Lohit.’ (It does not yield ‘*(is) Tezu in Lohit?’)

[both 140823_JO_JA_eli2 348]

7.3.1.1 Polar tag questions

Polar questions have only two possible answers, either yes or no. These constructions in Tawrã require a sentence-final question tag *kya/ye/a* or a confirmation particle CNF *âm*⁴. Rhetorical questions are built in the same manner, although these questions may take a specialized indefinite pronoun *palang* that has been discussed in §3.5.1.6. Polar questions have the same sentence structure and constituent order as declarative sentences except for a clause-final question tag: in Example 351a, a polar tag question is answered with the same structure in Example 351b.

(351) a. *bajar gũ bréidyibô kya?*

bajar gũ bréi-dyibô kya

market(<IA) LOC buy-PST.EGO Q

‘Did you buy (it) at the market?’

³It should be noted that informants know such questions without an interrogative element from neighboring Indo-Aryan languages.

⁴This particle is likely a (non-nasalized) existential verb *â + cliticized negative = *m* cf. §6.4.1.

b. *âm*, *bajar* *gũ* *bréidyibô*

âm bajar *gũ bréi-dyibô*

yes market(<IA) LOC buy-PST.EGO

‘Yes, I bought (it) at the market.’ [both 240122_AM_eli1 278]

The difference between the question tags *kya*⁵/*ye/a* is dialectal: *ye* is employed in the plains around Tajaw (Tezu) while *kya* is more common in the valleys of Anjaw.

Confirmation particle CNF *âm* also has an affirmative function, employed to ask for the speaker’s agreement, as in Example 352.

(352) *pra âm* ?

pra âm

good CNF

‘It’s good, right?’ [field notes 19102024]

Example 353 corresponds to a verbless or appositive clause (§7.1) in which the identity of the referent is questioned. A negative counterpart can be construed with the standard negative marker *yüm* (§6.4.1) after the two NPs and before the question particle, as in Example 354. In Example 355 the question tag is outside the main clause, almost like an afterthought, but it has scope over it.

(353) *yebe tayîng.bom àng kya?*

ewe tayîng.bom àng kya

DEM type.of.bee house Q

‘Is this a *tayîngbom*’s hive?’ [TA22_frog 66 314]

(354) *we tayîng.bom àng yüm kya?*

we tayîng.bom àng yüm kya

DEM type.of.bee house NEG Q

‘Isn’t this a *tayîngbom*’s hive?’ [230823_STBT_eli1]

⁵In the case of *kya*, a borrowing from (Arunachali) Hindi *kyā* seems unlikely, for the question particle shares *eidemic resonance* (cf. Bickel and Nichols 2007, 209f) with the question pronoun *shya* ‘who’. For a discussion on questions in Hindi-Urdu cf. Bhatt (2016, 521) for a more general overview also cf. Masica (1993, 388). A fronting of the question particle, as in Standard Hindi, is not allowed in Tawrā.

- (355) *nyu nepar khina bü laga, jüm kya?*
 [nyu nepar khì-na] bü la-ga] [jüm kya]_{Q.cl}
 2s a.country reach-BACK once hearsay-PL.I truth Q
 ‘You went to Nepal (and came back), they say. Is it true?’ [310322_JOJA_eli1
 1680]

7.3.1.1.1 Disjunction construction with *kya*

A sub-construction of the polar tag question involves the question particle *kya* after either two verbs (the second of which is negated) or two different adjectives. Table 7.4 illustrates the structure of the *kya* disjunction construction with verbal and verbless clauses. Its function is to express a doubt regarding two possible options. As such, this construction is commonly the argument of a (negated) verb of cognition ‘I don’t know whether X or Y’ (cf. complement clause §10.6.1): *I don’t know whether V or no V or whether X is Y or Z*.

Table 7.4: Structure of the *kya* construction on verbal and verbless clauses

verbal clause	V	<i>kya</i>	V <i>yüm</i>	<i>kya</i>
verbless clause	Adj1	<i>kya</i>	Adj2	<i>kya</i>

In Example 356, a verb occurs once with a positive and once with a negative polarity. In Example 357, two different adjectives occur with a question tag. In both cases, the *kya* construction is a verbal complement of the negated verb *kasâ* ‘know’. The focus morpheme *nyu* commonly occurs after each option (§8.3.2).

- (356) *bôbiyà kya bôbi yüm kya kasâ yüm*
 [bô-bi-ya kya bô-bi yüm kya]_{compl.cl} kasâ yüm
 go-POT.ALLO-ALLO Q go-FUT.ALLO NEG Q know NEG
 ‘I don’t know whether he will go or not.’ [060923_STBT_eli1 732]

- (357) *ewe yu talü nyu kya pâtlâ nyu kya kasâ yüm*
 [ewe yu talü nyu kya pâtlâ nyu kya]_{compl.cl} kasâ yüm
 DEM beer strong FOC Q thin(<IA) FOC Q know NEG
 ‘I don’t know whether this beer is strong or thin.’ [060923_STBT_eli1 725]

7.3.1.2 Content question

In a content question, the speaker expects to receive precise information, as opposed to a yes-or-no or polar question (cf. Payne 1997, 300-303). Tawrã has several question words that specify the type of information requested, such as *who*, *where*, etc., as discussed in §3.5.1.2. §7.3.1.2.1 discusses the use of the question particle *palang/(si)mlang*, which highlights the lack of information and evidence from the speaker when making a question.

In a pragmatically neutral situation, these question words occupy the same position as the questioned constituent, a feature also termed *in situ*: Example 358a is a ditransitive clause where all three arguments of the verb are overtly realized along with an oblique argument of location (‘in the house’). All four arguments are questioned *in situ* in Examples 358b, 358c, 358d and 358e. These sentences, although grammatically correct, are perceived as odd, as it is uncommon for all arguments to be overt.

- (358) a. *ma we a we àng kē gũ tapè hàngya*
 [ma we]_A [a we]_P [àng kē gũ]_{adju.loc} [tapè]_T [hàng-ya]_{pred}
 mother DEM child DEM house LOC.INS LOC food give-ALLO
 ‘The mother gave the child food in the house.’
- b. *shya a we àng kē gũ tapè hàngya*
 [shya]_{A/Q} [a we]_P [àng kē gũ]_{adju.loc} [tapè]_T [hàng-ya]_{pred}
 who child DEM house LOC.INS LOC food give-ALLO
 ‘Who gave the child food in the house?’

c. *mâ we shyà àng kě gũ tapề hàngya*
 [mâ we]_A [shyà]_{P/Q} [àng kě gũ]_{adju.loc} [tapề]_T [hàng-ya]_{pred}
 mother DEM who house LOC.INS LOC food give-ALLO

‘To whom did the mother give food in the house?’

d. *mâ we a we hanûk tapề hàngya*
 [mâ we]_A [a we]_P [hanûk]_Q [tapề]_T [hàng-ya]_{pred}
 mother DEM child DEM where food give-ALLO

‘Where did the mother give food to the child?’

e. *mâ we a we àng kě gũ ja hàngya*
 [mâ we]_A [a we]_P [àng kě gũ]_{adju.loc} [ja]_{T/Q} [hàng-ya]_{pred}
 mother DEM child DEM house LOC.INS LOC what give-ALLO

‘What did the mother give to the child in the house.’

However, as discussed in §7.2.1.1, a characteristic of Tawrã naturalistic speech is that clauses do not exhibit many overt arguments and it is the context that allows the hearer to reconstruct role and reference relations. Topic constructions also have the function of helping recognize the role and reference of arguments (§8.3). In Example 359, the topic is introduced through a topic construction while the interrogative clause does not have any overt arguments.

- (359) *tsâi.tûm tahû mane suru hanûk nyu hanaya gũ ?*
 [tsâi.tûm tahû mane suru]_{top} [hanûk nyu hana-ya gũ]_{content.Q}
 a.clan clan FILL(<ASM) beginning(<IA) where FOC come-ALLO NF
 ‘The Chaitom clan, I mean, (in the) beginning, where did it come from?’
 [TA22_chaitom_clan 35 200]

7.3.1.2.1 Rhetorical and wondering particle *palang*/(*sa*)*mlang*

The question particle *palang* or (*sa*)*mlang* indicates the speaker’s lack of knowledge or evidence that would help him/her answer a question (cf. Zemp 2018, 704). As such, it

reflects a *wonder* with which the speaker formulates a question. Furthermore, this effect can be exploited for rhetorical or otherwise sardonic questions. This particle thus occurs in interrogative clauses as a question tag, although it shares a similar pragmatic effect of modal particles (cf. §9). Question particles *palang*/(*sa*)*mlang* also occur as fillers and indefinite pronouns (cf. §3.5.1.6).

In Example 360 the speaker asks about the loss of someone's land. The use of the particle *mlang* highlights the speaker's wonder. Similarly, in Example 361 a speaker employs the particle *palang* in a question; this reflects his uncertainty about whether there is an umbrella at all.

- (360) *we cya khalai kadé bô we mlang we?*
we cya khalai kadé bô we mlang we
 DEM 3 field how go DEM WND DEM
 'How is it that his land is lost (lit. gone)?' [TA24_land_7 13]

- (361) *chata ã palang?*
chata ã palang
 umbrella(<IA) have WND
 'Would you have an umbrella (at all)?' [field notes 201024]

The inherent lack of knowledge or evidence can be exploited for other pragmatic effects. In Example 362, the use of the particle *palang* renders the questions rhetorical. Similarly, Example 363 is a fixed construction employed in a rhetoric/sardonic way.

- (362) *tapê thâdyibô palang?*
tapê thâ-dyibô palang
 food eat-PFV.EGO WND
 'Did I have food?' [060823_JOJA_voc 777]

(363) *kadé mlâ thâ palang!*

kadé mlâ thâ palang

how.many village eat WND

‘How slowly (lit. how many villages) you eat!’ [field notes 201024]

7.3.2 Command moods

Command moods are employed to “express directive speech acts, i.e. orders and requests, but also invitations, the giving of advice, warnings, wishes, instructions, etc.” (König and Siemund 2007, 303). Tawrã has specific verbal suffixes to express the imperative *-na* (and *-tè*) §6.3.5.0.1, the prohibitive *-yâq* §6.3.5.0.2, the hortative *-keq* §6.3.5.0.3 and the jussive *-tyêyu* §6.3.5.0.4. These are discussed as part of the TAMN verbal suffixes in §6.3.5. It should be noted that command moods do not display a different syntax from that of a declarative clause.

8 Semantics and pragmatics of clause constituent core arguments

This chapter describes the semantics and pragmatics of clause constituent core arguments in Tawrã. As such, it includes a description of the main argument selectors involved in grammatical relations, as well as a discussion on topic and focus. This broad discussion involves an inevitable overlap, e.g. between the addressee of imperative clauses (§8.2.2) and the syntax of imperative clauses in general (§6.3.5.0.1). Overlap and repetition are avoided by referring to relevant sections and by focusing on the alignment of arguments as much as possible.

As will become clear in this chapter, there is no evidence for a strong grammaticalization of morphosyntactic properties that relate an argument to a clause (e.g. as a Subject or Object) in Tawrã. Although some constructions show a strong tendency in favor of a S/A pattern, these are commonly overridden by semantico-pragmatic considerations. In this sense, the concepts of topic and focus offer a pragmatically oriented way to think about grammatical relations in Tawrã.

This chapter is organized as follows: §8.1 will introduce relevant general notions about grammatical relations. §8.2 presents a summary of the main findings of this chapter regarding grammatical relations, while the sub-sections are descriptions of particular constructions (here termed *argument selectors*) where grammatical relations are relevant: §8.2.1 discusses predicate marking. §8.2.2 reviews the addressee in imperatives. §8.2.3 is about constituent order. Marking of core argument noun phrases (henceforth NPs) is discussed in §8.2.4. Omitted arguments in multi-clause constructions are discussed in §8.2.5. §8.3 introduces the concepts of topic and focus. §8.3.1 describes the construction and morphemes associated with the topic, while §8.3.2 presents the concept of focus and

its associated morpheme.

8.1 Key concepts

Grammatical relations refer to how arguments are syntactically integrated into the clause, e.g. as Subject and Direct and Indirect Objects, in a particular construction. These concepts are generally held to be crucial for linguistic analysis, even if they are described as absent in a given language (Witzlack-Makarevich 2011, 1;29; Witzlack-Makarevich 2019, 2f). In general terms, the valence of a predicate determines the number of arguments that it governs. In this sense, transitivity can be thought of as a template or a projection of the clause head onto the clause in which a particular number of NPs may fit in.¹ How these NPs interact morpho-syntactically (i.e. who acts on whom) will be determined by the grammatical relations found in the language.

Arguments are semantically defined in relation to the specifications of the predicate, e.g. *killer* and *killed*, as well as on their own referential type (Bickel 2011, 403f). Furthermore, arguments are commonly distinguished from adjuncts, the latter being less tightly associated with the predicate - they provide additional information about the general setting (manner, place and time). The *optional* character of adjuncts (vis-à-vis arguments) has traditionally been taken as the main difference between these two. However, for some authors, the difference between adjuncts and arguments is considered to be largely language-specific (Witzlack-Makarevich 2011, 41f).

Verbs are categorized by their transitivity (as discussed in §3.3.2.1), which defines the number of participants. The semantic roles associated with these participants are specified by the verb (i.e. these are in principle verb-specific) and it will define its prepositional-semantic frame, as well as that of the state/event and the clause (Givón 2001a, 106). Following Van Valin and LaPolla (1997, 139), the link between the myriad of possible verb-specific semantic roles is bridged through *macro-roles*, a set of generalized roles distinguished by verbal valence, i.e. intransitive (valence 1), transitive (valence 2) and ditransitive (valence 3). These roles are thus defined: **S** is the only argument of

¹The term *argument structure* can be thought of as being the same theoretical principle but from a NP-oriented point of view (cf. Post 2007, 682).

an intransitive verb. **A** stands for the ‘most actor-like argument’ in a transitive verb, **P**² stands for the ‘not most actor-like argument’ in a transitive verb. In the case of the not actor-like arguments of ditransitives, **G** is the ‘most goal-like or ground-like’ and **T** for the other ‘not most goal-like, not most actor-like’ argument (Bickel 2011, 402f; Haspelmath 2011, 549). These roles are displayed in Table 8.1.

Table 8.1: Macro-roles (based on Bickel (2011: 402) and Haspelmath (2011))

valence 1	S		
valence 2	A	P	
valence 3		G	T

The study of grammatical relations attempts to isolate sets of arguments that are treated the same way in opposition to one or more arguments, i.e. by the use of the same case marker or by triggering the same agreement on the verb. The grammaticalization of argument *patterning* in one or another way will determine an alignment type for a specific construction (Bickel 2011, 399f; also Comrie 1982, 97; LaPolla 2023, 270f). In practical terms, a construction that discriminates one set of arguments over other(s), e.g. case marking that occurs on P (but not S and A) will thus be described as having an accusative alignment pattern. Crucially, a language may have more than one pattern, as different morphosyntactic properties (i.e. different constructions and criteria) may show different alignments that may not converge; they are regarded as construction-specific (Witzlack-Makarevich 2011, 29f). Furthermore, these constructions are seldom absolute (cf. the discussion on differential object marking in Spanish reviewed by Bossong 2021). Thus, the grammaticalization of specific construction as grammatical relations is often a matter of degree (Givón 2001a, 177; Givón 1997, 25f).

²There is a considerable amount of variation in both terminology and the analytical implications of the different initials used for the macro-roles, cf. Haspelmath (2011) for a comprehensive overview. Further discussions of the differences and implications of the main approaches to grammatical relations would fall outside the scope of this grammar. Hence the labels S, A, P, G and T will be employed.

8.2 Argument selectors

Based on an overview of all relevant constructions in Tawrã, it can be stated that the language mostly yields a neutral alignment, i.e. it shows little signs of grammaticalizing referent identification through a particular alignment (e.g. S/A). Solely from the domain of addressees in the imperative (§8.2.2) emerges a clear S/A alignment. An overview of grammatical relations is displayed in Table 8.2, which lists all constructions (argument selectors) considered in this chapter. Note that in ditransitive clauses ditransitive-A is subsumed under A, just as T and G are under P, as these never exhibit a differential syntactic behavior in the language.

Table 8.2: Overview of grammatical roles in Tawrã

Argument selectors	alignment (selected set of arguments)
Predicate marking	S/A/P (neutral)
Addressee in imperatives	S/A (accusative)
Constituent order	S/A/P (neutral)
Marking of core argument NPs	S/A/P (neutral)
Omitted argument	S/A/P (neutral)

The discussion from the subsections devoted to predicate marking, marking of core argument NPs, omitted arguments and constituent order will demonstrate that these domains of the grammar are not sensitive to grammatical relations. In other words, in those constructions there are no signs of a privileged alignment in one or another way - the language has not grammaticalized these constructions to express the category of Subject or Object. Rather, the interpretation of referents is heavily dependent on pragmatics. In this sense, as discussed by Enfield (2009), grammatical roles can often be assigned by our knowledge of the world, without relying, e.g. on constituent order or case marking (also cf. Chao 2011[1968], 95f. Givón 2001a, 200). Example 364 illustrates this point: both interpretations (*a person* killing an elephant and *an elephant* killing a person) are theoretically possible. However, humans, especially if alone, are more likely to be attacked

and even killed by elephants whereas the opposite i.e. people killing an elephant is not common (anymore).

(364) taméqĩng mề sêkũho pà

taméqĩng mề sê-kũ-ho pà

elephant.sp person kill-COMPL.II-STAT ALAS

‘Sadly, the elephant killed the person’ (also, albeit oddly: ‘Sadly the person killed the elephant’). [220122_JO_JA_eli1 410]

For this reason, semantic roles (briefly mentioned in §8.1) will be employed as pre-theoretical concepts, i.e. without assuming their grammaticalization into a specific alignment in this grammar. Although these roles are defined by each verb, *macro-roles* (§8.1) provide a general frame anchored to the valency of the verb.

8.2.1 Predicate marking

As mentioned in §8.2 (also §6.1), there is no obligatory agreement/argument cross-referencing on the predicate. Rather, TAMN (tense, aspect, mode and number) suffixes exhibit “person-based sensitivities” (Post 2007, 689) in the sense that they do support the identification of a grammatical relation that we might define as S/A alignment (§8.1). However, none of these suffixes can be straightforwardly described as “subject agreement suffixes”, as each of them is also sensitive to the semantico-pragmatic factors of animacy, engagement and speech act conditions. Particularly regarding those suffixes that are sensitive to egophoricity (§6.3.1.1), it is clear from semantic-pragmatic and distributional criteria that they are not a form of agreement (cf. Bergqvist and Kittilä 2019, 1f.).

An example of a suffix group that is commonly associated with S/A agreement is the non-singular suffixes *-ga* PL.I *-ring/-rẽ* PL.II and *-nega* DU (§6.3.2.0.2). In Examples 365 and 366, the non-singular suffix *-ga* stands for a non-singular S argument.

- (365) hayũ jũm we tũm mlã ... tũm mlã ki ìgade na
 [hayũ jũm [we tũm mlã]_{loc.adj} ∅_S [tũm mlã ki]_{loc.adj}
 past true DEM a.place village a.place village LOC.INS
 ì-**ga**-de na]_{main.cl}
 be.there-PL.I-HAB.EGO AFF
 ‘In the past we lived at Tom village, right?’ [TA23_chaitom 7 34]

- (366) we pêphrêgana mane
 we_S pê-phrê-**ga**-na mane
 DEM separate-away-PL.I-FACT FILL(<Asm)
 ‘They separated, I mean.’ [TA23_chaitom 7 34]

Similarly, in Examples 367 and 368 the verbal suffix *-ga* is associated with a non-singular A argument. In Example 367, the villagers cut one slice of meat for a criminal. In Example 368, the same villagers found the criminal again after he escaped.

- (367) wũw tyüwga
 ∅_A [wũw]_P tyüw-**ga**
 big.chunk.of.meat slice-PL.I
 ‘(They) sliced a big chunk of meat.’ [TA22_kutyemba 65 174]
- (368) kehîng gê katyânggana matyô
 kehîng gê ∅_A ∅_P katyâng-**ga**-na matyô
 CLF:day one find-PL.I-FACT maybe
 ‘One day (they) saw (him).’ [TA22_kutyemba 153 317]

The association between the non-singular suffix and S/A is, however, loose, as it may be overridden by animacy: non-singular verb suffixes commonly index referents that range high in animacy.³ In Example 369, there could be one or more monkeys chasing

³Tawrã has a small set of verb pairs that differ in imposing constraints of agency to core arguments. These have been discussed in §3.3.2.3.

one or more elephants - the lack of a plural verbal suffix is not enough cue to assess that. The alignment of the plural suffix is, in this case, neutral as the verb is effectively unmarked for neither argument. However, *animacy* is highly dependent on context. In Example 370, the dual marker *-nega* DU is triggered by the child and the dog. We may assume here that the dual marker stands in solidarity with the high agency of the dog in this particular story, even though it is marked as an adjunct.

- (369) taméqĩng tamyũm dyanga
taméqĩng tamyũm dyang-ya
 elephant monkey chase-ALLO
 ‘The elephant(s) chase(s) the monkey(s)’ (also: ‘The monkey(s) chase(s) the elephant(s)'). [170923_STBT_eli1 573]

- (370) tya a we kwâg dũ îngnegaho
tya a we [kwâg dũ]_{adju} îng-nega-ho
 now child DEM dog with sleep-DU-STAT
 ‘Now the child is sleeping with the dog.’ [TA22_frog1 21 89]

Furthermore, non-singular suffixes do not have to occur if the referent either has a collective suffix (§3.3.1.4) or is known to be non-singular from the context. Examples 371a and 371b illustrate this point. In Example 371a, the village people are pulling ropes. Both the A and P arguments are overt. The A argument (the villagers) also has a collective suffix and it is not marked by the plural suffix *-ga*. Similarly, in Example 371b the villagers are dropping more ropes and the verb does not take the *-ga* suffix, as it is understood that the referent is non-singular.

- (371) a. mlǎ mẽ we làng tarâw tatyế hrēmâ gũ (...)
[mlǎ mẽ we làng]_A [taráp tatyế]_P hrēmâ gũ
 village.MOD person DEM COLL type.of.creeper rope pull-ITER NF
 ‘The village people pulling *tarap* ropes (...)’ [TA22_Kutyemba_phla 230 456]

b. mane azamâ khrâk

∅_A ∅_P mane aza-mâ khrâk

FILL(<ASM) drop-ITER EVI

‘(They) were dropping (ropes).’ [TA22_Kutyemba_phla 234 461]

Lastly, plural suffixes may refer to the P argument (instead of S/A) in cases where the plurality of that argument is somehow relevant. In Example 372, the object is a herd of goats while the addressee is a singular referent (note that shepherding is normally undertaken alone). The plural suffix refers to the P argument (‘the goats’). Similarly, in Example 373 the man challenges the tiger to count a school of fish in a pond. The plural suffix refers to the school of fish, as there are no other referents available. Furthermore, the fact that the fish were many is relevant, as the tiger could not finish counting them. In both cases, the plural suffix refers to the referent that is most prominently plural, regardless of its semantic role and animacy (in these cases it is a P argument).

(372) kalyab kě bô chagulyei we lantang rôrêna!

[kalyab kě bô]_{rel.cl} chagulyei we lantang]_P rô-rê-na

field LOC.INS go goat DEM COLL catch-PL.II-IMP

‘Catch the goats that went to the field!’ [140522_JO]

(373) (...) “tangâ a abrabra wa tshâwa we we tsâinagana” la la

[tangâ a [abrabra wa tshâw-ya]_{rel.cl} we we]_P tsâi-na-ga-na

fish small manner.of.moving xxx move-ALLO DEM DEM count-FACT-PL.I-IMP

la la

say HSAY

“(…) count those moving fish!” he said, it is said.’ [TA22_tiger 28 84]

8.2.2 Addressee in imperatives

The addressee of an imperative can only be S/A (i.e. not P, G or T). In Example 374, the addressee is a S argument while in Example 375 it is an A argument. A full description of the imperative can be found in §6.3.5.0.1.

(374) “karàhwîbô ne hanaringna” la

karà-hwî-bô ne hana-ring-na la
rain-clear.up-PFV.ALLO only come-PL.II-IMP say

“‘Only come (when) the rain has cleared up!”, he said.’ [TA23_talyu_clan 30 182]

(375) pra hwéringna!

pra hwé-ring-na
good look-PL.II-IMP

‘Look after it!’ [TA23_fishstory 153]

8.2.3 Constituent order

Constituent order (also known as *word order*) has been criticized as a predictor for grammatical relations for, in many cases, it can provide but a tendency in the identification of grammatical roles. Furthermore, it has been argued that constituent order is most relevant for verb-medial languages, since e.g. in a A P V configuration S (from S V) can be grouped with both A and P; S and A are the first elements of a clause while S and P are placed before the verb (Witzlack-Makarevich 2019, 18; also Li and Thompson 1976; also cf. LaPolla 2006, 155f; also cf. the discussion in Enfield 2009).

Constituent order is not a predictor of grammatical relations in Tawrã since known referents are routinely ellipsed and their position in the clause is governed by pragmatic considerations. In other words, the language has not grammaticalized constituent order as a grammatical relation (cf. LaPolla and Poa 2006, 277f.). However problematic constituent order is for the study of grammatical relations, it can still be recognized in Tawrã that the most common and pragmatically less marked patterns are S V and A P V, as attested in the corpus and confirmed through elicitation. Example 376 is an intransitive clause with a S argument followed by the verb. Similarly, Example 377 is a transitive clause with an A and a P argument (in that order).

- (376) *ewe mambaw byúwa*
 [ewe mambaw]_S [byúw-ya]_V
 DEM fox run-ALLO
 ‘This fox runs.’ [TA22_chicken_thief_32 118]

- (377) *kutyêmba we mlá mẽ we la matyô*
 [kutyêmba we]_A [mlá mẽ we]_P [la]_V matyô
 a.name DEM village.MOD people DEM say maybe
 ‘This Kutyemba told these village people.’ [TA22_kutyemba_phla 35 125]

The pragmatically neutral constituent order S V/A P V is regarded as the most common configuration typologically and also as the most common constituent order in the Tibeto-Burman language family (cf. Dryer 2013; Dryer 2003, 43; including the TB languages of SEA, cf. Vittrant 2019, 101). However, some scholars working on the languages of South and Southeast Asia have recognized several problems with readily positing A P V as the default constituent order. The most common reasons that trigger changes in constituent order are i) argument dropping (ellipsis) and ii) pragmatically driven movements (Enfield 2007, 272f; Morey 2012, 514f).

Argument dropping in Tawra is discussed in §7.2.1.1. It suffices to mention here that S/A arguments are commonly ellipsed if they are understood or retrievable from the context. This is illustrated in Example 378, where a new argument (the wife) is introduced in the first clause and is the topic. The same referent is a P argument in the second and third clauses, where it first occurs overtly and is then replaced by a demonstrative. The A argument (‘they’) never occurs overtly.

- (378) *mîya we ne kasâpwi (...) mîyã we ahuya larêya la (...) we ahurê byê (...)*
 [[mîya we]_{Sj} ne kasâ-pwi]_{main.cl} [∅_A [mîyã we]_P ahu-ya la-rê-ya
 wife DEM only know-surely wife DEM ask-ALLO say-PL.II-ALLO
 la]_{main.cl} [∅_A [we]_{Pj} ahu-rê byê]_{main.cl}
 HSAY DEM ask-PL.II TEMP.SUB
 ‘Only (his) wife knows. They said to ask (his) wife. Having asked her (...)’
 [TA23_Tayang_clan 30 177]

The omission of both A and O arguments is also possible. Examples 379a, 379b and 379c are from the same story about the speaker's grandfather. In Example 379a, *Solum* is the topic. Thereafter, in Example 379b, the priest *Ajumso* is also introduced and becomes a topic. Thereafter, in Example 379c no argument is overt. The discourse deictic complex demonstrative *we cuwe* (§10.8.1), works as a thematic link between presuppositions (without being co-referential) and logically connects these two referents to the clause (i.e. *Solum* hiring the priest).

- (379) a. atya Solum atyu (...) phrêya mane
 a-tya *Solum* *atyu* *phrê-ya* *mane*
 1s.POSS-grandfather a.name up.there pioneer-ALLO FILL(<Asm)
 ‘My grandfather Solum up there (...), was the first (to reach the place).’
 [TA24_land 142 261]
- b. Ajumso (...) gwók Ajumso laga na
 Ajumso *gwók* *Ajumso* *la-ga* *na*
 a.name priest a.name say-PL.I INSIST
 ‘Ajumso, the priest’s name was Ajumso, right?’ [TA24_land 144 263]
- c. we cuwe mane kasemakáwshyüw la
 [*we* *cuwe* *mane* $\emptyset_A$ $\emptyset_P$ *kase-makáw-shyüw* *la*]_{main.cl}
 DEM DISC.DEIC.DEM FILL(<Asm) inquire-thoroughly-CAUS.V HSAY
 ‘So, (my uncle) had (the priest) inquire thoroughly (whether the area was
 suitable or not), they say.’ [TA24_land 146 267]

The pragmatically driven movements that may affect constituent order are subsumed under the *topic construction* discussed in §8.3. Therefore, this will only serve as a summary aimed at understanding grammatical relations. In a topic construction, a topicalized argument in a clause may either be fronted or backed. In Example 380, the topic is dislocated to the left and commonly takes the *byẽ* topicalizing morpheme. This clause, although topicalized, does not alter the canonical constituent order S V. Conversely, in Example 381 the topicalized argument is P (and not A), yielding a P A V constituent order.

Note that, without context, that example could be analyzed as having two P arguments: ‘as of the book, s/he didn’t find the child’, i.e. ‘s/he found the child, not the book’.

- (380) *ning byě macêi mathûw nyu ikûnaya*
 [*ning byě*]_{TOP/S} *macîk mathûw* *nyu ì-kû-na-ya*
 1P TOP water up.the.mountains FOC be.there-COMPL.II-FACT-ALLO
 ‘As for us, we lived up (in) the mountains.’ [TA22_origin 164 460]

- (381) *tako byě a we hwě.tyâng yüm*
 [*tako byě*]_{TOP/P} [*a we*]_{A/P} *hwě.tyâng yüm*
 book TOP child DEM find NEG
 ‘As of the book, the boy didn’t find it’ (alternatively: ‘S/he found the book, not the child’). [011024_STBT_eli2 1365]

Backed (also known as post-posed or right-dislocated) arguments resemble afterthoughts since they often occur after a pause. Syntactically, however, they are still inside the main clause (cf. Lambrecht 1994, 203). In Example 382, the topic occurs after the verb. Note that backed topics do not take a topicalizing morpheme.

- (382) *lulyang ìgana (...) we làng*
 [*lulyang ì-ga-na* [*we làng*]_{top}]_{main.cl}
 a.place be.there-PL.I-FACT DEM COLL
 ‘They live in Lulyang, these.’ [TA23_talyu_clan 44 272]

Summing up, constituent order is often subjected to pragmatic considerations. It is therefore most useful to employ the analytical tools provided by topic and focus (following Li and Thompson 1976) rather than Subject and Object in order to fathom the rationale of constituent order (cf. §8.3). There is a general tendency for Subjects to coincide with the topic and Objects to fall within focus and thus, in a language with topic-focus clausal organization, we expect an A P V to be the most common configuration. Further than that, there seem to be no grammatical principles governing the order of clause constituents in Tawrã and thus no *basic* word order.

8.2.4 Marking of core argument NPs

Tawrã does not employ case marking to distinguish core arguments, this includes the arguments of ditransitive verbs. The equal treatment of S/A/P (in this case, unmarked) has been termed *neutral* alignment (Comrie 2013). In Example 383, the S is unmarked. Similarly, in Examples 384 and 385 all arguments (A, P and T) are unmarked. Note that the demonstrative *we* is related to definiteness and does not express any grammatical function (§3.5.4).

(383) kalà lyó hanabô

[kalà lyó]_S hana-bô

officer white come-PFV.ALLO

‘The white officers had come.’ [TA22_clan 24 89]

(384) nyu mu îng plà dùwa byě ...

[nyu]_A mu [îng plà]_P dùw-ya byě

2S also a.bit salt need-ALLO COND

‘If you also need salt ...’ [TA22_kutyemba_phla 414 774]

(385) mîyaqa we tapüw bajei tó gê hàngya

[mîyaqa we]_A [tapüw]_P [bajei tó gê]_T hàng-ya

girl DEM old.man banana kilo one give-ALLO

‘The girl gave the old man one kilo of bananas.’ [200823_STBT_eli1 221]

Peripheral roles are distinguished from core arguments in that they may take different specialized postpositions according to their function (locative, instrumental, etc.), although this is not done consistently, cf. §7.2.1 for further details.

In sum, the marking of syntactic dependents follows a neutral pattern in which S/A/P are unmarked across valence classes. Lexical classes, TAM values and polarity do not play a role in this regard. Oblique and adjuncts, however, may be optionally flagged by a variety of postpositions (cf. §9.5).

8.2.5 Omitted arguments in multi-clause constructions

The interpretation of an ellipsed argument in multi-clause constructions and its logical (or obligatory) relation to the head of the main clause is widely considered to be an argument selector. A discriminating treatment of some arguments (vis-à-vis others) has been generally termed *pivot* (cf. LaPolla 2023, 273f). In some languages like English, a S/A co-reference is obligatory and this is taken as an argument for grammatical relations (also cf. the discussion on *conjunction reduction* in Witzlack-Makarevich 2019, 19f; LaPolla 2006, 156). In other languages, shared arguments are *preferably* expressed as S/A (instead of P) but, depending on the context, they can also be co-referent with other arguments that do not occur on the main clause (e.g. in Sazhin (Northeast Caucasian, Russia), cf. Forker 2019, 87f). As discussed in the case of Kharia (Munda, E and NE India) (Peterson 2010, 361), the tendency to maintain a S/A co-referentiality across coordinated clauses is bound to the logic of the narrative in which keeping the same topic (cf. §8.3) is crucial for coherence (also cf. Pu 1997). This situation corresponds to what I have observed in Tawrã, both in narrative texts and elicitation: co-referentiality in multi-clause constructions is subjected to narrative impositions and pragmatic considerations. This is translated in a tendency to have S/A co-referentiality which is, however, by no means obligatory. As such, it is not an argument selector.

In Example 386, the (anti-)hero *Kutyemba* tricks the village people into serving him food in exchange for him getting rid of a supposedly deadly rock. After eating, he tells the people to gather up. The same omitted A argument is co-referential in both clauses. This pattern is repeated in Example 387, where the same referent is both the ellipsed S argument of the subordinate clause and the A argument of the main clause. Similarly, in Example 388 the A argument in the main clause (a type of grass, referenced by the demonstrative in topic position) is co-referential with the S argument of the preceding subordinate clause.

(386) thâlibô gũ la la, mǎ̃ mề we la “chôq” !

[$\emptyset_i$ thâ-li-bô gũ]_{cosub.cl} [$\emptyset_i$ la la]_{false.start.clause} $\emptyset_i$ mǎ̃
 eat-REC.PST-PFV.ALLO NF say HSAY village.MOD
 mề_j we la chôq]_{main.cl}
 person DEM say let's!

‘After having eaten he_i said ... he_i said to the village people_j: “let’s do it!”’
 [TA22_kutyemba_phla 75 190]

(387) katyânga matyô gâqa kě ìho gũ

[$\emptyset_{Ai}$ katyâng-ya matyô [$\emptyset_{Si}$ gâqa kě ì-ho gũ]_{cosub.cl}]_{main.cl}
 hear-ALLO maybe near LOC.INS be.there-STAT NF

‘He must have heard it since he was close (lit. being close).’
 [TA22_kutyemba_phla 405 764]

(388) pübô gũ ali shî shyâbiya we

[$\emptyset_i$ pü-bô gũ]_{cosub.cl} ali shî shyâ-biya [we_i]_{top}]_{main.cl}
 bloom-PFV.ALLO NF like.this seed become-POT.ALLO DEM

‘Having bloomed, it will give seeds.’ [TA22_fishstory 156 272]

Co-referentiality is, however, not obligatory in multi-clause constructions in Tawrà. Examples 389 and 390 were prompted by visual stimuli: The hunter encounters the tiger and the *tiger* runs away. Then, the hunter again encounters the tiger but this time *the hunter* runs away. Both scenarios can be resolved by Example 389. This example illustrates that the omission of an argument in multi-clause constructions does not make a co-referential S/A reading the default interpretation. Conversely, in Example 390 the speaker wishes to be unambiguous about *the hunter* escaping (and not the tiger). He used an overt S argument *mề cya nyu* ‘he, the man’ that correlates with the topic (A argument) of the main clause and stands in opposition to the tiger from the previous sentence.

- (389) *tabrũ.tsũ we tamyâ we hwě-tyâng gũ byũwtyûwa*
 [[*tabrũ.tsũ we*]_i [*tamyâ we*]_j *hwě-tyâng gũ*]_{cosub.cl} [*∅*]_{i/j} *byũw-tyû-ho-ya*
 hunter DEM tiger DEM see NF run-away-STAT-ALLO
 ‘Having the hunter seen the tiger, the hunter/tiger runs away.’ [180923_STBT_eli1
 506]’

- (390) *tabrũ.tsũ we tamyâ we hwě-tyânggũ mề cya nyu byũwtyûwa*
 [[*tabrũ.tsũ we*]_i [*tamyâ we*]_j *hwě-tyâng gũ*]_{cosub.cl} [*mề cya nyu*]_i
 tiger DEM hunter DEM see NF person 3 FOC
byũw-tyû-ho-ya
 run-DIR-STAT-ALLO
 ‘Having seen the tiger, the hunter, s/he the person, runs away.’ [180923_STBT_eli1
 526]’

In Example 391, the head of the cosubordinate clause is an overt S argument ‘four days’ and it differs from the overt S argument of the main clause (the ‘*tatyâng* grass’). Similarly, in Example 392, the A argument of the cosubordinate clauses is not co-referential with the S argument of the main clause. As shown in these two examples, and as has been recognized in the literature, an overt referent in a multi-clause construction seems to break the iconicity of topic continuity (cf. Haiman 1981; also Aikhenvald 2023, 76f.), whereas ellipsis of the topic would commonly imply topic continuity.

- (391) “*kîya kaprêi hana gũ tatyâng khing bra lêzâbiya*” *la*
 [[*kîya kaprêi hana gũ*]_{cosub.cl} *tatyâng khing bra* *lê-zâ-biya*
 CLF:days four come NF a.grass one CLF:generic grow-DIR-FUT.ALLO
la]_{main.cl}
 say
 ‘“After four days a *tatyâng* grass will grow”, she said.’ [TA22_fishstory 143 256]

- (392) we mề laya gũ we ning khri hî hayũ ne khìzâ (...)
 [[we mề]_i la-ya gũ]_{cosub.cl} [we ning khri]_j hî hayũ ne khì-zâ]_{main.cl}
 DEM person say-ALLO NF DEM 1P a.clan EMPH first only reach-DIR
 ‘If anybody_i says (anything), only we, the Khri_j, got (there) first.’ [TA24_land 119
 216]

Example 393 further illustrates the complications of ellipsed arguments in multi-clause constructions and that they do not unequivocally signal topic continuity, especially if there are other potential referents. Crucially, an ellipsed argument here implies an argument that also includes preceding referents (cf. *partial reference* in Hampe and Lehmann 2013). This example is a summary of the main feats that led to the imprisonment of a thief: the narrator’s grandfather threw ash at the criminal’s face while the police officers and the narrator’s grandfather took hold of him. The A argument of the two cosubordinate clauses (‘tossing ash’ and ‘wrestling’) is the narrator’s grandfather. However, the A argument of the main clause (‘they caught him’) refers to the joint effort of the police and the narrator’s grandfather, i.e. it also contains the topic of the cosubordinate clauses. A clearer case is illustrated in Example 394, in which a tiger helplessly tries to count an ever-disappearing school of fish. The omitted A argument of the subordinate clause (the tiger) is not co-referential with the omitted S argument of the main clause (the school of fish).

- (393) we gũ we nya gũ pùga gũ tîya gũ rôrêya
 [we [∅_{Ai} gũ we nya gũ pùg-ya gũ]_{cosub.cl} [∅_{Ai} tî-ya gũ]_{cosub.cl}
 so ash DEM face LOC toss-ALLO NF wrestle-ALLO NF
 ∅_{Ai+j} ∅_P rô-rê-ya]_{main.cl}
 catch-PL.II-ALLO
 ‘So, by throwing_i this ash to his face, by wrestling_i, they_{i+j} caught him.’
 [TA22_CLAN2 93 186]

(394) “ekhün gâqing” la tsâitsâi we la gũthĩ tshâwhamâmâ khrâk

[$\emptyset_i$ e = khün kâqing la tsâi~tsâi we la gũ-thĩ]_{sub.clause} $\emptyset_j$

DEM = one two say count~PLURACT DEM say NF-CONC

tshâw-amâ-mâ khrâk]_{main.cl}

move.fast-get.lost-ITER EVI

‘(The tiger_i) kept counting “one, two”, (but the fish_j) kept moving and disappearing.’ [TA22_tiger 35 102]

In sum, the omission of arguments in multi-clause constructions does neither provide evidence for the grammaticalization of a privileged argument patterning nor for topic continuity, which is dependent on pragmatic factors.

8.3 Topic and focus

As discussed in §8.2, the language shows little signs of grammaticalizing referent identification through a particular alignment. One pragmatically-oriented way to think about grammatical relations in a language is through the concepts of *topic* and *focus*. *Topic* refers to those entities that are known to the hearer, the “center of attention” (Li and Thompson 1976, 464) while the *focus* is the missing information that the speaker assumes the hearer wants to know (cf. Andrews 2007a, 149f). As described in Figure 8.3, the topic occurs before the focus on natural discourse in Tawrã.

Table 8.3: Structure of natural discourse

Topic	Focus

In pragmatically neutral scenarios, i.e. when the speaker does not wish to *package* the information e.g. through a change in word order (Foley 2007, 363), but also in elicitation or where there is little shared knowledge, the topic can be readily associated with the S/A.⁴ According to Givón (2001a, 200), human language is biased to consider

⁴This is commonly discussed as *subject* in the literature (a *starting point*, cf. Chafe 1994, 83f. also Li and Thompson 1976, 459).

the (animate) actor-instigator of an event as the most topic-worthy in discourse (also cf. Lambrecht 1994, 131). The pragmatically unmarked configuration topic-focus is found in both verbless (Example 395) and verbal clauses (Example 396) in Tawrã.

(395) *hã mîyã manyúyã*

[*hã mîyã*]_{top} [*manyú-yã*]_{foc}

1s wife a.clan-FEM

‘My wife is Manyu (i.e. a female member of the Manyu clan).’ [251122_JO_eli2 2775]

(396) *kutyêmba we mlá mề we la matyô*

[*kutyêmba we*]_{top} [*mlá mề we*]_{foc} *la matyô*

a.name DEM village.MOD people DEM say maybe

‘Kutyemba told the village people.’ [TA22_kutyemba_phla 35 125]

It should be noted that a S/A that coincides with the topic is prone to be omitted, as it is understood from the context (the case of omitted arguments in multi-clause constructions was discussed in §8.2.5). In Example 397, the A argument *natya* ‘my grandfather’ is the topic of the second main clause. As it is known and retrievable, it is omitted.

(397) *we natya we ji asâng nyu îng dukan sukan bîsnes jiwina hî weli kamazâna*

[*we na-tya we ji*]_i *asâng nyu îng dukan sukan*

DEM 1s.POSS-grandfather DEM EMPH alive FOC a.bit store(<IA) ECHO

bîsnes ji-wi-na hî]_{main.cl} [*weli* *∅* *kama-zâ-na*]_{main.cl}

business(<E) do-AMB-FACT EMPH DISC.DEIC.DEM be.lost-DIR-FACT

‘(...) my grandfather, while he was alive, he did some business here and there. And like that, he passed away.’ [TA24_land 149 273]

Topic and S/A, however, do not necessarily coincide (cf. Lambrecht 1994, 118). Several East and Southeast languages⁵, and we may add Tawrã, make use of a construction

⁵Cf. Burmese (Jenny and Hnin Tun 2017, 252), Turung (Morey 2012, 240f), Japanese, Lisu (Li and Thompson 1976) or Lao (Enfield 2007, 96f; Enfield 2009). Equivalent constructions are found outside of Asia as well, e.g. in Qaqet (Hellwig 2019, 245).

in which the topic is outside the basic clause structure, i.e. it is not determined by the main verb but rather by pragmatic functions (Andrews 2007a, 154f; also cf. Do-Hurinville and Dao 2019, 410f; Foley 2007, 442f). These languages were described as being *topic-prominent*, i.e. where pragmatics, in the form of topic-focus constructions (also known as *topic-comment*), is the main drive to disambiguating role and reference of arguments instead of a grammaticalized alignment of arguments. Furthermore, these languages often share several typological features, such as being verb-final, having a *double subject* construction, i.e. a co-occurrence of the topic and a same-reference argument (e.g. ‘my uncle, he is the best’) and the absence of passive constructions (Li and Thompson 1976, 471f).

The remainder of this section will discuss two constructions that alter how a message is delivered to the hearer, i.e. how the semantic content relates to the general information structure of a clause: the topic constructions (§8.3.1) and the focus construction (§8.3.2).

8.3.1 Topic construction

There are two markers associated with the topic: the general topicalizing morpheme *byě* is discussed in §8.3.1.1. The contrastive topicalizing morpheme *ma* is discussed in §8.3.1.2.

8.3.1.1 Topic construction with *byě*

A speaker may employ the topic construction to raise a “matter of current concern” (cf. Lambrecht 1994, 150) or the “theme of the discourse” (Li and Thompson 1976, 464; Comrie 1989, 64). The topic position may be filled by an argument, an adjunct, a predicate phrase or a whole clause. This last point is relevant, as topicalizing a whole sentence implies breaking the iconic prosodic prominence of the topic argument (Lambrecht 1994, 234f). Topics lie outside the main clause’s structure in the sense that they do not have a semantic role concerning the main clause but rather have an extra-clause function related to discourse (cf. Givón 2001a, 198; also Andrews 2007a, 154f). Phonologically, however, there are no intonation cues that would suggest that the topic lies outside the main clause; it is either merely fronted or post-posed (cf. Widmer 2014, 671; also cf. Post

2007, 61-62).

The topic may be fronted or placed at the end of the clause (backed) (the latter is sometimes termed *antitopic* or *right-detachment*, cf. Lambrecht 1994, 204; also cf. Chafe 1976, 53). A fronted topic must take the topicalizing morpheme *byẽ* while a backed topic never takes this morpheme. The lack of a flagging device (morphological or prosodic) for backed topics might be due to the fact that shifting a constituent to a clause's final position is enough marking for the topic construction. Furthermore, it could also be the case that the topicalizing morpheme *byẽ* is by default interpreted as a conditional/temporal subordinator in clause-final position (cf. §10.4.1).

In Example 398, the topic is fronted and heralded by the topicalizing morpheme *byẽ*. As discussed in §8.3 and §8.2.3 the topic construction can alter the general tendency for S/A arguments to coincide with the topic. In this case, the fronted topic is a P argument (the A argument is ellipsed).

(398) *hã byẽ cângghanébim hâ!*

[*hã byẽ*]_{top/P} [$\emptyset_A$ *câng-hané-bi = m*]_{foc} *hâ*

1s TOP get-be.able-FUT.ALLO = NEG AFF

‘You’ll never get *me!*’ (lit. ‘as me, you won’t be able to get’) [TA22_kutyemba_phla 214 428]

The post-posed topic construction is illustrated in Example 399. It does not have a topicalizing morpheme. The topicalized arguments commonly have a demonstrative *we* (§4.3.3). This coincides with the statement by Li and Thompson (1976, 461) that a topic must be understood as definite, even if it is generic (also cf. Givón 2001a, 198, 472).

(399) *ning mâ bà gũ hadiga (...) we tükü we*

[*ning mâ bà gũ hadi-ga [we tükü we]*]_{NP-top}]_{main.cl}

1P mother father INS tell-PL.I DEM story DEM

‘Our parents told it, this story.’ (TA23_talyu_clan 6 29)

In Example 400, the identity of the ellipsed referents is made explicit by the use of the topic construction. Lambrecht (1994, 203) mentions that such forms of right-dislocation

should not be confused with an afterthought; the speaker is consciously exploiting a conventionalized construction that appends crucial propositional information that is not provided in the preceding sentence.

(400) brũ we kě nyu katyâga la mane (...) dâi we làng

[brũ we kě nyu katyâ-ga la mane [dâi we
burial.monument DEM LOC.INS FOC lurk-PL.I HSAY FILL(<I.A.) a.tribe DEM
làng]_{postposed-NP/Top}]_{main.cl}
COLL

‘They were lurking by the burial monument, those Idus.’ [TA22_CLAN2_122 230]

As mentioned above, the pragmatic effect of a topic construction is to highlight or contrast an argument. A topicalized argument becomes the center of discourse. Examples 401a and 401b illustrate this in a dialogue in which the speaker A) asks speaker B) about the past. The precise theme of the conversation is the relationship between the Tawrã and the incipient government after the Britishers left India. In Example 401a, speaker A) mentions both the Indian government (*sokarê*) and the Britishers (*kalà lyó*). In Example 401b, speaker A) initially says that the Britishers had reached the Tawrã area but then quickly goes back on his own words. He employs the topicalizing morpheme *byẽ* to contrast the potential referents: ‘as of the Britishers, they were not able to reach here (i.e. the Indian government *was* able to reach here)’.

(401) a. A): tya welü béqbé na we béqbé sokarê hanana (...) bíya kalà lyó ne hayũ
hana la

[tya welü béqbé na we béqbé sokarê hana-na]_{main.cl}
now and.then slowly INSIST DEM slowly government(<IA) come-FACT
[bíya kalà lyó ne hayũ hana la]_{main.cl}
past officer white only first come HSAY

‘And then slowly, right? Slowly the government came. In the past, only the Britishers (lit. white officers) came first, they say.’ [TA22_clan 20 78]

- b. A: (...) *kalà lyó hana-bô* (...) *kalà lyó byě ama khi cângm na*
 [kalà lyó hana-bô]_{main.cl} [[kalà lyó **byě**]_{top} ama khi câng = m
 officer white come-PFV.ALLO officer white TOP here reach be.able = NEG
*na*_{main.cl}
 INSIST
 ‘The Britishers (lit. white officers) had come (...) the Britishers (lit. white officers) could not reach here, right?’ [TA22_clan 24 89]

A similar case is illustrated in Example 402. The use of an overt topic highlights one referent (*bá.mề darang* ‘big clans’), an A argument, as the center of attention.

- (402) *bá.mề daráng byě bá.mề aqa pô mu kwâg mu bâ la*
 [bá.mề daráng **byě**]_{top} [bá.mề aqa]_{foc} pô mu kwâg mu bâ la
 clan big TOP clan small slave also dog also do HSAY
 ‘As of big clans, they would enslave (lit. like a slave like a dog) smaller clans, they say.’ [TA22_clan 12 56]

In Example 403, the topic (an S argument) is modified by an embedded relative clause. The pragmatic effect of topicalization is to highlight the fact that the child was born to a brother and a sister (a taboo).

- (403) *myúm mathi a mềnegadyibô byě ali nyu (...) asángna yüm*
 [myúm mathi a mề-nega-dyibô **byě**]_{topic/S(rel.cl)} ali
 elder.brother younger.sister son give.birth-DU-PFV.EGO TOP like.this
*nyu asáng-na yüm*_{main.cl}
 FOC grow-FACT NEG
 ‘As for the son that the brothers bore, (he) was like this, (he) did not grow.’
 [TA22_origin 42 125]

As discussed in §7.1, verbless clauses also follow a canonical topic-focus word order. Information structure allows a permutation of NPs in topic and focus slots. As an example, the topic and focus NP in Example 404a are exchanged in Example 404b. This mismatch

requires a somehow odd English translation to render and is pragmatically marked but grammatical in the language.

- (404) a. [azá byě]_{top} muhwa
king TOP rich
‘As for the king, he is rich.’
- b. [muhwa byě]_{top} azá
rich TOP king
‘As for (being) rich/the rich (person), it is the king.’ [both 220823_STBT_eli1 879]

8.3.1.2 Contrastive topic *ma*

The contrastive topicalizing morpheme *ma* flags a topic in contrast to other potential referents. Examples 405a, 405b and 405c are from the same story where a criminal is hiding in a cave. In Example 405a, the village people appoint a guard to play a gong in case the criminal would attempt to escape. In Example 405a, we learn that the criminal was aware of this. In Example 405c, the contrastive topicalizing morpheme is employed to mark the topic (the criminal) which stands in contrast to the guard.

- (405) a. tyûnagaya künüm mề khün agû khrâk
[tyû-na-ga-ya künüm]_{sub.cl} mề khing agû khrâk
play-FACT-PL.I-ALLO PRP person one guard EVI
‘There was one watchman in order to play (the gong).’ [TA22_kutyemba 393 740]
- b. ìho khrâk cya mề khün we kasâ matyô
[ì-ho khrâk]_{compl.cl} cya mề khing we kasâ matyô
be.there-STAT EVI 3 person one DEM know maybe
‘He knew that there was one person.’ [TA22_kutyemba 397 748]

c. cya ma gâqa ìho gũ (...)

[cya **ma** gâqa ì-ho gũ]_{cosub.cl}

3 CONTR.TOP near be.there-STAT NF

‘He being nearby (the entrance of the cave) (...)’ [TA22_kutyemba 398 750]

The contrastive topicalizing morpheme *ma* can be employed to contrast referents even if there are no other arguments available in discourse; one is left to reconstruct the alternative referent based on the context or on common or iconic pairs (e.g. ‘past and present’ or ‘father and son’). In Examples 406a and 406b, the speaker discusses what the appropriate rewards for a task used to be based on a story of his ancestors. In Example 406a, we learn that the reward was one blade and a nestful of eggs. In Example 406b, the speaker comments on the blade, since good blades used to be highly prized in the past. He uses the contrastive topicalizing morpheme for the past as it lies in contrast with the (unstated) present, i.e. *back then (as opposed to now)*.

(406) a. wrá.jù pla tyû mana ayüw hàng laga (...)

[wrá.jù pla]_{NP} [tyû mana ayüw]_{NP} hàng la-ga

type.of.dao CLF:blade chicken egg CLF:man.made.nests give HSAY-PL.I

‘(He) gave (him) one blade of a dao, one nest full of eggs, they say.’ [TA24_land 147 269]

b. natya làng hadiga ki mane wrá.jù pla gê bîya ma tyâng daráng na

[na-tya làng hadi-ga ki mane]_{main.cl} [[wrá.jù

1s.POSS-grandfather COLL tell-PL.I COUNT.EXP FILL(<Asm) type.of.dao

pla gê]_{NP} bîya **ma** tyâng daráng na]_{VLC}

CLF:blade one past CONT.TOP very big INSIST

‘My ancestors said that, I mean, one blade of a dao, in the past was very expensive (lit. big), right?’ [TA24_land 294 543]

8.3.2 Focus particle *nyu*

The focus particle *nyu* FOC indicates a pragmatic assertion, i.e. a proposition that differs from the presupposition. It signals a piece of information that cannot be taken for granted, as it is unpredictable or pragmatically unrecoverable (cf. Lambrecht 1994, 206f.) and one which the speaker presumes the hearer wants to know (cf. Andrews 2007a, 150f.). Commonly, this implies a form of new or controversial information.

Examples 407a and 407b are from a story in which a magical *takwikã* fish transforms into a woman after jumping out of a pot where it was being kept. The focus particle *nyu* in Example 407b indicates that the noun *mẽ* ‘person’ represents an unpredictable piece of information.

(407) a. *dûzâ takã khẽ gũ*

dû-zâ takã khẽ gũ

jump-DIR pot LOC.INS LOC

‘It jumped (out) of the pot.’ [TA23_FishStory 51 108]

b. *mẽ nyu shyâhobô tîng kazûm we làng kadâng gadyû pra*

[*mẽ nyu shyâ-ho-bô*]_{main.cl} [tîng kazûm we làng kadâng

person FOC become-STAT-PFV.ALLO cloth cloth DEM COLL excellent

gadyû pra]_{main.cl}

ECHO good

‘(But) it actually became a person! (With) beautiful clothes.’ [TA23_FishStory 54 112]

In Example 408, the speaker introduces a central character of the story. The focus particle highlights the main clause *pra yûm* ‘he (was) not good’. This is probably the case as one would not expect the main character of a story to be an antihero.

(408) we macimâ hî ìna la pra yüm nyu

[we maci-mâ hî ì-na la]_{main.cl} [pra yüm **nyu**]_{main.cl}

DEM harm-ITER EMPH be.there-FOC HSAY good NEG FOC

‘He was a harmful person, they say. He was not good.’ [TA23_Tayang_clan 19 108]

The focus/contrastive topic particle commonly occurs with location adjuncts (§7.2.1.2) and Goal/Source extended core arguments (termed *E* cf. §7.2.1.1). This occurrence of the focus particle could be explained by the fact that places are likely to be the focus of the sentence, i.e. an unpredictable (often *new*) piece of information. In Example 409, the focus particle occurs with the adjunct ‘house’. In Example 410, it occurs with a Source *E* argument while in Example 411 it marks a Goal *E* argument. Note that the language does not formally distinguish between Goal and Source *E* arguments (as discussed in §8.1).

(409) hã lâlhak phlê we àng nyu wesâ-dyibô

[hã [lâlhak]_{rel.cl} phlê we]_{NP.P} [àng]_{Adj} **nyu** wesâ-dyibô

1s trick bag DEM house FOC forget-PFV.EGO

‘I’ve forgotten my trickster bag at home.’ [TA22_JAstory 23 90]

(410) ning daireksan we akhã nyu hana mane

ning daireksan we [akhã]_E **nyu** hana mane

1P direction(<E) DEM south FOC come FILL(<ASM)

‘Our direction came (from) the south (i.e. we came from the south).’

[TA23_Chaitom_clan 36 210]

(411) abli gê shyûlyang nyu abli bô

abli gê [shyûlyang]_E **nyu** abli bô

CLF:half.group one a.place FOC CLF:half.group go

‘Some went to Shulyang, some went.’ [TA22_clan 89 279]

As discussed in §3.5.1.2, the focus morpheme *nyu* often occurs with interrogative pronouns. In some cases, the focus morpheme has completely merged with the interrogative pronoun: it is phonologically integrated into the pronoun and has also lost its pragmatic function to the extent that some interrogative pronouns seldom occur without *nyu*. In Example 412, the focus morpheme highlights the interrogative pronoun ‘who’. In Example 413, the focus morpheme has merged with the interrogative pronoun *kadé* ‘how’.

(412) *azá we shyá nyu?*

azá we shyá nyu

king DEM who FOC

‘Who is this king?’ [TA23_FishStory 114 212]

(413) *tshûnem nyu kadé.nyu shyâna hô mu ning ahuge*

[*tshû-na = m nyu kadé.nyu shyâ-na*]_{embedded.cl} *hô mu ning*

thrive-FACT=NEG FOC why happen-FACT AFF also 1P

ahu-ge

ask-go.and.back

‘‘How come you haven’t expanded (in numbers)’’, we went to ask (them).’

[TA23_talyu_clan 35 216]

As discussed in §7.3.1.1.1, the focus particle also occurs in the alternative construction *kya* (*either X or Y*), where it reinforces the contrast between two options in a question.

The focus particle *nyu* is homophonous with the 2nd person singular pronoun *nyu* (§3.5.1.1). This seems to match the development in Classical Tibetan, where the topicalizer *ni* is possibly related to the reconstructed form of the 2nd person singular pronoun for Proto-Tibeto-Burman **naŋ* (Matisoff 2003, 639).⁶ Note that in Classical Tibetan the 2nd person singular pronoun is *khyod*, seemingly unrelated to the mentioned form. A more striking example is Milang, where both the focus marker and the pronoun for 2nd person singular are *ni*. However, *ni* is also the form of the declarative (*factitive* for this grammar)

⁶I am thankful to Agnes Conrad from the University of Sydney for pointing this out to me.

clause-final particle in the language (cf. Modi 2017, 211, 272, 383). This isomorphism will be investigated in a future project.

9 Particles

Particles have been introduced in §3.5.2 and were defined as usually small, non-inflectional functional morphemes that are syntactically dependent (i.e. they cannot head phrases), occur at the right edge periphery and take multiple types of phrases under their scope. This broad definition includes morphemes with high functional diversity.

This chapter is organized in the following fashion: modal particles are described in §9.2. Fillers are overviewed in §9.3. Interjections are overviewed in §9.4. Postpositions are described in §9.5. Non-final markers are listed in §9.6, although their functions in multi-clause constructions are discussed in a dedicated chapter (§10). The description of modal particles will entertain the largest part of this chapter while the discussion on fillers and interjections is, comparatively, more superficial. Further descriptive work will fill that gap. §9.1 will present some introductory points regarding particles.

9.1 Preliminaries

Particles resemble clitics in their phrasal scope and categorical unrestrictedness (Bickel and Nichols 2007, 174). However, the phonological status of particles as independent words (i.e. they have tone/stress) distinguishes them from the common use of the term *clitic* (and *particle*) for phonologically deficient units (Aikhenvald 2003, 56).

Some particles occur in two different sub-categories, e.g. non-final marker/postposition *gũ* (§9.6). These can be considered to be the same morpheme, even though they have a different distribution, i.e. they head different phrases (cosubordinated clauses and NPs). In this grammar, the same particle is sometimes discussed in different subsections if it exhibits differentiated syntactic distribution. The problem of morpheme identity is partially solved through notes and cross-references to corresponding sections.

Concerning the scope of particles, their syntactic promiscuity is sometimes limited by distribution. In Example 414, the non-final marker *gũ* is incompatible with the modal particle *matyô* on the cosubordinated clause, either in the order *gũ matyô* or *matyô gũ*. Although both particles are from different subgroups (modal and non-final markers), their adjacent co-occurrence is problematic, as the non-final marker flags a clause as *non-final* while the modal particle must occur at the very end of the (*final*) main clause. In contrast, in Example 415, the non-final marker *gũ* marks the cosubordinated clause, while the main clause can be safely heralded by the modal particle *matyô* without causing interference, as in the previous example.

- (414) **cya iskul bôya gũ matyô/ matyô gũ àng ì yüm*
 [*cya iskul bô-ya gũ matyô*]_{cosub.cl} *àng ì yüm*]_{main.cl}
 3 school(<E) go-ALLO NF maybe house be.there NEG
 Intended: ‘As he maybe went to school, he was not at home.’

- (415) *prûli gũ hwé̃gana matyô*
 [*prû-li gũ*]_{cosub.cl} *hwé̃-ga-na matyô*]_{main.cl}
 burn-REC.PST NF see-PL.I-FACT maybe
 ‘After burning (the forest) they looked for him.’
 [TA23_kutyemba_phla 196 389]

As a comparative note, the term *particle* has been used to mean different things in the literature. For Sasse (2015, 171f.), *particle* stands in paradigmatic opposition to the *full word*, i.e. particles often “lack (...) concrete meanings and merely serve to mark certain syntactic functions, acting as operators of morphosyntactic categories and relations” (idem). Subtypes of particles, in the author’s typology, include conjunctions, complementizers, and adpositions (idem, 207f). Furthermore, the term is sometimes employed as a catch-all term for “items that do not fit easily into syntactic and semantic generalizations about the language” (Zwicky 1985, 290). As recognized by Post (2007, 612), the problem with such a broad definition of particles is that it admits a problematic functional diversity.

In the grammar writing tradition of South, Southeast and East Asia, particles commonly refer to uninflected, phonologically free units (Bickel and Nichols 2007, 173; Coupe 2008, 131). Many languages make use of *sentence-final particles* that have a fixed clause-final position and often have illocutionary force, mark evidentiality, modality, focus or sentence types and refine pragmatic meanings. They are entrenched in conversation and seldom found in written or formal speech (cf. Enfield 2007, 41; Law 2002; for an overview cf. Enfield 2021, 144f).

9.2 Modal particles

Modal particles have been briefly introduced in §3.5.2.1. They affect the epistemic and pragmatic reading of the affected clause/phrase and make “illocutionary and interactionally grounded distinctions” (Enfield 2007, 41f). As such, they display an important level of functional and semantic diversity with some degree of functional overlap, as will become clear in this section. Compared with *affixes* (particularly verbal TAMN suffixes, §6.3) as well as other particles like postpositions (§9.5) or non-final markers (§9.6), modal particles are relatively optional and grounded in informal speech.

Modal particles that mark epistemic categories such as inferential, reportative or evidential are associated with 3rd person arguments (i.e. allophoric forms, cf. §6.3.1.1) whereas an *assertor* who is conscious, involved and an experiencer of the actions commonly does not require ancillary devices to assert his/her experience in speech - unless it is pragmatically marked. Syntactically, most modal particles can only occur on a final clause (cf. §7) while others can also occur in phrases. Furthermore, there can be no more than one modal particle per clause/phrase.

This section will describe most modal particles found in the language so far. These are: *matyô* ‘maybe’ (§9.2.1), *dě* ‘probably’ (§9.2.2), *pà* ‘unfortunately’ ALAS (§9.2.3), *khrâk* ‘non-visual evidential’ EVI (§9.2.4), *thâ* ‘sadly/ for all I care’ (§9.2.4), *delé* ‘it seems’ (§9.2.6), *kho* ‘controversial’ (§9.2.7), *ki/kale/kalai* ‘counter-expectational’ COUNT.EXP (§9.2.8) *ta* ‘advisative’ ADVS (§9.2.9). The hearsay particle *la* will be discussed in §10.7.1. The observational negative *shyüw* has been discussed in §6.4.2. Table 9.1 displays a list of all modal particles so far found in the language. Further research focused on naturalistic

data is bound to find many more modal particles.

Table 9.1: Modal particles

Modal particle	Meaning	Reference
<i>matyô</i>	maybe	§9.2.1
<i>dě</i>	inferential evidential INFR	§9.2.2
<i>pà</i>	‘unfortunately’ ALAS	§9.2.3
<i>kh râk</i>	non-visual evidential EVI	§9.2.4
<i>thâ</i>	‘sadly/ for all I care’	§9.2.5
<i>delé</i>	‘it seems’	§9.2.6
<i>kho</i>	‘controversial’	§9.2.7
<i>ki/kale/kalai</i>	counter-expectational COUNT.EXP	§9.2.8
<i>ta</i>	advisative ADVS	§9.2.9
<i>la</i>	hearsay HSAY (also verb ‘say’)	§10.7.1
<i>shyüw</i>	observational negative	§6.4.2

Some modal particles have epistemic functions that fall under the domain of evidentiality (cf. Chafe and Nichols 1986). Evidential markers are often subdivided into *direct* and *indirect* in the literature (Bergqvist and Kittilä 2019, 6f; Aikhenvald 2006a, 25f): direct evidence is often associated with sensory perception (most prominently visual but also auditory) whereas indirect evidence may be a report or an inference. This distinction, however, is not particularly fortunate for Tawrã, as *kh râk* (§9.2.4) may convey both direct (non-visual) and non-direct (hearsay) evidence. Particle *dě* ‘probably’ (§9.2.2) conveys only indirect (inferential) evidence.

9.2.1 ‘maybe’ *matyô*

The modal particle *matyô* indicates a potential or irrealis mode. A clause with *matyô* is understood as a suggestion, a potential scenario or as the speaker assuming that the information provided is incomplete or imprecise.

In Examples 416 and 417, the speaker starts his account by saying that the story is

well-known and everyone has probably heard it already. This assumption is indicated by *matyô* ‘maybe’.

- (416) (...) *mề dũ la hanalyânabô matyô*
 [*mề dũ la hana-lyâ-na-bô matyô*]_{main.cl}
 person with HSAY come-BACK-FACT-PFV.ALLO maybe
 ‘Maybe this story has (already) come up among people.’ [TA22_clan2 2 21]

- (417) *katyânglyânabô matyô ning hadi*
 [*katyâng-lyâ-na-bô matyô [ning hadi]*]_{top}]_{main.cl}
 hear-back-FACT-PFV.ALLO maybe 1P story
 ‘Maybe it has been heard again, our story.’ [TA22_clan2 3 24]

Example 418 recounts how disputes were settled in the past. This is assumed to be a well-known fact and is not marked by *matyô* ‘maybe’. Conversely, in the same story, the speaker describes how Pashya, his clan’s mythological ancestor, would compensate the accuser party by taking from the accused. The uncertainty of this part is indicated by the use of the modal particle *matyô* in Example 419.

- (418) *ning gâmbra lûmkû he dângkûdûya we sokarêi bôshyüw yüm*
 [*ning gâmbra lûmkû he dâng-kû-dû-ya*]_{main.cl} [*we*
 1P gaonburha(<IA) inside EMPH finish-COMPL.II-need-ALLO DEM
sokarêi bô-shyüw yüm]_{main.cl}
 government(<IA) go-PERM NEG
 ‘We had to settle (lit. finish) (our disputes) under the gaonburha, we didn’t allow (them) to go to the government.’ [240122_clan 42 150]

- (419) *ning pâshya shìzâ labô matyô*
ning pâshya shì-zâ la-bô matyô
 1P a.name take-DIR HSAY-PFV.ALLO maybe
 ‘Our Pashya took (from them), they say, probably.’ [TA22_clan2 44 157]

In Examples 420a and 420b, the speaker recounts a family arrangement after a brother died and, according to custom, a clan's member, preferentially a brother, ought to remarry the widow. As it turned out, the person who remarried the widow was her classificatory son. This implausible arrangement is described using the modal particle *matyô*, as it escapes normality.

(420) a. (...) cya we jüm gũ (...) a we dũ pûmru nyu shyâ we nyu ma nyu harûmna
na

[cya we jüm gũ]_{VLC} [a we dũ pûmru nyu shyâ]_{main.cl}

3 DEM true INS child DEM with younger.brother FOC happen

[we nyu ma nyu harûm-na na]_{main.cl}

DEM FOC mother FOC remarry-FACT INSIST

‘(...) he, in truth, (...) this child was (his) brother and he remarried his mother, right?’ [TA24_land 173 319]

b. bâqa mîya nyu matyô âm

bâqa mîya nyu matyô âm

uncle wife FOC maybe AFF

It was the uncle's wife! [TA24_land 175 325]

As an epistemic marker of potential mode, *matyô* is suited for folk stories, for those events often transpire in a narrative world (Aikhenvald 2006a, 310). It should be noted, however, that *matyô* is not the only modal particle allowed in stories, as e.g. auditory evidential *khrâk* (§9.2.4) may also occur. In Example 421, the narrator refers to a mythological time, which is in line with the irrealis mode prompted by *matyô*.

- (421) (...) myêqẽ tadyâw khu tadyâw wegũ ma kadége mề kadâm.kara tawrã ìya matyô
myêqẽ tadyâw khu tadyâw wegũ ma kadége mề kadâm.kara
 beginning season FS season and.then CONTR.FOC how person proto.clan
tawrã ì-ya matyô
 a.tribe be.there-ALLO maybe
 ‘In the beginning, ehm, how many *Kadam-kara* Tawrãs might have been there.’
 [TA23_newmyth 4 15]

Finally, the end of that episode in the story is often heralded by *khú-biyà* ‘it will be enough’. The use of *matyô* in this case conveys the impression of the speaker that the story is not complete as there are probably other aspects of the story that weren’t told or could have been told differently.

- (422) ah we làng khúbiya we matyô
ah we làng khú-biya we matyô
 ah DEM COLL be.enough-POT.ALLO DEM maybe
 ‘Ah, these will be enough, maybe.’ [TA22_origin 45 162]

9.2.2 Inferential evidential *dề*

Modal particle *dề* indicates that the clause is an inference based on evidence, i.e. the speaker has some information to presume that the utterance is true (cf. Aikhenvald 2006a, 163). In Example 423, the speaker makes use of *dề* based on some sort of evidence, e.g. the house door might be locked.

- (423) bôbô đề
bô-bô đề
 go-PFV.ALLO INFR
 ‘He must have gone.’ [191024_JOJA_eli1 32]

In Example 424, the speaker makes use of *dề* to indicate that enough time has passed for the referent to have reached a certain place. He might not know this for sure; it is an inference.

- (424) cya tya Tajâw khîli dề
 cya tya Tajâw khî-li dề
 3 now a.place reach-REC.PST probably
 ‘He must have reached Tezu by now.’ [290924_STBT_eli1 146]

9.2.3 ‘unfortunately’ *pà*

The modal particle *pà* ‘unfortunately’ ALAS frames a clause as a sad or otherwise regrettable development. In Example 425, the speaker employs the modal particle to cat the main event (the dog getting his head stuck in a bottle) as unfortunate.

- (425) kwâg krû e bôtlo lûmkû nyu aprêho *pà*
 [kwâg krû e]_{NP} [bôtlo lûmkû nyu]_{NP} aprê-ho *pà*
 dog head DEM bottle(<E) inside FOC put.in-STAT ALAS
 ‘Unfortunately, the dog’s head is stuck in the bottle.’ [TA22_frog1 48 207]

9.2.4 Non-visual evidential *kh râk*

The modal particle *kh râk* is likely derived from the verb *takhrâk* ‘make noise’. As such, it provides direct evidence about a state of affairs (Aikhenvald 2006a, 26f.). This is commonly employed for auditory perception but it may also be employed for other non-visual senses, such as gustatory evidence. Furthermore, *kh râk* is also employed as a reportative without a direct reference to sensory perception (there is thus some overlap with hearsay *la*, cf. §10.7.1). For the sake of simplicity, it is glossed as EVI, although it conveys all non-visual perceptions.

In Example 426, the death of a pig is expressed using the auditory evidential, as its squeals are notoriously loud. In Example 427, the speaker cannot hear her uncle due to the inclement rain. Here the evidential stands for the inability to hear. In Example 428, the modal particle is employed as part of a story where the village people were shouting. In this case, there is no actual auditory evidence but rather this is extrapolated

to a story, since the shouting of the villagers would have triggered this particle. The use of the modal particle *khrâk* as a reportative will be discussed below.

- (426) blei sêi khrâk
 blei sêi khrâk
 pig die EVI
 ‘A pig is dying (it can be heard).’ [310322_eli1 280]

- (427) ja làng nakaw katyêngnam khrâk!
 ja làng na-kaw katyêng-na=m khrâk
 what COLL 1s.POSS-uncle hear-FACT=NEG EVI
 ‘I can’t hear anything, uncle!’ [140722_DI_eli2 1671]

- (428) cya “bônane” la khrâk âm
 cya [bô-na-ne]_{dir.speech} la khrâk âm
 3 go-FACT-INT say EVI yes
 ‘He said: “I will go”, yes.’ [TA24_land 239 446]

Apart from auditory perception, the evidential *khrâk* occurs with all other non-visual senses. In Example 429, it is employed with *nan̄gtyâ̄ng* ‘smell’. In Example 430, the villagers coil up a long rope from a cave where a trickster is hiding. This trickster has made them believe he is hiding very deep, where salt mines lie. When asked to rub salt on the tip of the rope to prove that he is truly in the depths of the cave, the trickster smears his own booger. One villager finally licks the tip of the rope and confirms it is salty, unaware that it was the trickster’s booger. In brief words, the evidential is employed with a verb of gustative perception. Finally, in Example 431 the evidential occurs with a verb of physical perception.

- (429) yu nan̄gtyâ̄nga khrâk
 yu nan̄gtyâ̄ng-ya khrâk
 rice.beer smell-ALLO EVI
 ‘It smells of rice beer.’ [070923_STBT_eli1 503]

- (430) (...) lyô hawrēya khrâk shyaw jümmâ khrâk
 [lyô hawrē-ya khrâk]_{main.cl} [shyaw jüm-mâ khrâk]_{main.cl}
 lick try-ALLO EVI be.salty be.real-ITER EVI
 ‘(...) he licked (the rope) (and) it was very salty.’ [TA22_kutyemba_phla 109 241]

- (431) hã khudâng khrâk
 hã khu-dâng khrâk
 1s have.fever-COMPL.I EVI
 ‘I have a fever.’ [111222_ST_BT_dict 616]

As discussed at the beginning of this section, the non-visual evidential *khrâk* occurs as a form of reportative, i.e. it conveys information as having been heard. As such, it can be found in stories¹ when the speaker wishes to portray something as ‘being heard’. In Example 432, the narrator talks about the story of the first *puja* performed by the Tawrã: the Tawrã did not know how to worship the gods at the dawn of time. Similarly, in Example 433 the speaker tells about the kinship ties between two people. The use of the non-visual evidential conveys it as being heard. Finally, the evidential *khrâk* may co-occur with the verb *la* ‘say’, which is commonly employed to report direct speech like the *la la* ‘double say’ construction where the verb ‘say’ co-occurs with semi-grammaticalized hearsay *la* (cf. §10.6.2), as in Example 428.

- (432) kûmya bwìya kasâdyüm khrâk
 [kûm-ya bwì-ya]_{compound.stem} kasâ-d=yüm khrâk
 bow.down-ALLO dance-ALLO know-EGO.HAB=NEG EVI
 ‘We didn’t know how to worship (lit. bow down and dance).’ [TA22_origin 119 347]

¹A similar case was described for *matyô*, the irrealis modal particle in §9.2.1.

- (433) (...) cya bi mru myûm mathî ga khrâk we
 cya [bi mru]_{sym.compound} [myûm
 3 elder.sister younger.brother[.male.persp.] younger.brother[.female.persp.]
 mathî]_{sym.compound} ga khrâk we
 younger.sister REFL EVI DEM
 ‘(...) these siblings were true siblings.’ [TA24_land 239 446]

- (434) we tayâga khrâk mlá mẽ we
 we tayâ-ga khrâk [mlá mẽ we]_{postposed.NP}
 DEM shout-PL.I EVI village.MOD person DEM
 ‘They were shouting, these village people.’ [TA22_kutyemba_phla 109 241]

9.2.5 ‘sadly/ for all I care’ *thâ*

The modal particle *thâ* has two semantic effects: on the one hand, it casts an action as detrimental, especially for the speaker. It may thus be used to express regret about an event. Furthermore, it expresses the speaker’s disappointment in someone else’s action, here termed ‘sadly/for all I care’.

In Example 435, the speaker frames a tragic event as negative and sad. The semantic effect of the particle *thâ*, however, is often read as being detrimental to the speaker. It thus signals a form of regret. In Example 436, the speaker regrets the fact that someone got a job; we may interpret e.g. the speaker hoped s/he him/herself would get the position. Similarly, in Example 437 the speaker regrets that someone got married - and not the speaker!

- (435) cya ma bà hîq a gû shîkû thâ
 [cya [ma bà hîq]_{sym.nom.compound}]_{NP} [a gû]_{sub.cl} shî-kû
 3 mother father EMPH be.a.child NF die-COMPL.II
 thâ]_{main.cl}
 sadly
 ‘Sadly, s/he lost both parents when s/he was young.’ [281124_BAB_eli1 1916]

- (436) kamho thâ
 kam-ho *thâ*
 work(<IA)-STAT sadly
 ‘Sadly, s/he got the job (i.e. and not me).’ [field notes 211024]

- (437) mîya bréiho thâ
 mîya bréi-ho *thâ*
 wife marry(male.prsp.)-STAT sadly
 ‘Sadly, he got married (i.e. and not me).’ [field notes 211024]

The modal particle occurs in a sub-construction [V-*ya* V *thâ*] that expresses disappointment or regret regarding someone’s action. In Example 438, the speaker is disappointed that someone wants to buy something.

- (438) bréiya bré thâ
 bréi-ya *bré thâ*
 buy-ALLO buy for.all.I.care
 ‘If you want to buy, just buy it! (I’m disappointed).’ [281124_BAB_eli1 1830]

The modal particle *thâ* is homophonous with the verb *thâ* ‘eat’. A semantic link is not easy to elucidate but could follow the metaphor that something that is eaten is gone, and that is a negative thing. There is a noteworthy development of the verb ‘eat’ in some Kiranti languages, cf. Yakkha, where the verb *ca* ‘eat’ has grammaticalized into a reflexive marker as well as an autobenefactive derivation (Schackow 2015, 303-307). Verbs of eating and drinking are prolific in figurative extensions, although such developments are not very common cross-linguistically (cf. Newman 2009).

9.2.6 ‘it seems’ *delé*

The modal particle *delé* is employed to express a lack of confidence in one’s statement. Similarly to *dě* (§9.2.2), it may express an inference although with a low level of

confidence from the speaker. In Example 439, the speaker is telling a story from the past in which dholes (*Cuon alpinus*) would hunt and take away rice bags. As he is not sure about the details of the story, he employs the modal particle *delé*.

- (439) *thâ.ming shìya dũwa dyüw delé*
thâ.ming shì-ya dũw-ya dyüw delé
 ration take-ALLO need-ALLO isn't.it it.seems
 'They had to take the rice bags, it seems.' [TA23_talyu_clan 51 318]

In examples 440 and 441 two speakers were asked to recount what they saw from a storybook (Mayer 1969). The use of the modal particle *delé* is consistent with the idea of providing some information without certainty, as they do not know the story and make observations based on what they see. In Example 440, a shining moon prompts the remark that it is getting late, while in Example 441 the body posture of the protagonist suggests that he is angry. In both cases, a lack of certainty is behind the use of the particle *delé*.

- (440) *lyâng kanúq delé*
lyâng kanú-q delé
 evening be.dark-CONT it.seems
 'It seems to be getting dark.' [TA22_frog1 17 73]

- (441) *a we khamyû mu jiho delé*
a we khamyû mu ji-ho delé
 child DEM anger also do-STAT it.seems
 'This child is also angry, it seems.' [TA22_frog1 59 271]

9.2.7 'unexpected/controversial' *kho*

The modal particle *kho* highlights a piece of information (a NP or a clause) as unexpected or controversial. It commonly precedes a negated clause, i.e. a negated action

that stands against one's expectations or common tenets (cf. Dixon 2009, 28f.). The particle *kho* carves out two syntactic niches within that semantic domain that are worth discussing: It highlights either a a) NP or b) whole clause, usually a complement clause.

When the particle *kho* flags a NP, it is understood as unexpected or controversial, in a sense similar to the effect of the focus marker *nyu* (cf. §8.3.2). The clause is commonly negated. In Example 442, the particle occurs within a NP headed by the discourse deictic demonstrative *weli*: 'contrary to this, it was not like that'.

(442) *weli kho ã yüm*

[*weli* *kho*]_{NP} *ã* *yüm*

DISC.DEIC.DEM even.though be.there NEG

'It was not like this (in spite of what you say).' [291024_STBT_eli2 817]

Example 443 requires some context: a child who does not go to school (as s/he is expected to) and someone asks his/her parents. They may reply with Example 443, where 'school' is controversial. Similarly, in Example 444, the use of *kho* with a language renders this lack of knowledge as unexpected; it would be appropriate if e.g. the speaker would know both Kaman and Kera'a, but not Tawrã, or if s/he were to reside in a Tawrã village for many years without learning the language.

(443) *we iskul kho bô yüm*

[*we* *iskul* *kho*]_{NP} *bô* *yüm*

DEM school(<E) even.though go NEG

'S/he does not go to school (although he should).' [281124_BAB_eli1 388]

(444) *tawrã kho kasâdyüm*

[*tawrã* *kho*]_{NP} *kasâ-d=yüm*

a.language even.though know-HAB.EGO = NEG

'I don't know Tawrã, though.' [field notes 211024]

The particle *kho* may also flag whole clauses if they are the argument of a verb, i.e. in a complement construction. Table 9.2 displays the structure of *kho* in a complement

construction: it includes the particle *kho* ‘even though’ modifying a complement clause (hence V_1). The complement clause is the argument of a complement-taking verb (hence V_{CTV}) which is commonly *ji* ‘do’, although any verb can be employed (cf. §10.6). The construction must be negated, as it expresses the *common sense* which contrasts with the preceding clause. The negator, however, does not have to occur on the complement clause; it is commonly raised and takes the clause-final position within the matrix (cf. Noonan 2007, 100f).

Table 9.2: Structure of the *V kho V yüm* construction

[[V_1	<i>kho</i>] _{compl.cl}	$V_{2/CTV}$	<i>yüm</i>]
	even though	do	NEG

In Example 445, the speaker expresses that there is unexpectedly no rain; we may assume that there are big clouds or other signs of imminent deluge - but no rain. Similarly, in Example 446 the speaker expected the police to hit the thief - but this was not the case. This clause also includes two arguments within the complement clause.

(445) *karazâ kho ji yüm*

[*kara-zâ kho*]_{compl.cl} *ji yüm*

rain-DIR even.though do NEG

‘It is not raining (although it would seem like it would).’ [190923_JOJA_eli1 187]

(446) *pulis we akâwya phôk kho ji yüm*

[*pulis we*]_{NP} [*akâw-ya*]_{rel.cl} $\emptyset$]_{NP} *phôk kho*]_{compl.cl} *ji yüm*

police(<E) DEM steal-ALLO beat even.though do NEG

‘The police did not beat the thief (although one would have expected them to do so).’ [190923_JOJA_eli1 314]

Example 447a and 447a are from the same story. In Example 447a, the interviewer (termed A) asks about dholes (*Cuon alpinus*) and the widespread belief that they could speak. The interviewee (termed B) in Example 447b affirms that they could not speak and employs the *kho V yüm* construction to oppose this information to a common belief.

- (447) a. A: âccha tya tyu taprûw lade mane (...) taprûw byẽ maro nyu kya kadé.nyu
kya meho xxx

âccha tya tyu [taprûw la-de mane]_{main.cl} [taprûw byẽ
alright(<Hin) now FS dhole say-HAB.EGO FILL(<ASM) dhole TOP
maro nyu kya]_{main.cl} [kadé.nyu kya]_{main.cl} me-ho xxx
talk FOC Q how Q rear-STATE (inaudible)

A: ‘Alright, now ehm, these (we) call *taphruw* (dhole) I mean (...) concerning the dholes, did they talk or how? Being reared (...)’ [TA23_talyu 46 290]

- b. B: taprûw a we maro nyu kho ji yüm

[taprûw a we maro nyu kho]_{compl.cl} ji yüm
dhole child DEM talk FOC CONTR do NEG

B: ‘The dhole pups did not speak, though.’ [TA23_talyu 47 297]

Example 448 displays a different complement-taking verb: *ã* ‘be there’: the speaker talks about changing marital practices within his clan. The use of the *kho V yüm* construction implies that the marriage did occur, against the common practice. Similarly, in Example 449 the main clause is headed by the verb *kasâ* ‘know’. The use of the *V kho V yüm* construction in this case is in line with information management, which places new, controversial information against the backdrop of a tenet, common sense or the available evidence.

- (448) (...) bréibrûga kho ã yüm

[bréi-brû-ga kho]_{compl.cl} ã yüm
marry-be.allowed-PL.I even.though be.there NEG

‘(...) they would not be allowed to marry, though.’ [TA23_talyu 74 459]

- (449) hanúk nyu hana kho kasâ yüm

[hanúk nyu hana kho]_{compl.cl} kasâ yüm
where FOC come even.though know NEG

‘I do not know whence we came, though.’ [190923_JOJA_eli1 162]

9.2.8 ‘counter-expectation’ *ki/kale/kalai/ke*

The modal particle *ki/kale/kalai/ke* indicates that a piece of information might be unexpected, unknown or otherwise relevant for the hearer. The alternative forms of this particle (*ki/kale/kalai/ke*) may be due to dialectal variation but that requires more study. In Example 450, the particle highlights the fact that the speaker knew something and the hearer seems not to have been aware of it. Similarly, in Example 451 the sentence includes the modal particle *ki* after the negated predicate as the speaker assumes that this information might not be known to the hearer. Finally, in Example 452 the fact that both the tiger and the person were able to talk is presented as something unexpected or unknown to the hearer.

(450) *hã kasâ ki nyu hã ahuwa goji (...)*

[*hã kasâ ki*]_{main.cl} *nyu hã ahu-ya goji*

1S know COUNT.EXP 2S 1S ask-ALLO do

‘I would have known, though, if you had asked me (...)’ [291124_BAB_eli1 253]

(451) *hayũ blüi tabrẽ thâdyüm ki*

hayũ blüi tabrẽ thâ-d=yüm ki

past pig meat eat-HAB.EGO=NEG COUNT.EXP

‘In the past, we didn’t eat pig meat, though.’ [190923_JOJA_eli1 800]

(452) *tamyá ma mẽ ma kapâq maroga ki la (...)*

[*tamyá*]_{NP} *ma* [*mẽ*]_{NP} *ma kapâq maro-ga ki la*]_{main.cl}

tiger ADD person ADD alone talk-PL.I COUNT.EXP HSAY

‘(...) only the tiger and the man talked (i.e. could talk), they say.’ [TA22_tiger 15 47]

9.2.9 Advisative *ta*

The modal particle *ta* highlights a piece of information as correct, preferable and advisable. It is often employed in contrast with something else. This modal particle also occurs as a predicate derivation (cf. §6.2.3.1). This particle may occur in imperative clauses ('you better do X!') but also as a statement, asking for the interlocutor's consent ('it is better to do X, right?'). In my corpus, this particle occurs within phrases (viz. NP, VP) rather than whole clauses. It remains to be studied whether it also occurs clause-finally, as most modal particles in the language.

In Example 453, the use of the modal particle *ta* suggests eating rice (e.g. as opposed to something else or not eating at all). The particle occurs within the incorporated NP. The effect in the information structure of this modal particle in this example slightly resembles the highlighting effect of the focus marker (§8.3.2), as it commonly signals an unexpected or new referent. However, the semantic effect of the modal particle is more specific, as it is circumscribed to things that are advisable rather than contrasting *per se*.² In Example 454, the particle occurs both in an adverbial phrase (*darêk* 'directly') and within the incorporated nominal *tako* 'book'. Its semantic effect is to frame the written complaint as the optimal way to resolve a problem.

(453) *tapẽ ta thâna*

[tapẽ ta]_{NP} thâ-na_{nom.incorp}

food ADVS eat-IMP

'(It is better to) eat rice (e.g. as opposed to something else).' [221024_JO_eli1 1673]

²In this line, it should be noted that the modal particle *ta* bears a resemblance with Indo-Aryan focus particles, cf. Hindi focus particle *to* and Assamese *tu* or *ta* [tɔ]. The Assamese open-mid back rounded vowel normally becomes /o/ in loanwords in Tawrã, cf. *bojar* (from Assamese *bajār* [bɔ'd̪ɜ̃r]) 'market'. If it were an Assamese loan, we would expect **to/tu* and not *ta*.

(454) darêk ta tako ta bâna

[darêk ta] [[tako ta]_{NP} bâ-na]_{nom.incorp}

directly(<E) ADVS book ADVS make-FACT

‘(It is better to) go directly (and) file a complaint (lit. make a book).’ [TA24_land 46 77]

9.3 Fillers

A filler is a type of particle that works as a hesitation word. A speaker may use it as a pause to what s/he is about to say, much like American English *like*. Fillers commonly occur at the end of phrases and clauses. The most common particle is *mane*, an Assamese loan. Indefinite pronouns like *(sim)thâ* ‘something’ (§3.5.1.6) may also occur as fillers. In Example 455, the speaker makes use of the filler *mane* which occurs after NPs and at the end of the main clause. Similarly, in Example 456 the filler *mane* occurs after the deictic demonstrative, which stands for a propositional content from a previous clause (cf. §10.8.1).

(455) webo ning natya we mane gôbmen cya mane ahwê mane ne

webo [ning natya we]_{NP} mane [[gôbmen]_{NP} [cya]_{NP}

and.then 1P grandfather DEM FILL(<Asm) government(<E) 3

mane ahwê mane ne]_{main.cl}

FILL(<Asm) send FILL(<Asm) INSIST

‘And then our grandfather, I mean, the government, I mean, sent him, I mean, right?’ [TA22_CLAN2 81 162]

(456) weli (...) weli gũ mane ning we hanapêgana

weli weli gũ mane ning we hana-pê-ga-na

DISC.DEIC.DEM DISC.DEIC.DEM INS FILL(<Asm) 1P DEM come-separate-PL.I-FACT

‘And this (...) this I mean, we came separately.’ [TA22_clan 96 300]

9.4 Interjections

Interjections are particles that function as whole sentences and express the speaker's attitude and intentions. Morphologically, interjections are uninflected. Syntactically they are complete utterances that lack syntactic clause constituents. They are commonly apposed to clauses. We can distinguish between four types of interjections: interrogative, declarative, imperative and exclamatory. Table 9.3 illustrates the four types of interjections along with an example. Further work with naturalistic data is bound to find many more interjections.

Table 9.3: Types of interjections

Type of interjection	Example
interrogative	<i>a</i> 'what?'
declarative	<i>âm</i> 'yes!'
imperative	<i>chôq</i> 'let's (e.g. go)'
exclamatory	<i>ah</i> 'ah!'

In Example 457, the interjection *hâ* AFF is embedded in the complement clause (part of direct speech). Conversely, in Example 458 the interjection *chôq* stands for the utterance of the topic. It occurs apposed to the main clause, although it is understood to be embedded in the verb *la* 'say'. Similarly, in Example 459 the interjection is apposed to a verbless clause.

- (457) “tashî we we hanâna welü tamûn hrûga mung hô” la la
 [tashî we we hanâ-na]_{main.cl} [welü [tamûng hrû-ga mung hô]_{compl.cl}
 testicle DEM DEM squeeze-IMP and.then fire burn-PL.I FS AFF
 la la]_{main.cl}
 say HSAY
 ‘He said: “squeeze (your) testicles (and) the fire (will) surely burn!”, they say.’
 [TA22_tiger 138 306]

(458) thâlibô gũ la la mǎ́ mề we la “chôq”

[*thâ-li-bô* *gũ*]_{cosub.cl} *la* *la* *mǎ́* *mề* *we* *la*
eat-REC.PST-PFV.ALLO NF say HSAY village.MOD people DEM say

[*chôq*]_{dir.speech}]_{main.cl}

let's!

‘Having eaten, he said to the village people “let’s (carry the stone)”, they say.’

[TA23_kutyemba_phla_075 190]

(459) cya ander tũw gê ne

[*cya* *ander* *tũw* *gê*]_{VLC} *ne*

3 under(<E) CLF:party one INSIST

‘He is the other party, right?’ [TA24_land 23 42]

9.5 Postpositions

Postpositions affect the whole NP and have a fixed, post-NP position. They often mark adjuncts (§7.2.1.2), i.e. optional NPs, although they are dependent on the semantics of the verb and the general sense of the clause (cf. Andrews 2007a, 157f). Syntactically, they stand outside the NP. In Example 460, the postposition *dũ* ‘with’ modifies the NP. Similarly, in Example 461 the instrumental postposition *gũ* occurs after a classifier phrase, which is the last constituent of the NP (cf. §4.1).

(460) (...) macêi daráng daráng dũ

[*macêi* *daráng* *daráng*]_{NP} *dũ*]_{adu}

water big big with

‘(...) with much, much water (lit. big big water).’ [TA22_origin_69 202]

- (461) tawrá pla gê gũ
 [tawrá [pla gê]_{CLP}NP gũ]_{adju}
 sword CLF:blade one INS
 ‘with one single sword’ [field notes 211024]

It should be noted that postpositions do not function as case markers as in related languages (cf. §8.2).

Prosodically, postpositions are p-words, as they have their own stress/tone, however, they do not occur independently from the noun phrase, irrespective of the absence of the head noun. Some postpositions also occur as non-final markers, i.e. they have scope over whole classes (cf. Morey 2012, 336). Most prominent in this regard is the non-final marker/locative/instrumental *gũ* (§10.3).

NP-modifiers with a locational sense have an overlapping function with relator nouns (§3.3.1.1) although they differ syntactically, since relator nouns are *nouns* and modify a head noun through nominal compounding. However, the discussion of *kẽ* LOC.INS (in §9.5.2) will illustrate that diachronic developments can foster an overlap between these categories.

This section will discuss four postpositions: general locative/instrumental *gũ* (§9.5.1), ‘inside’ locative *kẽ* (§9.5.2), ‘with’ *dũ* (§9.5.4) and ‘for’ (*a*)*ba* (§9.5.5) and two constructions: comparative ‘like’ *dahi* (§9.5.6) and NP quantifiers ‘each, every’ (§9.5.7). Additive *ma* has been discussed in §4.4.1, as it conjoins NPs. Table 9.4 displays all postpositions discussed in this section.

Table 9.4: Overview of postpositions

Postpositions	Reference	Notes
general locative/instrumental <i>gũ</i>	§9.5.1	also cosubordination marker
‘inside’ locative <i>kẽ</i>	§9.5.2	
‘atop’ locative <i>kyã</i>	§9.5.3	
‘with’ <i>dũ</i>	§9.5.4	also co-temporality marker
‘for’ (<i>a</i>) <i>ba</i>	§9.5.5	
comparative ‘like’	§9.5.6	
NP quantifiers ‘each, every’	§9.5.7	
additive <i>ma</i>	§4.4.1	

9.5.1 General locative/instrumental *gũ*

The general locative/instrumental marker *gũ* is transparently related to the non-final marker *gũ* (§10.3). As a locative postposition, *gũ* indicates location in a broad sense, i.e. without necessarily implying *within* or *atop* as opposed to more specialized markers and relator nouns. In Example 462, the speaker uses the locative postposition to indicate where the character is located. After the cosubordinated clause, the whole adjunct is mentioned again to specify the location. Similarly, in Example 463 the locative marker indicates the location where a snot was attached.

- (462) abang *gũ* tyěya *gũ* (...) kashũng *gũ* (...) áng kashũng *gũ* (...)
 [[abang *gũ*]_{adju} tyě-ya *gũ*]_{cosub.cl} [kashũng *gũ*]_{adju} [áng kashũng
 outside LOC wash-ALLO NF front.yard LOC house.MOD front.yard
gũ]_{adju}
 LOC
 ‘Washing (cloths) outside, at the front yard, at the house’s front yard (...)’
 [TA23_FishStory 108 203]

- (463) *mêqẽ kahîng tatyế cyâ mala gũ tadi madôhona*
 [[*mêqẽ kahîng tatyế cyâ mala*]_{NP} *gũ*]_{adju} [*tadi*]_{NP} *madô-ho-na*]_{main.cl}
 begin day rope lower tip LOC snot stick-STAT-FACT
 ‘A snot was attached to the tip of the rope of the first day.’ [TA22_kutyemba_phla 354 664]

The sense of location is also employed for time. In Examples 464 and 465, the locative indicates a temporal point. Suffixes and relator nouns that indicate location can readily be applied to a temporal sense; the limit between these two domains is porous in the language (cf. §3.3.1.1).

- (464) *bîya (...) câne tamyîm gũ (...)*
 [*bîya câne tamyîm gũ*]_{adju}
 past a.name time(<E) LOC
 ‘(In the) past, at Chane’s time (...)’ [TA22_tiger 27 1]

- (465) *hîng gũ bô yûm*
 [*hîng gũ*]_{adju} *bô yûm*
 day LOC go NEG
 ‘I don’t go during the day.’ [230122_JW_eli 995]

The postposition *gũ* also adds or highlights an instrument used in an action. In Example 466, the hammer used by *Drapkap* carries the instrumental marker. It should be noted, however, that instruments seldom take a postposition. This is illustrated in Examples 467a and 467b from a story in which a group of villagers are tricked into getting rid of a big stone. In Example 467a, the trickster tells the village people to tie up this big rock. In Example 467b, the instrument used for tying the rock (the *taraw* rope) is not marked by a postposition.

- (466) *drâpkap taplâ gũ khalâi tyûzâ*
drâpkap [*taplâ gũ*]_{adju} *khalâi tyû-zâ*
 a.name hammer INS land hit-DIR
 ‘Drapkap flattened (lit. hit) the land with a hammer.’ [201122_JA_eli1 1322]

(467) a. “ephlá pô ewe (...) rêbina (...)” la
e=phlá pô ewe rê-bi-na la
 DEM = stone.MOD rock DEM tie-EMPH-IMP say
 “‘Tie (...) this big rock!’ (...) he said.’ [TA22_kutyemba_phla 79 195]

b. we taráw tatyě rhě rêmâkũya
[we taráw tatyě]_{adju} rhě rê-mâ-kũ-ya
 DEM type.of.creeper rope pull tie-ITER-COMPL.II-ALLO
 ‘(They) pulled and tied (the rock) with a rope made of *taraw* creeper.’
 [TA22_kutyemba_phla 81 197]

The instrumental postposition occurs consistently with languages. In Example 468, one speaker tells another in which language he ought to tell a story using the postposition *gũ*.

(468) hwěya gũ ning tawrã gũ laya mane
[hwě-ya gũ]_{cosub.cl} [ning tawrã gũ]_{adju} la-ya mane]_{main.cl}
 see-ALLO NF 1P a.language INS say-ALLO FILL(<ASM)
 ‘By seeing it, you (should) say it in our Tawrã language.’
 [TA22_frog 5 31]

9.5.2 ‘inside’ locative *kě*

The morpheme *kě* is a locative postposition with a specialized meaning *inside*. In Example 469, the postposition indicates that the character is inside the stone’s crack. Similarly, in Example 470 the postposition implies that the action transpired within the village.

(469) we phlá krà kě mûtyû webo byě (...)
[we phlá krà kě]_{adju} mû-tyû webo byě
 DEM stone.MOD hole LOC.INS hide-DIR and.then TOP
 ‘He hid in the crack of the stone and then ...’ [TA22_kutyemba_phla 182 371]

(470) ning bá.mě kě (...)

[ning bá.mě kě]_{adju}

1P clan LOC.INS

‘In our clan (...)’ [TA22_clan 50 177]

Two features of the locative *kě* expose a hybrid membership within the postposition family. First, the locative may trigger a tonal change on the preceding noun (a characteristic of nouns, cf. §3.2.2), as in Example 471. This feature seems to occur irregularly among speakers. Furthermore, the postposition *kě* often occurs in a collocation preceding the general locative: *kě gũ* (LOC.ON LOC)³, as illustrated in Example 472. These two features suggest that the locative *kě* has been recruited from a noun, concretely a relator noun (§3.3.1.1), and still retains some nominal characteristics (cf. Jacques 2021, 363-364).

(471) (...) mlǎ mẽ mlǎ kě gāmra bâya bârěya

mlǎ mẽ [mlǎ kě]_{adju} gāmra bâ-ya

village.MOD person village.MOD LOC.INS gaonburha(<IA) make-ALLO

bâ-rě-ya

make-PL.II-ALLO

‘(...) a village person in the village, (they) designated (lit. made), they designated (lit. made) a gaonburha.’ [TA22_clan 37 132]

(472) (...) takǎ kě gũ macî mawiya gũ tshôna (...)

[takǎ kě gũ]_{adju} macî mawi-ya gũ tshô-na

pot LOC.INS LOC water fill.up-ALLO NF keep-FACT

‘(...) filling a pot with water he left it there (...)’ [TA23_FishStory_014 40]

³Such lexicalized forms are likely the origin of complex case/postposition combinations in the language family (cf. Schikowski in prep., 140-146).

9.5.3 ‘atop’ (*ta*)*kyã*

The postposition (*ta*)*kyã* (most commonly *kyã*) ‘atop’ specifies location as on the top of something. In Example 473, the speaker describes the odd formation of a group of animals (one atop the other) in a storybook (Rodriguez 2008). In Example 474, *kyã* is the preferred postposition for motion into water.

- (473) tamh we kyã mambaw ìho weli kyã tyû ìho
 [[*tahum we kyã*]_{adju} *mambaw ì-ho*]_{main.cl} [[*weli kyã*]_{adju} *tyû*
 bear DEM LOC.ON fox be.there-STAT DISC.DEIC.DEM LOC.ON chicken
ì-ho]_{main.cl}
 be.there-STAT

‘Atop the bear is the fox and atop it the chicken.’ [TA22_chicken_thief_30 114]

- (474) ja mu shyâ yûm delé macî kyã lyâ gû
 [*ja mu shyâ yûm delé* [[*macî kyã*]_{adju} *lyâ gû*]_{cosub.cl}]_{main.cl}
 what also happen NEG it.seems water LOC.ON fall NF
 ‘Nothing happened (to them), even after falling into water.’ [TA22_frog1_124 617]

9.5.4 ‘with’ *dũ*

The postposition *dũ* indicates an associated nominal referent, equivalent to the English preposition *with*, as well as possession. This postposition also exists as a non-final marker (§9.6). In Examples 475 and 476, the postposition *dũ* marks an adjunct of the clause.

- (475) pramu dũ ìde ye?
 [*pramu dũ*]_{adju} *ì-de ye*
 younger.brother with be.there-EGO.HAB Q
 ‘Do you live with your brother?’ [180122_AM_eli 1649]

- (476) pahwã dũ khilinega
 [pahwã dũ]_{adju} khili-nega
 frog with play(<IA)-DU
 ‘They play with the frog.’ [TA22_frog1 12 52]

The origin of this postposition is not clear. A link to the versatile verbs of position and location in Tani, cf. Proto-Tani *dùŋ ‘sit’, *dóŋ ‘lie down’ and *dàk ‘stand’ (Post 2008, 135) cannot be discarded. A diachronic study would have to make sense of a double development as a postposition and a non-final marker.

9.5.5 ‘for’ (a)ba

The postposition (a)ba expresses the beneficiary of an action, akin to the English *for*. In Example 477, the postposition (a)ba flags the NP which will benefit from the action; for whom it is intended.

- (477) hã nyu ba tashēne
 hã [nyu ba]_{adju} tashē-ne
 1s 2s for sing-INT
 ‘I’ll sing for you.’ [110823_STBT_eli2 25]

As discussed in §6.2.2.3, the bene/malefactive predicate derivation -hàng makes use of the postposition (a)ba ‘for’ to emphasize a *beneficiary* (i.e. writing **for** somebody who might not be able to write for him/herself) instead of a *recipient* argument.

9.5.6 Comparative ‘like’ dahi

The postposition *dehi* is employed to point out the likeness of two referents. It may occur on one NP (X like Y) or on both NPs (X like Y like). In Example 478, the postposition *dehi* occurs after the NP *kwâw* ‘dog’. In Example 479, the postposition occurs after both NPs involved in the comparison.

(478) *taphrû* lade we kwâw dahi cuwe (...)

[*taphrû la-de*]_{main.cl} [we kwâw dahi]_{adju} cuwe

dhole say-HAB.EGO DEM dog like DISC.DEIC.DEM

‘We call them *taphrû* (dhole), (they are) like dogs (...)’ [TA23_talyu_clan 10 52]

(479) *khûnyüm* dahi mề dahi mề na?

[*khûnyüm dahi*]_{adju} [mề dahi]_{adju} mề na

god like person like person INSIST

‘Gods were like people, (these) people, right?’ [TA22_clan 62 200]

9.5.7 NP quantifiers ‘each, every’

Three postpositions have the function of quantifying a NP referring to ‘each, every’: (*la*) *tsâi*, (*la*) *mang/müng* and *tyô*. The first two postpositions occur with an optional *la* particle, which is likely a grammaticalized form of the versatile verb *la* ‘say’. The postposition (*la*) *tsâi* goes back to a homophonous verb meaning ‘count’. The origin of (*la*) *mang/müng* is not clear. The postposition *tyô* also exists as an adjective meaning ‘last (in a row)’. The difference between these quantifiers is not clear, as they seem to occur in the same syntactic environment and do not have clear semantic restrictions. Example 480 illustrates that these three quantifiers are to a large extent interchangeable. In Example 481, the meaning ‘each, every’ is interpreted as ‘one month each time’ in the context. Lastly, Example 482 reveals the morphosyntactic status of *tsâi* as a postposition, for it does not trigger a tonal change on *mlă* ‘village’, as would be expected in N-N modification.

(480) halo (*la.*)*tsâi*/(*la.*)*mang/müng*/*tyô*

halo (*la.*)*tsâi*/(*la.*)*mang/müng*/*tyô*

month each

‘each month’

- (481) halo tyô ne ì cangya
 [halo tyô]_{adju} ne ì cang-ya
 month each only be.there be.able-ALLO
 ‘You may only stay one month each time.’ [161024_JOJA_eli 489]

- (482) mlă tsâi
 [mlă tsâi]_{adju}
 village each
 ‘each village’ [041024_STBT_eli1 355]

9.6 Non-final markers

Non-final markers are particles with scope over whole clauses. Their function is to turn clauses non-final, i.e. embedded in a multi-clause construction. The use of the term *marker*, although not a morphological label, aims at highlighting the abstract functionality of these particles as clause operators. Syntactically, we can distinguish between two structures: cosubordination (exemplified by the non-final marker *gũ*) refers to a flat structure (i.e. not nested) in which all non-final clauses modify the final clause, and subordination (heralded by most other non-final markers) entails a nested, hierarchical structure in which the subordinated clause can only modify its immediate superordinate clause (cf. §10.3). For this reason, non-final markers are best described amid the construction where they occur. Table 9.5 displays all non-final markers described in this grammar. It also includes the reference to the section where each construction is described, as well as a note on other functions of the particle (where available). Note that Table 9.5 does not display all multi-clause constructions, as some do not have dedicated a particle.

Table 9.5: Overview of non-final markers

Construction(s)	Device(s)	Reference	Notes
Cosubordination	<i>gũ</i>	§10.3	also locative/instrumental
Co-temporality	<i>dũ</i>	§10.3.4.1	also ‘with’
Temporal and conditional subordination	<i>bobyẽ</i>	§10.4.1	cf. <i>bô</i> ‘go’ and <i>byẽ</i> TOP
Unrealized alternative and contra-factual cond.	<i>(go)ji</i>	§10.4.2.3	<i>ji</i> ‘do, make’
- <i>thĩ</i> group			also emphatic marker
Temporal and hypothetical subordination	<i>bũ-thĩ</i>	§10.4.3.1	
‘as if’	<i>da-thĩ</i>	§10.4.3.2	
Concomitant action	<i>gũ-thĩ</i>	§10.3.4.2	

10 Multi-clause constructions

10.1 Introduction

This chapter describes multi-clause constructions in Tawrã, i.e. how individual clauses are combined to form *clusters of clauses* (Longacre 2007, 372). Clauses can be conjoined through apposition, i.e. without the use of a conjunction or similar device, cosubordination and subordination proper. This chapter also covers related topics: nominalization-based constructions, complementation, the extended functions of the verb *la* ‘say’ and complex demonstratives.

This chapter is organized in the following manner: Conjoined clauses are discussed in §10.2. Cosubordination, which mainly includes the non-final marker *gũ*, is discussed in §10.3. §10.4 on subordination includes a number of heterogeneous constructions: i) conditional and temporal subordination in §10.4.1, ii) three complex constructions under the label *other subordinating constructions* in 10.4.2: the ‘lest/not until’ construction (§10.4.2.1), the paraphrase *-ya kanüm* for purposive clauses (§10.4.2.2) and the unrealized alternative and contra-factual conditional construction (§10.4.2.3) and iii) two complex constructions that include the multi-clause marker *thĩ* (hence the *thĩ group*) in §10.4.3. Thereafter, §10.5 discusses nominalization-based constructions, including relative clauses (§10.5.2), event nominalization (§10.5.3) and the locative nominalizer (§10.5.4). Complement clauses and embedded discourse are discussed in §10.6. The extended functions of the verb *la* ‘say’ are discussed in §10.7, which includes the hearsay particle (§10.7.1), name labeling (§10.7.2) and discourse continuity (§10.7.3). Complex demonstratives, which have a discourse pragmatical function linking clauses, as discussed in §10.8).

It should be noted that the topic of omitted arguments in multi-clause constructions

have been discussed in §8.2.5, although it will be cross-referenced whenever relevant.

10.2 Conjoined clauses

Two independent clauses can be conjoined *asyndetically* in Tawrã, i.e. without a conjunction or a non-final device connecting two or more clauses; the clauses are simply apposed to each other (cf. Burling 2003, 313f). This construction is furthermore neutral in terms of coordination (X *and* Y) or contrast (X *or* Y), as these categories are understood from the context. In Example 483, two clauses are apposed. There is a pause in between the clauses (marked by the three dots) but this needs not be the case. In Example 484, the clauses *bô-mâ-kũ-ya* ‘it goes continuously (and completely)’ and *manya-mâ-kũ-ya* ‘it downpours continuously (and completely)’ are conjoined in a coupling (‘and’) fashion rendering ‘(the water) continuously goes **and** downpours completely’. Both clauses have a) the prosodic structure characteristic of a complete clause and b) the same suffixal chain.

- (483) *kadâge mu tyuamëna (...) kadâge mu dina (...)*
[kadâge mu tyu-amë-na]_{main.cl} [kadâge mu ji-na]_{main.cl}
 how.much also piece-from.behind-FACT how.much also do-FACT
 ‘As much as (they) pierce, as much as (they) do ...’ [TA22_origin_106]

- (484) (...) *we macîk karà bômâkũya manyamâkũya we dablang atyó laga (...)*
[we macîk karà bô-mâ-kũ-ya]_{main.cl} [manya-mâ-kũ-ya]_{main.cl} [we
 DEM water rain go-ITER-COMPL-ALLO downpour-ITER-COMPL.II-ALLO DEM
dablang atyó la-ga]_{main.cl}
 flood rise call-PL.I
 ‘Rainwater constantly flowing (and) constantly downpouring, we call it a rising flood.’ [TA23_newmyth_0005]

Example 485 has two clauses, each with its own S argument: *kalà lyó kalyu-mâ* ‘the white officer kept waiting’ and *güwri hana-lya-yüm* ‘the horse did not come back’. Note

that multi-clause constructions do not have to share core arguments, as discussed in §8.2.5. These clauses are conjoined in a contrastive fashion.

(485) *ka kalà lyó kalyumâ (...) güwri hanalyâm*

ka [kalà lyó kalyu-mâ]_{main.cl} [güwri hana-lyâ=m]_{main.cl}

FS officer white wait-ITER horse come-back = NEG

‘The white officer kept waiting (but) the horse never came back.’ [TA22_JAStory_035]

Note that discontinuous stems (§6.1.1.3) resemble conjoined clauses, since they seem to have two finite clauses. However, they form a syntactic unit. In Example 486, two verbs have a composed meaning (‘bow down’ and ‘dance’) ‘worship’.

(486) *ah alí gũ gûmna bwîna*

[ah alí gũ [gûm-na bwî-na]_{disc.stems}]_{main.cl}

ah like.this INS bow.down-IMP dance-IMP

‘Ah worship (lit. bow down (and) dance) like this!’ [TA22_origin_106]

10.2.1 Semantic functions of conjoined clauses

Other functions of conjoined independent clauses in Tawrã are related to iteration (with reduplicated clauses) and enumerating contemporaneous events. Example 487 corresponds to an iterative event signaled by a repetition of the predicate (cf. §6.1.1.4). As in English, repeating the clause (or even only the predicate) indicates that the action was repeated many a time, often in a tiresome manner. Example 488 makes use of the iterative suffix *-mâ*. The reduplication of the clause further intensifies the repetitive character of the action.

(487) *cyâga cyâga cyâga (...)*

[cyâ-ga]_{main.cl} [cyâ-ga]_{main.cl} [cyâ-ga]_{main.cl}

lower-PL.I lower-PL.I lower-PL.I

‘They lowered and lowered and lowered (the ropes).’ [TA22_kutyemba_phla_268]

(488) bômâ bômâ bômâ bôtalá khrâk

[bô-mâ]_{main.cl} [bô-mâ]_{main.cl} [bô-mâ]_{main.cl} [bô-talá khrâk]_{main.cl}

go-ITER go-ITER go-ITER go-still EVI

‘It kept going and going and going and still going, it seems.’

[TA22_kutyemba_phla_272]

Conjoined clauses can serve the purpose of naming or listing simultaneous events or referents. In Example 489, the speaker names all the main characters involved in the creation of the world.

(489) júm mề shyâdi ring shyâ tulù shyâ kapáw kabrûyâ shyâ

[júm mề shyâ-de]_{main.cl} [ring shyâ]_{main.cl} [tulù shyâ]_{main.cl} [kapáw

true person happen-EGO.HAB sun happen a.river happen a.place

kabrû-yâ shyâ]_{main.cl}

a.god-FEM happen

‘Humanity existed, there was the sun (god), the Tulu river, Kapaw Kabru goddess.’

[TA22_origin_04 24]

10.3 Cosubordination

This section discusses cosubordinating constructions with an emphasis on the non-final marker *gũ*. It starts with an overview of the main characteristics of cosubordination. The two main interpretations of the *gũ* clause are discussed in detail in §10.3.1 (sequential *gũ* construction) and in §10.3.2 (simultaneous *gũ* construction). Other pragmatic effects of the *gũ* construction include providing a reason or cause (§10.3.1.1) and recapitulating (§10.3.1.2). Non-canonical uses of this construction are discussed in §10.3.1.3. Negation of *gũ* clauses is discussed in §10.3.3. §10.3.4 discusses the occurrence of *gũ* along with a further marker (termed here *double* marker): co-temporal *dũ* and concomitant *thĩ*. §10.3.5 explores the grammaticalization path of the non-final marker *gũ*.

While this construction bears a certain resemblance with *clause chaining* and *converb* construction, as known from related languages from Northeast India (cf. Post and Modi

2024, for Tani languages; cf. for Mongsen Ao Coupe 2008, 410), there are some differences, especially regarding the TAMN suffixes that the co-subordinate clause allows. The non-final marker *gũ* makes a *final* clause *non-final*, i.e. it indicates that the marked clause is not the last one, although the clauses are thematically connected. Together, the non-final and final clauses make up a discourse episode (Givón 2001b, 355f; *operator dependence*, cf. Van Valin and LaPolla 1997, 454).

Olson (1981, 207f.) introduced the term *cosubordination* to refer to clauses that combine in a part-whole relation, even though they bear a relation of equivalence in the sense that both (if not more) parts constitute a whole. This term is here preferred to characterize these constructions in Tawrã, as it captures the *dependent* nature of non-final clauses without describing them as subordinate, offering a compromise between coordination and subordination (cf. Haspelmath 1995, 23f; Van Valin and LaPolla 1997, 451f). Table 10.1 illustrates a scheme that compares the structure of cosubordination and subordination. What is relevant for now is that cosubordination presupposes a potentially infinite sequence of cosubordinated clauses (here illustrated by the subordinator *gũ* NF) that conclude on a final clause (i.e. not marked by *gũ*). Cosubordinated clauses have a flat structure (i.e. not nested) and thus all modify the final clause. Conversely, subordination entails a nested, hierarchical structure in which the subordinate clause 3 can only modify clause 2, but never clause 1. Subordination will be further discussed in §10.4

Table 10.1: Comparison between cosubordination and subordination in Tawrã

Cosubordination		
[[... <i>gũ</i>] ₀ [... <i>gũ</i>] ₀ ...	final clause]	
Subordination		
[[[subordinate clause] ₁	subordinate clause] ₂	final clause] ₃

Non-final markers occur at the right edge periphery of the cosubordinate clause. Their function is to indicate that the clause is non-final, i.e. one should expect a main clause to follow (cf. DeLancey 1991, 3f; also discussed by Hyslop 2013, 159; cf. Post and Modi

2024). In this sense, the marker is semantically empty as it does not mean anything else than ‘more is to come’ (as opposed to subordination markers, cf. §10.4) (also cf. Genetti 2005, 50). All constructions described here are thus subject to interpretation based on the iconicity between a final and a cosubordinate clause. Furthermore, speakers make use of this implication for diverse pragmatic effects, as will become clear during this section.

The non-final marker *gũ*, like all other particles (§9), has scope over whole clauses. As such it may occur after the verbal predicate (with or without inflectional markers) in a verbal clause (§7.2) or after NP constituent in a verbless clause (§7.1). In Example 490, the non-final marker occurs after a verbless clause (VLC). In Example 491, it occurs after a verb.

(490) *cya kaman gũ tawrã mu kasâ*

[[[cya kaman]_{VCL} gũ]_{cosub.cl} tawrã mu kasâ]_{main.cl}

3 a.tribe NF a.language also know

‘Since he is Kaman, he also knows Tawrã.’ [301024_STBT_eli1 1561]

(491) “thrûga lya gũ kadyâng-rẽ-na!” *la*

[[[thrûga lya gũ]_{cosub.cl} kadyâng-rẽ-na]_{main.cl/embedded.cl} la]_{main.cl}

on.the.back lie NF push-PL.II-IMP say

‘He said: “push lying on your back!”’ [TA23_kutyemba_phla 99 226]

A cosubordinated *gũ* clause is interpreted as *sequential* or *simultaneous* to the main clause. The conflation of both a sequential (clause chaining) and simultaneous (adverbial, modifying function) interpretation in the same marker is a feature that occurs in related TB as well as some Indo-Aryan languages (cf. Coupe 2008, 409f; also Bickel 1998, 384). The distinction between these two functions rests on pragmatics and on the context, as well as on the suffixes found on the cosubordinated clause; the availability of a TAMN suffix on the verbal stem of the cosubordinated clause that is different from that of the main clause naturally triggers a *sequential* interpretation while the lack of TAMN suffixes (or an identical suffix in both clauses) usually triggers a *simultaneous* interpretation. A

sequential interpretation means that the events transpire in sequence while in a *simultaneous* interpretation the predicates are tightly interlocked and resemble the idea of *one event* and *adverbial modification* (cf. Givón 2001b, 330f). Example 492a corresponds to a *sequential* interpretation: the cosubordinated clause takes an inflectional suffix (here the suffix of recent past *-li*) and the actions are understood to transpire one after the other; a simultaneous reading of this example is not available. Conversely, Example 492b corresponds to a *simultaneous* interpretation: the clauses depict a single event, in this case, ‘go jumping’. Adding a TAMN suffix would break the simultaneity.

(492) a. kwâg dûli gũ bôya

[kwâg dû-*li* gũ]_{cosub.cl} bô-ya

dog jump-REC.PST NF go-ALLO

‘The dog jumped and then goes’ or ‘after jumping the dog goes’ (but not *‘it goes jumping’).

b. kwâg dû gũ bôya

[kwâg dû gũ]_{cosub.cl} bô-ya

dog jump NF go-ALLO

‘The dog goes by jumping’ (and not *‘it jump and goes’).

[both 290924_STBT_eli1 479]

Similarly, Examples 493a and 493b have the same verb *shêi* ‘die’ with both interpretations: Example 493a has a sequential interpretation (note the recent past suffix *-li* on the verb). Conversely, in Example 493b the cosubordinated clause is understood as manner modification (also *reason*).

(493) a. shêili gũ dângbô

[shêi-*li* gũ]_{cosub.cl} dâng-bô

die-REC.PST NF finish-PFV.ALLO

‘After (his) death it was finished.’

b. mề shêi gũ mlầ marấ shyâho

[mề shêi gũ]_{cosub.cl} mlầ marấ shyâ-ho

person die NF village abandon become-STAT

‘Through (also ‘because of’) the death of the people, the village became deserted.’ [both 301024_STBT_eli1 1620]

The cosubordinate clause is fully autonomous in its tense, aspect, polarity and illocutionary force (cf. Van Valin and LaPolla 1997, 452). The verb may either have no suffix on the stem or follow the inflectional suffixes of the final clause, as in Example 494 but it can just as well have a different TAMN suffix. In Example 495, the cosubordinate clause has a negative polarity while the final clause has a positive polarity. Negation in multi-clause constructions is further discussed in §10.3.3. Note that core arguments do not have to be shared in multi-clause constructions; co-referentiality is subjected to narrative impositions and pragmatic considerations although there is an iconic relation of S/A continuity between a cosubordinated and a final clause (cf. Haiman 1981; also Aikhenvald 2023, 76f.). Furthermore, there is no morpho-syntactic marking that would differentiate core arguments (cf. §8.2.5).

(494) thâde gũ yu mu tyûmde

[thâ-de gũ]_{cosub.cl} yu mu tyûm-de

eat-HAB.EGO NF wine also drink-HAB.EGO

‘When I eat I drink wine.’ [290924_STBT_eli1 513]

(495) nyu dyâng yûm gũ mu hã ne tsêq

[nyu dyâng yûm gũ]_{cosub.cl} mu hã ne tsê-q

2S comply NEG NF also 1S only be.right-CONT

‘Even if you deny it, I am right.’ [290924_STBT_eli1 536]

A high-level propositional category, such as egophoricity, need not be shared between the clauses involved in cosubordination. In Example 496, the final clause has a non-egophoric marker while the cosubordinate clause is marked with an egophoric marker *-de*

HAB.EGO. Similarly, in Example 497 the cosubordinate clause has an egophoric marking while the final clause has an allophoric marking that stands for a different referent.

- (496) habang bôdi gũ hã grũ pohwã krã kê thô
 [[habang bô-di gũ]_{cosub.cl} hã grũ pohwã krã kê thô]_{main.cl}
 jungle go-HAB.EGO NF 1S foot pit hole LOC.INS get.stuck
 ‘My foot got stuck in a pit while going to the jungle.’ [BAB 06062023]

- (497) hã kalyaw bôde gũ cya inga
 [hã kalyaw bô-de_i gũ]_{cosub.cl} cya ing-ya_j
 1S field go-HAB.EGO NF 3 sleep-ALLO
 ‘S/he sleeps while I go to the field.’ [290924_STBT_eli1 571]

10.3.1 Sequential interpretation

This section will discuss the semantic-pragmatic meaning of a sequential *gũ* interpretation of the cosubordinate clause. After going over the main morphosyntactic characteristics, three conventionalized, semantic uses of this construction are described: reason and cause (§10.3.1.1), recapitulation (§10.3.1.2) and other non-canonical uses (§10.3.1.3).

The sequential interpretation implies that one or more cosubordinate clauses (marked by *gũ*) canonically conclude in a *final* clause *without* the *gũ* marker. The use of the term *final* here implies that the clause does not take the marker *gũ* and that it is the end of the multi-clause construction. It does not imply that the non-final clauses are *non-finite*. The sequential interpretation shares a number of properties with *clause chain* constructions as discussed in the literature from languages from South Asian and Papua New Guinea (Haspelmath 1995, 20f; Dixon 2009, 12). However, clause chain constructions in Papuan languages usually consist of partially inflected *medial* verbs (Sarvasy 2015, 665; Longacre 2007, 374f;398; for an overview cf. Foley 2010, 28f) or non-final constructions in Macro-Tani languages (Post and Modi 2024, 11f). In these cases, the first and middle predicate

in the clause chain is partially inflected or not inflected at all while the final predicate is fully inflected¹.

The relation between the dependent and the final clause in the *sequential gũ interpretation* is one of temporal sequence “after doing X, Y”, which corresponds to the English conjunction/adverbs *and*, *then* and *and then* (Dixon 2009, 9). The order of the clauses commonly reflects the temporal order of actions, so the *gũ* clause(s) are canonically concluded by a final clause.

In Example 498, the cosubordinate clause *yu tapè pra-langka ji-ya gũ* ‘(having) served (him) food and drinks nicely’ occurs before the final clause *ahwênâ* ‘(they) sent (him) back’, as it depicts a chronological and logical sequence of actions.

(498) *bená bô gũ mane yu tapè pralangka jiya gũ ahwênâ*

<i>bená</i>	<i>bô</i>	<i>gũ</i>	<i>mane</i>	<i>[yu tapè pra-langka ji-ya</i>
the.next.day	TEMP.SUB	INS	FILL(<I.A.)	wine food good-ADV do-ALLO
<i>gũ]cosub.cl ahwê-nâ</i>				
NF	send-back			

‘The next day, I mean, they served (him) food and drinks nicely (and) sent (him) back.’ [TA22_CLAN2_174 330]

The canonical sequentiality of the *gũ* construction can be exploited to conventionalized pragmatico-semantic ends. The following subsections will discuss these uses in detail: §10.3.1.1 discusses *reason* and *cause*. §10.3.1.2 discusses *recapitulation*. Finally, §10.3.1.3 discusses non-canonical uses of the construction.

10.3.1.1 Reason-cause

As described by Post and Modi (2024, 12f.) for the Macro-Tani languages, clause chaining implies an iconic sequentiality which may give way to a causal interpretation²

¹Note that fully inflected subordinate clauses are not uncommon in the languages of South Asia (cf. Hock 2016a, 315).

²In the Papuan case, a causal relation through a medial verb is only available in cases of right-dislocation, i.e. as a non-iconic use (cf. Sarvasy 2015, 686). In the Tibeto-Burman group, causal relation in (co)subordination is available in a number of structures: the general non-final device e.g. in Tawrā or Macro-Tani (Post and Modi 2024, 11f.), specialized non-final devices (converbs), e.g. in Chantyal (Noonan 1999) or case markers such as the ergative (along with the infinitive marker) in Kiranti (cf. Bickel 1999c).

(also Noonan 1999, 7f): V1-NF **and** V2 > **because of** V1; V2. In this conventionalized use of the *gũ* interpretation, a sequential interpretation is seldom available (cf. §10.3.1).

In Example 499, the cosubordinate clause is understood as the reason of the action in the final clause. Similarly, in Example 500 the verb *komjuri ji* ‘(do) quarrel’ is marked as non-final and understood as the reason. Finally, in Example 501 the event of the cosubordinate clause is also the reason for the main clause.

- (499) *jâpmalu we kapâ hwě gã mề ìdyibô*
[jâpmalu we kapâ hwě gã]_{cosub.cl} mề ìdyibô
 a.god DEM alone look NF person live-PFV.EGO
 ‘The people (we) lived because Japmalu alone took care (of us).’ [TA22_origin 101 249]

- (500) *komjuri jina gã hwûwga khûga*
[komjuri ji-na gã]_{cosub.cl} hwûw-ga khû-ga
 quarrel(<IA) do-FACT NF argue-PL.I retaliate-PL.I
 ‘Because of quarreling they fought.’ [TA22_origin 101 249]

- (501) *ashyâ tabrékũ gã míngtyubô*
[ashyâ tabré-kũ gã]_{cosub.cl} míng-tyubô
 mithun hunt-COMPL.II NF become.extinct-start
 ‘Because of (completely) hunting the mithuns completely, they will become extinct.’ [150924_HM_eli 582]

As described by Genetti (2005, 50f), the semantic content of the (co)subordinator can seldom be unequivocally ascribed to one label (reason, adverbial subordination, etc.); the marker *per se* would be polysemous and depends on the construction, the whole semantic of the context and what is already known. In Example 502, the cosubordination chain can be read in a causal relation in which the cosubordinate clause is the reason for the final clause. However, a mere sequential interpretation (i.e. ‘one fell and the other came out’) is also possible. A similar case is Example 503, in which the two events can be

logically connected as one being a consequence of the other (i.e. ‘because they became tired, they left him’), but a mere sequential reading (i.e. ‘they get tired *and* left him’) is also available.

(502) bürü lyâ gũ kwâ gũahrúkũho

[bürü lyâ gũ]_{cosub.cl} kwâ gũ-ahrú-kũ-ho

down fall NF owl stick.out-suddenly-COMPL.II-STAT

‘The owl suddenly came out because (the dog) fell down’, also ‘(The dog) fell down and so the owl came out suddenly.’ [TA22_frog1_34 155]

(503) hwâ yüm bô gũ maráŕẽ laga

[hwâ yüm bô gũ]_{cosub.cl} mará-ring la-ga

feel.active NEG TEMP.SUB NF abandon-PL.II HSAY-PL.I

‘They left (him) because they became tired, they say’, also ‘They became tired and thus they left (him), they say.’ [TA22_kutyemba_phla 316 598]

10.3.1.2 Recapitulation

Non-final clauses may also be employed to repeat a relevant piece of information from the preceding sentence. This strategy links the new utterance with the larger context, providing discourse continuity. The repeated construction occurs before the new (main) clause (cf. Genetti 2005, 49; also cf. Dixon 2009, 8; Noonan 1999, 6). Other names for an equivalent construction are *tail-to-head linkage* and *bridging device*.

Examples 504a and 504b are from a story about the imprisonment of a criminal. Examples 504a ends one episode of the story. Example 504b is its direct continuation. Here the speaker begins a new episode in the story which is connected with the previous actions. This link is made explicit by the use of a non-final clause that repeats the previous verb. Similarly, in Example 505 the villagers prepare food for a trickster, as he said he would only help them get rid of a deadly stone once he was given the biggest pig to eat.

- (504) a. ace tyôringya mane
 we tyô-ring-ya mane
 DEM arrest-PL.II-ALLO FILL(<Asm)
 ‘They arrested him, I mean.’[TA22_CLAN2_103 205]
- b. we tyôya gũ mane dâi we mane (...)
 we tyô-ya gũ mane dâi we mane
 then arrest-ALLO NF FILL(<Asm) a.tribe DEM FILL(<Asm)
 ‘Having been arrested, I mean, this Idu, I mean’[TA22_CLAN2_105 208]
- (505) thâlibô gũ la la mǎ mề we la “chôq”
 [thâ-li-bô gũ]_{cosub.cl} la la mǎ mề we la chôq
 eat-REC.PST-PFV.ALLO NF say HSAY village.MOD people DEM say let’s!
 ‘Having eaten, he said to the village people “let’s (carry the stone)”, they say.’
 [TA23_kutyemba_phla_075 190]

Recapitulation may not need an initial sentence to be repeated if the context provides enough information for the hearer to reconstruct a logical sequence of actions. Examples 506a and 506b are from the same story. They are not immediately continuous as there is some repetition and digression that has been omitted here for the sake of clarity but they are connected within the main argument of the story. In Example 506a, we learn of the tiger going to the jungle. Example 506b employs the non-final marker on an event that was not explicitly mentioned before; it does not correspond to a *recapitulation* in a strict sense because the hearer must reconstruct a missing scene in which the tiger went hunting. This is, however, a logical development within the story. In this sense, *gũ* links the clause with a larger (albeit unstated) context. The narrative iconicity entailed by *gũ* will again become evident once the non-canonical uses of this construction are described in §10.3.1.3.

- (506) a. abang bônatyêyuke pütýô shyânaya künam
habang bô-na-tyêyu-ke pütýô shyâ-na-ya künam
 forest go-FACT-JUS-HORT different become-FACT-ALLO PRP
 ‘Let it (the tiger) go to the jungle in order to become different (i.e. wild)!’
 [TA23_kutyemba_phla_131 288]
- b. tabrê cângge gũ webô gũ (...)
tabrê câng-ge gũ webô gũ
 meat get-and.back NF and.then INS
 ‘After finding and bringing meat, then (...)’ [TA23_kutyemba_phla_136 302]

The recapitulation construction bears a functional similarity with the *perfective* in other languages³ (in the sense of “continuing present relevance”, cf. Comrie 1976, 52f) since the cosubordinated *gũ* clause implies an “inherent relationality in terms of time reference” (cf. Sarvasy 2015, 688f), i.e. the action of the main clause occurs only after the one from the dependent clause is finished. This is sometimes indicated in the English translation as ‘having done X’

10.3.1.3 Non-canonical use of cosubordinate clauses

This section discusses non-canonical, conventionalized uses of the *gũ* construction. Authors like Sarvasy (2015) and Post and Modi (2024, 21) have recognized the extended use of subordinated⁴ clauses that lie *stranded* or post-posed (after a final clause) and have a rhetorico-pragmatic function: Since cosubordinated clauses canonically precede a final clause, many languages exploit the break of iconicity between events and sentence structure (Berghäll 2015, 337). This conventionalized ellipsis of a main clause has diverse pragmatic and epistemic effects, such as signaling repeated actions, inconclusive statements, background information (Sarvasy 2015; Evans 2007, 367;340f; in Karbi cf. Konnerth 2014, 557f) or even reporting presupposed information that may not be challenged, e.g. from mythological texts (in Bulu Puroik, cf. Lieberherr 2017, 222f).

³As described by Sarvasy (2015, 688f) in Nungon, there is no morphological form for a final verb to express the perfective aspect; a post-posed non-final (medial) verb may be employed to indicate that an action has already been completed.

⁴This corresponds to *non-final* clauses in the Tani languages (cf. Post and Modi 2024, 21).

In Tawrã, a clause headed by a cosubordinated predicate may, on a basic level, merely imply that there is *more to come*. Example 507 is headed by a cosubordinate clause; although it has the prosodic structure of a ‘complete’ conditional clause one expects a final clause in the next sentence. This particular sentence is not completed by a final clause; the hearer has to complete it from the context.

(507) ephlăe nyu mề thâ gũ byẽ

e = *phlă* = *e* *nyu* *mề* *thâ* *gũ* *byẽ*

DEM = stone = DEM FOC person eat NF COND

‘If this stone eats people (...) (i.e. what is left for us!)’ [TA23_kutyemba_phla_46 142]

A speaker may exploit that expectation or iconicity to more elaborated pragmatic effects, namely to describe an event that is ongoing or that does not have a clear end. This often corresponds to the background of a main event (in the main clause), a common use for non-final clauses (cf. Sarvasy 2015, 699). This can be illustrated in Example 508, where the villagers want to capture a criminal who is hiding inside a cave. As part of their efforts to catch him, they tied many *taraw* creepers into a long rope. The cosubordinate clause creates a background upon which a final clause would set a main event.

(508) mấ mề we làng tarâw tatyế hrẽmâ gũ

[*mấ* *mề* *we* *làng* *tarâw* *tatyế* *hrẽ-mâ* *gũ*]_{cosub.cl}

village[.MOD] person DEM COLL type.of.creeper rope pull-ITER NF

‘The village people were pulling *taraw* ropes.’ [TA22_kutyemba_phla_0230 456]

In Example 509, the verbs constitute a discontinuous stem (with a semantic or otherwise iconic tie, cf. §6.1.1.3). The subordinator has thus scope over both predicates, making them non-final. The implication here is that the action of building kept going on as the fire consumed the improvised hut where the couple was hiding.

- (509) *halá lyûm tyô azê-ya plû-ya gũ*
 [halá lyûm tyô azê-ya plû-ya gũ]_{cosub.cl}
 CLF:layer eight each build.walls-ALLO build.floor-ALLO NF
 ‘He was building (walls and the floor) of eight layers each.’ [TA23_newmyth_025 155]

While the dependent clause most commonly precedes the main clause, an inversion of order, i.e. a non-final cosubordinate clause succeeding its main clause has the same effect as the ellipsed final clauses in a clause chain as described above. In Example 510, a non-final clause breaks the iconic order of non-final clauses and gives context to the event narrated in the main clause.

- (510) *we mẽ míngna harü na we myúm mathi ñegaya gũ*
 [we mẽ míng-na harü na [we myúm
 DEM person become.extinct-FACT at.the.time EMPH DEM elder.brother(fem.prsp.)
mathi ñ-nega-ya gũ]_{cosub.cl}]_{main.cl}
 younger.sister live-DU-ALLO NF
 ‘At that time people became extinct, as the elder brother and the younger sister lived.’ [TA22_origin_30 94]

10.3.2 Simultaneous interpretation

The simultaneous interpretation implies that the cosubordinate clause occurs at the same time as the main clause to the extent that they can be considered to form one action. This semantic relation has been sometimes termed *adverbial (subordination)*, i.e. to modify verbs and clauses or the *manner* in which an action is carried out (cf. Lieberherr 2017, 230f.). As such, they resemble a recursively hierarchical embedding of clauses (cf. Post and Modi 2024, 8; cf. Comrie et al. 2012, 184f; also discussed by Gast and Diessel 2012, 25f; Thompson et al. 2007, 249f; Haspelmath 1995, 7f; Dixon 2009, 35f).

In Example 511, the cosubordinated predicate *tashê gũ* ‘by singing’ is concomitant with the main clause’s predicate and modifies it. Similarly, in Example 512 the cosubordinate clause is understood as a manner/adverbial-modification to the main clause.

(511) *tashẽ gũ hanaya*

[*tashẽ gũ*]_{cosub.cl} *hana-ya*

sing NF come-ALLO

‘He came singing.’ [250122_JO_eli2 1826]

(512) “*thrũga lya gũ kadyâng-rẽ-na!*” *la*

[*[[thrũga lya gũ]*]_{cosub.cl} *kadyâng-rẽ-na*]_{main.cl/embedded.cl} *la*]_{main.cl}

on.the.back lie NF push-PL.II-IMP say

‘He said: “push lying on your back!”’ [TA23_kutyemba_phla 99 226]

In Example 513, the cosubordinate clause *wri-kũ gũ* ‘cheating (completely)’ modifies the predicate of the final clause. Similarly, Example 514 is part of a creation myth: the two non-final clauses indicate how the remains of a cursed child were pulverized before being blown away. These actions (marked by *gũ*) are perceived as simultaneous or at least to constitute a coherent action. The referent *a we* ‘this child’ remains the P argument of both cosubordinated clauses.

(513) (...) *kalà lyó güwri we (...) la wrikũ gũ shikũbô khrâk*

[*[kalà lyó güwri we la [wri-kũ gũ]*]_{cosub.cl} *shí-kũ-bô*

officer white horse DEM HSAY deceive-COMPL.II NF take-COMPL.II-PFV.ALLO

khrâk]_{main.cl}

EVI

‘(...) through deception, he took the white officer’s horse away, they say.’

[TA22_JAstory 57-59 191]

(514) *a we tyé gũ apyék gũ tapyek khalây tek pôwbô*

[*[a we tyé gũ]*]_{cosub.cl} [*apyék gũ*]_{cosub.cl} *tapyek khalai tek pwôk-bô*]_{main.cl}

child DEM cut NF spread NF sand soil FS blow-PFV.ALLO

‘By cutting and spreading this child, they blew it away to the sand and soil.’

[TA22_origin 58 168]

Simultaneous interpretations commonly share tense and aspect between the main and the cosubordinate clause. Example 515 depicts two simultaneous actions: taking out the cows and singing. Both the final and cosubordinate clauses take the egophoric habitual suffix *-de*. A difference in tense and aspect (i.e. a cosubordinate clause with an inflectional suffix that is different from that of the final clause) would commonly break the simultaneous interpretation as discussed in §10.3.1. Negation, however, does not need to be shared, as will be discussed in §10.3.3. In Example 516, the cosubordinate clause is negated even though the whole multi-clause has a simultaneous interpretation.

- (515) matsâp haláwde gũ tashẽ tashẽde
 [matsâp haláw-de gũ]_{cosub.cl} tashẽ tashẽ-de
 cow take.out-HAB.EGO NF song sing-HAB.EGO
 ‘I sing a song when I take the cows out (for grazing).’ [290924_STBT_eli1 1180]

- (516) tashẽ yüm gũ hanaya
 [tashẽ yüm gũ]_{cosub.cl} hana-ya
 sing NEG NF come-ALLO
 ‘He came without singing.’ [150924_HM_eli 701]

As mentioned in the introduction to this section (§10.3), the simultaneous interpretation discussed here shares a functional similarity with serial verb constructions in the sense that they portray tightly integrated events (often resembling *single events*) and they cannot take any TAMN suffix. However, this construction does not have a lexical status, does not form idiomatic combinations and, crucially, does not convey culturally significant concatenations of events as serial verb constructions do in other languages (Aikhenvald 2006b, 10f; cf. Schultze-Berndt 2000, 36f; also *iconic* combinations in the sense of DeLancey 1991, 15).

10.3.3 Negation in cosubordination

Negation occurs with the standard negator *yüm* (§6.4.1) directly attached to the negated predicate, i.e. before the non-final *gũ* marker. The scope of the negation does

not exceed the marked clause. In Example 517, the negation applies only to the non-final *gũ* clause. The negation of the non-final clause often has a ‘without doing X’ meaning (cf. Genetti 2005, 66).

- (517) *tama thâm phürü gã shêibô*
 [tama thâm phürü gã]_{cosub.cl} shêi-bô
 medicine eat=NEG yet NF die-PFV.ALLO
 ‘S/he died without having taken medicine. [150924_HM_eli 720]’

In Example 518, negation only applies to the final verb *bô-ga-na yüm* (go-PL.I-FACT NEG) ‘they won’t go’. In Example 519, the negated verb implies a manner (here *unaware*). Both final and cosubordinate clauses may be negated, as in Example 520.

- (518) *tabrê thâli gã mu bôgana yüm*
 [tabrê thâli gã mu]_{cosub.cl} bô-ga-na yüm
 meat eat-REC.PST NF also go-PL.I-FACT NEG
 ‘Even after eating meat they won’t go.’ [290924_STBT_eli1 1156]

- (519) *shêi habin.la kasâ yüm gã ìwithâ matyô*
 [sê ha.bin.la kasâ yüm gã]_{cosub.cl} ì-wi-thâ matyô
 kill POT know NEG NF be.there-AMB-mindlessly maybe
 ‘He was not aware that (they wanted to) kill (him).’ [TA23_tayang 31 187]

- (520) *thâm gã aphisî bô yüm*
 [thâm gã]_{cosub.cl} aphisî bô yüm
 eat=NEG NF office(<E) go NEG
 ‘S/he doesn’t go to the office without eating.’ [150924_HM_eli 735]

A conditional reading of the clause chain is also available by negating the *gũ* construction (cf. Dixon 2009, 14f). As such, there is some functional overlap with the conditional suffix *bô/byê* (cf. §10.4.1). In Example 521, the negated subordinated *gũ* clause implies a condition while the matrix is the apodosis.

- (521) *cya jüm tükü maro yüm gũ hã bôkũne*
 [*cya jüm tükü maro yüm gũ*]_{cosub.cl} *hã bô-kũ-ne*
 3 true matter talk NEG NF 1S go-COMPL.II-INT
 ‘If he does not tell the truth, I will go.’ [150924_HM_eli 802]

As mentioned at the end of §10.3.1, a particular concessive ‘even if’ sense is achieved when the dependent clause takes the adverb *mu* ‘also’. This sub-construction implies a strong expectation that the apodosis holds true, as opposed to a mere negation of the *gũ* cosubordinate clause⁵. Examples 522 and 523 display this sub-construction with the adverb *mu* ‘also’.

- (522) “shyâm shyâm” la gũ mu hã lade ne tsêq
 [*shyâm shyâm la gũ mu*]_{cosub.cl} *hã la-de ne tsê-q*
 no! no! say NF also 1S say-HAB.EGO only be.true-CONT
 ‘Even if you say “no no!”, only what I say is right.’ [290924_STBT_eli1 1135]

- (523) *chêi hawrẽ yüm gũ mu chei hawrẽ da thĩ wê shìbiya*
 [*chêi hawrẽ yüm gũ mu*]_{cosub.cl} *chei hawrẽ da thĩ wê shì-biya*
 walk try NEG NF also walk try as.if CONC think be.able-POT.ALLO
 ‘Even if I can’t walk I think that I can try.’ [290924_STBT_eli1 1035]

10.3.4 Double marker with *gũ*

Two non-final particles may occur along with the general cosubordinate marker *gũ*: co-temporality marker *dũ* and concomitant marker *thĩ*. These two markers have a similar meaning, although they differ in their syntactic distribution: The co-temporality marker *dũ* may modify a clause alone, although it occurs with the non-final marker *gũ*, preceding it. Conversely, the morpheme *thĩ* can only occur coupled to a further particle (cf. §10.4.3). Table 10.2 schematizes these *double marker* constructions within a *gũ* co-subordinated clause.

⁵Note a similar development in Galo: the concessive subordinator *-dakkòm* is composed of a ‘change of state’ inflection suffix element *-dak* and the additive *-kom* (Post 2007, 616f).

Table 10.2: Syntax of double markers with *gũ*

co-temporal marker	[V	<i>dũ</i>	<i>gũ</i>]
concomitant actions	[V	<i>gũ</i>	<i>thĩ</i>]

It should be noted that the co-subordinator *gũ* also occurs with the subordinator *bô*, as will be discussed in §10.4.1.

10.3.4.1 Co-temporality marker *dũ*

The marker *dũ* indicates simultaneous actions in a way similar to the marker *thĩ/-hĩ* (discussed in §10.3.4.2). This marker commonly occurs along with the main non-final marker *gũ*, although in elicitation it may occur alone: In Example 524, the marker modifies a verb in a way similar to the non-final marker *gũ*.

- (524) *bô dũ (...)*
bô dũ
 go CO.TEMP
 ‘When going (...)’ [021023_STBT_eli3 243]

Most commonly, the marker *dũ* occurs along the general non-final marker *gũ* in order to emphasize the simultaneity of the main and the non-final clause. Examples 525a, 525b and 525c are from the same story where the villagers are trying to catch a trickster who is hiding in a narrow cave. After the villagers placed a trap at the exit of the cave, the trickster convinced the appointed guard to stick his hand down the hole of the cave, thus activating the trap. In Example 525a, the speaker uses both non-final markers to emphasize that the actions that are to follow occurred immediately after the fingers of the guard touched the criminal’s hand. In Example 525b, there is a further non-final clause that describes how the guard was pulled down from the fingers to be finally caught by the trap in Example 525c.

- (525) a. *abrâng rô tyângya dũ gũ (...)*
abrâng rô tyâng-ya dũ gũ
 finger catch touch-ALLO CO.TEMP NF
 ‘Just as (his) fingers touched (him) (...)’ [TA22_kutyemba_phla 464 851]
- b. *agûya mề we magũ hazàkũ gũ (...)*
[agû-ya]_{rel.cl} mề we_{NP} magũ hazà-kũ gũ
 guard-ALLO person DEM pull.back drop-COMPL.II NF
 ‘The guard being pulled down (...)’ [TA22_kutyemba_phla 468 856]
- c. *ríng gapô sêkũya we we*
[ríng gapô]_{NP} sê-kũ-ya we we
 sun.MOD trap catch-COMPL.II-ALLO DEM DEM
 ‘The sun trap caught him.’ [TA22_kutyemba_phla 470 858]

The co-temporal marker *dũ* also exists as a postposition meaning ‘with’, as discussed in §9.5.4. In a case similar to the cosubordinator *gũ*, the co-temporal subordinator *dũ* seems to have grammaticalized from a nominal to a predicate operator.

10.3.4.2 Concomitant action *gũ thĩ*

The concomitant construction is composed of *gũ* + *thĩ* (only *thĩ* in fast speech) and forces a simultaneous interpretation of the cosubordinated clause. This is a helpful construction if the verb in the *gũ* cosubordinated clause takes a TAMN suffix, which would be otherwise interpreted as sequential, as discussed in §10.3.1. By adding the focus particle *thĩ* to the subordinated clause, a break in temporality (entailed by sequentiality, i.e. X after Y) is avoided. Furthermore, the simultaneous reading forced by the *gũ thĩ* construction does not necessarily imply *adverbial modification*, as the subordinator *gũ* on a bare stem would usually do (cf. §10.3.2), but the actions of the main and the subordinated clause are understood as transpiring in a simultaneous or near immediate fashion: *while doing X*. In Example 526, two actions occur simultaneously: watching T.V. and falling

asleep. The verb of the cosubordinate clause has a habitual suffix *-de* (§6.3.4.7). Similarly, in Example 527 two actions, performed by different referents, co-occur in the same chain of events: picking up a frog and being looked at. In Example 528, the subordinate clause has the stative marker *-ho*. The *gũ thĩ* construction forces a concomitant reading and avoids a sequential reading (having shined/ having reached a state of shining).

(526) tibi hwêde gũ thĩ ãngkũya

[tibi hwê-de gũ thĩ]_{cosub.cl} ãng-kũ-ya

T.V.[<E] watch-HAB.EGO NF CONC sleep-COMPL.II-ALLO

‘I fell asleep while watching TV.’ [250122_JO_eli2 924]

(527) cya làng we dẽkũbô la (...) hwẽamẽ gũ thĩ

cya làng we dẽ-kũ-bô la [hwẽ-amẽ gũ thĩ]_{cosub.cl}

3 COLL DEM pick.up-COMPL.II-PFV.ALLO HSAY see-from.behind NF CONC

‘They picked up (the frog), they say, while being seen from behind.’ [TA22_frog1 139 752]

(528) halo gũngho gũ thĩ àng hanadyibô

[halo gũng-ho gũ thĩ]_{cosub.cl} àng hana-dyibô

moon shine-STATE NF CONC house come-PFV.EGO

‘I came home while the moon was shining.’ [290924_STBT_eli1 991]

The *gũ thĩ* construction can be employed to highlight the concurrent nature of two events (as with co-temporality marker *dũ* CO.TEMP, cf. §10.3.4.1). In Example 529a, the simultaneous *gũ* interpretation does not imply an immediate sequence as does the *gũ thĩ* construction in Example 529b.

(529) a. bréiya gũ mlã hanabô

bréi-ya gũ mlã hana-bô

buy-ALLO NF village come-PFV.ALLO

‘S/he came to the village through marriage (lit. buying).’

b. bréiya gũ thĩ mlǎ hanabô

bréi-ya gũ thĩ mlǎ hana-bô

buy-ALLO NF CONC village come-PFV.ALLO

‘S/he came to the village immediately after marriage (lit. buying).’ [both 290924_STBT_eli1 980]

10.3.5 Grammaticalization path

The non-final marker *gũ* also exists as an instrumental (§9.5.1) and instrumental/locative marker (§9.5.1). A syncretism between a clause chaining/(co)subordinator device and the locative marker is not unheard of in the Tibeto-Burman language family. Commonly, however, there is a stronger tie between nominalizing devices, agentive/ergative (which is also an instrumental marker in many languages) and clause chaining/subordinator devices (cf. Bickel 1999c), particularly in Tibeto-Burman languages of Eastern Nepal and the Indo-Burmic group (Coupe 2017, 16f) but also beyond (e.g. in Wadu Pumi (Qiangic) Daudey 2014, 218f). In Macro Tani, the dative marker (along with the genitive/ablative) is the adverbial subordinator but the locative marker also plays a role in subordination, albeit usually compositional with other markers (Post and Modi 2024; Post 2009, 76f; Modi 2017, 446f).

As discussed by Coupe (2017, 16f) and Bickel (1999c), the metaphorical extension of nominal morphology to clause chaining, i.e. from a referential to predicative function in the Tibeto-Burman language family and beyond might reflect a shared cognitive basis. The situation in Tawrǎ stands out in this regard as the cosubordinator and the instrumental particle are the same morpheme, as opposed to other related languages where the subordinator is either a) historically derived from an instrumental/ergative marker/nominalizer or b) composed of an instrumental/ergative marker/nominalizer and another element.

The cosubordinator *gũ* has scope over whole clauses and often appears connected to a post-predicate demonstrative. This is illustrated in Example 530, where the cosubordinator is attached to a demonstrative *we gũ*. Probably due to their juxtaposition, collocations like *we gũ* (DEM NF) have crystallized into *discourse connectives* along with *webô*(*byē*) and

webôgũ (all ‘and then’), acquiring a function as discourse operators and often heralding the commencement of a speech act. Complex demonstratives will be discussed in §10.8.

- (530) *khalai dũ phlă dũ shyâzâ we gũ kadâm.kara la*
 [khalai dũ phlă dũ shyâ-zâ we gũ]_{cosub.cl} kadâm.kara la]_{main.cl}
 field with stone with become-DIR DEM NF a.place say
 ‘Because it was (perhaps *emerged*) out of field and stone, it’s called *Kadam-kara*.’
 [TA23_phlakam_clan_013 61]

10.4 Subordinating constructions

Subordination is here understood as the hierarchical embedding of a clause into a superordinate clause (Thompson et al. 2007, 237f; Haspelmath 1995, 8) as opposed to an ahierarchical or flat structure of co-subordination (cf. §10.3). §10.4.1 describes conditional and temporal subordination (particles *byě*, *bô* and *bobyě*). §10.4.2 amalgamates four complex subordinating constructions under the label *other subordinating constructions*: lest/until (§10.4.2.1), purposive *kanüim* (§10.4.2.2) and unrealized alternative and contra-factual conditional (§10.4.2.3). §10.4.3 describes the *thĩ* group, which includes two complex subordinating constructions that make use of the multi-clause marker *thĩ*. Note that the relative clause is discussed in §10.5.2 and complementation in §10.6. Table 10.3 presents an overview of the main constructions discussed under subordination.

Table 10.3: Overview of subordination markers and constructions

Construction(s)	Device(s)	Reference
Conditional & temporal subordination	<i>bobyě</i>	§10.4.1
‘lest/until’	<i>yüim tapwi</i>	§10.4.2.1
Purposive	<i>(-ya) kanüim</i>	§10.4.2.2
Unrealized alternative and contra-factual cond.	<i>(go)ji</i>	§10.4.2.3
<i>thĩ</i> group		
Concomitant action	<i>bũ thĩ</i>	§10.3.4.2
‘as if’	<i>da-thĩ</i>	§10.4.3.2

10.4.1 Temporal and conditional subordination *bô*, *bobyẽ* and *byẽ*

The set of non-final particles *bô*, *bobyẽ* and *byẽ* mark both temporal ('when') and conditional ('if') linking, i.e. the flagged clause is understood as the apodosis of a non-marked, main clause, signaling semantic dependency. The morphemes involved in this construction are: *bô*, *byẽ* and their combined form *bobyẽ*. The conflation of temporal and conditional subordination is a typologically common feature and their distinction rests on the semantics and pragmatics of discourse (Dixon 2006a, 14; also cf. Breugel 2014, 511-514,).

A brief digression on the composition and etymology of the morphemes involved is in order. The particle *byẽ* also marks topic (cf. §8.3), a common configuration in the languages of the world (cf Haiman 1978; also cf. Thompson et al. 2007, 245f.). The morpheme *bô* goes back to a homophonous versatile verb meaning 'go' which has been prolific in acquiring grammatical meaning in the language (as discussed in §6.3.4.8). When occurring without the *byẽ* segment, the *bô* morpheme is sensitive to egophoricity (cf. §6.3.1.1), as will be shown presently. The form *bobyẽ* (also cf. *webobyẽ* §10.8.2) shows signs of phonological integration, as the first segment loses its independent tone (presumably *bô* > *bo*).

For the sake of clarity, temporal and conditional subordination will be treated in two different subsections bearing in mind that all three suffixes express both meanings.

10.4.1.1 Conditional subordination

The temporal relation between the subordinated and the main clause is a definer for the type of linkage between the clauses (cf. a similar discussion in §10.3). In Examples 531 and 532, the relation between the clauses triggered by *byẽ* is one of condition ('if'); it signals a possible scenario without a clear temporal relation. The main clause represents the consequence (the apodosis).

- (531) (...) “nyu hayû tsâidângdyibô ka byě hã we (...) adyângra gũ krângna” la la
 [nyu hayû tsâi-dâng-dyibô ka byě]_{cond} hã we adyângra gũ
 2S first count-COMPL.I-PFV.EGO xxx COND 1S DEM back.of.neck LOC
 krâng-na la la]_{main.cl}
 bite-IMP say HSAY

“If you finish(ed) counting first, then, bite me in the back of the neck”, he said, it is said.’ [TA22_tiger 43-44 118]

- (532) “ane hwěya wêrēdyibô byě tatyě cyârēzâna” la la
 [ane hwě-ya wê-rē-dyibô byě]_{cond} tatyě cyârē-zâ-na la
 2P see-ALLO want-PL.II-PFV.EGO COND rope hang.down-PL.II-DIR-IMP say
 la]_{main.cl}
 HSAY

“If you wanted to see (me), hang down some ropes”, he said, it is said.’
 [TA22_kutyemba_stone 221 440]

The morpheme *bô* has an allomorph *dyibô* that is sensitive to the main actant of the clause in the same manner as the pattern described for egophoricity in §6.3.1.1 (which also applies to the perfective suffix pair *bô/dyibô* allophoric/egophoric). In Example 533, the *bô* particle signals conditional subordination. Conversely, in Example 534 the conditional device is *dyibô*, since the main actant is a 1st person singular.

- (533) bülei ã bô bülei
 [bülei ã bô]_{cond} bülei]_{main.cl}
 pig be.there COND.ALLO pig

‘If there is pig (available), (we offer) pig.’ [field notes 261024]

- (534) mề ahwêna dyibô cya we “gâmbra we ahwêtyu” labing
 [mề ahwê-na dyibô]_{cond} cya we gâmbra we ahwê-tyu
 person send-FACT COND.EGO 3 DEM gaonburha DEM send-already
 la-bing]_{main.cl}
 say-POT.ALLO
 ‘If I send somebody, (they) might say “(we) already sent the *goanburhas*).”’
 [TA24_land 119 216]

The combined form *bobyẽ* (*bô* + *byẽ*) occurs in Example 535, followed by a cosub-ordinate clause.

- (535) dra maciya marẽ ewe (...) “e ì bobyẽ ja thâ gũ lübiya”
 [dra maci-ya marẽ ewe e ì bobyẽ]_{cond.cl} ja [thâ gũ]_{cosub.cl}
 worthless harm-ALLO harm DEM DEM be.there COND what eat NF
 lü-biya]_{main.cl}
 survive-POT.ALLO
 ‘(...) this worthless (person) did (only) harm (...). (They said:) “If he stays (with us), how will we survive?”’ [TA23_tayang 24 140]

10.4.1.2 Temporal subordination

As discussed above, the non-final particles *bô*, *byẽ* and *bobyẽ* also express an inter-clause relation of temporal dependency (‘*having done X ...*’ or ‘*when*’). In Example 536, the subordinate clause takes the *bô* non-final marker. Semantically, the subordinate clause is understood to be temporally prior to the main clause.

- (536) we ashâ wrínêzâ bô we gũ ja shyâbô mane
 [we [ashâ wrí-nê-zâ bô]_{sub.cl} we gũ ja
 DEM plain.area come.down-ASSO-DIR TEMP.SUB DEM INS what
 shyâ-bô mane]_{main.cl}
 happen-PFV.ALLO FILL(<Asm)
 ‘Having come down to the plains, what happened, I mean.’ [TA24_land 280 517]

A modal particle (here *la* HSAY) is normally not accepted in a subordinate clause, although in Example 537 this is probably a confusion since the (omitted) verb in the main clause is also *la* ‘say’.

- (537) hwèshyü bô laga mlá mẽ we làng we (...)
 [hwè-shyü bô la-ga]_{sub.cl} mlá mẽ we làng we]_{main.cl}
 tighten-CAUS.V TEMP.SUB HSAY-PL.I village.MOD people DEM COLL DEM
 ‘Having had someone tighten it up, they say, (he said to) the people (...)’
 [TA22_kutyemba_phla 90 211]

It remains to be established whether the subordinator *bô* also has a *dyibô* counterpart when used for temporal subordination.

The morphemes *byẽ* and *bobyẽ* can also flag temporal subordination but this is less common than *bô*. In Example 538, the morpheme *byẽ* indicates temporal subordination. Similarly, in Example 539 this is flagged by the complex morpheme *bobyẽ*.

- (538) ráptatângkũnatyû byẽ macêi bôm
 [ráp-tatâng-kũ-na-tyû byẽ]_{sub.cl} macêi bô = m
 float-accumulate-COMPL-FACT-DIR TEMP.SUB water go = NEG
 ‘As (the debris) accumulated, the water could not flow. [TA23_newmyth 09 51]

- (539) lünega bobyẽ (...) mẽ we (...) kade.nyu tshûnaya (...)
 [lü-nega bobyẽ]_{sub.cl} mẽ we kade.nyu tshû-na-ya
 survive-DU TEMP.SUB person DEM how expand-FACT-ALLO
 ‘Having survived, these people, how (would they) thrive (...)?’ [TA23_newmyth 33 198]

Similarly to the sequential *gũ* marker discussed in §10.3.1, the iconicity of a sequence of events can give way to a causal relation between the subordinated and the main clause. In Example 540, the clausal subordinator *bô* indicates the reason for the event in the main clause.

- (540) (...) hanâmâ hanâmâ hrû yüm bô khümyûa (...)
 [hanâ-mâ]_{main.cl} [hanâ-mâ]_{main.cl} [[hrû yüm bô]_{sub.cl} khümyû-ya]_{main.cl}
 squeeze-ITER squeeze-ITER burn NEG TEMP.SUB be.angry-ALLO
 ‘(...) he squeezed and squeezed (his testicles). Since (the fire) did not burn, he became angry.’ [TA22_tiger 116 262]

A final note is in order regarding the common collocations of subordinators and the cosubordinator *gũ*, e.g. *bô gũ* (TEMP.SUB NF): In Example 541, the subordinate clause occurs with a non-final marker *gũ*. Similarly, in Examples 542 and 543 the subordinator *byẽ* occurs along the non-final marker *gũ*. These examples can be thought of as instances of incomplete grammaticalization: the morphemes involved also occur in other contexts and may still carry structural aspects of their original use when occurring as non-final markers. In the words of Givón (2018, 21): “(...) diachronic change, much like biological evolution, begins with functional re-analysis, and that structural re-adjustment invariably lags behind”. Here we can distinguish two patterns: the cosubordinator *gũ* occurs after *bô* (i.e. *bô gũ*) and before *byẽ* (i.e. *gũ byẽ*).

- (541) hwâ yüm bô gũ marấrẽ laga (...)
 [hwâ yüm **bô** **gũ**]_{sub/cosub} marấ-rẽ la-ga]_{main.cl}
 want NEG TEMP.SUB NF leave-PL.II HSAY-PL.I
 ‘(After) getting tired (lit. did not want), they left (him), they say. (...)’
 [TA22_kutyemba_phla 316 598]

- (542) ephlâe nyu mề thâ gũ byẽ (...)
 [e=phlầ=e nyu mề thâ **gũ byẽ**]_{cond}
 DEM=stone=DEM FOC people eat NF COND
 ‘If this stone eats people (...)’ [TA22_kutyemba_stone 46 142]

- (543) nyu mu plà shìya gũ byẽ (...)
 [nyu mu plà shì-ya **gũ byẽ**]_{cond}
 2S also salt take-ALLO NF COND
 ‘If you also (want to) take salt (...)’ [TA22_kutyemba_stone 432 803]

10.4.2 Other subordinating constructions

This section includes an eclectic group of subordinating constructions that share the use of more than one device (e.g. a particular verb plus the negative morpheme) and have highly specialized pragmatic effects. §10.4.2.1 describes the *lest/until* construction. §10.4.2.2 describes the purposive construction. Finally, the ‘unrealized alternative’ and ‘contra-factual conditional’ are described in §10.4.2.3.

10.4.2.1 Lest/until NEG *tapwi* construction

The *lest/until* construction is composed of the subordinator *tapwi* attached to a negated verb. Table 10.4 illustrates the structure of this construction.

Table 10.4: Structure of the *lest/until* construction

[V *yüm tapwi*]_{sub.cl}

The subordinate clause is understood as a *possible consequence*, i.e. the action in the main clause should be performed to prevent the event in the subordinate clause (Dixon 2009, 23f.). The polarity of the main clause remains independent.

In Example 544, the villagers want to kill the mythical *Pâshyã*. The subordinate clause corresponds to an (undesired) development while the main clause is understood as the actions taken by the villagers to ensure that *Pâshyã* would not resuscitate once more. Similarly, in Example 545 the subordinate clause expresses an unwanted event that hinders the realization of the main verb. Since the main verb is negated, the whole clause is understood as a condition: ‘if you leave, I’ll talk’.

(544) lêlyâ yüm tapwi la gũ bü lyûm sênâ mu

[lê-lyâ yüm tapwi la gũ]_{sub.cl} bü lyûm sê-nâ mu
resuscitate-back NEG until say NF CLF:times eight kill-AGAIN also

‘(They) killed (him) eight times, lest he would resuscitate.’ [TA23_tayang_clan 32 193]

- (545) abang bô yüm ja mu la yüm
 [abang bô yüm tapwi]_{sub.cl} ja mu la yüm
 outside go NEG until what also say NEG
 ‘I won’t say anything until you leave.’ [150924_HM_eli 813]

10.4.2.2 Purposive *kanüm*

A purposive clause ‘in order to’ is a type of dependent clause formed with a periphrasis that commonly involves a predicate marked with the allophoric suffix *-ya* and the particle *kanüm* (sometimes *künüm* and *knüm*) ‘in order to’. The marked clause is understood as the cause or the explanation of the main clause.

In Examples 546 and 547, the purposive clause precedes the main clause. A negative purposive clause (‘in order not to X’) is achieved through the general negative suffix *yüm* (cf. §6.4.1), as illustrated in Example 548.

- (546) sîk thâya kanüm dîhoq
 [sîk thâ-ya kanüm]_{sub.cl} dî-ho-q]_{main.cl}
 kernel eat-ALLO PRP sit-STATE-CONT
 ‘(The bird) sits in order to eat kernels.’ [220222_RI_eli1 83]

- (547) tyûnagaya künüm mề khing agû khrâk
 [tyû-na-ga-ya künüm]_{sub.cl} mề khing agû khrâk]_{main.cl}
 play-FACT-PL.I-ALLO PRP person one watch EVI
 ‘One person guarded in order to play (the gong), it seems (i.e. in case the trickster were to escape).’ [TA23_kutyemba_phla 393 740]

- (548) maroya yüm künüm cya thalyâng.na krangtyû
 [maro-ya yüm künüm]_{sub.cl} cya thalyâng.na krang-tyû]_{main.cl}
 speak-ALLO NEG PRP 3 tongue bite-DIR
 ‘S/he bit his own tongue in order not to speak.’

The origin of the non-final particle is unclear. There is a homophonous noun meaning ‘reason, cause’ but this can be a later development. A possible pathway for its grammaticalization could be along the lines of the contra-factual alternative construction *ji(-na) yüm* (§10.4.2.3.2), which often occurs as *jina = m*. Similarly, in the case of the purposive a potential verb **ka* could have been recruited as a purposive subordinator and it might have taken along the factitive suffix *-na* and the general negator *yüm* (or rather its cliticized form *=m*), rendering *ka-na = m* and eventually *kanüm/künüm*. A suspect verb is *ka* ‘untie’; it could have formed a metaphor like *not being untied* which was eventually re-analyzed as a purposive particle. The alternative pronunciation *künüm* and *knüm* would reflect a further stage of phonological integration: the construction would have been re-analyzed as one morpheme with stress/tone falling on the last syllable (iambic foot, cf. §2.2.1.1). Light syllables (§2.2.5) undergo various forms of erosion like centralization (/a/>/ü/) and loss of the vocalic nucleus (/ü/>∅).

10.4.2.3 Unrealized alternative and contra-factual conditional with *(go)ji* and *ji*

Two constructions employ the (related) non-final markers *(go)ji* and *ji* to express unrealized alternatives and contra-factual conditional. While the marker *ji* is related to the verb *ji* ‘do’, the origin of the optional *go-* segment is unknown.

10.4.2.3.1 Unrealized alternative *(go)ji*

The non-final marker *(go)ji* indicates unrealized potential scenarios, termed here *ir-real*. This construction commonly occurs with the modal particle *ki/kale/kalai/ke* (§9.2.8). In Example 549, the subordinate clause takes the *goji* marker, rendering it an unrealized scenario.

(549) *hã kasâ ki nyu hã ahuwa goji tatyâya ki*

[*hã kasâ ki*]_{main.cl} [[*nyu hã ahu-ya goji*]_{sub.cl} *tatyâ-ya ki*]_{main.cl}

1S know COUNT.EXP 2S 1S ask-ALLO IRR tell-ALLO COUNT.EXP

‘I would have known, if you would have asked me, I would have told you.’

[291124_BAB_eli1 253]

The subordinate clause is autonomous in its polarity. The negator occurs between the main predicate and the non-final marker *(go)ji*. The subordinate clause in Example 550 is understood as a potential event and also as the reason that didn't allow the apodosis to take place. Conversely, in Example 551, the apodosis, here a hypothetical event is negated (the bird *not* flying away).

- (550) *cya làng yu tyû yüm ji kalyab bôbôya ki*
 [cya làng yu tyû yüm ji]_{sub.cl} kalyab bô-bô-ya ki
 3 COLL beer drink NEG do field go-PFV.ALLO-ALLO COUNT.EXP
 'If they hadn't drunk beer, they would have gone the field.' [011024_STBT_eli1 485]

- (551) *takhrâkbôm ji pîya bábôya yüm*
 [takhrâk-bô = m ji]_{sub.cl} pîya bá-bô-ya yüm
 maike.noise-PFV.ALLO = NEG do bird fly-PFV.ALLO-ALLO NEG
 'If (we) hadn't made any noise, the bird wouldn't have flown away.'
 [011024_STBT_eli1 350]

10.4.2.3.2 Contra-factual alternative *ji(-na) yüm*

The contra-factual alternative construction also involves the verb *ji* 'do' (cf. §10.4.2.3.1). This construction involves the verb/particle *ji*, which commonly takes an optional factitive suffix *-na* (§6.3.6.2.1) and the whole clause is negated. This is a further example of a hybrid non-final particle (§9.6), since *ji* still operates as a (auxiliary) verb in hosting verbal suffixes and being negated.

Example 552 illustrates the contra-factual alternative construction, which syntactically and semantically subordinates the flagged clause to the main clause. The subordinated cause is an unrealized scenario. In Example 553, the speaker refers to the widespread belief that *dholes* (*Cuon alpinus*) could talk. He employs this construction (without the factitive *-na*) to signal that, even though they could not speak, they were deemed useful.

(552) puriya jinam kalyab bâ

[*puri-ya* *ji-na = m*]_{sub.cl} *kalyab bâ*]_{main.cl}

study(<IA)-ALLO do-FACT=NEG field do

‘Instead of studying s/he works the land.’ [291024_STBT_eli2 1060]

(553) (...) weli cya taprûw a we nyu marone ji yüm we weli mé kam (...) jiya la

weli *cya taprûw a* *we nyu maro-ne ji yüm we*

DISC.DEIC.DEM 3 dhole pup DEM FOC speak-INT do NEG DEM

weli *mé kam* *ji-ya* *la*

DISC.DEIC.DEM all work(<IA) do-ALLO HSAY

‘(...) even though the dhole pups could not speak, they (...) were all useful (lit. did work), they say.’ [TA23_talyu_clan 53 330d]

The marker of this complex construction (*ji-na = m*) superficially resembles the purposive particle *kaniüm* (§10.4.2.2). As hinted in §10.4.2.2, these subordinators might share a diachronic origin, in which a verb, the factitive suffix and the general negator can be reanalyzed as a subordinator: *V-na yüm* > *V-na = m*. This development is not complete in the construction *ji-na = m*, as the constituent parts still retain their meanings and syntax.

10.4.3 Secondary non-final markers: the *thĩ* group

The multi-clause marker *thĩ* cannot modify a clause by itself; it occurs coupled with a preceding morpheme in three complex non-final markers. This section discusses temporal and hypothetical subordination *bü thĩ* (§10.4.3.1) and the ‘as if’ *da thĩ* construction (§10.4.3.2). The concomitant action *gũ thĩ* has been discussed in §10.3.4.2. Due to its broad range of uses, a generic label CONC (i.e. *concomitant*) is preferred over a number of specific labels.

10.4.3.1 Temporal and hypothetical subordination *bü thĩ*

The temporal and hypothetical subordinator *bü thĩ* indicates that the action in the subordinate clause occurs prior to the main clause. It may also be employed in a hypothetical sense, e.g. *whenever you go, do X*. This construction employs the adverb/classifier

bü ‘once, time(s)’ and the multi-clause marker *thĩ/hĩ*. In Example 554, the subordinate clause occurs prior to the main clause and is also its logical antecedent.

(554) hana bü thĩ maromâ

hana bü thĩ maro-mâ

come once CONC talk-ITER

‘When(ever) s/he came, s/he talked too much.’ [100622_JA_BAT_eli2 2815]

10.4.3.2 ‘As if’ *da thĩ*

A subordinator *da* occurs along with the multi-clause marker *thĩ* in a construction that signals the efforts of the S/A argument to imitate, i.e. to do *as if* somebody else (*hypothetical manner* for Dixon 2009, 35). This construction is often used humorously. The particle is likely related to the versatile verb *dá* ‘return, pay off’ (also as a predicate derivation ‘in exchange, in return’, cf. §6.2.3.5) and was stripped of its suprasegmental features as part of its grammaticalization.

The clause marked with *da thĩ* includes the imitated referent and is understood to be subordinated to the main clause. Examples 555a and 555b illustrate this construction.

(555) a. cya kalà da thĩ chêk

cya [kalà da thĩ]_{sub.cl} chêk

3 officer as.if CONC walk

‘S/he walks as if s/he was a (police) officer.’ 290924_STBT_eli1 882]

b. cya cha da thĩ jiya

cya [cha da thĩ]_{sub.cl} ji-ya

3 Assamese as.if CONC do-ALLO

‘S/he behaves (lit. does) as if s/he was Assamese.’ [290924_STBT_eli1 856]

10.5 Nominalization based constructions

This section discusses three constructions based on nominalization in the language. After introducing relevant typological considerations in §10.5.1, this section will introduce the relative clause in §10.5.2, event nominalization in §10.5.3 and the locative nominalization with *blúm* in §10.5.4. While the latter has a specialized device to mark the nominalization, the former two constructions do not have a dedicated morpheme. It should be noted that cross-class membership, i.e. the ability of a root to occur as different word classes (sometimes termed *zero-derivation*) is discussed in §3.2.1 and is not considered a productive form of nominalization.

10.5.1 Theoretical considerations

Nominalization is here understood as a process by which a nominal expression is derived from a term that cannot be otherwise treated as a noun in the language (cf. Post 2011a, 257; Yap et al. 2011, 3). In turn, a noun in Tawrã can be minimally defined as having the ability to stand as head of an argument NP (cf. §3.2.2).

For many Asian languages, it has been noted that nominalization is involved in a myriad of constructions and functions, including relativization, genitive constructions, complementation and focus constructions. Such heterogeneous constructions seem to converge under the umbrella of nominalization in many languages, something that has been termed the *Standard Sino-Tibetan Nominalization* (cf. Bickel 1999a; cf. DeLancey 2002, 56f. cf. Post 2011a, 256f. also cf. Genetti 2011).

In Tawrã, we can distinguish three constructions that involve nominalization: the relative clause (§10.5.2), the locative nominalization (§10.5.4) and the event nominalization (which also includes nominal complements; §10.5.3). One way to organize these constructions is by the type of nominalization: while the relative clause active nominalization are unmarked, the locative nominalization has a specialized morpheme *blúm*. The lack of a specialized device for nominalization is also known as *zero nominalization* and is shared by certain morphologically simplified languages of the Tibeto-Burman family, such as Turung (Morey 2012, 498) and Old Chinese (LaPolla 2017, 93f). However, it

is also found in morphologically complex languages like Japhug (*finite relative clauses*, cf. Jacques 2021, 1053-1055, 1258) or Rawang (LaPolla 2008a, 800). One possibility is that Tawrã, along with the languages aforementioned, would have kept this morphologically simple strategy, while many other members of the language family adopted innovative nominalizing devices (for an overview cf. Post 2011b), including the sister-language Kera'a, which has a nominalizer/nominative case marker *-me* (possibly grammaticalized from the word 'person'; Peck 2023, p.c.). Alternatively, Tawrã could have developed a nominalizing device which was then lost over time. There is no evidence at present to discard any of these possibilities. In any case, unmarked nominalization has remained largely unexploited as compared to other nominalizing strategies in related languages which have extended their functions to other domains of the language (cf. Post 2011a; Noonan 1997; Bickel 1999a). Regarding the relation between nominalization and other aspects of the language, it should be noted that the structure [nominal₁][nominal₂] (i.e. unmarked apposition) is also the form of nominal modification in the language (cf. §4.2).

10.5.2 Relative clause

A relative clause is a subordinate clause (§10.3) that narrows down the reference of a NP by providing some type of information (cf Andrews 2007b, 206). It therefore belongs to the grammar of the NP as a modifier of the head noun (Givón 2018, 126). As discussed at the end of §10.5.1, nominalizations and relative clauses are syncretic in Tawrã. The *nominalization-relativization* syncretism in many members of the language family likely goes back to a tendency to have a [nominal + nominal] structure in which the preceding nominal (in this case, a nominalized verb) modifies the following nominal (LaPolla 2008a, 807). According to LaPolla (2017, 93f), this construction would also be the origin of the genitive in the language family.

A nominalized clause is identical to a main clause in morphological terms, i.e. unrestricted in the type of morphology that the verb can take. However, it differs from a main clause in that it includes a syntactic gap; the relative clause does not have a head NP argument and is rather syntactically contained within the noun phrase it relativizes (cf. a similar analysis for Rawang, cf. LaPolla 2008a, 800f; LaPolla 2008b, 50). Table 10.5 illustrates the structure of a relative clause in Tawrã. The relative clause (rel.cl)

usually precedes the NP head but is syntactically within it. The gap in the relative clause (that is co-referential with the nominal head of the NP) is signaled by the null morpheme $\emptyset$. As will become clear in this section, the function of the relative clause can be that of either a subject or object; no nominalization strategy is associated with a particular role.

Table 10.5: Structure of the relative clause

$$[[\emptyset_i \text{ predicate}]_{\text{rel.cl}} \text{ nominal head}_i]_{\text{NP}}$$

Examples 556 and 557 include relative clauses: In Example 556, *mîyaqa* ‘girl’ is the head of the noun phrase that is relativized. Within the relative clause, it has the function of a P argument. The structure is thus straightforward: the head noun of the clause ‘the girl has died’ is relativized by adding a clausal modifier (i.e. the unmarked relative clause) inside the NP itself. In the relative clause, there is a syntactic gap that stands for the grammatical function of the relativized NP - here a P argument (cf. Andrews 2007b, 206). Similarly, in Example 557 the gap in the relative clause is co-referential with the head of the NP. The function of the relativized nominal head within the relative clause is that of an A argument. There are no restrictions regarding the semantic function of the relativized referent in the relative clause *vis-à-vis* the higher clause.

(556) *hã pra wêde mîyaqa we shèkūnaya*

<i>hã</i>	$\emptyset$	<i>pra</i>	<i>wê-de</i>		<i>mîyaqa</i>	<i>we</i>		<i>shè-kū-na-ya</i>
1S		good	think-HAB.EGO		girl	DEM		die-COMPL.II-FACT-ALLO
[	$\emptyset_i$		V	]	rel.cl	N_i	]	NP_{mat}
							[V	]
								pred

‘The girl that I loved passed away.’ [030622_JO_eli1 469]

(557) *tawé ãya jâp we làng iskul bôga*

$\emptyset$	<i>tawé</i>	<i>ã-ya</i>		<i>jâp</i>	<i>we</i>	<i>làng</i>	<i>iskul</i>	<i>bô-ga</i>
	mind	be.there-ALLO		a.tribe	DEM	COLL	school(<E)	go-PL.I
[	$\emptyset_i$	N	V	]	rel.cl	N_i	]	NP_{mat}
							[N	V
								pred

‘The Kaman who have a mind (i.e. are intelligent) go to school.’ 290924_STBT_eli1 836]

It should be noted that a T argument (‘not most goal-like, not most actor-like’, cf. §8.1) can also be relativized, as in Example 558. Adjuncts, however, cannot be relativized, as in Example 559a. This example can be rephrased with an embedded purposive paraphrase, as in Example 559b (cf. §10.4.2.2). Complementation is discussed in §10.6.1.

(558) cithi zûzâ dotpen we hàngna

[$\emptyset_i$ cithi zû-zâ]_{rel.cl} [dotpen_i we]_{head.nom}]_{NP} hàng-na
 letter(<IA) write-DIR pen(<IA/E) DEM give-IMP

‘Give me the pen with which (you) wrote the letter.’ [150924_HM_eli 139]

(559) a. *cya sêbô knüm kasâ yüm

cya sê-bô knüm kasâ yüm
 3 kill-PFV.ALLO reason know NEG

Intended: ‘I don’t know the reason why he killed him’.

b. ja knüm cya sêtyeqîng kasâ yüm

[ja knüm cya sê-tyeqîng]_{compl.cl} kasâ yüm
 what reason 3 kill-POT know NEG

‘I don’t know the reason why he killed him.’ [301024_STBT_eli1 1348]

Normally, the relative clause precedes the argument. This strategy corresponds to a *pre-nominal* relative clauses (Bickel 1999a; Andrews 2007b, 208f). Tawrã thus follows a common pattern in the Tibeto-Burman family in which relative clauses precede NP heads while adjectives follow them (Andrews 2007b, 210; e.g. in Manange cf. Hildebrandt 2004, 174f) (cf. §3.3.3).

This is, however, not the only strategy available in the language: the relativized argument may also be embedded within the relative clause, i.e. surrounded by material that belongs to the relative clause and not to the matrix (Bickel 1999a, 273). In Example 560, the relativized argument *àng* ‘house’ is *within* its own relative clause since the surrounding elements belong to the relative cause and not to the matrix. There is in this sense no syntactic *gap* in the embedded sentence (cf. Comrie 1989, 147).

- (560) blehîng àng hwě.tyâng we (làng) kě tashěne
 [blehîng àng hwě.tyâng we làng]_{NP/rel.cl} kě tashě-ne
 yesterday house see DEM COLL LOC.INS sing-INT
 ‘I’ll sing in the house that I saw yesterday.’ [290924_STBT_eli1; modeled after
 Lehmann (1984:111) and Comrie (1989[1981] 145)]

This alternative strategy has been termed *internally headed relative clauses* (Bickel 1999a, 273; Andrews 2007b, 212) and it is mostly associated with North American languages such as Mohave, Diegueño and Navajo. In these languages, internally headed relative clauses are often in complementary distribution with other forms of relativization (cf. Lehmann 1984; Vries 2018, 3; for an example in Pesh, a Meso-American language cf. Chamoreau 2021, 293f). Internally headed relative clauses have also been described for the Tibeto-Burman language family (for Galo, cf. Post 2007, 763f; in Kiranti, cf. Bickel 1999a; in Mongsen Ao, cf. Coupe 2008, 219f; in Japhug, cf. Jacques 2021, 2171f), although a large, comparative analysis of this construction in the language family is warranted.

In the case of Tawrã, internally headed relative clauses are rare and have only been attested in elicitation. The functional and pragmatic distribution between these two strategies (as in the case of Japhug, cf. Jacques 2021, 2171f) is not yet understood and requires further investigation.

Tawrã relative clauses can be headless, i.e. a relativized NP without a head noun is also allowed in Tawrã. As discussed in §4.1.3, headless NPs are often retrievable from the context. Examples 561a and 561b illustrate two configurations of the relative clause: The relative clause can be placed before the relativized NP (pre-nominal), as in Example 561a (the most common pattern). It may also occur after the head NP of the main clause *mě* ‘person’, as in Example 561b (internally headed).

(561) a. kalyaw bâya mề bôbô

[$\emptyset_i$ kalyaw bâ-ya]_{rel.cl} mề_i]_{NP} bô-bô
 field do-ALLO person go-PFV.ALLO

b. mề kalyaw bâya bôbô

[mề kalyaw bâ-ya]_{NP/rel.cl} bô-bô
 person field do-ALLO go-PFV.ALLO

Both: ‘The farmer (lit. field work person) is gone.’ [both 290823_STBT_eli2 523]

A relative clause can also be headless, i.e. without a nominal head. In Example 562, the head of the relative clause is the collective suffix (*làng*), which has certain nominal properties but is not a noun *sensu stricto* (cf. §3.3.1.4). The NP may be completely headless, as in Example 563.

(562) caina boder (...) boder nyu shyâzâ làng (...) khìga na

[caina boder boder nyu shyâ-zâ]_{rel.cl} làng]_{NP} khì-ga na

China(<E) border(<E) border(<E) FOC happen-DIR COLL reach-PL.I INSIST

‘The ones that were at the Chinese border, the border they reached (there), right?’
 [TA22_clan 128 372]

(563) [$\emptyset$ kalyaw bâ-ya]_{rel.cl}]_{NP} bô-bô

field do-ALLO go-PFV.ALLO

‘The farmer (lit. field worker) is gone.’ [290823_STBT_eli2 523]

A brief note on the role of the demonstrative *we* is in order. As discussed in §3.5.4, the demonstrative *we* serves a pragmatic function of focusing the attention on relevant participants. It is not *per se* involved in nominalization/the relative clause, as the probably cognate nominalizer/demonstrative in Rawang *wē* (LaPolla 2008a) and other, potentially cognate devices in the languages of the area (indicative nominalizer *ve*, cf. Matisoff 1973a, 406f; *wa* in Turung, cf. Morey 2012, 346). In Example 564a, the NP occurs in a

verbless clause and is relativized. The function of *we* is that of a demonstrative and not that of a nominalizer, as it can be omitted. In Example 564b, the headless relative clause does not have a demonstrative.

- (564) a. *tabrê thâdyibô we pra yüm*
 [*tabrê thâ-dyibô*]_{rel.cl} *we*_{NP} *pra yüm*
 meat eat-PFV.EGO DEM good NEG
 ‘The meat that I ate is/was not good.’
- b. *bréidyibô bra gê kamaho*
 [*bréi-dyibô*]_{rel.cl} *bra gê*_{NP} *kama-ho*
 buy-PFV.EGO CLF:generic one be.lost-STAT
 ‘The one I brought got lost.’ [both 150924_HM_eli 190]

10.5.3 Event nominalization

Event nominalization refers to the nominalization of an action as opposed to nominal constituents that function as arguments with referential status within a clause (Yap et al. 2011, 3f.). In Tawrã, this nominalization is unmarked. As discussed in §10.5.1, a nominal expression derived from nominalization acquires certain *nominal* properties (cf. Andrews 2007b, 232; Comrie 1989, 145). In the case of event nominalization, the derived expression allows certain nominal suffixes, such as locative, collective and possessive markers. In Example 565, the nominalized verb is possessed by *cya* - ‘his/her’. Similarly, in Example 566 a predicate (*înghoya*) is marked by a locative postposition.

- (565) *cya maro we kapú shì*
cya [*maro*]_{rel.cl} *we kapú shì*
 3 talk DEM disgust take
 ‘His talking annoyed me.’ [150924_HM_eli 300]

- (566) (...) mathi we myûm ãnghoya kě dûhoa la
 mathi we myûm [ĩng-ho-ya]_{rel.cl} kě
 younger.sister DEM elder.brother(fem.prsp.) sleep-STAT-ALLO LOC.INS
 dû-ho-ya la
 jump-STAT-ALLO HSAY
 ‘(And then) the sister jumped to (her) brother’s sleep (place), they say.’
 [TA23_newmyth 44 267]

Event nominalizations can be employed to modify a further NP. This strategy, also known as *nominal complementation* (cf. Dixon 2006a, 15f.), mirrors the standard form of nominal compounding (§4.2) in which the modifying nominal (here the event nominalization) precedes the head of the compound. Note that clausal complementation is discussed in §10.6.

In Example 567, the modifying clause (i.e. the nominalized clause) modifies the nominal head *hadi* ‘story’. Note that there is no gap in the nominalized clause and it is not subordinated to a higher clause as a relative clause (§10.5.2) (cf. Genetti 2006, 145f.). Finally, it should be noted that this strategy has only been found in elicitation; it does not seem to be prominent in the language (a similar observation has been reported for Galo, cf. Post 2007, 761).

- (567) kutyemba mlá mẽ lâhlak hadi
 [kutyemba mlá mẽ lâhlak]_{mod.cl} hadi
 a.name village.MOD person cheat story
 ‘The story of Kutyemba cheating the village people.’ or ‘The Kutyemba-cheating-the-village-people story.’ [011024_STBT_eli2 685]

10.5.4 Locative nominalizer *blûm*

The locative nominalizer *blûm* is related to the independent noun *blûm* ‘place’ (a common grammaticalization process, cf. Kuteva et al. 2019, 334f.). When modifying verbs, it refers to places where the action of the verb may take place e.g. ‘place to

sleep’. The modified verb disallows any verbal suffix on the stem. This differs from the unmarked nominalization, e.g. in relative clauses, which has no restriction in this regard (cf. §10.5.2). In Example 568, a relative phrase nominalized by *blúm* is inside a headless (adnominal) noun phrase. Conversely, adding a verbal suffix to the modified verb is not allowed, as displayed in Example 569.

- (568) *cya lāng blúm we kē gũ*
[cya lāng blúm]_{rel.cl} we kē gũ_{NP}
 3 enter LOC.NMLZ DEM LOC.INS LOC
 ‘At the place where he entered.’ [TA22_kutyemba_phla 276 528]

- (569) **lāng-ya blúm*
[lāng-ya blúm]_{rel.cl}_{NP}
 enter-ALLO LOC.NMLZ
 Intended: ‘where he enters/entry-place.’

This form of nominalization is also a productive way to craft new nominals, as in Example 570a and 570b.

- (570) a. *îng blúm*
 sleep LOC.NMLZ
 ‘bed (lit. sleeping place)’
 b. *mế thāng prũ blúm*
 person.MOD body burn LOC.NMLZ
 ‘cemetery (lit. place where human bodies are cremated)’
 [both 011024_STBT_eli2 701]

Several TB languages have nominalizers with varying degrees of specialization that go back to nouns meaning ‘place’, ‘person’, etc. (also termed ‘secondary nominalizers’, cf. Yap et al. 2011, 5f; LaPolla 2008a, 806f; Post 2007, 757; LaPolla and Huang 2003,

223f). We may assume that there is a functional pressure for nouns that commonly occur in certain constructions to undergo grammaticalization. However, it is noteworthy that Tawrã has not grammaticalized an agent nominalizer (*nomen agentis*) as have a large number of related languages, including Kera'a (-*mī*) or Galo (-*nà* 'nomen agentis', cf. Post 2007, 754). A candidate for grammaticalization is undoubtedly *mě* 'person', as in *kam bâ-ya mề* (work(<IA) make-ALLO person) 'worker'. However, there are no signs, neither phonological nor syntactic, that this process would occur, as it did in the sister language Kera'a, in Dzongkha and Qiang (LaPolla and Huang 2003, 131).

10.6 Complement clause and embedded discourse

Complementation refers to the strategy by which certain verbs take a whole clause (the *complement clause*) instead of a NP as an argument (Dixon 2006a, 1). As such, it corresponds to a form of subordination (§10.3). This section will describe a general strategy of apposition (§10.6.1) as well as the restrictions of particular verbs (*wê* 'think, want' §10.6.1.1; and *dù* 'need, must' §10.6.1.2). Direct speech complements and the grammaticalization of *la* 'say, call' will be discussed in §10.6.2. Embedded cited discourse will be discussed in §10.6.3. A nominalized clause can be employed to modify a nominal. This has also been termed *nominal complementation* and has been discussed in §10.5.3.

10.6.1 General strategy for complementation

Verb complementation is achieved through the direct apposition of the main clause to its argument, without any complementizer, infinitive or nominalizing device. Both the main and the complement clauses can take verbal suffixes and structurally resemble two apposed main clauses. Syntactically, however, the complement clause is inside the main clause. The main verb in the complement clause is autonomous in the morphology it can take. As an example, the complement clause is autonomous regarding egophoricity, which stands for a high-level propositional category (§6.3.1.1). In Example 571, both the complement and the main clause have an egophoric marker on the predicate (cf. the intentional suffix in §6.3.4.6). In Example 572, the predicate in the complement clause

has an allophoric marker whereas the predicate in the main clause has an egophoric marker.

(571) tahwîne wêde

[*tahwî-ne*]_{compl.cl} *wê-de*]_{main.cl}

teach-INT think-HAB.EGO

‘I want to teach (you).’ [TA23_pineapple_07 25]

(572) (...) hwěya wēdyibô byě

[*hwě-ya*]_{compl.cl} *wē-dyibô* *byě*]_{main.cl}

see-ALLO think-PFV.EGO COND

‘If you wanted to see (me) (...)’ [TA23_kutyemba_phla_227 450]

The *clause status* of the complement clause is not self-evident and requires a brief explanation, since arguments are commonly omitted in the language if understood from the context (cf. §8.2.5). Furthermore, many complement-taking verbs in other languages correspond to predicate derivations in Tawrā (cf. §6.2, as well as the case of the verb *dù* as briefly discussed at the end of §10.6.1.2). However, in Example 573, each clause has its own main predicate and a distinct S argument: the main clause has *ning* ‘we’ (1P) and the complement clause ‘(my) father’.

(573) ning katyângde we gwók we talyũ gwók we (...) bü bà we gwók ga bô we

ning katyâng-de [*we gwók we talyũ gwók we bü bà we gwók*

1P hear-HAB.EGO DEM priest DEM a.clan priest DEM FS father DEM priest

ga bô we]_{compl.cl}

self TEMP.SUB DEM

‘We have heard that this priest, the Talyũ priest, (...) (my) father was himself a priest.’ [TA23_Talyu_clan_86 524]

Most predication verbs allow different arguments in the complement clause than in the main clause. In Example 574, the A argument of the main clause differs from the omitted S argument of the complement clause.

- (574) cya àng bôna hwě.tyângya la
 cya_i [∅]_j àng bô-na]_{pred.cl} hwě.tyâng-ya la
 3 house go-FACT see-ALLO HSAY
 ‘He_i saw him_j go home, it is said.’ [150924_HM_eli 350]

The verbs *wê* ‘think, want’ and *dù* ‘need, must’ have an idiosyncratic different-S/A-restriction in complement clauses. These are discussed separately in §10.6.1.1 and 10.6.1.2 respectively.

10.6.1.1 Different-subject restriction of *wê* ‘think, want to’

The verb *wê* ‘think, want to’ differs from other verbal complement-taking verbs in that it doesn’t allow a different S/A argument in the complement clause. Instead, to express a different S/A argument with the verb *wê* ‘think, want to’, the complement clause must be marked by the hortative (§6.3.5.0.3), thus turning the construction into a wish or thought: ‘may you go away! I think’. In Example 575a, the same S/A argument in the complement clause is allowed while a different S/A argument in the complement clause is not, as in Example 575b. An equivalent sense, however, can be expressed through a workaround construction with the hortative *-ké*, as in Example 575c.

- (575) a. hã bôya wêde
 hã [bô-ya]_{compl.cl} wê-de
 1S go-ALLO think-HAB.EGO
 ‘I want to go’ or ‘I am thinking of going.’
- b. *hã nyu bôya wêde
 hã nyu bô-ya wê-de
 1S 2S go-ALLO think-HAB.EGO
 Intended: ‘I want you to go.’

c. *hã nyu bôbingké wêde*

hã [nyu bô-bing-ké]_{compl.cl} wê-de

1S 2S go-POT.ALLO-HORT.PL think-HAB.EGO

‘I want you to go’ (Lit: ‘I think: “may you probably go!”’).

[all three 071023_JOTHA_eli1 668]

10.6.1.2 Same subject restriction of complementation modal verb *dũ* ‘need, must’

The modal verb *dũ* ‘need, must’ can fill its argument slot with a nominal P arguments as well as by a complement clause. It differs from other verbs in that it does not allow a different argument in the complement clause. In Example 576, the A argument of the complement and the higher clause is the same. Conversely, a different A argument in the complement clause is banned, as in Example 577a. The modal verb demands the same A argument in the complement clause, as in Example 577b. There is no alternative construction like with the verb *wê* ‘think, want’, as discussed in §10.6.1.1.

(576) *ning gâmbra lûmkũ hî dângkũ dũwa*

[ning gâmbra lûmkũ hî dâng-kũ]_{compl.cl} dũw-ya

1P gaonburha(<I.A.) inside EMPH finish-COMPL.II need-ALLO

‘We had to settle (any quarrels) within our gaonburhas.’ [TA22_clan_42 150]

(577) a. **hã nyu thâyâ dũwa*

hã nyu thây-ya dũw-ya

1S 2S eat-ALLO need-ALLO

Intended: ‘I need you to/that you take (medicine).’ [071023_JOTHA_eli1 1232]

b. *hã tama thâyâ dũwa*

hã [tama thây-ya]_{compl.cl} dũw-ya

1S medicine eat-ALLO need-ALLO

‘I need to take medicine.’ [071023_JOTHA_eli1 1202]

10.6.2 Direct speech complement

Direct speech is embedded *verbatim* as a P argument by a verb of utterance like ‘say’ or ‘ask’. There is no way to report an utterance, such as *indirect speech* in other languages. A negated verb of cognition can also subordinate a whole clause, as will be discussed in §10.6.3.

In Example 578 a declarative clause is the P argument of the verb *la* ‘say, call’. In Example 579, direct speech (an imperative clause) is the P argument of the matrix headed by *ahu* ‘ask’.

(578) we lanabô “ah nyu aye we we”

we la-na-bô [ah nyu a=e we we]_{compl.cl}

DEM say-FACT-PFV.ALLO ah 2S child=DEM DEM DEM

‘(Then) he said: “Oh this child of yours!” [TA22_origin 52 153]

(579) (...) “tya hã ing lâlhak hanekûna” kalà lyó ahuwa la

[tya hã ing lâlhak hane-kû-na]_{compl.cl} kalà lyó ahu-ya la

now 1S more trick try-COMPL.II-IMP officer white ask-ALLO HSAY

‘The white officer asked: “(...) trick me now!”, it is said.’ [TA22_JAstory_62 15]

In some cases, a question is embeded both in a verb ‘ask’ and a verb ‘say’ (in either order). In Example 580, the complement clause is a direct quotation (a question), as indicated by the verb *la* ‘say’ and ‘ask’. Similarly, in Example 581 the verb ‘ask’ is followed by ‘say’.

(580) “hã jadi pram wêhãdi” la ahuwa la

[hã jadi pra=m wê-hã-de]_{compl.cl} la ahu-ya la

1S what good=NEG think-FRUS-HAB.EGO say ask-ALLO HSAY

‘She asked: “why do you hate me?”, it is said.’ [TA23_fishstory_69 133]

- (581) (...) mîyã we ahuya larêya la
 mîyã we ah-u-ya la-rê-ya la
 wife DEM ask-ALLO say-PL.II-ALLO HSAY
 ‘(...) they asked the wife, it is said.’ [TA23_tayang_clan 30 177]

The verb *la* ‘say, call’ has taken a complex grammaticalization path that will be discussed in §10.7.

10.6.3 Embedded cited discourse or clause subordinated to a negative verb of cognition

Interrogative clauses with question words can be embedded in a main clause without any conjunction or special device. Commonly, the verb of the main clause is of negative polarity. Figure 10.6 describes the typical structure of this construction: it consists of two final clauses, one of which carries a question word and is subordinated to the main clause. Even though the embedded clause has a question word, the illocutionary force of the construction is declarative (cf. Coupe 2018, 9).

Table 10.6: Structure of embedded cited discourse

[[embedded interrogative clause with question word	]	Main clause]

This construction functions as an answer to the embedded question and can only be headed by a verb of cognition and perception, such as ‘know’ or ‘see’, on the main clause. Examples 582 and 583 are two cases of embedded questions. Example 584 further includes a question tag in the embedded question. Lastly, Example 585 has a verb ‘see’ embedding the question.

(582) hã kadényu bônane kasâdyüm

[hã kadényu bô-na-ne]_{embedded.Q} kasâ-d=yüm

1s how go-back-INT know-HAB.EGO = NEG

‘I don’t know how I’ll go back (lit. how I’ll go back, I don’t know).’

[140823_JOJA_eli2 436]

(583) hanûk nyu hanabô kasâdyüm

[hanûk nyu hana-bô]_{embedded.Q} kasâ-d=yüm

where FOC come-PFV.ALLO know-HAB.EGO = NEG

‘I don’t know where he came (from) (lit. where we came, I don’t know).’

[190923_JOJA_eli1 162, 150924_HM_eli 438]

(584) asûng kadé shyâ kya mu kasâna yüm

[asûng kadé shyâ kya]_{embedded.Q} mu kasâ-na yüm

generation how.many become Q also know-FACT NEG

‘I also don’t know how many generations have happened.’ [TA23_thadu_clan_016 92]

(585) hanûk nyu hana kho hwě.tyâng yüm

[hanûk nyu hana kho]_{embedded.Q} hwě.tyâng yüm

where FOC come even.though see NEG

‘I didn’t see where he was coming from.’ [290924_STBT_eli1 676]

As a final note, this construction can be framed in a general discourse strategy prevalent in Asian languages in which the topic first announces the *kind* of information that its speech act intends before transmitting the actual message (the focus or new information), cf. §8.3. The use of a rhetorical question, here as embedded cited discourse, seems to be prevalent in South Asian languages (cf. Bickel 1999b, 25) and bears a resemblance to relative-correlative strategies in IA languages⁶.

⁶Relative-correlative constructions are well attested in IA languages where they embed cited discourse,

10.7 Extended functions of the verb *la* ‘say’

The verb ‘say’ *la* is prolific in the language in having grammaticalized extended functions. This section will review the grammaticalization of the versatile verb *la* as a hearsay particle (§10.7.1), as name labeling (§10.7.2) and as a marker of discourse continuity (§10.7.3).

The verb *la* ‘say’ is likely related to Proto-Tani **lu* ‘say’ (Sun 1993, 279). Further afield, cf. Chintang (Kiranti, Nepal) *lus-* ‘sing, dance, speak, snore, feel, taste’, also an auxiliary for reciprocal actions (Rai et al. 2011, 138f.).

10.7.1 Hearsay particle

The *hearsay* particle reports an event as being heard (*it is said*). It is, therefore, closely related to the homophonous verb *la* ‘say’. As with other particles (§9), the *hearsay* occurs after the main verb. The use of the hearsay is particularly common in stories, as the source of the information pertains to oral tradition. In Example 587, the outcome of a story is described as being heard ‘it is said’. Similarly, in Example 588, the growth of a child is reported as part of the story.

purpose structures (Hock 2016c, 32) and are also a relativizing strategy (Hock 2016b; cf. Andrews 2007b, 214). Many IA languages have a *jV* relative pronoun that may or may not be correlated by a *correlative* pronoun on the main clause. Nepali represents an interesting case: although it has a *jo* relative pronoun, it may embed a question without the pronoun in a manner very similar to TB languages, including Tawra. In Example 586, the embedded question clause occurs as part of the declarative statement. There is no relative pronoun in this case.

(586) Nepali (personal knowledge)

ke kurā gar-thi-yo ma-lāi thāhā-thi-en-a
 what matter do-HAB.PST-PST.3S 1S-DAT know-AUX.PST-PST.NEG-3S
 ‘I didn’t know what they (used to) say’ or ‘what they (used to) say, I don’t know.’

There is a long list of studies that discuss whether non-Indo-Aryan languages from South Asia have borrowed the relative-correlative construction from IA or whether it is native (for an overview cf. Hock 2016a; also Coupe 2018). The consensus is that these constructions in non-Indo-Aryan languages have often a double origin: a native construction that has been influenced by IA (cf. Schackow 2015, 424). On top of that, many languages in intense contact with IA have also borrowed relative pronouns therefrom (e.g. in Kham, cf. Watters 2009, 165; also in Kharia cf. Peterson 2010, 408).

- (587) we thâthâ hi gũ tamyâ abang shyâkũdyâ la
 [we thâ~thâ thi gũ]_{cosub.cl} tamyâ abang shyâ-kũ-dyâ la
 DEM eat~PLAC EMPH NF tiger wild become-COMPL.II-AWAY HSAY
 ‘By eating and eating (raw meat), the tiger became wild, it is said.’ [TA22_tiger 125 277]
- (588) tya a we lê la kadé.gũ manûmnaya
 [tya a we lê la]_{main.cl} [kadé.gũ manûm-na-ya]_{main.cl}
 now child DEM grow HSAY how wash-FACT-ALLO
 ‘Now this child grew, they say. How would they wash it?’ [TA23_tayang_clan 56 346]

The hearsay particle does not only occur with direct speech or after a clause; it can also occur after a sound. In Example 589, the hearsay particle (with a plural suffix) occurs after *gru* ‘a sound’. In this case, the perception of the sound is only indirectly reported by the speaker.

- (589) we bupâ gru laga dũ (...) tots tyô wra rôho gũ mane grá khrâk
 [we bupâ gru la-ga dũ]_{cosub.cl} [tots tyô wra
 DEM suddenly sound HSAY-PL.I CO.TEMP torchlight(<E) turn.on sword
 rô-ho gũ]_{cosub.cl} mane grá khrâk
 hold-STAT NF FILL(<ASM) cry EVI
 ‘(And) suddenly as a sound came, he having the torchlight one and the sword ready, I mean, there was a cry. [TA22_CLAN2 167 313]

The grammaticalization path say>hearsay is not unproblematic, as these *new* meanings are tightly related to the lexical origin of the verb (‘say, call’) and a clear-cut distinction between the lexical and the grammatical meaning cannot be always established.

The hearsay/name labeling often occurs along the verb ‘say’ in a *double say* construction. In Example 590, an order is embedded by the verb ‘say, call’ in a *double say*. The hearsay element indicates the source of the information.

- (590) “tatyé cyârêzâna” la la
 tatyé cyâ-rê-zâ-na la la
 rope lower-PL.II-DIR-IMP say HSAY
 ‘He said: “lower the ropes!”, it is said.’ [TA22_kutyemba_phla_228 453]

As discussed by Genetti (2005, 47), the *double say* construction provides a fascinating conundrum, as there are no good arguments to say unequivocally that, e.g. Tawrã *hearsay* particle *la* is completely grammaticalized: it can always be argued that it has the lexical content of the verb (‘it is said’), even if it co-occurs along with the verb *la* ‘say, call’ (‘he said ... it is said’). Furthermore, in Tawrã there are also no traces of phonological erosion and, although particles commonly do not take verbal suffixes, the hearsay does sometimes take a plural suffix. This is illustrated in Example 591. Lastly, the clause-final position of the grammaticalized elements cannot be told apart in the *la la* ‘(...) he said, it is said’ collocations. In sum, there is a complex relation between the lexical and the grammatical development of the verb ‘say, call’. This liminal stage, however, seems to be stable in time (cf. Genetti 2005, 48), and it is perhaps anchored in the double say construction (‘he said, it is said’) and further fostered by a similar development in neighboring languages (cf. Saxena 1988; for an overview, cf. Bashir 2016; for Macro-Tani cf. Post and Modi 2024, 9).

- (591) ni we abrêya laga
 ni we abrêi-ya la-ga
 nipple DEM be.double-ALLO HSAY/say-PL.I
 ‘The nipple was bifurcated, it is said/they say.’ [TA22_tiger_09 35]

10.7.2 Name labeling

The verb *la* ‘say’ can also be employed to refer to the name of a referent, the way it is called. This use is closely related to the original meaning of the verb *la* ‘say, tell, call’. This grammaticalization path from the verb ‘say’ is a typologically common development (cf Kuteva et al. 2019, 380; cf. Saxena 1988, 380; for a concrete example in Goenmai, cf. Hellwig 2011, 445f).

In Example 592, the speaker refers to the name of a trap ('sun trap') by use of the name labeling *la*. Similarly, in Example 593 a speaker employs the verb *la* for naming referents. However, the use of the non-final marker *gũ* reminds us of the tension between the grammatical and the lexical meaning, as was discussed in §10.6.2.

(592) *ríng gapô la la*

ríng gapô la la

sun.MOD trap say HSAY

'They used to call it *ring gapo* (lit. sun-trap), it is said.' [TA22_kutyemba_phla_458 845]

(593) *bũ.dũ mazênu la gũ dâi nyu hana*

[*bũ.dũ mazênu la gũ*]_{sub.cl} [*dâi nyu hana*]_{focus}

a.name a.name say NF a.tribe FOC come

'So-called Bũdũ and Mazênu, (they) came (from) the Idu (side).' [TA23_dimha_clan 7 38]

10.7.3 Discourse continuity with *la* 'say'

The discourse continuity function of the verb *la* 'say' logically connects clauses with the preceding context. The association often corresponds to a causal conjunction in other languages (cf. Saxena 1988; Matisoff 1991, 400). This function of the verb 'say' commonly occurs with the marker for cosubordination *gũ* (§10.3), as in *la gũ* ('say NF'). Example 594 displays the use of the discourse continuity morpheme *la*, which occurs along the homophonous *hearsay la* (§10.7.1).

(594) thâlibô gũ la la mǎ̃ mề we la “chôq”

[*thâ-li-bô* *gũ la la*]_{topic/sub.cl} *mǎ̃* *mề* *we la*
eat-REC.PST-PFV.ALLO NF say HSAY village.MOD people DEM say

[*chôq*]_{dir.speech}]_{main.cl}

let's!

‘Having eaten, he said “let’s (carry the stone)”, they say.’ [TA23_kutyemba_phla_075 190]

In Example 595, the speaker reveals the hidden plan of a nephew: he intended to kill the speaker’s grandfather. This is central to the *arabesque* of the story, although not a new idea therein, since the grandfather had sent the nephew’s father to jail. The narrator emphasizes this act within the larger point of the story through the use of *la*.

(595) hadyû we ning nyu mǎ̃ na bô mane âm (...) sêya la

[*hadyû we ning nyu mǎ̃ na bô mane âm*]_{main.cl} [*sê-ya*
nephew DEM 1P FOC village INSIST go FILL(<I.A.) AFF kill-ALLO

la]_{main.cl}

say

‘This nephew went (to) our village, yes, to kill.’ [TA22_CLAN2_114 221]

Examples 596a and 596b are from a story prompted by a storybook (Mayer 1969). In Example 596a, the child and the dog realize that the frog is not there. In Example 596b, the speaker links the absence of the frog to the development of the story with the use of *la*. In this case, the linker *la*, along with the non-final marker *gũ*, works as a causal conjunction, which is here translated as ‘because’.

(596) a. a we hrûnegaho

[*a we hrû-nega-ho*]_{main.cl}

child DEM be.surprised-DU-STAT

‘This child (and the dog) are surprised.’ [TA22_frog1 32 150]

b. (...) pahwā we yüm la gũ

[pahwā we yüm la gũ]_{cosub.cl}

frog DEM NEG say NF

‘(...) because the frog is not there.’ [TA22_frog1 33 152]

The use of a non-final marker *gũ* with *la* in Example 596b is noteworthy, as it resembles the pragmatic effect of topicalization by *setting the scene* (sometimes called *old information*, cf. §8.3.1). Similarly, some of the pragmatic effects of cosubordination alone are to recapitulate (cf. §10.3.1.2) and to provide a reason or cause (§10.3.1.1), i.e. in both cases, it highlights a piece of information which is considered to be relevant for the understanding of the story. The pragmatic relation between these two strategies should be further explored.

Lastly, it is also worth mentioning that, during elicitation, a Nepali translation of Example 596b was volunteered by a Tawrā informant, cf. Example 597. Note that the Nepali topicalizing morpheme *bhane* is itself grammaticalized from the verb *bhannu* ‘say’ plus a non-finite (converbal) form, which is functionally similar to the Tawrā cosubordinator *gũ* (thus *la gũ*). In Nepali, the verb ‘say’ also has a causal function (cf. Saxena 1988, 375; cf. Bickel 1999b, 25; Masica 1993, 402f). Areal or even cognitive explanations, as hinted by some scholars (Saxena 1988, 375; Genetti 2005, 47f; for a review cf. LaPolla 2001, 235f. Lieberherr 2017, 235f.), should be further explored in order to understand the parallel development in these languages and the larger Himalayan region.

(597) (Mishmi) Nepali

[bhyāgutō ch-aina **bhanera**]_{NP-top}

frog be.there-3S.NEG TOP(lit. say-CNV)

‘(...) because the frog was not there.’

10.8 Complex demonstratives (discourse operators)

Complex demonstratives serve the purpose of organizing discourse. Although they are built on the demonstrative *we* (§4.3.3), but do not function as demonstratives in

that they are not co-referential with an NP in the same clause. The position of complex demonstratives in the clause is governed by pragmatic considerations, although they commonly take a sentence-initial position. The forms discussed here are *weli* and *we ciwe* (§10.8.1) and *webo* (§10.8.2).

10.8.1 Discourse deictics *weli* and (*we*) *cuwe*

A discourse deictic complex demonstrative draws the hearer’s attention to an NP or to the propositional content of a whole sentence previously mentioned without being co-referential. As such, the discourse deictic demonstratives have a discourse pragmatic function: it serves as a *thematic link* between two prepositions (cf. Diessel 1999, 100f; “pure textual deixis”, cf. Lyons 1977, 668). There are two devices for that end in Tawrã: *weli* (< *we* DEM + *-li*, possibly a suffix of recent past, cf. §6.3.4.9) and (*we*) *cuwe* (**cu-* + *we* ‘DEM’). The latter commonly occurs preceded by a demonstrative (here in parentheses). There is no perceivable difference between these two morphemes.

In Example 598, a group of animals are arranged one atop the other. The complex demonstrative *weli* makes it clear that the intended reference is the preceding NP, in this case, the bear. Discourse deictic demonstratives are not normally employed to retrieve relevant NPs; that is undertaken by the demonstrative in its simple form (*we*, cf. §4.3.3.2).

- (598) tamh we kyã mambaw ìho welì kyã tyû ìho
 [[tamh we kyã] [mambaw]_S [ì-ho]_V]_{main.cl} [[weli kyã]
 bear DEM LOC.ON fox be.there-STAT DISC.DEIC.DEM LOC.ON
 [tyû]_S [ì-ho]_V]_{main.cl}
 chicken be.there-STAT
 ‘Atop the bear is the fox and atop it (i.e. the fox), the chicken.’
 [TA22_chicken_thief_30 114]

In Example 599a, the speaker mentions a relevant episode in his clan’s history. In Example 599b, (right after it), he retrieves this episode through the discourse deictic device *cuwe*.

- (599) a. *tulu ning tuluqa gũ lê pêga we asûng kadê shyâ (...)*
 [tulu ning tuluqa gũ lê pê-ga]_{main.cl} [we asûng kadê
 FS 1P a.place LOC migrate separate-PL.I DEM generation how.many
shyâ]
 happen
 ‘Tulu... (since) we separated from Tulu, how many generations (have passed)
 (...)’ [TA23_talyu_clan 61 368c]
- b. *we cuwe ning kas ahuhona yûm*
we cuwe ning kas ahu-ho-na yûm
 DEM DISC.DEIC.DEM 1P FS ask-STAT-FACT NEG
 ‘This (i.e. how many generations), we never, ehm, asked.’ [TA23_talyu_clan
 61 368d]

Examples 600a and 600b are from the same story in which a cheater tricks the white officer. In Example 600a, the trickster asks the officer for his white horse to retrieve his trickster bag. In Example 600b, *weli* refers to the imperative clause uttered by the trickster. In this case, *weli* has a clause-internal position.

- (600) a. “*güwri lyó we tya.gê hangâbîna hã lâlhak phlê we shîngadyô*” *la gũ (...)*
 [güwri lyó we tya.gê hangâ-bî-na]_{imp.cl} [hã [lâlhak]_{rel.cl} phlê we
 horse white DEM right.now lend-EMPH-IMP 1S trick bag DEM
*shî-nâ-gadyô la gũ]*_{cosub.cl}
 take-BACK-ahead say NF
 ‘Saying “lend me your white horse (and) I’ll go ahead and bring back my
 trickster bag (...)”’ [TA22_JAstory 27 97]
- b. *webo kalà lyó weli la dũ (...)*
*webo [kalà lyó weli la dũ]*_{sub.cl}
 and.then officer white DISC.DEIC.DEM say CO.TEMP
 ‘And then, (while) this was said, the white officer (...)’ [TA22_JAstory 29
 107]

In Example 601, the speaker utters a clause that is immediately corrected. The discourse deictic demonstrative *weli* refers to a statement priorly mentioned (omitted here for the sake of clarity). Its pragmatic function is to reinforce the important thing that the wife had said. As discussed by Diessel (1999, 102), discourse deictic demonstratives refer to discourse that either precedes or succeeds the clause; *weli* could therefore not be coreferential with a NP in the same clause (as in *‘this wife, I told her’).

(601) we gũ la mĩyã we weli ladyibô (...)

[we gũ la]_{main.cl} [mĩyã we **weli** la-dyibô]_{main.cl}

DEM INS SAY wife DEM DISC.DEIC.DEM say-PFV.EGO

‘Having said this, having I said this to this wife (...)’ (and not *‘having said this, I told this to the wife (...)’ [TA24_land 105 185]

The discourse deictic complex demonstrative is, however, often employed to separate episodes within the same story. Examples 602a and 602b are from the same story: the villagers are tying ropes for days. In Example 602a, we learn that 4-5 days have passed. The use of *weli* heralds a new event (Example 602b) that is logically related to the preceding actions. As such, it resembles the function of the sentence connector *webo* (§10.8.2).

(602) a. kaprâ hĩng mangâ hĩng (...)

[kaprâ hĩng mangâ hĩng]_{main.cl}

four CLF:day five CLF:day

‘Four five days (...) [TA22_kutyemba_phla 264 504]

b. weli (...) tatyẽ we malà harũm mathômathô gũ (...)

weli tatyẽ we malà harũm [mathô~mathô gũ]_{cosub.cl}

DISC.DEIC.DEM rope DEM tip join tie.together~PLAC NF

‘Joining and tying tips from ropes (...)’ [TA22_kutyemba_phla 266 507]

Furthermore, a discourse deictic complex demonstrative may be employed as a strategy to signal to the hearer the end of a bracket in discourse and the return to the main

point. Similar to English *anyways...* In Example 603a, the speaker introduces his grandfather to the story. Then, in Example 603b, he mentions a priest, who also plays a role in the story. This is a tangent of the main story. Thereafter, in Example 603c, the speaker resumes the main story. This is indicated by the discourse deictic *cuwe*. Note that in Example 603c neither argument of overt; topic continuity (§8.3) resolves and potential ambiguity.

- (603) a. atya Solum atyu (...) phrêya mane
a-tya Solum atyu phrê-ya mane
 1s.POSS-grandfather a.name up.there pioneer-ALLO FILL(<Asm)
 ‘My grandfather Solum up there (...), was the first (to reach the place).’
 [TA24_land 142 261]
- b. Ajumso (...) gwók Ajumso laga na
Ajumso gwók Ajumso la-ga na
 a.name priest a.name say-PL.I INSIST
 ‘Ajumso, the priest’s name was Ajumso, right?’ [TA24_land 144 263]
- c. we cuwe mane kasemakáwshyüw la
[we cuwe mane Ø_A Ø_P kase-makáw-shyüw la]_{main.cl}
 DEM DISC.DEIC.DEM FILL(<Asm) inquire-thoroughly-CAUS.V HSAY
 ‘Anyways, (my uncle) had (the priest) inquire thoroughly (whether the area was suitable or not), they say.’ [TA24_land 146 267]

The discourse deictic complex demonstrative *weli* is also part of a conventionalized paraphrase employed to conclude stories: *weli ge.ne*⁷ ‘this much’. In Example 604, the narrator signals the end of the story by using this paraphrase.

- (604) welī ge.ne pwi simthâ mu yüm
weli ge.ne pwi simthâ mu yüm
 DISC.DEIC.DEM only different something also NEG
 ‘This much only, there is nothing else.’ [TA23_phlakam_clan_27 141]

⁷The adverb *ge-ne* is likely a grammaticalized form from *gê ne* (‘one’ plus ‘only’).

10.8.2 Sentence connector *webo* ‘and then’

The complex demonstrative *webo* (often along with the topicalizing morpheme *webo byẽ*) has the function of signaling a semantic relation between clauses (often conjoined clauses, cf. §10.2) (Diessel 1999, 125f.). *Webo* is composed of the demonstrative *we* and a **bo* syllable, likely a grammaticalized form of the verb *bô* ‘go’ (also cf. §6.3.4.8).

The use of this connector is illustrated in Examples 605a and 605a, where the narrator’s grandfather awaits his guest’s attempt to assassinate him. We learn in 605a that the grandfather heard a noise (of a trap) and was promptly ready with his torchlight and sword. In 605b, the deceiving guest shouts that he is a sleepwalker (and was therefore up late at night). The logical connection between the clauses is prompted by *webo*.

- (605) a. *we bupâ gru laga dũ (...) tots tyô wra rôho gũ mane grá khrâk*
 [we bupâ gru la-ga dũ]_{cosub.cl} [tots tyô wra
 DEM suddenly sound HSAY-PL.I CO.TEMP torchlight(<E) turn.on sword
*rô-ho gũ]*_{cosub.cl} *mane grá khrâk*
 hold-STAT NF FILL(<ASM) cry EVI
 ‘(And) suddenly as a sound came, he having the torchlight one and the sword ready, I mean, there was a cry.’ [TA22_CLAN2 167 313]
- b. *webo cya laya we “dyûmkũtyûm” “dyûmkũtyûm” laga gũ mane*
webo cya la-ya [we dyûmu-kũ-tyûm
 and.then 3 say-ALLO DEM sleepwalk-COMPL.II-unintentional
*dyûmu-kũ-tyûm la-ga gũ]*_{cosub.cl} *mane*
 sleepwalk-COMPL.II-unint. HSAY-PL.I NF FILL(<ASM)
 ‘And then he said, saying “I’m a sleepwalker” “I’m a sleepwalker”, I mean.’
 [TA22_CLAN2 169 318]

Examples 606a and 606b are from a story about a cursed child who never grew. In Example 606a, we know that the child was born (from two siblings). Thereafter, in Example 606b, we learn that it never grew more than the size of a mole’s pup. The use

of *webo byẽ* (*webo* plus conditional/topicalizing marker *byẽ*) again connects the sentence to the preceding context.

(606) a. a hîng mềnaya

a hîng mề-na-ya

child one bear-FACT-ALLO

‘(They) had a child.’ [TA22_origin 32 102]

b. webo byẽ a we lyüm náng (...) kalyô la ìya kacék kalyô la ìya

[*webo byẽ a we lyüm náng* $\emptyset_V$]_{main.cl} [*kalyô la ì-ya*]_{main.cl}

and.then TOP child DEM eight year child HSAY be.there-ALLO

[*kacêk kalyô la ì-ya*]_{main.cl}

mole.sp offspring HSAY be.there-ALLO

‘And then this child (for) eight years (...) was a pup, they say, a mole’s pup, they say.’ [TA22_origin 34 104]

11 Lexicon

The following is a lexicon of approximately 1,650 entries across all word classes: it includes nouns (*n*), verbs (*v*), adjectives (*adj*), adverbs (*adv*), classifiers (*clf*), interjection (*intj*) and grammatical morphemes (*gr*). This lexicon, which was collected during my field trips (2022-2024), came about thanks to the help of many members of the community, particularly Shri Johakso Manyu and Shri Jabralum Chaitom, although all mistakes herein are solely mine. It should be noted that all classifiers that are listed in this lexicon are also listed in Appendix 13.

The lexicon is based on the dialects from the plains (Tezu circle, Lohit district) but inevitably includes some entries from Chagla dialect. Regarding the final codas (§2.2.4.1), I tried to stick to the conservative, consonantal form whenever possible but some inconsistencies would require further lexicographic work.

The intrinsic homophony of certain entries was resolved with a small number (e.g. ¹) after the entry. If an entry has more than one sense, this is also indicated with a small number. This approach is practical, e.g. for versatile verbs and their grammaticalized forms (cf. §6.2) but it has the downside that it clumps together homophonous words of potentially different origin. It should be noted that the decision to make two entries out of one is often arbitrary, cf. the two meanings of the verb *brék*: ‘buy’ and ‘marry’. In this case, I decided to make two entries since the second meaning is a) derived from the first (‘marry’ from ‘buy’) and b) it implies something different, although at some basic level they are the same. A different criteria could have merged these into one all-encompassing meaning.

This lexicon follows the main guidelines proposed in §1.7. However, some points ought to be brought up again: those phonological alternations that have been described in §2.2.5 are omitted (e.g. t/d alternation) while unpredictable phonological variation

is marked by a vertical line (e.g. *ali/adi* ‘like this’). The few loanwords of this list are indicated with the source language in parenthesis, e.g. (IA). Class terms (§3.3.1.3) and conventionalized N-V combinations (§6.5) are in parenthesis, cf. (*tako*) *zû* ‘write’. For a few verbs, there is information on valency and animacy restrictions: A transitive frame is marked (vt) and an intransitive (vi). Animacy restrictions are indicated in parenthesis (i.e. *as of humans*).

One field of the lexicon that has been particularly neglected is that of flora and fauna. In many cases, there is a simple translation for common animals, e.g. *tyû* ‘chicken’. Where Tawrã has a generic term (in the sense of Berlin 1992, 60-64), I’ve added the note *generic*, e.g. *thô* ‘hornbill (generic)’. Whenever I am certain about the term of a more specific species, I have added the common name (e.g. *talyâng* ‘cardamom’) and/or the scientific name. I employ the paraphrase ‘a kind of’ (following Wierzbicka 1985) when I am unsure about the identification of the species and when there are no common names available.

The lexicon follows an alphabetical order displayed in Table 11.1. This order is along the lines of *Pāṇini*’s phonological analysis of Sanskrit, which has become a standard in the linguistic studies of South Asian languages. Tone is ordered in the following manner: mid (unmarked, M), high <́> (H), high-falling <̂> (HF) and low <̀> (L).

Table 11.1: Alphabetical order of this lexicon

a	i	u	e	o	ü	ai	ei	aw	üw
k	kh	g	ng						
c	ch	j	y	sh					
t	th	d	ts	tsh	z	n	r	l	s
p	ph	b	m	v	w				
h	q								

A

a¹ *n* child

a² *gr* diminutive

a³ *gr* question tag

a⁴ *v* be a child

a⁵ *v* bear a child

akâ *n* granary

akâp *v* steal

akê *v* show
akrâw *v* take out, remove
akû *v* knock (e.g. a door)
akâw *v* rest, be at ease
akyã *gr* atop, by
akhà *n* funnel trap
akhî *n* co-wife, other wife
akhrû ¹ *n* river bank
akhrû ² *n* dry field (e.g. for dry rice)
akhrûng *n* a kind of banana
agã *v* set up (as of a trap)
agûm *n* village
agâw *v* crouch
agrâi *n* common giant flying squirrel
(Petaurista petaurista)
agrû *n* freshwater prawn
aglûm *v* feel cold
acîng *n* navel, belly button
acûne *n* firefly
ajîn-dyâw *n* rich woman
ayêk *v* be dizzy
ayô *n* a kind of millet
ayüw ¹ *n* man-made chicken's nest
ayüw ² *clf* man-made chicken's nest or
eggs therein
ashâ *n* plain area
ashei *v* demand, beg
ashyâ *n* mithun
atî *v* leave, be lost
atyã *adj* lazy
atyê-sei *clf* measurement: one little
finger
atyó *v* rise (of water)
atyô *n* younger daughter

atyô-ka *n* palm (of hand)
atyô-pha *n* surface of hand
atyu *n* north (towards the mountains)
atyûm *n* rack (over *côri*)
atyâw *n* vagina
athê/athû *v* hail
athûk *v* notch
athyúw *n* bladder
adâ *n* balcony
adi *cf.* *ali*
adè *v* be reluctant, be adamant
adâityang *n* scorpion
adyângra *n* nape, back of neck
atsâw *v* appoint
azê *n* metal bracelet
ana ¹ *n* morning
ana ² *clf* weight: 6.25 paisa coins
anaya *n* early morning (*cf.* *ana*
'morning')
anâtsi/tatsî *n* snail
ane *pro* 2P
anêqing *pro* 2D
anûhing *adv* day after tomorrow
anya *n* front part, face
anyú *n* family
anyû-a *n* nephew
aram ¹ *n* rest (IA)
aram ² *adj* easy (IA)
arêqto *n* a kind of eel (probably spotted
snakehead)
alâ-talûm *n* a kind of banana
alaî-kà *n* crossbow
alaî-taphyo *n* bow
alâng-pa *adv* together
ali/adi *adv* like this
alú *adv* up there
alû *n* big mortar

alô¹ *n* garlic
alô² *n* hanging bridge (one single rope)
alyüm *n* road, way
alhîng *n* liver
alhîng-apó *n* main organs
alhûng *n* nail
asâ *n* artery, vein
asang *n* trunk
asáng *v* grow (e.g. children)
asâng-ra *adj* alive, raw
asû-grũ *clf* measurement: chest to fingertip
asûng¹ *n* generation
asûng² *clf* generations
asûm *n* taro
asûw *n* nest
apáw *n* thread
apâwgra/kaprâw *n* spider (generic)
apu *n* ceiling
apu-khalà large hanging rack
 (below the ceiling)
apû *n* fence
apó *n* lungs
apó-pum *n* heart (cf. *pum*)
apû *n* arrow
apwî¹ *n* dry area
apwî² *v* dry (e.g. cloths)
apûw *v* sweep, clean
aprûw *v* pluck (e.g. feathers)
aba *gr* for
abe *n* morning
abe lyâng midmorning
abe-rã *adv* soon
abü *v* carry away (as of the wind)

abûw *clf* bundles e.g. of firewood
abrêi *v* be bifurcated, be rugged, be double
(tamüng) ablá¹ *n* flame
ablá² *v* turn on (e.g. light)
am-lang¹ *pro* something
am-lang² *gr* question tag
ama *n* mineral lick
amã¹ *n* down slope
amã² *n* south (prob. 'following the river')
amang *n* name
amû *adj* false, worthless
amê *v* follow
amê *v* invite, call
amêng *v* be windy
awa/uwa *n* loan
awâ *n* bruise, wound
awâk *v* throw (water), sprinkle
awê *n* a kind of garland
awé *v* spin (vt)
ahô *clf* volume: circumference around arms
ahu *v* ask
ahû¹ *v* sweat
ahû² *n* sweat
ahwê *v* send (someone), see off
ahrûng *v* roast
awrei *n* snow, ice
a-yã *n* daughter
a-yĩ *n* grandson
a-yûwak *n* son, male young relative
a-rûzang *n* friend
a-pu¹ *n* womb

a-pu² *v* be pregnant

a-mě *n* family

á *v* hatch, grow

ám *n* cloud

âm *intj* yes

àng *n* house

ã¹ *intj* yes (cf. *âm*)

ã² *gr* be there (inanimate)

I

ì¹ *v* be there, have, live (animate)

ì-kha sustain oneself

ì² *v* spear (e.g. fish)

îng¹ *v* sleep

îng² *adv* also, some, at least, a bit

îng³ *clf* bits, times

îng-dũ¹ *pro* something

îng-dũ² *adv* more, a little

îng-îng *adv* a bit

U

urüng *v* be shocked

uwa *v* cf. *awa*

ûm¹ *n* hair (of body), fur, feather

ûm² *adj* old (of clothes or houses)

ûm-khrai *n* a kind of edible plant

ûrei *v* pinch

(yu) ùm *v* be drunk

E

e = cf. *we*

O

ô *v* shoot

ôm *v* bury

ôm blûm cemetery

Ü

üdyû *n* peacock

úm-tyù *clf* volume: fistfuls

K

ka¹ *v* feed (child), prechew

ka² *clf* bits of food, portions of food, mouthfuls

ka³ *v* untie

kacang *v* urinate

kacêk/kacî *n* a kind of mole

(a) kacîng *n* umbilical cord

kacu *v* judge

kayangyê *v* have diarrhea

kashûng *n* front yard

kashyâ-îng *n* pity

katûwa-grei *n* lizard

katê *v* swim

katô¹ *v* click one's tongue (as a sign of warning)

katúm *n* mortar

katúw *v* pincer grasp, pick up (with two fingers)

katyâng *v* perceive, find, lurk, hear

katyàw *n* crab (generic)

katyê¹ *adj* small, a bit

katyê² *clf* little amounts, bits

katyê-ne *less*
katyê *n* tick found on big animals
katyú *v* carry (over the head)
katyûqu *adj* short
kathâw *n* offering of raw meat
kathê *v* point with finger
kathîng *n* a kind of ant
kathù *v* take care
kadí *qw* when
kading *n* star
kadâng (kadyù) *adv* delicious, tasty
kadé *qw* how
kadé-ge *qw* how much
kadyâng *v* kick, push (with the legs)
kadyêsyem *adv* with difficulty
katsûm *n* chest
kanáng¹ *n* year
kanáng² *clf* years
kanú *adj* dark
kane-hîng *adv* two days after tomorrow
kanüm¹ *n* reason, cause
kanüm² *gr* purposive, in order to
kanúng *n* opium (origin unclear)
(a) kará *v* carry (child on a cloth in front)
karà¹ *v* rain
karà² *n* rain
karà³ *adj* wild (e.g. of a mithun)
(a) karà⁴ *adj* fatherless (child)
karâ *n* a kind of drum
karàng *adv* clear (as in 'burn clear' an area)
karâp *v* burn (in a swidden field)
karô *adv* quick

kalà *n* officer, lord, white person
kaláng¹ *n* door
kaláng² *adj* long
kalyab *n* field
kalyâng *clf* evenings
kalyù *v* wait
kalyô *n* infant, pup
kasâ *v* know
kasang *num* three
kasahîng *adv* three days after tomorrow
kasô *n* a kind of forest rat
kapa *adj* fool, dumb, deaf
(tâ) kapá *n* banded gourami
(Trichogaster fasciata)
kapâq *adj* alone, same, only
kapé *n* cotton
kapê *n* leech
kapû *n* edge (e.g. of a field)
kapô *v* store, hide
kapú *v* hate, disgust
kapei *clf* handfuls (of rice or mustard seeds)
kapûng *n* bat
kaprâw *cf.* *apâwgra*
kaprêi *num* four
kaplâ *n* shelter trap
kapyêk *clf* volume: third of a clenched fist (e.g. of flour)
kapyù *adj* next
kapyù kanang next year
kaphrûlyu *n* butterfly, moth
kabê *v* mediate (in a conflict)
kabü *cf.* *kawü*
kam¹ *n* work (IA)
kam² *v* have a job (IA)
kama¹ *v* miss, be lost
kama² *v* die

kamáng <i>n</i> leader	kili-tulu <i>n</i> a kind of rash
kamu <i>n</i> son-in-law	kîya <i>clf</i> periods of 24h
kame <i>n</i> a kind of large flying squirrel	kurú-kaw <i>adj</i> bold
kambâ/kampâ <i>adv</i> for free, already	kurù ¹ <i>n</i> head
kawâ <i>v</i> yawn	kurù ² <i>n</i> upper part (of a tree)
kawü/kabü <i>v</i> hold (e.g. a sword)	kû ¹ <i>v</i> dry (meat by the fire)
kawrà <i>adj</i> naked	kû ² <i>n</i> plateau
kawrí <i>v</i> tie up	kũ-cẽ <i>n</i> sickle (Asm)
kahã ¹ <i>n</i> node, knot, joint	ke <i>v</i> ride, drive (e.g. a car)
kahã ² <i>clf</i> nodes, knots or joints (e.g. in a hand)	kese <i>cf.</i> mese
kahîng ¹ <i>n</i> day (only in compounds)	kê <i>n</i> paddy, rice
kahîng ² <i>clf</i> daytimes	kê-kàw <i>n</i> cooked rice
kahum <i>n</i> bamboo spear	kẽ <i>gr</i> locative (inside)
kahéi <i>n</i> boundary	ko ¹ <i>v</i> make roof
kaqing ¹ <i>num</i> two	ko ² <i>n</i> skin, leather, hide
kaqing ² <i>clf</i> weight: 20 <i>tula</i> coins	(masang) ko ³ <i>n</i> bark
kahlâi <i>n</i> intestines, gut	ko-wrá <i>n</i> peel (e.g. of banana), cover, shell
kahrûw ¹ <i>v</i> help, respond (in an emergency)	kotsarüw <i>n</i> school (prob. Asm)
kahrûw ² <i>v</i> be scared, be surprised, scare, surprise	kôlu <i>cf.</i> khûlu
kahrûk <i>n</i> ribs	kü <i>v</i> wear (blankets or shawls)
kâ ¹ <i>v</i> be bitter (e.g. of bitter gourd)	künyâng <i>num</i> nine
kâ ² <i>clf</i> halves (e.g. of fruit)	kûm-pó <i>n</i> door frame, entrance
kâri <i>n</i> male animal	kài <i>cf.</i> tài
tyû kâri cock, male chicken	kâw <i>n</i> maternal uncle
kârü ¹ <i>n</i> stranger, guest	kâw-bùm maternal line
kârü ² <i>n</i> brother-in-law	kya ¹ <i>gr</i> question tag
kà <i>n</i> crocodile (difference to <i>tanûng</i> ‘crocodile’ not clear)	(tã) kya ² <i>n</i> Indian flying barb (<i>Esomus danrica</i>)
kã-zõ <i>n</i> striped squirrel (<i>Tamiops maclellandii</i>)	krâng <i>v</i> bite (as of humans)
ki <i>gr</i> though	krã <i>n</i> hole, crack
	krî <i>v</i> ruminate
	krú-na <i>n</i> ear
	krú-pe <i>n</i> a kind of jewelry

klûm *v* worship
klûm-dyâ *n* wealth
klâi *n* feces, excreta
klâi-pom *n* intestines, guts
klaw *clf* sips
kwâg *n* dog
kwâ-kũ *n* owl
kwâra *n* clearing
kwí *v* dig a hole
kwê *v* distill

Kh

kha *v* lie down
khanyâng *n* rolling pin
(khalâi) khanyùm *n* earthquake
khalai *n* ground, land, field, earth, soil
khalâi *n* spit
khalyâng-pà *n* shoulder
khalyâng-pà-hapyê *clf* volume: water up to the shoulders
khalyû ¹ *n* shell (hollow, e.g. of a snail)
khalyû ² *v* pull out (a snail out of its shell)
khamtâi *n* Khamti (language and people)
khamyû *v* get angry
khawag *n* a kind of jewelry
khá *clf* steps
kháng *n* Nocte (language and people)
khârai *n* gunpowder
khi *n* Chinese (language and culture)
khika-khika *adv* in each
khing(-gê) *num* one (normal form)

khili *v* play (IA)
khì *v* reach
khu(-dyâng) *v* have fever, have hangover
khumphâ *n* bribe
khûlu/kôlo *n* ladle
khũ *clf* bits of food
khû ¹ *v* fight, revenge
khû ² *n* vengeance, retaliation
khû ³ *v* be enough
kho *gr* even though
khünyüm *v* rattle, shake
khûlyu *n* eggshell
khünyüm *n* god, ghost
(tamüng) khaw *n* smoke
khyü *v* reflect light, shine
khra-khra *adv* manner of being dirty
khram-shâng *n* small bow used in *puja*
kh râk ¹ *n* sound
kh râk ² *gr* evidential
khrai *adj* paralytic
(tamüng) khre *n* matches
khro *v* cry
khru *v* lock (e.g. door)

G

ga ² *n* silk
ga ³ *gr* plural (PL.I)
ga(-tyâng) ¹ *gr* self, own
gá-ga individually
gang *v* capture, claim (e.g. land)
gâqa *adj* close
gatyô ¹ *adv* differently
gatyô ² *adj* different
gazêqe *adv* quick

ganyu *adv* despite, even though
garêi¹ *n* car (IA)
garêi² *clf* carfuls (e.g. of bananas)
galai-kù *n* turban
gabâng *n* head chain
gamâ *v* roam
gapô *n* trap
gawà *n* a kind of bird (prob. *Upupa epops*)
gá *v* break, crack
gângkhi *n* milk (Asm)
gânta *clf* hours (IA)
gâmbra *n* gaonburha, leader of the village (Asm)
gãwã¹ *adj* round
gãwã² *v* be round
gîharak *n* gun
guwi *v* roam
gû *n* ash, hearth
gũ¹ *gr* non final marker
gũ² *gr* instrumental, locative
gũ³ *v* shine
gû *v* stick out
(nya) gé *n* nostril
gê *v* ferry, cross (by boat)
gê *num* one (classifier form, cf. *khing(-gê)*)
gê¹ *v* circle (vi, of a vulture)
gê² *v* coil (vt)
gü *n* a kind of hat
gürâ *n* drum
güwri *n* horse (cf. Asm *ghūra*)
gûm *v* bow down, worship
gyâi *v* be tired

grã *v* grope (search, touch without seeing)
gru *v* burn (vi)
grũ¹ *n* leg, foot, hand
grũ² *clf* volume: clenched fist
grũ³ *v* grab, hold, lift (with whole hand)
grũ-bülüm *n* sinus tarsi (eye of the foot)
grũ-bülüm-hapyê *clf* volume: water up to feet
grũ-shya-hapyê *clf* volume: water up the calf
grú-khünyüm *n* shoe
grú *v* be strong
grûw *n* crowd
glâi *v* carry (on back)
glyã¹ *n* bamboo container
glyã² *clf* containers
glyu *n* soy bean
gwí *v* hug, attach
gwí-shîk hug, embrace, hold on
gwók¹ *n* priest, shaman, *pujari*
gwók² *v* perform *puja*

NG

ngâ *v* fell (a tree)

C

câng¹ *n* cage (for small animals)
câng² *v* get, find, be available, be able
câng-ru get rich
câmbri *n* a kind of bird (prob. *Dicrurus leucophaeus*)
cici *intj* move!
cîng *n* chick, baby bird

tyû cing baby chicken
cuhwa *n* rust
côri *n* hanging rack (above the hearth)
cûro *v* make a gesture of scorn
céi *n* juice, broth
cya *pro* 3
cyaqîng *pro* 3D
cyâ *v* hang down, lower (e.g. rope)
cyâng *n* bed
cyû *v* remove (shoes)
cyûw *v* be wet

Ch

cha *n* Assamese (language and people)
châgulyei *n* goat (Asm)
chî¹ *n* a kind of fish (prob. Asian stinging catfish (*Heteropneustes fossilis*))
chî² *adj* pure (of rice beer)
che *n* oil
chêk *v* walk
(tyû) chề *n* pullet
chôq *intj* let's go!
(îng) chüm *v* fall asleep
chûwen *n* louse

J

ja¹ *pro* something
ja-ja whatever, anything
ja² *qw* what
ja-ba *qw* why
jakwí *n* drop door fish trap
jadi¹ *pro* something

jadi² *qw* what, why
jâp *n* Kaman, Miju (language and people)
jâp-dabrüm *n* a kind of offering wrapped (in a bamboo leaf)
ji¹ *v* do
ji² *v* be about
ji³ *gr* causative
ji⁴ *v* be good, be healthy (only in 'not feeling well')
jîng *v* join (e.g. a hearth)
jîng¹ *n* end
jîng² *adj* finished
jorâ¹ *n* pair (Asm)
jorâ² *clf* pairs (Asm)
jû *v* crawl (supported by buttocks and legs)
jû-kûra *n* freshwater mussel
jüm¹ *n* truth
jüm² *n* direction (only in 'the eight directions', meaning 'everywhere')
jüm³ *adj* true, real
jüm⁴ *v* be true, be real
(tawá) jûm *n* honey

Y

ya¹ *n* night
-ya² *gr* allophoric
-yá *gr* superlative
yáng *v* fan (e.g. the fire)
yâng *v* insert, attach (e.g. bait on a hook)
yâmu¹ *v* dream
yâmu² *n* dream
-yâq *gr* prohibitive

-yâ *gr* feminine
 yu¹ *n* rice beer
 yú thâ leftover fermented rice after making beer
 yú tapê food and drinks (as in a feast)
 yu (ùm)² *v* get drunk
 yû *v* be overripe, be spoiled (of food)
 yù *v* shrink, melt
 yú *gr* comparative
 yê *gr* question tag
 yo *v* cross (a water body)
 yéi *v* slip
 yâw *n* mother-in-law
 yüw *clf* side of face
 yüm/=m *gr* general negator
 yúm *v* fly
 yúw-mlâ *n* place, region

Sh

shá *n* fishing net
 shà *adj* wild
 tyû shà wild chicken, fowl
 shâ-brã *n* testicles
 shí *adj* red
 shî/shêk *v* hang (on a hanger, a wall)
 shîm-shim *adv* manner of seething
 shî¹ *v* take
 shîzâ bring
 shî² *v* feel, think, consider
 shî³ *v* be allowed, be possible
 shî⁴ *n* claw (of birds)
 shî⁵ *n* inside of ribs (after the intestines have been taken out)

shû-duwa *n* a kind of winnower (round)
 shü-tawê *n* good nature
 shêi¹ *v* die
 shêi² *v* get caught
 shya *gr* who, whoever
 shyá-îng someone
 shyá-la whoever
 shya mu nobody
 shya-brê *n* a kind of short sword
 shyâ¹ *v* happen, become
 khümyû shyâ get angry, become upset
 qwîng shyâ become old
 shyâ² *v* be over
 shyâ³ *v* have to, must
 shyâ⁴ *v* be useful (only used negatively: be unuseful)
 shyâ-brê *n* young man
 shyâ-kama *n* orphan
 shyâm *intj* no!
 shyû *v* climb
 shyû *v* be dry
 thrôka shyû be thirsty
 shyo *v* kick
 shyûm *v* be stuck
 shyaw *v* be tasty, be sweet, be salty
 shyüw¹ *n* breath
 shyüw² *v* breath, live
 shyüw³ *v* allow, appoint, let
 shyüw⁴ *gr* volitional causative
 tahwî-shyüw learn
 shyüw⁵ *gr* observational negative

T

- ta**¹ *gr* advisative
ta² *gr* inopinative
ta³ *gr* hortative singular
takang *n* takin, gnu goat
takã *n* a kind of pot
 pôwüŋ takã a kind of jewelry
 (resembling a small pot)
takũ-kwě *n* wooden cutting board used
 for meat
take *n* dirt
tako *n* book
takaw *n* storage room
takyã *n* on, atop
takla *clf* steps when walking
takwâg *n* wolf
takwĩkã *n* a kind of fish
takwing *n* mushroom (generic)
takwûi *n* monitor lizard
takhâng *n* supporting log (used for
 chopping wood)
takhyã *n* table
takhi *n* sack (e.g. of rice)
takhrù *v* roll (vt) (e.g. a *beedi* or a
 cigarette)
takhrê-brúw *n* stick, rod
takhrâi *n* snare trap
takhrêi *n* weaver
takhrâk *v* make noise, be noisy
tago *n* needle
tagra¹ *n* spirit of a dead person
tagra² *adj* late, deceased
-tagra³ *gr* ‘in disbelief’
tagrû *adj* fast, strong, handsome
taglú *n* provision
tangâ/tã *n* (generic) fish
tacông *n* furniture
 tacông bâya mề carpenter
tacû *n* scales (weighting machine)
tachà *n* a kind of banana
tachêi-gülüm *n* a bamboo figures used
 in *pûja*
taji *n* (fishing) bait
taya *v* tonight
tayã *v* shout
tayîng-bom *n* a kind of bee
tayû *n* a kind of bee
tashîk¹ *n* fruit
tashîk² *n* testicles
tashîk³ *n* appetite, hunger
tashîm-ling *n* a kind of insect
tashê¹ *n* song
tashê² *v* sing
tashyâ *v* stab, scar (with an instrument)
tashyâ-tã *n* camel
tatâng *v* stagnate, accumulate
tatûm *n* dust
tatyâ¹ *v* tell
tatyâ² *n* requirement, toll, tax
tatyâng *n* a kind of grass
tatyû *n* pestle
tatyû-pũ *n* stick, *lathi*
tatyù *n* rule, law
tatyú *v* problem
tatyé *n* rope, cord
tatyê *n* a kind of hammer
tatyû *v* beat up
tatháng *v* support (e.g. for clearing a
 field)
tathâ *n* crop (generic)

tathyû *n* a kind of bee

tadi *n* mocus

tadâw *n* eagle

tadyá *n* black giant squirrel

tadyâw ¹ *n* season, time

yu bâya tadyâw beer making season

tabrê bôya tadyâw time to go

hunting

tadyâw ² *adj* seasonal

pyâ tadyâw seasonal bird

tadyáng *v* knock the door (during *puja*)

tatsî *cf.* *anâtsi*

tazû *v* a kind of basket

tazû *v* be smart, be extrovert

tazêi *adj* elder

tanà *n* vegetable (generic)

tanûng *n* crocodile, gharial (also *cf.* *kà*)

tanûm *n* thorn, bur

tanyâng-tazô *n* a weaving pattern

tará *cf.* *tawrá*

tarã *n* a kind of grass

tarap ¹ *n* sister-in-law

tarap ² *clf* herds

taráp *n* a kind of creeper

tarum *clf* areas for grazing

tarüm ¹ *n* flock

tarüm ² *clf* flocks

tala ¹ *adv* still

-tala ² *gr* permansive

talá *clf* coins of gold and silver

talâ *n* a kind of deer (probably *Moschus leucogaster*)

talâ-lyě *n* volcano, lava

talà *n* male bird

tyû talà cock

talyâ *n* witch

talyâ ² *n* ammunition

talyâng *n* cardamom

talyâng *n* evening

talû *n* a kind of jewelry

talû *n* the Dibang river

taló ¹ *n* funerary ceremony

taló ² *v* perform *taló*

talõ *n* wing

talü *adj* strong (of rice beer)

tasî-bû *n* a kind of plant

tasîng *n* a kind of porcupine

tasé *n* weeder

tasê *v* clear (the jungle)

tapa *adj* first (as in first child or first plant to bloom)

tapâ *n* spear

tapâw *v* weave, plait

tapak ¹ *n* elder, respected person

tapak ² *v* get old

tapak ³ *adj* elder

(tâ) tapî *n* a kind of fish

tapûm *n* insect (generic), worm

tapûm-darû *n* altar (square structure outside of a Tawrã house where *puja* is performed)

tapû *n* bamboo shoots

tapê *n* place in between

tapê *n* rice, food

yú-tapê food (as in a feast)

tapé-tanà food

tapô-thelei *n* a kind of bird (probably *Aethopyga nipalensis*)

tapü *n* flower (generic)

tapwi *gr* habitual, 'surely'

tapwî *n* civet

tapûng ¹ *v* grab, snatch (of birds)
tapûng ² *clf* volume: half clenched fist
tapûtsüng *n* a kind of shrew
(tâ) taprâw *n* school (of fish)
taprô-tyu *n* stairs, ladder
taprûw *n* dhole (*Cuon alpinus*)
taplá ¹ *n* spoon
taplá ² *clf* spoons
tapyéi *n* female animal
 matsâp tapyéi cow
tapyêk *n* sand
(macîk) taphra *n* river, stream
taphû *v* be full (of stomach)
taphrû *n* bridge
taphrû-sã *n* hanging bridge
tabí *v* slap (inner side of hand)
tabû *n* maize
tabû *v* touch lightly, push
tabâw *n* snake (generic)
 (tâ) tabâw a kind of eel
tabri ¹ *n* a kind of bamboo basket
tabri ² *clf* baskets (often with dried fish)
tabri-tagra *n* a big hunt
tabrû-tsû *n* hunter
tabrê ¹ *n* meat
tabrê ² *v* hunt
tabró *adj* colorful (of animals)
tablüw *n* incomppliance
tabyê *n* honey
tama *n* medicine
 tama kalà doctor
tamû *n* a kind of wild banana
tamê *v* refuse
tamé *v* be crazy

tamê *n* animal
taméqĩng *n* elephant
tamüng *n* fire
tamyá *n* tiger
tamyâ-bodà *n* lion
tamyûm *n* monkey (generic)
tawà *n* bee (generic)
tawâ *n* jug, pouch
 apû tawã arrow pouch
tawâ-rei *n* a kind of wasp
tawì *n* lake, pond
tawê *n* mind, heart, opinion
tawó *n* hoe
tawra/tara *n* sword, *dao*
 tawrá-bro a kind of sword
 tawrá-hablûm a kind of (large)
 sword
 tawrá-jù a kind of sword
 tawrá-tyò a kind of sword (dull,
 used as a tool)
 tawrá-za knife
tawrã *n* Tawrã, Digaru (language and
 people)
tawrî *n* mud
tawrêi *n* flattened bamboo
tawrûw *n* daily wage
tahĩng *adv* today
tahû ¹ *n* group, clan, sub-clan
tahû ² *clf* groups
tahum *n* bear
tahrô *num* six
tahwa *n* news
tahwî *v* teach
tahwê *n* strap, tumpline
tâshi *v* war, fight, quarrel
tâzi *n* upper part

tâplă *n* hammer
tâmě *n* boar
tâwa ¹ *n* jaw
tâwa ² *n* way, pass
tâwrüw *v* build (a house)
tâ-khala *n* shadow
(kará) tã *n* interval without rain
tâ/tangâ *n* fish (generic)
tâqa *adj* light (weight)
tíng *v* spring up (e.g. of mushrooms)
tî *v* wrestle
tîng ¹ *n* shirt
tîng ² *v* wear (a shirt), put on
tíya *n* owner, lord
tukũ *v* set on fire (vt)
tudyâ *n* chin
tûri *n* mustard (IA)
tûlu *n* the Lohit river
tûpla *v* fold (of fingers)
te ¹ *v* be heavy
te ² *v* cough
tekhê *v* lean on
têka *v* chop (firewood in small pieces)
tê ¹ *v* hit, crash, crack open
tê ² *v* knot
tohô *n* a kind of plant
tohwâ *v* wave
tó *clf* measurement: one kilo
tükü *n* matter, thing, language, word
tükü-bra *n* letter, text, story, matter
tüning *clf* measurement: two fingers
tûla *clf* weight: one *rupee* coin
tüsang *clf* measurement: three fingers
tüm *n* powder

tài/kài *adj* hot
tûw *clf* halves
tya ¹ *n* grandfather
tya ² *n* ancestor
tya ³ *adv* now
tyá ¹ *adv* once
tyá ² *clf* shots from a gun
tyáng ¹ *n* thread of wool
tyáng ² *clf* threads
tyáng ³ *clf* plots on land
tyâk *v* cover (under a cloth), hide, put lid on
tyâng ¹ *adj* green
(nyâ) tyâng ² *v* spin (fiber into yarn)
tyâng ³ *v* touch, discover
tyâng ⁴ *adv* very
tyângso *n* gums
tyâng-ra *adj* moist, fresh
tyàng ¹ *n* body
tyàng ² *clf* bodies of humans and large animals (e.g. pigs, cows)
tyã *v* carry (an animal or a heavy load on a stick)
tyû ¹ *n* chicken
tyû shà wild fowl
tyû ² *v* hit, play, pierce, flatten, dig (with a tool), harvest
(nâshik) tyûng *v* be hungry
tyûpi *v* fold (cloth)
tyûm *v* drink
tyûmya mề drinker
tyû ¹ *v* throw, dump (with both hands)
tyû ² *n* a kind of basket (carried by men)
tyé *v* cut (e.g. animals), cultivate, plough, hit
tyě *v* wash (clothes)
tyê ¹ *n* shortage

tyê² *v* lack, be incomplete
tyề *v* bind
tyô¹ *v* close (e.g. a window)
tyô² *v* turn on, take a picture
tyô³ *adj* last (in a row)
pramû tyô youngest brother
tyô⁴ *gr* each
(khalâi) tyô⁵ *v* spit
tyô⁶ *v* weave (on a handloom)
tyü *v* snatch (of a dog)
tyû *v* sell
tyù *v* start
(grũ) ti *n* heel
tyüw *v* cover (from rain)
tyúw *v* slice, cut, break (wood)
twĩ *n* pond
tri *n* a kind of bird (*Saxicola ferreus*)
trû¹ *v* meet, find
trû² *gr* 'get to V' (e.g. get to sit together)

Th

thang *n* corpse, dead body
tharâng *v* listen
thalyâng-na *n* tongue
thâ¹ *v* eat, take
thâ² *n* food, rice (only in compounds)
thâ³ *gr* sadly, for all I care
thâ-ming ration (of food)
thâk *v* bite (of animals)
thàng *n* hair (of head)
thîya *n* mountain, hill
thû *v* bind, trap (e.g. with a snare)
thê¹ *v* be cold

thê² *adj* cold (as felt by the hands or mouth)
thô¹ *n* flesh
thô² *n* hornbill (generic)
thô *v* get stuck
thünù *n* lip
thürûw *adj* obedient
thû *clf* volume: pinch, drop
thûkrü *n* hips
thûkrü-hapyê *clf* volume: water up to the hips
thâi¹ *n* poison
thâi² *v* be poisonous, be spicy
tháw¹ *n* traditional Mishmi garment (worn by women)
tháw² *v* wear (a tháw garment)
thâw *v* put a pole
thrû-ka *n* throat
thrû-taplá *n* wooden spoon
thrûga *adv* on the back
thrúm-bum *n* mouth
thrúm ûm mustache

D

da¹ *n* wire, cable
da² *gr* as if
(bülüm) dagaw *n* eye bags
dara *adj* insignificant, worthless
darâi *pro* something
ja darâi mu nothing
darák *n* sparrow
daráng *adj* big, expensive
dabû *adj* middle
yâ dabû midnight
dablang *v* flood

dabyûm <i>n</i> jackfruit	dế <i>v</i> stand
dá <i>v</i> return, pay back	dề ¹ <i>gr</i> inference
dâng ¹ <i>v</i> finish, be clear	dề ² <i>v</i> pick up (with all fingers)
dâng ² <i>adv</i> miserably, poor	do ¹ <i>v</i> cough
dâng ³ <i>adv</i> very, completely	do ² <i>v</i> pop out
dâi <i>n</i> Kera'a, Idu (language and people)	doq-doq <i>adv</i> manner of ascending (of smoke or steam)
dâi-klâi <i>n</i> pheasant	dôlak <i>n</i> swamp
di <i>v</i> sit	dô/dũ <i>n</i> gong, metal disk
ding <i>n</i> Singpho (language and people)	dũthê <i>clf</i> measurement: thumb to pointing finger
dú <i>gr</i> with, near	dũ ¹ <i>n</i> large <i>puja</i> performed for prosperity
dũ ¹ <i>v</i> jump	dũ ² <i>v</i> perform <i>dũ puja</i>
dũ ² <i>v</i> suck	dũ <i>pro</i> many
dũ ³ <i>v</i> smoke	dũw <i>v</i> do, happen
dũ ¹ <i>gr</i> associative, with	dũw <i>v</i> need, be required
dũ ² <i>clf</i> large, man-made structures (e.g. houses, granaries)	drâk <i>n</i> pineapple
dũ/dũ-dubâ ³ <i>clf</i> times a traditional Mishmi houses has been refurbished	drã <i>n</i> copper vessel
dũ ¹ <i>n</i> jail	drũw <i>n</i> a kind of bird (prob. <i>Passer montanus</i>)
dũ ² <i>cf.</i> <i>dô</i>	dyang <i>v</i> chase, pursue
dũk-suk <i>n</i> problem, sadness (IA)	dyâ ¹ <i>adj</i> far, foreign
dula <i>n</i> a kind of turtle	dyâ ² <i>adj</i> old, past
dũphru <i>v</i> puddle	dyâ ³ <i>v</i> be far away
(áng) dũwã <i>n</i> courtyard	dyâ ⁴ <i>adj</i> rich, wealthy
de <i>adv</i> at least	dyâng ¹ <i>v</i> be full (of moon)
(tyàng) dedyá/hadyá <i>v</i> bath, wash oneself	dyâng ² <i>v</i> obey, comply
dedyâ <i>v</i> dominate, rule	dyu <i>n</i> kodo millet (<i>Paspalum scrobiculatum</i>)
delé <i>gr</i> it seems	dyuka-naka <i>adv</i> terribly (IA)
delei <i>clf</i> weight: coin of 50 paisa	dyú-gê <i>adv</i> just like
deming <i>n</i> key (Tib)	dyûmu <i>v</i> sleepwalk
dêqe <i>adv</i> before	dyù <i>v</i> fall off (of a branch or body part)
dẽ-katã ¹ <i>v</i> stand on tiptoes	dyê <i>v</i> protrude, landslide
dẽ-katã ² <i>clf</i> volume: water level up to the head (while standing on tiptoes)	dyê <i>v</i> pick up (whole hand), catch
	dyo-ka <i>n</i> Western side (Roing, Tinsukia)

dyó¹ *v* be different
dyó² *n* difference
dyü¹ *adv* like, similar
dyü² *v* copy, imitate
dyû *n* corner, border, ridge
(thâ/tyûm) dyûm *v* be full (of eating/drinking)
dyûm-hya *clf* weight: 12.5 paisa coins
dyáw *n* Mishmi *dal* (a kind of lentil)

Ts

tsâng *v* have a rush
tsâp-ka *n* a kind of flat pan
tsi *v* tear (e.g. paper)
tsu¹ *n* blister
tsu² *v* touch, find (of something that is in a high altitude)
tsu³ *v* make things worse
tsu-tsu *adv* manner of being loose (of dry land)
tsû-wã *adj* bright
tsekwi *v* comb
tsê *v* be right
tsûm *v* name, be called
tsûm *v* tighten (when making a wall)
tsûng¹ *n* a big rain (a bad omen)
tsûng² *v* wear (shoes)
tsâi¹ *v* count
(la) tsâi² *gr* each
tsâi³ *n* a kind of bird (prob. *Anthus hodgsoni*)
(tamûng) tsûw *n* charcoal

Tsh

tshâk¹ *v* move fast
tshâk² *n* tick (found on people and chicken)
(karà) tshâng¹ *v* be wet (as of rain)
tshâng² *v* rot
tshâng³ *adj* rotten
tshâng⁴ *v* turn pale (when ashamed)
tshâp *v* punch
tshâm-nûm *n* a kind of catfish
tshã-da *adv* manner of being clean
tshû (masâng) *n* a kind of pine tree
tshû *v* enlarge, multiply
tshô *v* keep, stay
tshûm *v* blow out, blow nose
tshwîng *n* flea

Z

zá *n* juncture
-zâ *gr* directive
zâk *gr* atop
zâng *n* magical being
zã¹ *v* pack (in a bag)
zã² *v* be pregnant (of animals)
(tako) zû *v* write
zû-lô *v* read
zú¹ *v* be cold (as felt on the body)
zú² *adj* cold
zé *v* be tight, be close
zè *v* be exhausted
zó *v* embroider
zó katûwa tailor

zô *n* bud
zõ *v* increase, expand
zú *v* inset (e.g. hairpin in hair)
züm-pâ *adv* together

N

na-¹ *gr* 1s possessive
-na² *gr* imperative
-na³ *gr* factitive
na⁴ *n* leaf, sheet of paper
na⁵ *clf* foldables
nakapîk-hapyê *gr* measurement of quantity: water up to the chest
nang *n* younger brother-in-law, male relative (wife's side)
náng¹ *v* smell, have bad smell
náng² *n* year (only when modified)
kapwi náng next year
nâng *v* avoid eating (due to taboo)
ning *pro* 1p
ne¹ *v* be stuck
ne² *n* breast, nipple, udder, milk
ne³ *adv* only
nyamûng *n* fuel, gas, kerosene
nyáng *n* last room (in a Mishmi house)
nyâ¹ *v* boil, cook
nyâ² *v* thread
nyâ-blâng *n* a kind of seed (similar to sesame)
nyâ-pô *n* sesame
nyâng *v* step on, push
nyâra *v* be sick
nyâhing *v* have epilepsy

nyà *v* be in pain, hurt
nyàng *n* daughter-in-law
nyî *v* cut, butcher, saw, harvest
nyîm *v* grind (in a mortar)
nyu¹ *gr* 2s
nyu² *gr* contrastive focus
nyu³ *v* be ready (of food)
nyû¹ *v* dance (of the priest)
nyû² *v* wipe
nyûmke *n* Mishmi altar
nyêto *n* nipple
nyûm¹ *v* hiccup
nyûm² *v* pierce, fish (with a spear), pierce, pound (e.g. rice), wrestle (as of horned animals)
nyûm-ûm *v* be soft

R

ra *n* (cutting) edge
rajâbam *n* a kind of eel
râ¹ *v* carry (in pocket)
râ² *v* wear (a bag)
ráng¹ *n* cackle, sound of hens
ráng² *clf* stalks of trees and plants
ráp¹ *v* flow, float
ráp² *clf* streams and tributaries
ráb *n* horn
râk¹ *n* fear
râk² *v* be afraid
râg *v* scrape (e.g. the scales of a fish)
(bülüm) rá *v* be red (of eye)
râ¹ *adv* late
râ-râ late
râ² *adj* old, late
ring¹ *n* sun

-ring/-rê² *gr* plural (PL.II)
rîng-kwôk *n* rabbit
rîng-tapüm *n* leprosy
rîwa *v* be confused
rîng *v* be diminished (of water bodies)
ru¹ *v* get together, help each other, make love, befriend
ru² *v* make sense
ru³ *n* alliance
ru⁴ *n* landslide
ru-su *n* marrow
ruzàng *n* place in between, place among
rû *v* sew
 kazüm rûya mề tailor
rû-atyâi *n* kidneys
rũ *v* make fun of
rũ *v* be sexually aroused
re *v* pile up (e.g. stones)
rê *cf.* *hrê*
rê *v* twist, tie up
 (talyâ) rê prepare ammunition on a gun
rô *n* catch, take hold of
rûbung *n* bone
râi-zô *n* a weaving pattern (used on the Mishmi coat)
rûw-da *v* scatter
rwê *v* take off (clothes)

L

la¹ *v* say, name, call
la² *gr* hearsay
lakâ *n* chestnut

-langka *gr* adverbializer (only with *pra* 'good')
labũ *n* debris
labûya *clf* measurement: four fingers
lacrâp *n* elbow
la-mûng *gr* periodically
 hîng la-mûng daily
lâka *n* armpit
lâng¹ *v* enter, go in
lâng² *n* entry, entrance
lâng³ *adj* warm (of water)
(tamüng) lâng⁴ *v* bask by the fire
lâpa *n* loincloth
lâhlak *v* trick
lâ-manyâng *v* be happy, make happy
làng *gr* collective marker
lầ *n* tooth, tusk, pincer (of a crab)
lî *v* escape, survive, rescue
lu *v* poke, insert, stick (finger)
lû *v* pour, sweat, extinguish (fire)
lủ-wing *n* big pestle
lũ *n* hive
lê¹ *v* rub, roll (e.g. tobacco or a *roṭi*)
lê² *gr* all over
lê¹ *v* exit, go away, leave, migrate, resuscitate
lê² *v* stick out (e.g. the tongue)
lê³ *adj* big, biggest
 bũlei lê the biggest pig
 talâ lê a big, mythical rain
lowá *n* shackle
lô *v* pierce, thread (a needle), make a hole
lũ¹ *v* spare, save, survive, remain
(thâ) lũ² *n* leftover food
lũ³ *v* climb, cross by a rope
lũküng *n* spoon (made of a dried gourd)

lüm¹ *n* big seed (e.g. of a jackfruit)

lüm² *adv* these days

lümúng *n* tail

lũm(-kũ)¹ *n* inner part, lower part

lũm(-kũ)² *gr* at the time, meanwhile

lai *n* necklace

lai-hayáw necklace made of coins

lâi¹ *v* be half grown (e.g. of the moon)

lâi² *adj* immature, first (harvest of rice)

lâi³ *v* make a fence, make a partition

(khalâi) léi *v* shake (e.g. earthquake)

lêi *v* slither (e.g. as of a snake or a baby)

lya¹ *v* return

lya² *gr* back

lyáng *v* plant (e.g. rice)

lyâ *v* throw (with one hand), fall down

lyâng¹ *n* halfway, place in-between

lyâng² *clf* halves (generic)

lyâng³ *clf* measurement: fathom

lyâng⁴ *v* be evening, be dark

lyâng(-kap)⁵ *n* plain area

lyù *n* value, price

(macík) lyû *n* wave

lyó¹ *v* be white

lyó² *adj* white

lyó³ *n* egg white

lyûm *num* eight

lyüw *v* come up (e.g. as of fire)

lyûw *v* defecate

S

samplang *gr* what-to-say

sâkwã *n* anus

sâng *n* shadow

sáng *cf.* *masáng*

sâ *v* step over, trample

(kaprâw) sâ *n* spiderweb

sâ-pû *n* footprint

sîgri *n* cigarette (IA)

síng¹ *v* wear (shoes)

síng² *v* inhale (e.g. powder)

sîk *v* request, ask

sîni *n* sugar (IA)

sû¹ *v* boil

sû² *n* liquid (only in compounds)

krû-sû earwax

sûnggu *n* a kind of pot

sû-lyâng *adv* tomorrow evening

sû-hing *adv* tomorrow (when counting days)

sê *v* scratch

sâi *n* iron

sâi tyûya mề blacksmith

sêi *v* beg

sêipahê *n* soldier (IA)

so¹ *v* comb

so² *n* grease, fat

sokarêi *n* government (IA)

sôpa *adj* clean, clear

sümthâ *gr* question tag

sûmbe *n* a kind of earring

P

pa¹ *v* fetch water, draw water

pa² *gr* place of many

pacû *n* a kind of bean

pacyâ¹ *n* relative

pacyâ² *v* be related

pachâ *v* whisper
pachà *n* flint
pathi *cf.* *mathi*
pada *n* Adi (language and people)
patsâk *n* chili pepper
(bülüm) patsüng *v* blink
para/puri *v* read, study (IA)
palai *clf* small plots of land (often used as garden)
pasûng shî *n* pomelo
pawüng/pûwüng ¹ *n* silver
pawüng/pûwüng ² *n* money
pawüng/pûwüng ³ *clf* weight: a silver coin
pahwã *n* frog
pá ¹ *n* yeast
(kanûng) pá ² *n* opium cloth
páp *clf* pots (containing food or liquids)
pâ *v* cross (e.g. a river), go across
pâng ¹ *v* close (e.g. door)
pâng ² *v* make tea
pâtla *adj* thin, watery (IA)
pâlo *n* oath
pâlyu *n* non-tribal
pâhwüw *n* a kind of shield bug (consumed by the Mishmi and neighboring tribes)
pà ¹ *gr* unfortunately
pà ² *n* ax
pà taphyù a kind of ax
pâ *v* be curved
pi ¹ *clf* foldings
(bülüm) pi ² *n* eyelid
pi ³ *v* pack food in leaves

pi ⁴ *v* peck, ax
pis *clf* piece (E)
píg *v* squeeze
pî *v* roast meat lightly
pîsha *n* money, coin (IA)
pînya *v* fish (by diverting fish to a shallow area)
pu ¹ *n* elder brother, elder person, brother-in-law (respectful)
pu ² *v* punch (downwards with one's fist, e.g. hitting the table)
putyâ *adv* on the back
(masâng) putyû *n* stump
puri *v* study, read (IA)
pum *clf* round objects (e.g. fruits, birds also cars)
pû ¹ *v* be bitter
(phlá) pû ² *n* rock
pû-mru *n* male siblings
pû-pramú-pathì *n* siblings
pũ *n* remains
thâ pũ food rests
pũ *v* pick (the thread through the wrap while weaving)
pe *clf* rooms in a house
(tâ) pelé *n* a kind of fish, Pool bard (*Puntius sophore*)
pé *n* a kind of winnower
pêmpe *adv* anything, whatever
pê *v* scatter, divide, separate, share
pokyo *n* sapsapon (*Crassocephalum crepidioides*)
pohwã *n* pit
pó *v* open (e.g. a door)
pô ¹ *v* hang (e.g. clothes)
pô ² *v* increase cost
pô ³ *n* slave

pô ⁴ *n* be a slave
pôdu *cf.* *phûdu*
pômto *n* papaya (prob. IA)
pôw-tung *n* bronze pot gifted during *talo*
pò *n* jail
pü *v* bloom
pûyüm *n* a kind of bird (probably *Psilopogon virens*)
püdai *n* gold
pünyüm *n* brains
püring *n* copper pot (gifted during *talo*)
püsîng *n* a kind of bird
pütyô *adj* different, separated
pütsí *gr* portions of food
püg *v* toss (e.g. ashes or powder)
püw *n* pan (generic)
pûw *v* carry (in a basket)
pûwüing *cf.* *pawüing*
pra *adj* good, beautiful
pramru *n* younger brother (male persp.)
prâ *v* gulp, swallow
prû *v* feed
prê ¹ *adj* strong
prê ² *adv* strongly
prê *v* pull down, press down
práw *v* dig (as of animals), make a burrow
prûw *v* set ablaze, light, burn (vt)
pla *clf* large, flat and sharp objects (e.g. planks, blades, money bills)
plâ *v* burn slightly (e.g. bamboo leaves, in order to harden them)
plâng *v* swell

plà *n* salt
plã *v* lie down (face down)
plû *v* make a floor (for a traditional Mishmi house)
plûng *n* back, back side
plai *n* seed (generic)
plûw *v* blink
pyâ *n* bird (generic)
pyâ-kalà *n* crow
pyâ-thá *n* a bamboo figure resembling a bird used in certain *pujas*
(khalai) pyéi *v* spit
pwi ¹ *v* be allowed, may
pwi ² *adj* other, different

Ph

pha *gr* bunches (e.g. of ginger, banana)
phati *adv* suddenly
phalâm *n* tea
phák *v* shake (vi) (as of the priest)
phâbum *n* knee
phâbum-hapyê *clf* volume: water up to the knees
phã *v* defend, protect
phûdu/phûndu/pôdu *n* eggplant
phûm *n* a kind of tree
(masáng) phû ¹ *n* root
(bá-mề) phû ² *n* clan
phû ³ *clf* roots
phû ⁴ *clf* clans (e.g. in a congregation)
phoko *v* stomach
phó ¹ *clf* bundles or fistfuls (e.g. of mustard leaves)
phó ² *v* make bundles (e.g. of mustard leaves)
phôk *v* hit (with back of hand)

phri *v* clean, polish, sharpen
phrêng *v* break
phrê¹ *v* build (e.g. a road or a bridge)
phrê² *v* pioneer, find a way, come through
phrűw *adv* yet (as in ‘not yet’)
phlâ *v* be bulky (e.g. of grains when ripe)
phlă-brõ *n* stone
phlũ¹ *n* kettle
phlũ² *clf* kettlefuls
-phlũ³ *gr* abundantly
phlẽ *n* bag (particularly the Mishmi bag)

B

ba *v* carry (on back)
ba-pũng *v* carry (on back)
baji *clf* time (IA)
bajêi *n* banana (generic)
baràng *n* firewood
bamî *n* a kind of eel (prob. barred spiny eel (*Macrogathus pancalus*))
baqba *adv* manner of being long (of hair)
bá¹ *v* flow (of liquids)
bá² *clf* shallow and slow parts of a river that enable crossing
bá³ *v* fly
-bá⁴ *gr* already
bá-mẽ *n* clan (lit. father’s people)
bâ *v* make, build
bâpak *n* uncle

bâqa¹ *adj* shallow, thin
bâqa² *n* uncle
bà *n* father
bi¹ *n* elder sister, elder female
-bi² *gr* allophoric potential marker
bigrú *n* a kind of bird
bilyâi *n* a kind of earring
-bî *gr* emphatic
bîya¹ *adj* old
bîya² *n* past
bîyũ¹ *adv* straight, directly
bîyũ² *n* side, line (e.g. family line)
bu *n* pus
buthâ *n* wet field
bupâ *adv* suddenly
bum-kahum *n* trapping pit
bum-plâ *adv* here and there
bûm *v* fill up, cover (with earth)
(masang) bùm¹ *n* stem, original place, foundation
bùm² *n* maternal uncle
nabúm-nama maternal uncle
búm-mîyaqa maternal uncle’s daughter (i.e. potential spouse)
bùmýă female from maternal uncle’s side
bú *clf* log-like things
(tatháng) bũ *v* support
bêqe *adv* slowly
bẽ *v* dig out
bó¹ *v* blast, burst
bó² *n* pop (as in popcorn)
bô¹ *v* go, flow
bô² *gr* perfective allophoric
bô³ *gr* conditional
bôtlo *n* bottle (IA)
bom *n* first step (in a ladder)

bôm¹ *n* edge
bôm² *v* rain heavily
bü *clf* times (e.g. that something happens)
bügrî *n* a kind of bird (Psilopogon asiaticus)
büra *n* thunder
bûra *adv* earlier
bürî¹ *adv* low, under
bürî² *n* under, lower part
bürü *n* cheek
bürüm *n* underpart (of a house)
bülik *n* pig
bülüm *n* eye
bü-dû *adv* again
büwrà *n* storm
bû *clf* handfuls (e.g. of flour)
bûqû *adv* just recently
búw *n* thirst
bra *n* a kind of big bamboo (used in constructions)
bra¹ *clf* small round objects (e.g. beads)
bra² *clf* generic
bra³ *n* father-in-law
bra-tyô *n* pipe
brã *clf* rope-like things (e.g. straps)
brá *clf* ways and roads to reach a place
brú-tshû *n* increase
brû *v* be enough
brû *v* burial monument
breyâ *n* gecko
brék¹ *v* buy
brék² *v* marry
brè *n* enemy, war

brô *v* be born (as of animals)
brü/brik *clf* bamboo-shaped things
brú¹ *v* be generous
brú² *v* ooze, come out (e.g. liquids)
blâng *v* be full, fill up
blehíng *adv* yesterday
blúm¹ *v* flow
blúm² *n* place
blúm³ *gr* locative nominalizer
byě *gr* conditional, topic marker
byêk *n* types of grass
byüw *n* (white) ashes
byúk *v* run, fly, dash
bwì *v* dance

M

= **m** *cf.* *yüm*
ma¹ *n* mother
ma² *v* become dull, become senseless
ma³ *gr* comitative
ma⁴ *gr* confirmation
ma⁵ *gr* obsolesced negative marker (only in lexicalized forms)
bülüm ma blind
tamě ma crazy
ma⁵ *v* close (eyes)
ma-ma *adv* close
makáng *v* seize, take control (of an area or building)
makáw¹ *adv* thoroughly, enough, sufficiently
makáw² *gr* thoroughly
makum *n* resting hut (on a field)
magîn *n* cucumber
mago *v* establish, found
magû *v* pull (a rope)

mang <i>gr</i> each	la mane convince
maci <i>v</i> harm, do damage	mane ² <i>gr</i> filler (Asm)
macîk <i>n</i> water	manê <i>v</i> stick (e.g. hot plastic to a surface)
macyû <i>n</i> sambar deer (<i>Rusa unicolor</i>)	manûm <i>v</i> wash (one's hand or a child)
majarei <i>n</i> cat (IA)	(karà) manyâ <i>v</i> downpour (rain)
majéi <i>n</i> buffalo	manyu <i>v</i> cook
mayã <i>adj</i> foreigner, outsider	mará <i>v</i> laugh
mayî <i>n</i> egg yolk	mará <i>v</i> abandon, give up
mayû <i>v</i> shrink, melt	marẽ <i>v</i> poison, harm
matú <i>n</i> back (body part)	maro <i>v</i> speak, talk
matyâ <i>v</i> release, set free	marũ <i>n</i> horse
matyâng <i>n</i> villager, neighbor	marũne <i>n</i> metal balls (used in puja)
matyû <i>v</i> be quick	mala <i>v</i> dry (on the sun, e.g. rice or clothes)
matyũ <i>v</i> bend (as of fingers)	malâ <i>v</i> search, find
matyũ <i>n</i> floor, room, place	malà <i>n</i> tip, edge (of a rope)
matyô <i>gr</i> maybe	(macîk) malai <i>n</i> spring (fountain)
mathâng <i>v</i> be stiff, be tight (as of meat or stone)	malò <i>n</i> a kind of palm
mathi/pathi <i>n</i> younger sister	malûm <i>num</i> hundred
mathô <i>v</i> join, tie together	malú ¹ <i>n</i> a kind of sickness
mathûw ¹ <i>n</i> up in the mountains, the source of the rivers	malú ² <i>n</i> a rite to cure the <i>malú</i> sickness
mathûw ² <i>n</i> north	malûng <i>v</i> fish, hunt
madô <i>v</i> apply, stick (sticky substances, like a bugger)	malyú <i>n</i> rhinoceros
madyâ <i>v</i> spread clothes or grains to dry	masáng/sáng <i>n</i> tree, wood
madyû <i>v</i> help, serve	masó <i>v</i> scratch (e.g. with the snout, of a pig)
matsâp <i>n</i> cow, beef, ox	masü <i>v</i> have a cold
mazúm <i>n</i> deer (generic)	mapak <i>n</i> aunt (father's elder sister)
mazâi <i>gr</i> force, compel	maprà <i>v</i> sweep, keep in good condition
(tyû) mana <i>n</i> egg	mapré <i>v</i> strain, scold
kru-mana earlobe	maputyâk <i>adv</i> on all fours
manû <i>v</i> work	mapyâ <i>v</i> graze
mane ¹ <i>v</i> bait, gain trust	mabaw <i>n</i> fox
	mamîng <i>v</i> erase, massacre
	mamyú ¹ <i>adv</i> after, later, second
	mamyú ² <i>n</i> behind, hinder part

mamyú³ *n* late time
mawi *v* fill up (e.g. a bottle in a pond)
-maha *gr* desperately, aimlessly
mahrâi *v* scare, scold (e.g. of a dog)
má¹ *v* be black
má² *adj* black
má³ *adj* Indian
kalà má Indian officer, Indian government
-mâ *gr* iterative
mâqa *n* aunt (father's younger sister)
mâng *v* roast
mânga *num* five
mâlai *v* scatter
mâmbi *n* sheep
màng *v* be unconscious
misha *clf* weight: coin of 25 paisa
míng¹ *v* become extinct, disappear
míng² *v* be yellow
mîng *adj* foreign, someone else's
mîyã *n* wife
mîyaqa *n* girl, woman
mu *adv* also, even
muwra *n* a kind of priest
muhwa *adj* rich
mú *v* hide (e.g. from an animal)
múgaya hide and seek (a game)
mûwa *n* husband
mûwaqa *n* man
me *v* rear (animals)
methê *n* sweet, candy (IA)
melyû *n* dowry
mese/kase *v* inquire
mesê¹ *v* kill

mesê² *v* catch
mesê³ *gr* destructive
mé¹ *v* vomit
mé² *gr* whole, all
mêqẽ¹ *adj* new, first
mêqẽ² *n* beginning
mêqẽ³ *adv* first
mẽ¹ *v* give birth
mẽ² *clf* kilometers
mê¹ *n* person
mé mẽ everybody
mé yù adult, grown man
mê² *n* wife (only in 'your wife (person)')
múng *v* blow (e.g. fire)
myã *clf* measurement: one palm
myu *gr* direction, through
myûm *n* younger brother (female perspective)
myûm-mathî siblings
mlã *n* village, place
mlu *n* penis

V

-(vi)dyô *gr* certain future

W

wa *v* scratch
wá *v* tilt
wáyi *v* be upside down
wi¹ *v* tie up (vt)
-wi² *gr* around
wing/weng/wüng¹ *n* cup
wing/weng/wüng² *clf* volume: cupfuls
wí *v* hold, take, embrace

wî *v* overflow (of water)
we/e= *gr* demonstrative
wetyâng *v* remember
weli/welü *gr* deictic demonstrative
wesâ *v* forget
webo *gr* and then
wê¹ *v* think, want
 wê-hă misunderstand
wê² *num* seven
wê³ *v* be tired, be well (only in negative, i.e. 'be not well')
wêlang *v* understand
wêmara *v* forgive
wô *v* grab (with all fingers), hold, hook
wúng *v* bark (as of dogs)
wúw *n* big chunk of meat
wra *n* vacant place
 áng wra unused house
wráng *n* steep, slope
wrà¹ *n* boat
wrà² *v* aim (a gun)
wrai *v* underline
wrái *n* bamboo pattern (used for walls)
wri *v* cheat
 (karà) **wrí** *v* end (as of rain)
wrí *v* grind (using a rolling pin)
wrûm¹ *n* depth
wrûm² *v* be deep

H

haká *v* mix (e.g. soil and clay)
hakû *v* complain, accuse
hagă *n* first room in Mishmi longhouse

hagû¹ *n* guard
hagû² *v* guard
hagaw *v* crawl
haglêi *v* send (things)
haglûm *v* feel cold, shiver
hangâ *v* hire, lend
hajîn *n* ginger
hayâ *v* cook (food)
hayû¹ *n* front part
hayû² *n* past
hayû³ *adv* in front, first
hatû *v* help
hatûm *n* sky
hatûm-garêi *n* airplane (lit. sky car)
hatûw *v* cut (e.g. branches)
hatyâ *v* be lazy
hatyâng¹ *v* shout, cry
hatyâng² *v* cause trouble
hatyâng³ *n* thumb
hatyâng⁴ *clf* measurement: one thumb (including grease in animals)
hatyû *v* fry (with sand)
hatyêi *n* egg
hadâ *v* borrow, loan
hadi¹ *n* story
 hadí-atyà story
hadi² *v* tell a story
hadi³ *v* give space to sit
hadûm¹ *clf* groups of plants, trees, thickets of bamboo
hadûm² *clf* groups of same gender siblings
hadâi *v* place, raise
hadéi *v* make an oath, promise
hadyá *cf.* *dedyá*
hadyû *n* nephew (sister's side)

hadyũ *v* wake up (vt)
(macík) hadyûm *v* fall (of water)
hadyó ¹ *v* be extra
hadyó ² *gr* on (as in ‘pass on’)
hadyó ³ *adv* one after the other
hadyô *v* place, put, set (for long or round objects)
hadyû ¹ *v* resemble, be like
hadyû ² *v* draw
hazá *n* king, officer
hazâ *v* drop
hazúm *clf* bunches of objects that can be tied or skewered together (e.g. meat, lychee or grapes)
hazê ¹ *n* bamboo wall (in a traditional Mishmi house)
hazê ² *v* build bamboo walls
hana *v* come
hanâ ¹ *v* squeeze
hanâ ² *v* milk
hanûk *qw* where
hané ¹ *v* be able
hané ² *v* try
hanyâ-gum *n* nose
hanyâ-kwê *n* nostril
hanyû *v* sneeze
hara *v* be crooked
harák *v* pile up (firewood in the hearth)
harang *n* otter
haráng ¹ *n* wind
haráng ² *v* blow (of the wind)
haráng-da *n* phone
harí-hera *adv* zigzag (manner of movement)

hare *n* banyan tree
(masâng) haro *n* branch
harũ *n* spine, backbone
harûm ¹ *v* be together, join
harûm ² *v* remarry (male persp.)
(masâng) harèi ¹ *n* root
harèi ² *n* relative, kin
halá ¹ *n* inner part, lower part
halá ² *clf* foldables (e.g. leaves, sheets of paper or cloths)
halâi *n* red porcupine
haláw *v* take out to graze
halák *v* spy, observe
halâk ¹ *n* shame, embarrassment
halâk ² *v* be ashamed
halâng *num* ten
halí *v* see off
halîng *v* swim
halú *v* insert into cylindrical object (e.g. a key into the lock)
halo ¹ *n* moon
halo ² *clf* months
halü *v* keep, save
halû *v* discard, throw away (e.g. food)
(tâ) halêi *n* (fish) scale
halyá *v* fell, topple (e.g. tree or stone)
pi halyá *ax* down
halyú *v* change
halyo ¹ *n* happiness
halyo ² *v* be happy
halyo ³ *adj* happy
hapá *v* depart
hapi *v* distribute, share
haprê *v* fill up, put in (e.g. gunpowder in a gun)
haprû *v* arm, hand
hapyê *adv* straight

habá *v* hit, whip
habâ *n* a kind of bamboo
habai *v* be able
habang¹ *n* outside
habang² *n* forest
habang³ *adj* wild
habûm¹ *n* roof
habûm² *v* stretch, pull
habüw *v* carry (on shoulder)
(thang) habúw *v* hair fall
habra *n* buckwheat
habrâng *clf* measurement: breadth of a finger
habrâng-brü/habrang-bra *n* finger, leg (of a crab)
habrûk *clf* measurement: two handfuls (e.g. of flour)
hablâng¹ *v* flood, deluge
hablâng² *n* flood
hablâng³ *n* a kind of banana
hablí¹ *pro* some
hablí² *clf* groups, halves
hablü *v* lightning
hamâ *v* be lost, lost count
hamà¹ *v* compete
hamà² *n* smoky smell
hamáng *n* mosquito
hamâng *v* extinguish
hamû *v* hide
hamûng *v* be ripe (of grain)
hawa *v* waste time, delay
hawá *n* porridge, food
hawrí¹ *v* fall down
hawrí² *v* migrate to the plains

hawring *n* hook (used for fishing)
hawrû *v* receive, welcome, host
harûya mẽ host (*n*)
haqâng *v* be ripe (of fruit)
haqûwa *v* sweat
hawrê *v* try, be careful
hawrêi *n* blood
hawrüng tâpum *n* fishing rod
hám *v* rest
hâ *intj* affirmative
hâpta *clf* weeks (H)
hâqa *adv* without
hâqma *v* hide (e.g. money)
hàng¹ *v* give
hàng² *gr* benefactive, malefactive
hã *pro* 1s
hîng *n* day
hulû *n* golden langur (*Trachypitecus geei*)
hum *v* visit, rest
húm *clf* measurement: cubit (elbow to fingertips)
húm-atyû *clf* measurement: elbow to fist
hû-ganan *n* vulture (probably *Gyps himalayensis*)
hû-cẽ *v* a kind of bamboo
hû-tapû *hû-cẽ*-bamboo shoots
helo *n* underpart, lowerpart (e.g. of a house)
hêra¹ *adj* lazy
hêra² *v* be lazy
hügâ *clf* measurement: thumb to middle finger
hâw *v* pick out, insert (hand or finger in any a kind of container, e.g. a bag, a

pocket, a hole, even in water, to pick out something) **qyá-raw** *n* young woman

hrě/rě *v* pull (e.g. a rope, or a blanket out of a hanging rope)

hrâw *v* burn, light (vi)

hrûw *v* be sour (e.g. rice beer)

hwa *v* want, be active

hwâ *v* be wrong

hwâng-ra *adj* raw (food)

hwâ-wrã ¹ *n* clearing, cleared land

hwâ-wrã ² *adj* clear, open (of a forest)

hwi *n* hive

tawá hwi bee hive

(karà) hwî *v* clear up (as of rain)

hwe *v* pull down (e.g. cloths)

hwè *v* tighten

hwé-tyâng ¹ *v* look, find

hwé-tyâng ² *v* look after, care for

hwû *n* a kind of bamboo

hyâ *v* prepare (food)

Q

-q *gr* continuative

qrû *v* pry

qwã *v* be able

qwîng *v* be old, be ruined, be rotten
(e.g. of wood)

(mîya) qwúw *n* second wife

qwrû *v* lever

12 Appendix I: Texts

This appendix includes excerpts from three stories: Story 12 is *Kutyemba's stone*, Story 12 is a clan story of the Tayang clan and Story 12 is the clan story of the Taly

Text I: Kutyemba's stone (excerpt)

This story was recorded in 2022 in Sunpura Assam. The speakers are Shri Satolum Talampai and Shri Batamso Talampai. The former is the main narrator. The few interventions of Shri Batamso are indicated by the tag *Second narrator*.

Kutyemba's stone is part of the many stories associated with Kutyemba, a mythic trickster. Kutyemba is keen on exploiting the naivety of villagers, as well as that of the Britishers.

(607) mlă gũ alí bîya mề (...)

mlă gũ alí bîya mề

village LOC like.this past person

‘In a village, like this, the people from the past (...)’ [TA22_kutyemba_phla 2 70]

(608) mlă kềq gũ phlă pû daráng khün ã la

mlă kềq gũ phlă pû daráng khing ã la

village LOC.INS LOC stone.MOD rock big one be.there HSAY

‘In a village there was a big rock, they say.’ [TA22_kutyemba_phla 4 73]

- (609) *phlǎ́ pû thũya kě ali tyâng daráng phlǎ́*
stone.MOD rock mountain LOC.INS like.this very big stone
‘A rock in the mountain, like this, a very big stone.’ [TA22_kutyemba_phla 6 78]
- (610) *webo gũ (...)*
webo gũ
and.then INS
‘And then (...)’ [TA22_kutyemba_phla 10 84]
- (611) *kûmtyemba we la la*
kûmtyemba we la la
a.name DEM say HSAY
‘Thus said Kutyemba.’ [TA22_kutyemba_phla 12 86]
- (612) *ewe phlǎ́ ewe (...)* *ane mlǎ́ (...)*
ewe phlǎ́ ewe ane mlǎ́
DEM stone DEM 2P village
‘This stone (...) your village (...)’ [TA22_kutyemba_phla 14 89]
- (613) *mamingzâya wêq adi ephlǎ́ ewe gũ ane (...)*
maming-zâ-ya wê-q adi e=phlǎ́ ewe gũ ane
destroy-DIR-ALLO think-CONT like.this DEM=stone DEM INS 2P
‘It wants to destroy (you), this stone, you (...)’ [TA22_kutyemba_phla 18 93]
- (614) *mề thâ dũ phlǎ́ la*
mề thâ dũ phlǎ́ la
person eat do stone HSAY
‘They say this stone eats people.’ [TA22_kutyemba_phla 20 97]

- (615) ephlăe ãho
e = phlă = e ã-ho
 DEM = stone = DEM be.there-STAT
 ‘This stone is there.’ [TA22_kutyemba_phla 21 99]
- (616) kûmtyemba we (...)
kûmtyemba we
 a.name DEM
 ‘This Kutyemba (...)’ [TA22_kutyemba_phla 23 101]
- (617) we tyû mlă mẽ we la we (...) “tya hîng ephlăe glâikûvidyô!”
we tyû mlă mẽ we la we tya hîng e = phlă = e
 DEM north village.MOD people DEM say DEM now day DEM = stone = DEM
glâi-kû-vidyô
 carry-COMPL.II-CERT.FUT
 ‘Thus he spoke to the village people from the north: “today I’ll carry this stone away!”’ [TA22_kutyemba_phla 25 103]
- (618) we mu ephlăe glâikûya knüm gũ (...)
we mu e = phlă = e glâi-kû-ya knüm gũ
 DEM also DEM = stone = DEM carry-COMPL.II-ALLO PRP INS
 ‘This, in order to carry this stone away (...)’ [TA22_kutyemba_phla 27 111]
- (619) tya ane mlă mẽ (...)
tya ane mlă mẽ
 now 2P village.MOD people
 ‘Now, you village people (...)’ [TA22_kutyemba_phla 28 114]

- (620) *mlǎ mẽ ne (...)*
mlǎ mẽ ane
 village.MOD people 2P
 ‘You village people (...)’ [TA22_kutyemba_phla 29 116]
- (621) “ble (...) ble lê khün mûnbi (...) mûnya gũ thâ günei glâine” la la
ble ble lê khing mûng-bi mûng-ya gũ thâ günei glâi-ne la la
 pig pig big one roast-EMPH roast-ALLO NF eat SEQ carry-INT say HSAY
 “‘The pig (...) roast the biggest pig. After roasting and eating it I will carry (the stone) away’ he said, it is said.’ [TA22_kutyemba_phla 33 120]
- (622) Second narrator: kûtyemba we
kûtyemba we
 a.name DEM
 ‘this Kutyemba’ [TA22_kutyemba_phla 35 125a]
- (623) âm kûtyemba we mlǎ mẽ la matyô
âm kûtyemba we mlǎ mẽ la matyô
 yes a.name DEM village.MOD people say maybe
 ‘Yes, this Kutyemba told the village people.’ [TA22_kutyemba_phla 35 125b]
- (624) ble lê we (...) mûnbina ethâ (...)
ble lê we mûng-bi-na e = thâ
 pig big DEM roast-EMPH-IMP DEM = eat
 ‘Roast this biggest pig and (after) eating it (...)’ [TA22_kutyemba_phla 37 128]

- (625) *thâli gũne ane mlă kě (...)*
thâ-li gũne ane mlă kě
 eat-REC.PST SEQ 2P village LOC.INS
 ‘After eating, at your village (...)’ [TA22_kutyemba_phla 39 131]
- (626) “phlă pû we glâikũhuvidyô” la la
phlă pû we glâi-kũ-hu-vidyô la la
 stone.MOD rock DEM carry-COMPL.II-FS-CERT.FUT say HSAY
 “‘I will carry (away) this rock”, he said, it is said.’ [TA22_kutyemba_phla 40 134]
- (627) *webo (...)* kahîng gê (...) weli nyu bobyě (...)
webo kahîng gê weli nyu bobyě
 and.then CLF:daytimes one DISC.DEIC.DEM FOC TOP
 ‘And then (...) one day (...) as of this (...)’ [TA22_kutyemba_phla 44 140]
- (628) *mlă mẽ la ruga wegũ “ephlăe nyu mẽ thâ gũ byě” (...)*
mlă mẽ la ru-ga wegũ e = phlă = e nyu
 village.MOD people say get.together-PL.I and.then DEM = stone = DEM FOC
mẽ thâ gũ byě
 person eat NF TOP
 ‘The village people got together to talk and (said:) “as of this stone eating people (...)”’ [TA22_kutyemba_phla 46 142]
- (629) “ja tathâ tacî ja mu shyâ yüm”
ja tathâ tacî ja mu shyâ yüm
 what crop ECHO what also happen NEG
 “No crops (will grow), nothing will grow (lit. happen).” [TA22_kutyemba_phla 48 147]

- (630) “ephlãe gũ düi delé” la
e = phlã = e gũ düi delé la
 DEM = stone = DEM INS do it.seems HSAY
 “‘This stone will do it, it seems”, they say.’ [TA22_kutyemba_phla 49 149]
- (631) ewe weli ganyu xxx chôq
ewe weli ganyu xxx chôq
 DEM DISC.DEIC.DEM but.then (inaudible) let’s.do.it
 “And then, let’s do it!” [TA22_kutyemba_phla 51 153]
- (632) mlã mề ruga gũ (...)
mlã mề ru-ga gũ
 village.MOD person get.together-PL.I NF
 ‘The village people getting together (...)’ [TA22_kutyemba_phla 52 154]
- (633) büle tyàng gê (...) mûnrengya khrâk
büle tyàng gê mûng-reng-ya khrâk
 pig CLF:large.bodies one roast-PL.II-ALLO EVI
 ‘To roast (...) one whole pig.’ [TA22_kutyemba_phla 53 156]
- (634) büle lê (...) ewe büle.e me ja shyâq
büle lê ewe büle = e me ja shyâ-q
 pig biggest DEM pig = DEM rear what happen-CONT
 ‘The biggest pig (...) what is the use of rearing this pig?’ [TA22_kutyemba_phla 56 159]

- (635) “glâikũshyûrêtakéq” labô
glâi-kũ-shyû-rê-ta-kéq *la-bô*
 carry-COMPL.II-CAUS.V-PL.II-permanently-HORT.PL HSAY-PFV.ALLO
 “‘Let’s have (him) carry (the stone) away for good”, they said.’ [TA22_kutyemba_phla
 60 163]
- (636) webo kahîng gê mũnya gũ (...)
webo kahîng gê mũng-ya gũ
 and.then CLF:daytime one roast-ALLO NF
 ‘And then, one day roasting (...)’ [TA22_kutyemba_phla 60 163]
- (637) thâ prôwa la büle lê we
thâ prô-ya la büle lê we
 food feed-ALLO HSAY pig biggest DEM
 ‘The biggest pig, (they) fed (him), they say.’ [TA22_kutyemba_phla 63 171]
- (638) wúw tyúwga
wúw tyúw-ga
 big.chunk slice-PL.I
 ‘They sliced big chunks.’ [TA22_kutyemba_phla 65 174]
- (639) cya zâdâmhangâ matyô
cya zâ-dâm-hàng-ya matyô
 3 pack-in.abundance-BEN-ALLO maybe
 ‘(They) might have packed (meat) for him in abundance.’ [TA22_kutyemba_phla
 67 176]

- (640) *kûtyemba we cya we (...) kûtyemba cya we büle thâmyuho matyô (...) tabrê*
kûtyemba we cya we kûtyemba cya we büle thâ-myu-ho matyô tabrê
 a.name DEM 3 DEM a.name 3 DEM pig eat-desire-STAT maybe meat
 ‘This Kutyemba he, this Kutyemba he wanted to eat pig maybe. (He wanted pig)
 meat.’ [TA22_kutyemba_phla 70 181]
- (641) “thâmyuho we” la matyô
thâ-myu-ho we la matyô
 eat-desire-STAT DEM say maybe
 “‘He wanted to eat”, they said.’ [TA22_kutyemba_phla 72 184]
- (642) webobyê (...) thâlibô gũ la la mlá mẽ we la “chôq”
webobyê thâ-li-bô gũ la la mlá mẽ we la
 and.then eat-REC.PST-PFV.ALLO NF say HSAY village.MOD people DEM say
chôq
 let’s.do.it!
 ‘And then, after eating, the village people said: “let’s (do it)!”’ [TA22_kutyemba_phla
 75 190]
- (643) “ephlá pô ewe (...) rêbina (...)” la
e=phlá pô ewe rê-bi-na la
 DEM=stone.MOD rock DEM tie-EMPH-IMP say
 “‘Tie (...) this big rock!” (...) he said.’ [TA22_kutyemba_phla 79 195]
- (644) we taráw tatyé hrê rêmâkūya
we taráw tatyé hrê rê-mâ-kū-ya
 DEM type.of.creeper rope pull tie-ITER-COMPL.II-ALLO
 ‘(They) pulled and tied the rock with a rope made of *taraw* creepers.’
 [TA22_kutyemba_phla 81 197]

(645) xxx weli gũ tahwẽ malâ gũ (...)

xxx weli gũ tahwẽ malâ gũ

(inaudible) DISC.DEIC.DEM INS strap search NF

‘And then, looking for straps (...).’ [TA22_kutyemba_phla 84 201]

(646) ali hwêshyûwga khrâk phlă pû we

ali hwêshyûw-ga khrâk phlă pû we

like.this tighten-CAUS.V-PL.I EVI stone.MOD rock DEM

‘This stone, like this they tightened it.’ [TA22_kutyemba_phla 86 205]

(647) eglâiya knüm

e = glâi-ya knüm

DEM = carry-ALLO PRP

‘In order to carry it.’ [TA22_kutyemba_phla 88 208]

(648) hwêshyûbô laga mlă mẽ we làng we

hwêshyû-bô la-ga mlă mẽ we làng we

tighten-CAUS.V-PFV.ALLO HSAY-PL.I village.MOD person DEM COLL DEM

‘They tighten (the rock), they say, these village people.’ [TA22_kutyemba_phla 90 211]

(649) “enya nyu byẽ (...) hă nu nyu byẽ ìrěyâq”

e = nya nyu byẽ hă nu nyu byẽ ì-rě-yâq

DEM = front FOC TOP 1S near FOC TOP be.there-PL.II-PROH

‘(He said:) “as of (my) front, do not stay near me.”’ [TA22_kutyemba_phla 92 214]

- (650) “alú ane plũng nyu (...) phlă pũ we thrũga lya gũ (...)”
alú ane plũng nyu phlă pũ we thrũga lya gũ
 up.there 2P back.side FOC stone.MOD rock DEM on.the.back lie NF
 “Up there, behind you (...) lying on your back, this big rock (...)” [TA22_kutyemba_phla 96 222]
- (651) “ane mé.mé (...) thrũga lya gũ kadyângrẽna” la
ane mé.mé thrũga lya gũ kadyâng-rẽ-na la
 2P all on.the.back lie NF push.with.legs-PL.II-IMP say
 ‘He said: “all of you (...) push with your legs while lying on your back.”’
 [TA22_kutyemba_phla 99 226]
- (652) “kadyâng halyázâna wegũ (...) hã glâikũdyô” lazâ la
kadyâng halyá-zâ-na wegũ hã glâi-kũ-dyô la-zâ
 push.with.legs topple-DIR-IMP and.then 1S carry-COMPL.II-ahead say-DIR
la
 HSAY
 ‘(He) said: “push with your legs and topple it! And I will carry (the stone) ahead”,
 it is said.’ [TA22_kutyemba_phla 100 228]
- (653) wegũ “kadyângde dũ gũ (...) bülüm màho tayãrẽna” la la
wegũ kadyâng-de dũ gũ bülüm mà-ho
 and.then push.with.legs-HAB.EGO CO.TEMP NF eye close-STAT
tayã-rẽ-na la la
 shout-PL.II-IMP say HSAY
 ‘He said: “and then while we are pushing with our legs (...) shout with your eyes
 closed”, it is said.’ [TA22_kutyemba_phla 105 235]

- (654) *bülüm màho kadyângho (...) we tayàga khrâk mǎ́ mề we*
bülüm mà-ho kadyâng-ho we tayà-ga khrâk mǎ́
 eye close-STAT push.with.legs-STAT DEM shout-PL.I EVI village.MOD
mề we
 person DEM
 ‘With their eyes closed they pushed with their legs and shouted, these village people.’ [TA22_kutyemba_phla 109 241]

- (655) *we lûmkũ (...) phati (...) byúw atîkũna la glâinam thĩ gũ*
we lûmkũ phati byúw atî-kũ-na la glâi-na = m
 DEM meantime suddenly run escape-COMPL.II-FACT HSAY carry-FACT = NEG
thĩ gũ
 CONC NF
 ‘In the meantime, suddenly, he run away, they say, without carrying (the stone).’
 [TA22_kutyemba_phla 115 248]

Text II: Tayang clan story (excerpt)

This story was recorded in 2022. Shri Jabralum Chaitom interviews Shri Juhailum Tayang about the story of his clan. He tells the story of his grandfather and how he captured a dangerous criminal. The interviewer is marked as *I*.

(656) *katyânglyânabô matyô ning hadi*

katyâng-lyâ-na-bô matyô ning hadi

hear-back-FACT-PFV.ALLO maybe 1P story

‘It was heard again, our story.’ [TA22_clan 3 24]

(657) *nyu pram bûqü ashi tashi we làng hadilyâna hadina ja shyânabing*

nyu pra=m bûqü ashi tashi we làng hadi-lyâ-na hadi-na ja

2s good=NEG FS fight ECHO DEM COLL tell-back-FACT tell-FACT what

shyâ-na-bing

become-FACT-POT.ALLO

‘You are not right, ehm, these fights are told again, told again, what will become?’

[TA22_clan 4 27]

(658) *I: bîya ali dyângga tyéga we lana tyé (...)*

bîya ali dyâng-ga tyé-ga we la-na tyé

past like.this chase-PL.I cut-PL.I DEM HSAY-FACT cut

‘In the past, like this, they chased and cut (each other), they say. (They) cut (...)’

[TA22_clan 6 36]

(659) *I: bîya byē (...) bîya jüm (...) ning emh*

bîya byē bîya jüm ning emh

past TOP past true 1P emh

‘As of the past, the true past, our ehm (...)’ [TA22_clan 7 40]

- (660) I: sakarêi hanam pürü gana
sakarêi hana=m pürü gana
 government(<IA) come=NEG yet SEQ
 ‘The government had not yet come.’ [TA22_clan 9 50]
- (661) I: ning tawrã bá.mề kade ìdi
ning tawrã bá.mề kade ì-di
 1P a.tribe clan how.many be.there-HAB.EGO
 ‘How many our Tawrã clans were!’ [TA22_clan 10 51]
- (662) I: bá.mề daráng byě bá.mề aqa pô mu kwâg mu bâ la
bá.mề daráng byě bá.mề aqa pô mu kwâg mu bâ la
 clan big TOP clan small slave also dog also do HSAY
 ‘Big clans would enslave (lit. like slave like dog) smaller clans, they say.’ [TA22_clan 12 56]
- (663) I: weli gũ hatai gũ tsôga dedyâ gũ tsôga la
weli gũ hatai gũ tsô-ga dedyâ gũ tsô-ga la
 DISC.DEIC.DEM INS xxx NF keep-PL.I dominate NF keep-PL.I HSAY
 ‘They would dominate them, they say.’ [TA22_clan 13 61]
- (664) we tyâng dyâ (...) hadi we ning kasânam
we tyâng dyâ hadi we ning kasâ-na=m
 DEM very old story DEM 1P know-FACT=NEG
 ‘This very ancient, story of ours I don’t know.’ [TA22_clan 15 66]

- (665) *tapái we (...) hadi we*
tapái we hadi we
 middle DEM story DEM
 ‘This old (...) this story’ [TA22_clan 16 69]
- (666) I: *bá.mề ruzàng brè (...) dyâng tashi dyângga*
bá.mề ruzàng brè dyâng tashi dyâng-ga
 clan among animosity chase fight chase-PL.I
 ‘There was animosity among the clans. They chased and fought (each other),
 they chased (each other).’ [TA22_clan 18 74]
- (667) I: *tya weli béqbé na (...) we béqbé sokarê hanana*
tya weli béqbé na we béqbé sokarê hana-na
 now DISC.DEIC.DEM slowly INSIST DEM slowly government(<IA) come-FACT
 ‘Now, this, slowly, slowly the government came.’ [TA22_clan 20 78]
- (668) I: *bîya kalà lyó ne hayũ hana la*
bîya kalà lyó ne hayũ hana la
 past officer white only first come HSAY
 ‘At the time, only the white officers (i.e. Britishers) came first, they say.’ [TA22_clan
 22 83]
- (669) I: *weli bene kalà lyó na*
weli bene kalà lyó na
 DISC.DEIC.DEM xxx officer white INSIST
 ‘Then these white officers (i.e. Britishers), right?’ [TA22_clan 23 86]

- (670) I: *kalà lyó byě ama khìcâng yüm na*
kalà lyó byě ama khì-câng yüm na
 officer white TOP here reach-be.able NEG INSIST
 ‘As of the white officers (i.e. Britishers), they could not get here, right?’ [TA22_clan 25 91]
- (671) I: *ama hĩ hanaya (...) shyâ wegũ hĩ we*
ama hĩ hana-ya shyâ wegũ hĩ we
 here EMPH come-ALLO become and.then EMPH DEM
 ‘To come here (...) it would happen’ [TA22_clan 27 96]
- (672) I: *ning ama (...) ning india sokarêi na*
ning ama ning india sokarêi na
 1P here 1P Indian(<E) government(<IA) INSIST
 ‘Our here, our Indian government, right?’ [TA22_clan 29 100]
- (673) *azá má làng*
azá má làng
 king black COLL
 ‘the black kings (i.e. the Indian government)’ [TA22_clan 30 104]
- (674) I: *azá má làng hanana laga*
azá má làng hana-na la-ga
 king black COLL come-FACT HSAY-PL.I
 ‘The black kings (i.e. the Indian government) came, they say.’ [TA22_clan 30 104]

- (675) I: *webo la má bá.mề (...) mềqmềq gũ konsila bâya na*
webo la má bá.mề mềqmềq gũ konsila bâ-ya na
 and.then say black kind beginning LOC counselor(<E) make-ALLO INSIST
 ‘The black kind (i.e. the Indian government) made counselors at the beginning, right?’ [TA22_clan 32 108]
- (676) I: *tayâng kalà làng bârěya na*
tayâng kalà làng bâ-rě-ya na
 a.clan officer COLL make-PL.II-ALLO INSIST
 ‘They appointed (lit. made) Tayang officers, right?’ [TA22_clan 34 115]
- (677) I: *webo béqbé (...)*
webo béqbé
 and.then slowly
 ‘And then slowly (...)’ [TA22_clan 36 119]
- (678) I: *ning ali (...) mấ mề mấ kễ gâmra bâya (...) bârěna na?*
ning ali mấ mề mấ kễ gâmra
 1P like.this village.MOD person village.MOD LOC.INS gaonburha(<IA)
bâ-ya bâ-rě-na na
 make-ALLO make-PL.II-FACT INSIST
 ‘We, like this, appointed (lit. made) gaonburha (from) village people, right?’
 [TA22_clan 37 123]
- (679) I: *konsila nabo aphishi kotsariüw mu shyâya*
konsila nabo aphishi kotsariüw mu shyâ-ya
 counselor(<E) xxx office(<E) school(<Asm?) also become-ALLO
 ‘The counselor, there was an office *kotsariüw*’ [TA22_clan 39 133]

- (680) I: we cya làng gũ ning (...) alí gau mlá mề khika.khika gâmbra mu ì dũ-q la gâmbra mu bârěya (...)

we cya làng gũ ning alí gau mlá mề khika.khika
 DEM 3 COLL INS 1P like.this village(<IA) village.MOD person in.each
gâmbra mu ì dũ-q la gâmbra mu
 gaonburha(<IA) also be.there need-CONT HSAY gaonburha(<IA) also
bâ-rě-ya
 make-PL.II-ALLO

‘They, we, in each (village) there must be a village person, a *gaonburha*, they say. They also appointed (lit. made) *gaonburha*.’ [TA22_clan 40 137]

- (681) I: we já.sũ tashi xxx ali tükü aqa daráng byě

we já.sũ tashi xxx ali tükü aqa daráng byě
 DEM whatever fight (inaudible) like.this matter small big COND
 ‘Whatever quarrel, be it big or small (...)’ [TA22_clan 41 144]

- (682) I: ning gâmbra lûmkũ hîq dângkũ dũya we sokarêi bô shyũ yûm

ning gâmbra lûmkũ hîq dâng-kũ dũw-ya we
 1P gaonburha(<IA) inside EMPH finish-COMPL.II need-ALLO DEM
sokarêi bô-shyũ yûm
 government(<IA) go-allow NEG

‘We had to settle (lit. finish) under the *gaonburha*, we didn’t allow (them) to go to the government.’ [TA22_clan 42 150]

- (683) bü (...) ning pâshya shìza labô matyô

bü ning pâshya shì-za la-bô matyô
 once 1P a.name take-DIR HSAY-PFV.ALLO maybe

‘Once, our Pashya took (it), they say, maybe.’ [TA22_clan 44 157]

- (684) I: ah we làng khúbiya we matyô
ah we làng khú-biya we matyô
 ah DEM COLL be.enough-POT.ALLO DEM maybe
 ‘Oh, that will be enough, maybe.’ [TA22_clan 45 162]
- (685) I: we weli mu (...) we hadi mu hadina
we weli mu we hadi mu hadi-na
 DEM DISC.DEIC.DEM also DEM story also tell-IMP
 ‘This, this also (...) also tell this story!’ [TA22_clan 47 168]
- (686) we hadi ì shìzâ gũ (...) we ning hanúk ìzâ na?
we hadi ì shì-zâ gũ we ning hanúk ì-zâ na
 DEM story be.there take-DIR NF DEM 1P where be.there-DIR INSIST
 ‘This story, taking a place (...) (of) where we stayed, right?’ [TA22_clan 48 171]
- (687) we ma (...) ning bá.mề pà ne shyânabô
we ma ning bá.mề kapâk ne shyâ-na-bô
 DEM CONTR.TOP 1P clan alone only become-FACT-PERF.ALLO
 ‘This (...) (how) our clan alone came to be.’ [TA22_clan 49 174]
- (688) ning bá.mề kě ning (...)
ning bá.mề kě ning
 1P clan LOC.INS 1P
 ‘In our clan, our (...)’ [TA22_clan 50 177]
- (689) tayâng bá.mề kasâshikũna
tayâng bá.mề kasâ shì-kũ-na
 a.clan clan know take-COMPL.II-IMP
 ‘Consider the Tayang clan.’ [TA22_clan 51 180]

(690) I: wegũ hadi kasâ gũ (...)

wegũ hadi kasâ gũ

and.then story know NF

‘By knowing this story (...).’ [TA22_clan 53 183]

(691) I: ekẽ hanana sokarêi

e=kẽ hana-na sokarêi

DEM=LOC.INS come-FACT government(<IA)

‘The government came here.’ [TA22_clan 54 184]

(692) we ma bîya gũ

we ma bîya gũ

DEM CONTR.TOP past INS

‘The government came here.’ [TA22_clan 55 186]

(693) I: pâshya (...) ama sokarêi hanam phrũ gũ (...)

pâshya ama sokarêi hana=m phrũ gũ

a.name here government(<IA) come=NEG yet NF

‘Pashya (...) the government hadn’t reached here yet (...)’ [TA22_clan 56 188]

(694) I: we mamyú ne sokarêi hana

we mamyú ne sokarêi hana

DEM later only government(<IA) come

‘Only later did the government come.’ [TA22_clan 58 194]

(695) I: we mamyú (...) kadége ganei kalà lyó hana matyô

we mamyú kadége ganei kalà lyó hana matyô

DEM later how.much SEQ officer white come maybe

‘Later, how much later did the white officers come.’ [TA22_clan 60 196]

- (696) I: *pâshyã (...)* *bîya tükü na*
pâshyã bîya tükü na
 a.name past matter INSIST
 ‘Pashya (...) an old matter, right?’ [TA22_clan 61 199]
- (697) I: *khûnyüm dahi mẽ dahi mẽ na?*
khûnyüm dahi mẽ dahi mẽ na
 god like person like person INSIST
 ‘Gods were like people, (these) people, right?’ [TA22_clan 62 200]
- (698) *ning bîya ning bîya gũ (...)* *tayâng hana we hi*
ning bîya ning bîya gũ tayâng hana we hi
 1P past 1P past LOC a.clan come DEM EMPH
 ‘In our past (...) the Tayang came.’ [TA22_clan 64 207]
- (699) *dâi mlă nyu hana dahi ning we la (...)* *ashí nyu hana mane*
dâi mlă nyu hana dahi ning we la ashí nyu hana
 a.tribe place FOC come like 1P DEM HSAY horizontal.axis FOC come
mane
 FILL(<Asm)
 ‘They came from the Idu villages, they say. From a the west (horizontal axis), I mean.’ [TA22_clan 65 212]
- (700) *we hi nyu hanazâ wegũ (...)* *hanahanahana we*
we hi nyu hana-zâ wegũ hana~hana~hana we
 DEM EMPH FOC come-DIR and.then come~PLAC~PLAC DEM
 ‘Like this they came. They came, came, came, this.’ [TA22_clan 66 215]

- (701) eh wrâng dyu hana gũ thîya nyu tyû
eh wrâng dyu hana gũ thîya nyu tyû
 eh steep border come NF mountain FOC xxx
 ‘They came along the edge of the mountains.’ [TA22_clan 67 219]
- (702) kadé.nyu plang mlầ we?
kadé.nyu plang mlầ we
 how FILL place DEM
 ‘What place was that?’ [TA22_clan 71 229]
- (703) wesâhona xxx shyâm atyu ning mlầ
wesâ-ho-na xxx shyâm atyu ning mlầ we
 forget-STAT-FACT xxx no north 1P place DEM
 ‘I’ve forgotten (lit. it’s forgotten). No! (Up to the) north, our village.’ [TA22_clan 73 232]
- (704) krõ làng mlầ takyã (...) kadâ.nyu plang (...)
krõ làng mlầ takyã kadâ.nyu plang we
 a.clan COLL place LOC.INS how FILL DEM
 ‘At the village of the Kro, how was it (...)’ [TA22_clan 75 242]
- (705) ah we takyã alủ...
ah we takyã alủ
 ah DEM LOC.ATOP up.there
 ‘Ah, up there (...)’ [TA22_clan 77 247]

(706) I: cîlyang

cîlyang

a.place

‘At Chilyang’ [TA22_clan 77 247]

(707) cîlyang takyã alú shyâ

cîlyang takyã alú shyâ

a.place LOC up.there become

‘It was Chilyang, up there.’ [TA22_clan 78 249]

(708) sâga mlâ!

sâga mlâ

a.clan village

‘(It’s the) Saga village!’ [TA22_clan 80 254]

(709) I: sâga mlâ oh!

sâga mlâ ho

a.clan village oh

‘Oh, Saga village!’ [TA22_clan 80 254]

(710) we nyi atyû saglama sâga mlâ ìxxxna âm

we ning atyû saglama sâga mlâ ì-xxx-na âm

DEM 1P north FS a.clan village be.there-xxx-FACT yes

‘They were in our northern (...) Saga village, yes.’ [TA22_clan 81 256]

(711) *sâga mlâ ìzâ wegũ ning kadége tyûsei asã kadége we ìzâna*
sâga mlâ ì-zâ wegũ ning kadége tyûsei asã kadége
 a.clan village be.there-DIR and.then 1P how.many a.clan a.clan how.many
we ì-zâ-na
 DEM be.there-DIR-FACT
 ‘Being at Saga village, how many were (clans) were there, (like) Tyusei, Asa.’
 [TA22_clan 82 259]

(712) *we ìrênazâ wegũ mane wegũ mane (...)*
we ì-rê-na-zâ wegũ mane wegũ mane
 DEM be.there-PL.II-FACT-DIR and.then FILL(<Asm) and.then FILL(<Asm)
 ‘These being there, I mean, this, I mean (...)’ [TA22_clan 83 265]

(713) *wegũ mane ning we (...)*
wegũ mane ning we
 and.then FILL(<Asm) 1P DEM
 ‘And then, I mean, we (...)’ [TA22_clan 84 269]

(714) *we pêphrêgana mane (...)*
we pê-phrê-ga-na mane
 DEM separate-away-PL.I-FACT FILL(<Asm)
 ‘These separated, I mean (...)’ [TA22_clan 84 269]

(715) *I: pêga*
pê-ga
 separate-PL.I
 ‘(They) separated.’ [TA22_clan 84 269]

- (716) âm, pêna wegũ hanûk xxx bôkũna ahi
âm pê-na wegũ hanûk xxx bô-kũ-na aci
 yes separate-FACT and.then where xxx go-COMPL.II-FACT horizontal.axis
 ‘Yes. Having separated, where (not clear) (they) went, (to the) west.’ [TA22_clan 84-6 269-272]
- (717) shirẽ nyu mu bôbô erẽ nyu mu hanabô
shi-rẽ nyu mu bô-bô we-rẽ nyu mu hana-bô
 there-towards FOC also go-PFV.ALLO DEM-towards FOC also come-PFV.ALLO
 ‘Gone here and there.’ [TA22_clan 87 275]
- (718) welibô ning ewe (...)
welibô ning ewe
 DISC.DEIC.DEM 1P DEM
 ‘This our (...)’ [TA22_clan 88 277]
- (719) abli gê shyûlyang nyu (...) abli bô
abli gê shyûlyang nyu abli bô
 CLF:some one a.place FOC CLF:some go
 ‘Some (went) to Shyulyang, some went (there).’ [TA22_clan 89 279]
- (720) ning tahũ abli gê malũ.agũm abli gê hana
ning tahũ abli gê malũ.agũm abli gê hana
 1P group CLF:some one a.place CLF:some one come
 ‘(From) our group (sub-clan), some (went) to Maluagum, some went.’ [TA22_clan 90 282]

- (721) *we abli gê we hî cilyang la gũ abli gê hana*
 DEM CLF:some one DEM EMPH a.place say NF CLF:some one come
 ‘Some, some went to so-called Chilyang.’ [TA22_clan 91 285]

(...)

- (722) *we hazá lyó ì yüm tüpwi wegũ hazá má ne ìya*
 DEM king white be.there NEG until and.then king black only be.there-ALLO
 ‘The Britishers (lit. the white kings) being not there anymore; only the Indian government (lit. the black king) were there.’ [TA22_clan 121 357]

- (723) *hazá má làng we ayũ tükü na*
hazá má làng we ayũ tükü na
 king black COLL DEM past matter INSIST
 ‘This Indian government (lit. the black king) is an old matter.’ [TA22_clan 123 361]

- (724) *webo hazá má làng hana-bô gũnei webo gũnei ewe ning (...)*
webo hazá má làng hana-bô gũnei webo gũnei ewe ning
 and.then king black coll come-PFV.ALLO SEQ and.then SEQ DEM 1P
 ‘And then the Indian government (lit. the black king) had come and then, then this, our (...)’ [TA22_clan 125 364]

- (725) *nabá làng we*
na-bá làng we
 1S-father.MOD COLL DEM
 ‘These (my) fathers’ [TA22_clan 126 367]

- (726) *caina boder (...) boder nyu shyâzâ làng (...) khiga na*
caina boder boder nyu shyâ-zâ làng khi-ga na
 China(<E) border(<E) border(<E) FOC happen-DIR COLL reach-PL.I INSIST
 ‘The ones that were at the China border, they reached (there), right?’ [TA22_clan 128 372]
- (727) *cya làng tyu khiga we (...) khi mlă nyu boder (...) khiya gũ mane (...) glâiringkibô*
cya làng tyu khi-ga we khi mlă nyu boder khi-ya gũ
 3 COLL FS reach-PL.I DEM China place FOC border(<E) reach-ALLO NF
mane glâi-ring-ki-bô
 FILL(<Asm) carry-PL.II-xxx-PFV.ALLO
 ‘They reached (...) bringing (goods) to the Chinese border (...) having carried it.’
 [TA22_clan 129 377]
- (728) *aglú glêga mane*
aglú glê-ga mane
 provision carry-PL.I FILL(<Asm)
 ‘They carried provisions, I mean.’ [TA22_clan 131 389]
- (729) *we glêga mề khing.gê kamáng gũ wê tyô lyûm tyô glêya xxx bôga mane*
we glê-ga mề khing.gê kamáng gũ wê tyô lyûm tyô
 DEM carry-PL.I person one leader INS seven maybe eight maybe
glêi-ya xxx bô-ga mane
 carry-ALLO xxx go-PL.I FILL(<Asm)
 ‘This one carrying person, (with?) the leader, would carry seven or eight xxx they go, I mean.’ [TA22_clan 132 391]

(730) we bôbôbô we ning we

we bô~bô~bô we ning we

DEM go~PLAC~PLAC DEM 1P DEM

‘This one went, went and went, this our (person).’ [TA22_clan 133 395]

(731) atyû we kadé.ge khiringge we (...) hatyû nyu glênaringzâ ja la dũ (...) atyû nyu
bôq atyû nyu xxx we glâinazâ enyu.e sâman nu glâigana

atyû we kadé.ge khi-ge we hatyû nyu glê-na-ring-zâ

north DEM how.much reach-and.back DEM north FOC carry-FACT-PL.II-DIR

ja la dũ atyû nyu bô-q atyû nyu xxx we glâi-na-zâ

what say CO.TEMP north FOC go-CONT north FOC xxx DEM carry-FACT-DIR

e=nyu=e sâman nu glâi-ga-na

DEM=FOC=DEM goods FOC carry-PL.I-FACT

‘How many reached (and came back) to this norther (place). They carried (goods) to the north. What to say. They go to the north, to the north xxx they carried goods, they carried.’ [TA22_clan 135 404]

(...)

Text II: Talyu clan story (excerpt)

The Talyu clan story was recorded by Shri Johakso Manyu in 2023. Chalalum Talyu tells the migration story of his clan to the plains. He also talks about the forgotten practice of rearing dholes (*Cuon alpinus*). The interviewer is marker as *I*.

- (732) ning talyũ we suruwat we tuluqagũ shyâ la kahani katyângdi la hadi we
 ning talyũ we suruwat we tuluqa gũ shyâ la kahani
 1P a.clan DEM beginning(<IA) DEM a.place LOC become HSAY story(<IA)
 katyâng-di la hadi we
 hear-HAB.EGO HSAY story DEM
 ‘Our Talyu (clan), at the beginning, came to be (from) Tulu-a (i.e. small Tulu),
 they say. We heard the story, this story.’ [TA23_talyu_clan 5 20]

- (733) ning mâ bà gũ hadiga we tükü we
 ning mâ bà gũ hadi-ga we tükü we
 1P mother father INS tell-PL.I DEM matter DEM
 ‘Our parents told this thing.’ [TA23_talyu_clan 6 29]

- (734) jũm acîk bîya gũ dyó gũ pwi gũ shyâzâ hanazâ tükü byẽ ning katyângdüm tapwi
 jũm acîk bîya gũ dyó gũ pwi gũ shyâ-zâ hana-zâ tükü byẽ
 true xxx past LOC different INS other INS happen-DIR come-DIR matter TOP
 ning katyâng-d=üm tapwi
 1P hear-HAB.EGO=NEG until
 ‘In the true (...) past, as of something being or coming (from) a different (place)
 (or) otherwise, we never heard (of it).’ [TA23_talyu_clan 7 34]

- (735) ning we mâ bà gũ kahani adi atyà we tuluqa gũ shyâ-zâ (...) tuluq agũ ning pûmru
 kaprêi ìgagabo la shyâ laga
ning we mâ bà gũ kahani adi atyà we tuluqa gũ
 1P DEM mother father INS story(<IA) story ECHO DEM a.place LOC
shyâ-zâ tuluqa gũ ning pûmru kaprêi ì-ga-ga-bô
 become-DIR a.place LOC 1P younger.brother four be.there-FS-PL.I-PFV.ALLO
la shyâ la-ga
 HSAY become HSAY-PL.I
 ‘We, (our) parents told (about) Tuluqa (lit. small Tulu). At Tuluqa (lit. small Tulu) there were four brothers of ours, they say. It happened, they say.’
 [TA23_talyu_clan 8 39]

- (736) pûmru kaprêi ìya gũ bîya thü ning we cuwe (...)
pûmru kaprêi ì-ya gũ bîya thü ning we cuwe
 younger.brother four be.there-ALLO NF past ABL 1P DEM DISC.DEIC.DEM
 ‘There were four brothers in the past, this ours (...)’ [TA23_talyu_clan 9 48]

- (737) tya ning taprûw lade na?
tya ning taprûw la-de na
 now 1P dhole say-HAB.EGO INSIST
 ‘Now, we say *taprûw* (for dhole), right?’ [TA23_talyu_clan 10 52]

- (738) I: taprûw
taprûw
 dhole
 ‘dhole’ [TA23_talyu_clan 10 52]

- (739) *taprûw lade we kwôw dahi cuwe taprûw lade ning (...) ning talyũ dú gũ ì khrâk
bũ gũ*
taprûw la-de we kwôw dahi cuwe taprûw la-de ning
 dhole say-HAB.EGO DEM dog like DISC.DEIC.DEM dhole say-HAB.EGO 1P
ning talyũ dũ gũ ì khrâk bũ gũ
 1P a.clan with LOC be.there EVI once LOC
 ‘We call them *taprûw* (dhole), (they are) like dogs, we call them *taprûw* (dhole)
 our (...) they lived with our Talyu clan, once.’ [TA23_talyu_clan 10 52]

- (740) I: oh mehoring mane
oh me-ho-ring mane
 oh rear-STAT-PL.II FILL(<Asm)
 ‘Oh, so you reared them, I mean?’ [TA23_talyu_clan 11 60]

- (741) *ã meho taprûw bra kaqîng ìya la*
ã me-ho taprûw bra kaqîng ìya la
 yes rear-STAT dhole CLF:generic two be.there-ALLO HSAY
 ‘Yes, (they) reared them, there were two dholes.’ [TA23_talyu_clan 11 60]

- (742) I: accha
accha
 good(<H)
 ‘I see.’ [TA23_talyu_clan 11 60]

- (743) bra kaqîng ìya wegũ pûmru we (...) cyang (...) tamêga shyâ mu thûde jiya yûm bum.kamh jiya yûm taprûw a shêibiya la

bra kaqîng ì-ya wegũ pûmru we cya làng
 CLF:generic two be.there-ALLO and.then younger.brother DEM 3 COLL
tamê-ga shyâ mu thû-de ji-ya yûm bum.kamh ji-ya
 refuse-PL.I happen also trap-HAB.EGO do-ALLO NEG trap.in.i.pit do-ALLO
yûm taprûw a shêi-biya la
 NEG dhole pup die-POT.ALLO SAY

‘Being there two individuals (dhole), this younger brother told them not to trap (them) and not to make trapping pits, as they would die.’ [TA23_talyu_clan 11 60]

- (744) tamêga ning we tabrê tathâ we mane (...) taprûw we gũ zâ (...) tyége gũ ning we kambâ thâga khrâk (...) webo pûmru jinaya (...)

tamê-ga ning we tabrê tathâ we mane taprûw we gũ zâ
 refuse-PL.I 1P DEM meat ECHO DEM FILL(<Asm) dhole DEM NF DIR
tyé-ge gũ ning we kambâ thâ-ga khrâk webo pûmru
 cut-and.back NF 1P DEM for.free eat-PL.I EVI and.then younger.brother
ji-na-ya
 do-FACT-ALLO

‘They disagreed. We, the meat (...) the dhole itself would go and hunt and we would eat for free, it seems. Then this younger brother (...)’ [TA23_talyu_clan 12 70]

- (745) gadyü âm (...) ta tablüw jina gũ (...)

ga-dyü âm ta tablüw ji-na gũ
 self-like yes FS incomppliance do-FACT NF

‘They (...) without listening (...)’ [TA23_talyu_clan 14 81]

- (746) takhrâizâkûna webo taprû a bra gê shêikûbi
takhrâi-zâ-kû-na *webo* *taprû* *a* *bra* *gê*
 trap.in.a.snare.trap-DIR-COMPL.II-FACT and.then dhole pup CLF:general one
shêi-kû-bi
 catch-COMPL.II-PFV
 ‘(They) trapped (one). And then one dhole pup was caught.’ [TA23_talyu_clan 15 83]
- (747) I: taprû a
taprû *a*
 dhole pup
 ‘(A) dhole pup’ [TA23_talyu_clan 15 83]
- (748) taprû a bra gê shêikû
taprû *a* *bra* *gê* *shêi-kû*
 dhole pup CLF:general one catch-COMPL.II
 ‘One dhole pup was caught.’ [TA23_talyu_clan 15 83]
- (749) shêikûbi
shêi-kû-bi
 catch-COMPL.II-PFV
 ‘It was caught.’ [TA23_talyu_clan 15 83]
- (750) âm, webo gû ja shyâna webo gû la khûganabi gû
âm *webo* *gû* *ja* *shyâ-na* *webo* *gû* *la* *khû-ga-na-bi* *gû*
 yes and.then INS what happen-FACT and.then INS say fight-PL.I-FACT-PFV NF
 ‘Yes, and then, what happened? And then they had a discussion.’ [TA23_talyu_clan 15 83]

- (751) *pû-mru kaprei kê gũ ashâ nyu abli gê bôkũbô*
pû-mru kaprei kê gũ ashâ nyu abli
 elder.brother-younger.brother four LOC.INS LOC plain.area FOC CLF:part
gê bô-kũ-bô
 one go-COMPL.II-PFV.ALLO
 ‘Among the four brothers, (to) the plains went some.’ [TA23_talyu_clan 16 93]

- (752) *pû-mru bra gê bra gê ashâ bô gũ bra gê we dâi mlâ nyu bra gê thîya bô gũ bô bra*
gê we tuluqa kê ìya la
pû-mru bra gê bra gê ashâ bô
 elder.brother-younger.brother CLF:general one CLF:general one plain.area go
gũ bra gê we dâi mlâ nyu bra gê thîya bô
 NF CLF:general one DEM a.tribe place FOC CLF:general one mountain go
gũ bô bra gê we tuluqa kê ì-ya la
 NF go CLF:general one DEM a.place LOC.INS be.there-ALLO HSAY
 ‘One brother, one went to the plain area, one went to the land of the Idus,
 one went to the mountains and then one is at Tulu-a (small Tulu), they say.’
 [TA23_talyu_clan 17 96a]

- (753) *I: ashâ*
ashâ
 plain.area
 ‘The plains’ [TA23_talyu_clan 17 96b]

- (754) *weli hadigabô xxx*
weli hadi-ga-bô xxx
 DISC.DEIC.DEM tell-PL.I-PFV.EGO xxx
 ‘This they told (us) xxx’ [TA23_talyu_clan 17 96c]

(755) we talyu bá.mě̀ (...)

we talyu bá.mě̀

DEM a.clan clan

‘This Talyu clan (...)’ [TA23_talyu_clan 18 106a]

(756) I: thîya nyu bôya we tya hanûk mlằ ?

thîya nyu bô-ya we tya hanûk mlằ

mountain FOC go-ALLO DEM now where place

‘Those who went to the mountain, where are they now?’ [TA23_talyu_clan 18 106b]

(757) we thîya nyu we tya (...)

we thîya nyu we tya

DEM mountain FOC DEM now

‘The ones at the the mountain now (...)’ [TA23_talyu_clan 18 106c]

(758) asil jüm lana bô thîya ning we pwi we ji hanûk gũ mu ìhona yüm tüpwi na cipru

ma welibo mathûw we jîngkũ.chuwâkũ

asil jüm la-na bô thîya ning we pwi we

true(<IA) true say-FACT TEMP.SUB mountain 1P DEM different DEM

ji hanûk gũ mu ì-ho-na yüm tüpwi na cipru ma

EMPH where INS also be.there-STAT-FACT NEG until INSIST a.place FOC

welibo mathûw we jîngkũ.chu.wâkũ

DISC.DEIC.DEM up.there DEM a.place

‘To tell the truth, we were everywhere in mountains, in Chipru and also up there in Jinku-Chuwaku.’ [TA23_talyu_clan 19 112]

(759) I: mmh

mmh

mmh

‘mmh’ [TA23_talyu_clan 20 122a]

(760) we ne khing talyũ mlǎ̃ tapwi we b̃ya h̃i (...)

we ne khing talyũ mlǎ̃ tapwi we b̃ya h̃i

DEM only one a.clan place from DEM past EMPH

‘This from only one Talyu village, in the past (...)’ [TA23_talyu_clan 20 122b]

(761) I: oh b̃ya h̃i

oh b̃ya h̃i

oh past EMPH

‘Oh in the past.’ [TA23_talyu_clan 20 122c]

(762) âm we talyũ mlǎ̃ tapwi

âm we talyũ mlǎ̃ tapwi

yes DEM a.clan place from

‘Yes, from this Talyu village.’ [TA23_talyu_clan 20 122d]

(763) welibo gũ cipru we welibo tohwa nyu hwagum lade

welibo gũ cipru we welibo tohwa nyu hwagum

DISC.DEIC.DEM INS a.place DEM DISC.DEIC.DEM a.place FOC a.place

la-de

call-HAB.EGO

‘And then this Chipru and then Towha, called Hwagum.’ [TA23_talyu_clan 21 128a]

(764) I: hwagum

hwagum

a.place

‘(in) Hwagum.’ [TA23_talyu_clan 21 122b]

(765) we ning mlâ

we ning mlâ

DEM 1P place

‘This our place.’ [TA23_talyu_clan 21 122c]

(766) we hwagum lade we

we hwagum la-de we

DEM a.place call-HAB.EGO DEM

‘We call it Hwagum.’ [TA23_talyu_clan 22 134]

(767) talyũ mlâ tapwi we we (...) weli ekẽ gũ hanûk e kẽ gũ tuluqa gũ lêbô hanûk ekẽ
gũ ejêk (...)

talyũ mlâ tapwi we we weli e=kẽ gũ hanûk

a.clan place until DEM DEM DISC.DEIC.DEM DEM=LOC.INS LOC where

e=kẽ gũ tuluqa gũ lê-bô hanûk e=kẽ gũ

DEM=LOC.INS LOC a.place LOC leave-PFV.ALLO where DEM=LOC LOC.INS

ejêk

exactly(<E)

‘Up to Talyu village, (they came through) here, and here, from Tulua they left
(but) where exactly (...)’ [TA23_talyu_clan 22 134]

- (768) *i ì-nahibô (...) we byě we (...) we kho minîng kasâ yüm na*
i ì-na-hi-bô we byě we we kho
 FS be.there-FACT-stay.behind-PFV.ALLO DEM TOP DEM DEM even.though
minîng kasâ yüm na
 meaning(<E) know NEG INSIST
 ‘(He) stayed back. As of him, I don’t know the truth (lit. *meaning*), right?’
 [TA23_talyu_clan 24 141a]
- (769) *I: we ta ma bra atyu nyu bônabô bra (...)*
we ta ma bra atyu nyu bô-na-bô
 DEM ADVS CONTR.TOP CLF:generic up.there FOC go-FACT-PFV.ALLO
bra
 CLF:generic
 ‘As of this, one (person) went up there, one (...)’ [TA23_talyu_clan 24 141b]
- (770) *I: bra gê we dâi mlă (...)*
bra gê we dâi mlă
 CLF:generic one DEM a.tribe place
 ‘One (went) to the Idu’s place (...)’ [TA23_talyu_clan 25 142a]
- (771) *âm weli kahani jiga*
âm weli kahani ji-ga
 yes DISC.DEIC.DEM story(<IA) do-PL.I
 ‘Yes, this is what they told (us).’ [TA23_talyu_clan 25 142b]
- (772) *I: dâi mlă we tya cya làng cya ja zûgasâ*
dâi mlă we tya cya làng cya ja zû-ga-sâ
 a.tribe place DEM now 3 COLL 3 what write-PL.I-POT
 ‘Those (living) on the Idu’s place, they, what do they write (for their clan/surname)?’
 [TA23_talyu_clan 25 142c]

- (773) we talyũ hî la zûga
 we talyũ hî la zû-ga
 DEM a.clan EMPH still write-PL.I
 ‘They still write *Talyu*.’ [TA23_talyu_clan 25 142d]
- (774) I: we ane (...) ane bü kadé trûgasâ we?
 we ane ane bü kadé trû-ga-sâ we
 DEM 2P 2P CLF:times how.many meet-PL.I-POT DEM
 ‘(And) you, how many times do you meet?’ [TA23_talyu_clan 26 156a]
- (775) ning we dêqe (...)
 ning we dêqe
 1P DEM earlier
 ‘Earlier we (...)’ [TA23_talyu_clan 26 156b]
- (776) bîya nang hayũ büw kaqing khìge
 bîya nang hayũ büw kaqing khì-ge
 past year last CLF:times two reach-and.back
 ‘Last year I went there two times.’ [TA23_talyu_clan 27 160a]
- (777) I: hanûk mlâ ìhoring?
 hanûk mlâ ì-ho-ring
 where place be.there-STAT-PL.II
 ‘Where were you?’ [TA23_talyu_clan 27 160b]

(778) we ewe tyû ta mlâ we nyu tya hã wesâkûhona dyâ (...) mayodyali gamu dyâ
 we ewe tyû ta mlâ we nyu tya hã wesâ-kû-ho-na
 DEM DEM up.there FS place DEM FOC now 1S forget-COMPL.II-STAT-FACT
 dyâ mayodyali gamu dyâ
 far.away a.place ABL far.away
 ‘This, up there, ehm, I now forgot the name of the place (but) it is far (...) far
 from Mayodyali.’ [TA23_talyu_clan 28 165a]

(779) I: oh italin nyu
 oh italin nyu
 oh a.place FOC
 ‘Oh at Italin.’ [TA23_talyu_clan 28 165b]

(780) weli nyu dyûw we shìya tyu mathûw tyâng dyâ
 weli nyu dyûw we shì-ya tyu mathûw
 DISC.DEIC.DEM FOC seem DEM perceive-ALLO up.there up.in.the.mountain
 tyâng dyâ
 very far.away
 ‘It seems like that, up there, so far away.’ [TA23_talyu_clan 28 165c]

(781) we cya làng amu ali khalai phlâ bréiya gũ ìga roing gũ ìga
 we cya làng amu ali khalai phlâ bréi-ya gũ ì-ga
 DEM 3 COLL like.this like.this ground stone buy-ALLO NF be.there-PL.I
 roing gũ ì-ga
 a.place LOC be.there-PL.I
 ‘They, like this, bought land (lit. land-stone) and they are living in Roing, they
 are there.’ [TA23_talyu_clan 29 175a]

(782) I: Oh xxx

oh xxx

oh xxx

‘Oh (inaudible)’ [TA23_talyu_clan 29 175b]

(783) *bajar kě ìhoring*

bajar kě ì-ho-ring

marker(<IA) LOC.INS be.there-STAT-PL.II

‘They are staying at the *bajar*.’ [TA23_talyu_clan 29 175c]

(784) *we mlă jüm tyu ma jüm bôya la la alyüm karàzâ tatyũ la gũ byě bü pwi alyüm*
(...)

we mlă jüm tyu ma jüm bô-ya la la alyüm karà-zâ tatyũ

DEM place true up.there FOC true go-ALLO say HSAY road rain-DIR problem

la gũ byě bü pwi alyüm

say NF TOP CLF:times different road

‘As of going to that true place up there, they say, (they) said there was a problem in the road (due to the) rain. (Taking) a different road (...)’ [TA23_talyu_clan 30 182a]

(785) “*karà hwî bô ne (...) hanaringna*” *la maroga gũ ning khîlyâna yüm*

karà hwî bô ne hana-ring-na la maro-ga gũ ning

rain clear.up TEMP.SUB only come-PL.II-FACT HSAY talk-PL.I NF 1P

khî-lyâ-na yüm

reach-back-FACT NEG

‘Saying “come only when the weather clears up” we did not go back.’

[TA23_talyu_clan 30 182b]

- (786) cya làng mu talyũ hĩ zũga
cya làng mu talyũ hĩ zũ-ga
 3 COLL also a.clan EMPH write-PL.I
 ‘They still write *Talyu*.’ [TA23_talyu_clan 31 192]
- (787) I: accha talyũ hĩq zũga cya làng mu
accha talyũ hĩq zũ-ga cya làng mu
 I.see(<H) a.clan EMPH write-PL.I 3 COLL also
 ‘Alright! they write *Talyu*.’ [TA23_talyu_clan 32 197a]
- (788) talyũ hĩq zũga
talyũ hĩq zũ-ga
 a.clan EMPH write-PL.I
 ‘They write *Talyu*.’ [TA23_talyu_clan 32 197b]
- (789) I: cya làng cya làng mu hadi we îdu hana la
cya làng cya làng mu hadi we îdu hana la
 3 COLL 3 COLL also tell DEM a.tribe come HSAY
 ‘(So) they, they also recount of coming to the Idu side.’ [TA23_talyu_clan 32 197c]
- (790) ah ma cya làng mu ji la tuluqa nyu hana hĩ la ning mu ahuga maroga we cya làng
 we ning tuluqa nyu shyâzâ (...)
ah ma cya làng mu ji la tuluqa nyu hana hĩ la ning mu
 ah FOC 3 COLL also EMPH say a.place FOC come EMPH say 1P also
ahu-ga maro-ga we cya làng we ning tuluqa nyu shyâ-zâ
 ask-PL.I speak-PL.I DEM 3 COLL DEM 1P a.place FOC emerge-DIR
 ‘Ah, they also say ‘yes, we (too) came from Tulua!’. We also asked and talked
 (and) they (say) we also emerged from Tulua (...)’ [TA23_talyu_clan 32 197d]

- (791) *wegũ* “ning ali hana” la
wegũ ning ali hana la
 and.then 1P like.this come HSAY
 ‘And then “like this we came” they say.’ [TA23_talyu_clan 33 206]
- (792) *weli* adiga cya làng mu
weli adi-ga cya làng mu
 DISC.DEIC.DEM tell-PL.I 3 COLL also
 ‘Like this they recount.’ [TA23_talyu_clan 34 208a]
- (793) I: mmh
mmh
 mmh
 ‘mmh!’ [TA23_talyu_clan 34 208b]
- (794) *wegũ* *pũ-mru* bá.mề kade.nyu (...)
wegũ pũ-mru bá.mề kade.nyu
 and.then elder.brother-younger.brother clan how.many
 ‘And then, how many brothers (...)’ [TA23_talyu_clan 34 208c]
- (795) *tshũnem* nyu kade.nyu shyâna hâ mu ning ahuge we mu
tshũ-na =m nyu kade.nyu shyâ-na hâ mu ning ahu-ge
 expand-FACT = NEG FOC how.come happen-FACT EMPH also 1P ask-and.back
we mu
 DEM also
 ‘They would not expand (i.e. reproduce). How come it became (like this), we went to ask (them).’ [TA23_talyu_clan 35 216a]

- (796) I: ane jabâ nyu tshûnem
ane jabâ nyu tshû-na = m
 2P why FOC expand-FACT = NEG
 ‘Why don’t you expand.’ [TA23_talyu_clan 35 216b]
- (797) mmh (...) ane kadé.nyu we nyu tuluqa gũ lê pêgana we tya khrû nyu (...)
mmh ane kadé.nyu we nyu tuluqa gũ lê pê-ga-na we tya
 mmh 2P how.come DEM FOC a.place INS exit separate-PL.I-FACT DEM now
khrû nyu
 until FOC
 ‘Mmh, how come you came out of Tulua, you separated (and) up to now (...)’
 [TA23_talyu_clan 35 216c]
- (798) katyê inade
katyê i-na-de
 small be.there-FACT-HAB.EGO
 ‘(They say: we) are few.’ [TA23_talyu_clan 36 224]
- (799) tya ning tyû pê nyu byẽ tulu tuluqa gũ pêga we thîya nyu talyũ ali gê ìya
tya ning tyû pê nyu byẽ tulu tuluqa gũ pê-ga we
 now 1P up.there TOP FOC TOP FS a.place LOC separate-PL.I DEM
thîya nyu talyũ ali gê ì-ya
 mountain FOC a.clan like.this one be.there-ALLO
 ‘Now we, who left, who left Tulua, at the mountains there is Talyu (i.e. good numbers).’ [TA23_talyu_clan 37 225]
- (800) ashâ nyu mu lê ìya
ashâ nyu mu lê ì-ya
 plain.area FOC also growth be.there-ALLO
 ‘In the plains there is also growth.’ [TA23_talyu_clan 38 230a]

(801) I: hanûk we tshûna?

hanûk we tshû-na

where DEM thrive-FACT

‘Where did you thrive?’ [TA23_talyu_clan 38 230b]

(802) lê talyû ìya byě “kadé.nyu shyâna yûm tshûnam” ha-la we ahuga we mu (...)

lê talyû ì-ya byě kadé.nyu shyâ-na yûm

growth a.clan be.there-ALLO COND how.come happen-FACT NEG

tshû-na=m ha.la we ahu-ga we mu

thrive-FACT=NEG DUB DEM ask-PL.I DEM also

‘If there is growth (among) the Talyu, “how come it (the growth) has not happened, you have not thrived”, (we) asked (them) (...)’ [TA23_talyu_clan 38 230c]

(803) we cya làng we laga “ning byě” (...)

we cya làng we la-ga ning byě

DEM 3 COLL dem say-PL.I 1P TOP

‘(Then) they said: “as of us (...)”’ [TA23_talyu_clan 38 230d]

(804) bú.bú gũ cya làng we ali matyô

bú.bú gũ cya làng we ali matyô

earlier INS 3 COLL dem like.this maybe

‘Earlier it was like that.’ [TA23_talyu_clan 39 239]

(805) komjuri jina gũ webo hwûwga khûga weli gũ pûmru hî abli sêgakûbî khrâk

komjuri ji-na gũ webo hwûw-ga khû-ga weli gũ

quarrel(<IA) do-FACT NF and.then argue-PL.I fight-PL.I DEIC.DEM.I INS

pû-mru hî abli sê-ga-kû-bî khrâk

elder.brother-younger.brother EMPH some kill-PL.I-COMPL.II-PFV EVI

‘They had a fight and they argued and fought and some siblings killed (each other).’ [TA23_talyu_clan 40 241]

- (806) I: oh we cya làng
oh we cya làng
 oh DEM 3 COLL
 ‘Oh them!’ [TA23_talyu_clan 41 248a]
- (807) a.rūzang hî (...) webo a mîya bréi yüm abli gê tapügkûbî (...)
a.rūzang hî webo a mîya bréi yüm abli gê
 friend EMPH and.then child wife marry NEG CLF:some one
tapüg-kû-bî
 be.old-COMPL.II-PFV
 ‘(These) friends, never married, some became old (...)’ [TA23_talyu_clan 41 248b]
- (808) tapüg pêga
tapüg pê-ga
 old.man scatter-PL.I
 ‘The old people scattered.’ [TA23_talyu_clan 41 248c]
- (809) âm weli gũ ìya gũ mane tshû yüm mane
âm weli gũ ìya gũ mane tshû yüm mane
 yes DISC.DEIC.DEM INS be.there-ALLO NF FILL(<Asm) thrive NEG FILL(<Asm)
 ‘Yes, that’s why they are, I mean, they did not thrive, I mean.’ [TA23_talyu_clan 41 248d]
- (810) we tya günei we tya aqâyi tya ning aqâyĩ güne we ayĩ yûwak grûw ìga (...) tya (...)
we tya günei we tya aqâyi tya ning aqâyĩ güne we
 DEM now ABL DEM now grandchildren now 1P grandchildren ABL DEM
ayĩ yûwak grûw ì-ga tya
 grandchildren offspring multitude be.there-PL.I now
 ‘Now, currently (we have) grandchildren, we have many grandchildren (...) now (...)’ [TA23_talyu_clan 42 256a]

- (811) *tya tshû we tya dûw shyâ*
tya tshû we tya dûw shyâ
 now thrive DEM now many become
 ‘Now you thrive, now you have become many.’ [TA23_talyu_clan 42 256b]
- (812) *ning we ali khìgetyabô na marogagetyabô*
ning we ali khi-ge-tyabô na maro-ga-ge-tyabô
 1P DEM like.this reach-and.back-ahead INSIST talk-PL.I-and.back-ahead
 ‘We got there on our own, we talked to them on our own.’ [TA23_talyu_clan 43 264a]
- (813) *I: we bra tuluqa ìya tuluqa ìya tahû bônasâ we*
we bra tuluqa ì-ya tuluqa ì-ya tahû
 DEM CLF:generic a.place be.there-ALLO a.place be.there-ALLO group
bô-na-sâ we
 go-FACT-POT DEM
 ‘This one living at Tuluqa, the group living at Tuluqa, where might they have gone?’
 [TA23_talyu_clan 43 264b]
- (814) *we tuluqa ìya we làng we cya làng mamyú enyu ashâ nyu hanahonaring na*
we tuluqa ì-ya we làng we cya làng mamyú e=nyu
 DEM a.place be.there-ALLO DEM COLL DEM 3 COLL later DEM=FOC
ashâ nyu hana-ho-na-ring na
 plain.area FOC come-STAT-FACT-PL.II INSIST
 ‘(Oh,) those living at Tuluqa, those came to the planes later on, right?’
 [TA23_talyu_clan 44 272a]

(815) ezâ hanana

e = zâ hana-na

DEM = DIR come-FACT

‘They came.’ [TA23_talyu_clan 44 272b]

(816) âm (...) hanahoringna elulyang ìgana we làng

âm hana-ho-ring-na e = lulyang ì-ga-na we làng

yes come-STAT-PL.II-FACT DEM = a.place be.there-PL.II-FACT DEM COLL

‘Yes, they have come, they are living in Lulyang.’ [TA23_talyu_clan 44 272c]

(817) enyu zâ hanahoringna (...) tuluqa we kē ning we shyà mu ìna yüm lêatîkû lê-takûhoring mamýú bôbyě

e = nyu zâ hana-ho-ring-na tuluqa we kē ning we shyà mu

DEM = FOC DIR come-STAT-PL.II-FACT a.place DEM LOC.INS 1P DEM who also

ì-na yüm lê-atî-kû lê-ta-kû-ho-ring

be.there-FACT NEG migrate-leave-COMPL.II migrate-permanently-COMPL.II-STAT-PL.II

mamýú bôbyě

later TEMP.SUB

‘Yes, they have come, those in Tulua, we, everyone left, left permanently, later on.’ [TA23_talyu_clan 45 281a]

(818) oke

oke

ok(<E)

‘Alright.’ [TA23_talyu_clan 45 281b]

- (819) I: *accha tya tat hî taprûw lade mane taprûw bye maro nyu kya kadényu kya meho mane (...)?*

accha tya tat hî taprûw la-de mane taprûw bye maro
 I.see(<H) now FS EMPH dhole call-HAB.EGO FILL(<Asm) dhole TOP talk
nyu kya kadényu kya me-ho mane
 FOC Q how Q rear-STAT FILL(<Asm)

‘Alright, now this so-called *taprûw* (dhole), I mean, could the dhole talk or what? How? (were they) reared? I mean (...).’ [TA23_talyu_clan 46 290]

- (820) *taprûw a we maro nyu kho ji yûm mane ali nyu na taprûw a ali nyu la*

taprûw a we maro nyu kho ji yûm mane ali nyu
 dhole pup DEM talk FOC even.though do NEG FILL(<Asm) like.this FOC
na taprûw a ali nyu la
 INSIST dhole pup like.this FOC HSAY

‘However, the dhole pups could not speak. The dhole pups like this, like this, they say.’ [TA23_talyu_clan 47 297]

- (821) *taprûw malûngeya we kalê*

taprûw malûng-ge-ya we kalê
 dhole hunt-and.back-ALLO DEM xxx

‘The dholes would go and hunt.’ [TA23_talyu_clan 48 301a]

- (822) I: *mmh*

mmh

mmh

‘Mmh’ [TA23_talyu_clan 48 301b]

- (823) *kîya kaqîng kîya kasang kîya mlâ gũ taprûw cya we tabrê tyého bobyẽ (...) àng hanana gũ (...)*

kîya kaqîng kîya kasang kîya mlâ gũ taprûw
CLF:daytimes two CLF:daytimes three CLF:daytimes village LOC dhole

cya we tabrê tyé-ho bobyẽ àng hana-na gũ

3 DEM meat cut-STAT TEMP.SUB house come-FACT NF

‘After three four days in the village, when the dholes would cut meat (and) they would come home (...)’ [TA23_talyu_clan 48 301c]

- (824) *thâ.ming takhi we làng (...)*

thâ.ming takhi we làng

ration sack DEM COLL

‘These rice bags (...)’ [TA23_talyu_clan 49 311]

- (825) *nánga thâ.ming takhi tyüga la*

náng-ya thâ.ming takhi tyü-ga la

smell-ALLO ration sack snatch-PL.I HSAY

‘They would smell (them) (and) snatch the rice bags, they say.’ [TA23_talyu_clan 50 313a]

- (826) I: mmh

mmh

mmh

‘Mmh’ [TA23_talyu_clan 50 313b]

- (827) *webo gũ mề we we kasâ xxx oh ewe (...)*

webo gũ mề we we kasâ xxx oh ewe

and.then INS person DEM DEM know xxx oh DEM

‘And then people knew, they (...)’ [TA23_talyu_clan 50 313c]

- (828) *thâ.ming shìya dũwa dyüw delé*
thâ.ming shì-ya dũw-ya dyüw delé
 ration take-ALLO need-ALLO isn't.it it.seems
 'They had to take the rice bags, it seems.' [TA23_talyu_clan 51 318a]
- (829) I: *âm s thâ.ming shì*
âm s thâ.ming shì
 yes FS ration take
 'Yes, they would take rice bags.' [TA23_talyu_clan 51 318b]
- (830) *thâya thâ.ming shìya dyi dũwga dyâ delé la we wegũ kasâ la*
thâ-ya thâ.ming shì-ya dyi dũw-ga dyâ delé la we wegũ
 eat-ALLO ration take-ALLO FS need-PL.I far it.seems HSAY DEM and.then
kasâ la
 know HSAY
 'They had to take the rice bags far away, it seems, they say. And then they knew.'
 [TA23_talyu_clan 51 318c]
- (831) *webo tabrê we daráng mesêho bô ali tabri.tagra daráng hî (...)*
webo tabrê we daráng mesê-ho bô ali tabri.tagra daráng
 and.then meat DEM big catch-STAT COND.ALLO like.this big.hunt big
hî
 EMPH
 'And then, when (they) would catch big meat, like this, a big catch (...)'
 [TA23_talyu_clan 52 325a]
- (832) I: *aha*
aha
 aha
 'Aha' [TA23_talyu_clan 52 325b]

(833) *webo we tyü-pwi khrâk*
webo we tyü-pwi khrâk
 and.then DEM snatch-HAB EVI
 ‘And then, (they) allowed (them) to snatch.’ [TA23_talyu_clan 52 325c]

(834) I: oh!
oh
oh
 ‘Oh!’ [TA23_talyu_clan 53 330a]

(835) *náng tyüwa*
náng tyü-ga
 smell snatch-PL.I
 ‘They would smell and snatch it.’ [TA23_talyu_clan 53 330b]

(836) I: *we daráng bô daráng hî*
we daráng bô daráng hî
 DEM big TEMP.SUB big EMPH
 ‘When there was (a) big (rice bag), it was big.’ [TA23_talyu_clan 53 330c]

(837) *âm (...) weli cya taprûw a we nyu marone ji yüm we weli mé kam (...) jiya la*
âm weli cya taprûw a we nyu maro-ne ji yüm we
 yes DISC.DEIC.DEM 3 dhole pup DEM FOC speak-INT do NEG DEM
weli mé kam ji-ya la
 DISC.DEIC.DEM all work do-ALLO HSAY
 ‘Yes (...) even though the dhole pups could not speak, they (...) were all useful
 (lit. did work), they say.’ [TA23_talyu_clan 53 330d]

- (838) *weli kahani ded adigabô mane ning mât bà gũ*
weli kahani ded adi-ga-bô mane ning mât
 DISC.DEIC.DEM story(<IA) FS tell-PL.I-PFV.ALLO FILL(<Asm) 1P mother
bà gũ
 father INS
 ‘This story they told (us), our fathers and mothers.’ [TA23_talyu_clan 54 338]
- (839) *we tya ane jenereshyon kadâge shyâ we kho*
we tya ane jenereshyon kadâge shyâ we kho
 DEM now 2P generation(<E) how.many happen DEM even.though
 ‘However, how many generations have happened?’ [TA23_talyu_clan 55 344a]
- (840) *we hã we weli kasâdyüm jenereshyon asũng we nyu byẽ*
we hã we weli kasâ-d=yüm jenereshyon asũng
 DEM 1S DEM DISC.DEIC.DEM know-HAB.EGO=NEG generation(<E) generation
we nyu byẽ
 DEM FOC TEMP.SUB
 ‘As of this generation, I do not know (how many).’ [TA23_talyu_clan 55 344b]
- (841) *asũng kadê halâng gê shyâ we nyu kasâdyüm*
asũng kadê halâng gê shyâ we nyu kasâ-d=yüm
 generation how.many CLF:generation one happen DEM FOC know-HAB.EGO=NEG
 ‘How many generations, one generation happen, I do not know.’ [TA23_talyu_clan 56 349]

(842) *kadê asûng we nyu shyânazâ we nyu kho we ning (...)*
kadê asûng we nyu shyâ-na-zâ we nyu kho
 how.many generation DEM FOC happen-FACT-DIR DEM FOC even.though
we ning
 DEM 1P
 ‘How many generations have happened (i.e. I should know), we (...)’ [TA23_talyu_clan 60 365]

(843) *ning ji bü nama nabà we ahuhona yüm na (...) weli we làng we (...)*
ning ji bü na-ma na-bà we ahu-ho-na yüm na
 1P EMPH time 1.POSS-mother 1.POSS-father DEM ask-STAT-FACT NEG INSIST
weli we làng we
 DISC.DEIC.DEM DEM COLL DEM
 ‘We didn’t ask our father and mother, right (...) they (...)’ [TA23_talyu_clan 61 368a]

(844) *I: âm weli (...)*
âm weli
 yes DISC.DEIC.DEM
 ‘Yes, this (...)’ [TA23_talyu_clan 61 368b]

(845) *tulu ning tuluqa gũ lê pêga we asûng kadê shyâ (...)*
tulu ning tuluqa gũ lê pê-ga we asûng kadê
 FS 1P a.place LOC migrate separate-PL.I DEM generation how.many
shyâ
 happen
 ‘Tulu (...) (since) we separated from Tulua, how many generations (...)’
 [TA23_talyu_clan 61 368c]

(846) we cuwe ning kas ahuhona yüm

we cuwe ning kas ahu-ho-na yüm

DEM DISC.DEIC.DEM 1P FS ask-STAT-FACT NEG

‘This, we never, eh, asked.’ [TA23_talyu_clan 61 368d]

(847) tuluqa we macîk matyô

tuluqa we macîk matyô

a.place DEM water probably

‘This Tulua is a river, right?’ [TA23_talyu_clan 62 378a]

(848) ã we tuluqa we macîk we mlâ hî tapwi (...) we gũ we cuwe gũ lê pêngazâ ning

ã we tuluqa we macîk we mlâ hî tapwi we gũ we

yes DEM a.place DEM water DEM village EMPH surely DEM INS DEM

cuwe gũ lê pê-ga-zâ ning

DISC.DEIC.DEM INS migrate separate-PL.I-DIR 1P

‘Yes, this Tulua is a river (and) it is also a village. From there we migrated and separated.’ [TA23_talyu_clan 62 378b]

(849) we ning amu tya ali a lûmkû talyũ tapwi lûmkû tya talyũ bü lade we

we ning amu tya ali a lûmkû talyũ tapwi lûmkû tya talyũ

DEM 1P like.this now like.this FS inside a.clan ADD inside now a.clan

bü la-de we

once call-HAB.EGO DEM

‘We, like this, within the Talyu clan, we also call (ourselves) Talyu.’ [TA23_talyu_clan 63 386]

(850) ah talyũ karâ la-ga-de we karâ mlă nyu ì-na-ge-ya (...) karâ phlă karâ matyô we
cuwe karâ laga

ah talyũ karâ la-ga-de we karâ mlă nyu

ah a.clan a.clan call-PL.I-HAB.EGO DEM a.clan village FOC

ì-na-ge-ya karâ phlă karâ matyô we

be.there-FACT-and.back-ALLO a.name stone a.name maybe DEM

cuwe karâ la-ga

DISC.DEIC.DEM a.name call-PL.I

‘Ah, we call (ourselves) Talyu Kara, we lived in this Kara village. Kara seems to
be a stony place, we call it Kara.’ [TA23_talyu_clan 64 391a]

13 Appendix II : Classifiers

This appendix includes a list of all classifiers found so far in the language, which amount to approximately one hundred. The morpho-syntax and semantics of classifiers has been discussed in §5.1. It should be noted the the Lexicon (§11) also lists the classifiers presented here.

Some classifiers have more than one meaning and I have included a small number (e.g.¹) to tell them apart. Some classifiers are better understood in terms of the dimensions they refer to, e.g. *bra* ‘3-D small round things (e.g. beads)’ (Grinevald 2003, 97).

Tables 13.1, 13.2 and 13.3 include sortal classifiers. Table 13.4 includes mensural classifiers of measurements and distances. Tables 13.5 and 13.6 include mensural classifiers of volume. Finally, Table 13.7 includes mensural classifiers of weight. All tables include information on the origin of the classifier (if known) and some include etymological notes. Data from Karbi comes from Konnerth (2014), Dai Chin from So-Hartmann (2009), Proto Palaung from Sidwell 2015. Data from Kaman and Kera’a are from my field notes.

Table 13.1: Sortal classifiers

Classifier	Meaning	Origin (if known)	Notes
<i>abûw</i>	tied bundles (e.g. firewood)		
<i>ayüw</i>	chicken nests, basket-full of eggs	<i>ayüw</i> ‘man-made chicken’s nest’	
<i>asûng</i>	generations	<i>asûng</i> ‘generation’	
<i>bá</i>	1-D shallow, flat and slow parts of a river that enable crossing	cf. verb <i>bá</i> ‘flow (of liquids)’	Kaman <i>ba</i>
<i>bra</i> ¹	3-D small round things (e.g. beads)		Kera’a <i>bra</i> ‘CLF:small round’ (not a generic class.)
<i>bra</i> ²	generic (e.g. vegetable, humans)		may replace other sortal classifiers. Cf. Kaman <i>brat</i>
<i>brã</i>	1-D long and thin things		
<i>brǎ</i>	ways or roads to reach a place (including rivers and tributaries)		
<i>brík/brúw</i>	bamboo-shaped things e.g. bananas, fingers (also people for some speakers)		Kaman <i>brü</i>
<i>bű</i>	wooden logs and the like (including human feces)		
<i>bûm</i>	stems or roots	<i>bûm</i> ‘root, stem’	
<i>bü</i>	time(s), occurrence		numeral ‘one’ is often omitted

Table 13.2: Sortal classifiers, cont.

Classifier	Meaning	Origin (if known)	Notes
<i>dũ/dõ</i> ¹	large man-made structures e.g. houses and granaries		
<i>dũ(-dubâ)</i> ²	times a house has been refurbished		
<i>hadûm</i> ¹	group of plants and trees, thickets of bamboo		
<i>hadûm</i> ²	by extension: group of people of the same gender (e.g. group same-gender siblings)		
<i>halá</i>	2-D stackables		Kaman (<i>ha</i>) <i>lap</i>
<i>hazúm</i>	bunches that can be tied or skewered together		
<i>jorâ/jurâ</i>	pairs (e.g. sandals, oxen)	Asm. <i>jora</i> ‘pair’	
<i>kâ</i>	halves (of fruits)		
<i>kahã</i>	nodes, joints (e.g. in a bamboo)	<i>kahã</i> ‘nodes, joint’	
<i>kapyek</i>	handful (of seeds)		
<i>ká</i>	steps (in ladder)		Kaman <i>gal</i>
<i>khũ</i>	bits (of food)		
<i>na</i>	2-D foldables	<i>na</i> ‘leaf’	PTB <i>*s-nas</i> ‘leaf’ (Matisoff 2003, 432)
<i>palai</i>	small plots of land usually a garden		
<i>pe</i>	rooms in a house		
<i>pi</i>	foldings		
<i>pla</i>	flat objects (like swords or planks), also jewel and money		Cf. Proto Palaungic: <i>*plaa?</i> ‘blade’. Cf. Kaman <i>fla</i>

Table 13.3: Sortal classifiers, cont.

Classifier	Meaning	Origin (if known)	Notes
<i>pum</i>	3-D round objects like fruits, birds and cars		Cf. Karbi <i>púm</i> CLF:round Daai Chin <i>puum</i> CLF:3-D.big, cf. Kaman <i>pop</i> (note it does not include cars).
<i>pha</i> <i>phũ</i>	bunch (of bananas, ginger) roots of trees/plants, also yams, also nr. of clans in a congregation	<i>phũ</i> ‘root’	Kaman <i>fogn</i>
<i>ráng</i> <i>ráp</i>	stalks of trees and plants streams and tributaries of a river		
<i>tahũ</i> <i>takla</i>	groups of people steps	<i>tahũ</i> ‘group’	
<i>talá</i> <i>taplá</i>	coins of gold and silver spoonfuls	<i>taplá</i> ‘spoon’	Kaman <i>tala</i>
<i>tarap</i> <i>tarum</i>	herds (of animals) areas for grazing		Kaman <i>that</i>
<i>tarüim</i> <i>tũ</i>	flocks (of sheep) halves (generic)	<i>tarüim</i> ‘flock’	
<i>tyá</i>	times, shots by a gun		E.g. used as <i>tyá gê</i> ‘once’
<i>tyáng</i> ¹ <i>tyáng</i> ²	threads plots of land	thread of wool	Kaman <i>taploq</i>
<i>tyàng</i> <i>yüw</i>	bodies of large animals (in- cluding people) side (of face)	<i>tyàng</i> ‘body’	

Table 13.4: Mensural classifiers: measurements and distances

Classifier	Meaning	Origin (if known)	Notes
<i>ahô</i>	circumference around arms e.g. around a tree		
<i>asũ.grũ</i>	from the chest to the middle finger		
<i>atye-sei</i>	small finger (breadth)		
<i>dũthê</i>	distance from thumb to point- ing finger		
<i>habrâng</i>	finger (breadth)	<i>habrâng</i> ‘finger’	
<i>hatyâng</i>	thumb (breadth)	<i>hatyâng</i> ‘thumb’	also used to measure amount of grease in an animal
<i>húm</i>	cubit (elbow to fingertips)		standard measure- ment for size of mithuns
<i>húm-atyú</i>	elbow to fist		
<i>hũga</i>	distance from thumb to mid- dle finger		
<i>katyê</i>	small amounts	<i>katyê</i> ‘small, a bit’	
<i>labũya</i>	four fingers (breadth)		
<i>lyâng</i>	fathom		
<i>mẽ</i>	kilometers		origin unclear
<i>myã</i>	one palm (breadth)		standard measure- ment for pigs and horns
<i>tũning</i>	two fingers (breadth)		
<i>tũsang</i>	three fingers (breadth)		

Table 13.5: Mensural classifiers: Volume

Classifier	Meaning	Origin (if known)	Notes
<i>bũ</i>	handful		
<i>dẽ-katã</i>	water level up to the head while on tiptoes	<i>dẽ-katã</i> ‘stand on tiptoes’	
<i>garêi</i>	car load (e.g. of bananas)	<i>garêi</i> ‘car’ (n) (<IA, cf. Hindi <i>gāḍī</i>)	
<i>glyã</i>	bamboo container (of liquid)		
<i>grũ</i>	clenched fist	cf. <i>grũ</i> ‘hand, leg, arm’	
<i>grũ-bũlũm-</i>	water level up to the feet	lit. to the eye of the foot	
<i>hapyê</i>			
<i>grũ-shya-</i>	water level up to the calf		
<i>hapyê</i>			
<i>habrũk</i>	two handfuls		
<i>ka</i>	mouthful, a bit (of food)	<i>ka</i> ‘feed child (by chewing), prechew’ (v)	
<i>kapyêk</i>	third of a clenched fist		
<i>klũb</i>	sips		
<i>khalyâng-</i>	water level up to the shoulders		
<i>pà-hapyê</i>			
<i>nakapĩk-</i>	water level up to the chest		
<i>hapyê</i>			
<i>páp</i>	3-D pots of food		

Table 13.6: Mensural classifiers: Volume, cont.

Classifier	Meaning	Origin (if known)	Notes
<i>phâbum-hapyê</i>	water level up to the knees		
<i>phlû</i>	kettlefuls (of liquid)	<i>phlû</i> ‘kettle’	
<i>tapûng</i>	half clenched fist		
<i>thû</i>	pinch, drop		
<i>thûkrü-hapyê</i>	water level up to the hips	<i>thûkrü</i> ‘hips’	
<i>weng/</i> <i>wûng/ wing</i>	cup (of liquid)		

Table 13.7: Mensural classifiers: Weight

Classifier	Meaning	Origin (if known)	Notes
<i>ana</i>	weight of 6.25 <i>paisa</i> (coins)		
<i>dyûm-hya</i>	weight of 12.5 <i>paisa</i> (coins)		two <i>ana</i>
<i>dülei</i>	weight of 50 <i>paisa</i> (coins)		
<i>kaqîng</i>	weight of 2 <i>tûla</i>	<i>kaqîng</i> ‘two’	
<i>misha</i>	weight of a coin of 25 <i>paisa</i>		Kaman <i>musha</i>
<i>pîsa</i>	weight of 1 <i>paisa</i>	IA <i>paisa</i>	
<i>püwûng</i>	weight of a silver coin		
<i>tó</i>	one kilo		
<i>tûla</i>	weight of 1 <i>rupee</i> or 100 <i>paisa</i>		

Table 13.8: Mensural classifiers: Time

Classifier	Meaning	Origin (if known)	Notes
<i>baji</i>	hours	IA <i>baje</i> ‘hour’	synonym with <i>gânta</i> (also IA)
<i>gânta</i>	hours	IA <i>ghaṇṭā</i> ‘hour, clock’	
<i>halo</i>	month	<i>halo</i> ‘moon’	
<i>kahîng</i>	daytimes (not 24h)	<i>hîng</i> ‘day’	
<i>kalyâng</i>	evenings	<i>lyâng</i> ‘half’ (i.e. ‘half (of the day)’)’	
<i>kanáng</i>	years	<i>náng</i> ‘year’	
<i>kîya</i>	nights, periods of 24h		

14 References

- Abbi, A. (2018). Echo formations and expressives in South Asian languages. In: *Non-prototypical reduplication*. Ed. by A. Urdze. Berlin/Boston: De Gruyter Mouton, pp. 1–34.
- Adams, K. L. and N. F. Conklin (1973). Toward a theory of natural classification. In: *Papers from the 9th Regional Meeting of the Chicago Linguistic Society*. Ed. by C. Corum, T. C. Smith-Stark, and A. Weiser. Vol. 9. Chicago Linguistic Society, pp. 1–11.
- Adams, K. L. (1989). *Systems of numeral classification in the Mon-Khmer, Nicobarese and Asian subfamilies of Austroasiatic*. Department of Linguistics, Research School of Pacific Studies, Australian National University.
- Aikhenvald, A. Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford University Press.
- (2003). Typological parameters for the study of clitics, with special reference to Tariana. In: *Word: A Cross-linguistic Typology*. Ed. by R. M. W. Dixon and A. Y. Aikhenvald. Cambridge University Press, pp. 42–78.
 - (2006a). *Evidentiality*. Oxford University Press.
 - (2006b). Serial verb constructions in typological perspective. In: *Serial verb constructions: A cross-linguistic typology*. Oxford University Press, pp. 1–68.
 - (2007). Typological distinctions in word-formation. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 3. Cambridge University Press, pp. 1–65.
 - (2012a). ‘We can’t say it with one word’: multiverb constructions. In: *The Languages of the Amazon*. Oxford University Press, pp. 304–325.
 - (2012b). Possession and ownership: a cross linguistic perspective. In: *Possession and Ownership*. Oxford University Press.

- Aikhenvald, A. Y. (2023). A medial clause does it all: Coherence, continuity, and addressee involvement in Manambu. In: *Celebrating Indigenous Voice: Legends and Narratives in Languages of the Tropics and Beyond*. Ed. by A. Y. Aikhenvald et al. Berlin/Boston: De Gruyter Mouton, pp. 73–94.
- Akter, M. Z. (2022). A comprehensive description of Pangkhua: An endangered Tibeto-Burman language of Bangladesh. PhD thesis. The University of Sydney.
- Ama, A. (2024). *Talo/Tülu: Taraon/Kaman Death Ritual: The Journey to the next World*. Self-published.
- Andrews, A. D. (2007a). The major functions of the noun phrase. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 1. Cambridge University Press, pp. 132–223.
- (2007b). Relative clauses. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 2. Cambridge University Press, pp. 206–236.
- Andvik, E. (2010). *A Grammar of Tshangla*. Leiden: Brill.
- Bashir, E. (2016). Evidentiality and mirativity in Iranian, Nuristani, Indo-Aryan, Burushaski, and Dravidian. In: *The Languages and Linguistics of South Asia: A Comprehensive Guide*. Ed. by H. H. Hock and E. Bashir. Berlin/Boston: De Gruyter Mouton, pp. 584–589.
- Bashir, E. and T. J. Connors (2019). *A Descriptive Grammar of Hindko, Panjabi, and Saraiki*. Berlin/Boston: De Gruyter Mouton.
- Becker, A. L. (1975). A linguistic image of nature: The Burmese numerative classifier System. In: *Linguistics* 165, pp. 109–121.
- Becker, L. (2021). *Articles in the World's Languages*. Berlin/Boston: De Gruyter.
- Beckwith, C. I. (1993). Class nouns and classifiers in Thai. In: *Papers from the Third Annual Meeting of the Southeast Asian Linguistics Society*, pp. 11–26.
- Belletti, A. and L. Rizzi (1988). Psych-verbs and θ -theory. In: *Natural Language & Linguistic Theory*, pp. 291–352.
- Bellwood, P. (1993). Southeast Asia before History. In: *The Cambridge History of Southeast Asia*. Ed. by N. Tarling. Cambridge University Press, pp. 51–136.
- Bellwood, P. (2021). Homelands and dispersal histories of Mainland Southeast Asian language families. In: *The Languages and Linguistics of Mainland Southeast Asia: A com-*

- prehensive guide*. Ed. by P. Sidwell and M. Jenny. Berlin/Boston: De Gruyter Mouton, pp. 33–45.
- Benedict, P. (1992). Austroasiatic loanwords in Sino-Tibetan. In: *Mon-Khmer Studies* 18–19.
- Benedict, P. K. (1967). Austro-Thai Studies 3: Austro-Thai and Chinese. In: *Behavior Science Notes*, 2, no. 4.
- (1972). *Sino-Tibetan: A Conspectus*. Princeton/Cambridge Studies in Chinese Linguistics. Cambridge University Press.
- Bentley, D. (2016). Copular and existential constructions. In: *Manual of Romance Morphosyntax and Syntax*. Ed. by A. Dufter and E. Stark. Berlin/Boston: De Gruyter, pp. 332–366.
- Berghäll, L. (2015). *A Grammar of Mauwake*. Berlin: Language Science Press.
- Bergqvist, H. and S. Kittilä (2019). Epistemic perspectives: Evidentiality, egophoricity, and engagement. In: *Evidentiality, egophoricity and engagement*. Ed. by H. Bergqvist and S. Kittilä. Berlin: Language Science Press, pp. 1–22.
- Berlin, B. (1992). *Ethnobiological classification*. Princeton University Press.
- Bhatt, R. (2016). Minimalist approaches to South Asian syntax. In: *The Languages and Linguistics of South Asia: A Comprehensive Guide*. Ed. by H. H. Hock and E. Bashir. Berlin/Boston: De Gruyter Mouton, pp. 506–529.
- Bickel, B. (1996). *Aspect, mood, and time in Belhare: studies in the semantics-pragmatics interface of a Himalayan language*. Zürich: ASAS.
- (1998). Review article: Converbs in cross-linguistic perspective. In: *Linguistic Typology* 2(3), pp. 381–397.
- (1999a). Nominalization and focus constructions in some Kiranti languages. In: *Topics in Nepalese Linguistics*. Ed. by Y. P. Yadava and W. G. Glover. Royal Nepal Academy, pp. 271–296.
- (1999b). *Principles of event framing: genetic stability in grammar and discourse*. Available at: <http://https://dr.ntu.edu.sg/handle/10356/89999?mode=simple/>.
- Bickel, B. (1999c). From ergativus absolutus to topic marking in Kiranti: a typological perspective. In: *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 25, pp. 38–49.

- (2000a). Deictic transposition and referential practice in Belhare. In: *Journal of Linguistic Anthropology* 10, pp. 224–247.
 - (2000b). On the syntax of agreement in Tibeto-Burman. In: *Studies in Language* 24, pp. 583–609.
 - (2003). Referential density in discourse and syntactic typology. In: *Language* 79.4, pp. 708–736.
 - (2011). Grammatical relations typology. In: *The Oxford Handbook of Language Typology*. Ed. by J. J. Song. Oxford University Press, pp. 399–444.
- Bickel, B. and J. Nichols (2007). Inflectional morphology. In: *Language typology and syntactic description*. Ed. by T. Shopen. 2nd ed. Vol. 3. Cambridge University Press, pp. 169–240.
- (2020). Linguistic Typology and Hunter-Gatherer Languages. In: *The Language of Hunter-Gatherers*. Ed. by T. Güldemann, P. McConvell, and R. A. Rhodes. Cambridge University Press, pp. 67–75.
- Bickel, B. and F. Zúñiga (2017). The ‘word’ in polysynthetic languages: phonological and syntactic challenges. In: *The Oxford Handbook of Polysynthesis*. Oxford University Press, pp. 158–185.
- Bisang, W. (1993). Classifiers, quantifiers and class nouns in Hmong. In: *Studies in Language. International Journal sponsored by the Foundation “Foundations of Language”* 17.1, pp. 1–51.
- (1999). Classifiers in East and Southeast Asian languages: Counting and beyond. In: *Numeral Types and Changes Worldwide*. Ed. by J. Gvozdanović. Berlin/New York: De Gruyter Mouton, pp. 113–186.
 - (2017). Classification between Grammar and Culture: a Cross-Linguistic Perspective. In: *Classification from Antiquity to Modern Times: Sources, Methods, and Theories from an Interdisciplinary Perspective*. Ed. by T. Pommerening and W. Bisang. Berlin/Boston: De Gruyter, pp. 199–230.
- Bisang, W. (2023). Word Classes in Classical Chinese. In: *The Oxford Handbook of Word Classes*. Oxford University Press.
- Blench, R. (2015). Proposal for a writing system for Tawra, a language of Arunachal Pradesh. Unpublished.

- (2017). *A dictionary of Tawra, a language of Arunachal Pradesh*. URL: <https://rogerblench.info/Language/NEI/Mishmi/Tawra/Tawra%20Dictionary.pdf>.
 - (n.d.). The classification of Meyor (Zakhring). Available at: https://www.academia.edu/15754880/The_classification_of_Meyor_Zakhring.
- Blench, R. and S. Kri (n.d.). *A dictionary of Kman [= Miju], a language of Arunachal Pradesh*. URL: <http://www.rogerblench.info/Language/NEI/Isolates/Kman/Ling/Kman%20dictionary%202019.pdf>.
- Blench, R. and M. W. Post (2014). Rethinking Sino-Tibetan phylogeny from the perspective of North East Indian languages. In: *Trans-Himalayan Linguistics - Historical and Descriptive Linguistics of the Himalayan Area*. Ed. by T. Owen-Smith and N. Hill. Berlin/Boston: De Gruyter Mouton, pp. 71–104.
- Bodt, T. A. (2020). *Grammar of Duhumbi (Chugpa)*. Leiden/Boston: Brill.
- Boersma, P. and D. Weenink (2024). *Praat: doing phonetics by computer [Computer program: Version 6.4.07]*.
- Booij, G. (2000). Inflection and derivation. In: *Encyclopedia of Language & Linguistics, Second Edition*. Ed. by K. Brown. Oxford: Elsevier, pp. 654–661.
- (2005). *The Grammar of Words: An Introduction to Linguistic Morphology*. Oxford University Press.
- Boro, K. (2017). *A Grammar of Hakhun Tangsa*. PhD thesis. University of Oregon.
- Bossong, G. (2021). DOM and linguistic typology: A personal view. In: *Differential Object Marking in Romance: The Third Wave*. Ed. by J. Kabatek, P. Obrist, and A. Wall. Berlin/Boston: De Gruyter, pp. 21–36.
- Bourdin, P. (2008). On the grammaticalization of ‘come’ and ‘go’ into markers of textual connectivity. In: *Rethinking Grammaticalization: New perspectives*. Ed. by M. J. López-Couso and E. Seoane. Amsterdam/Philadelphia: John Benjamins, pp. 37–59.
- Bradley, D. (2017). Space in Lisu. In: *Himalayan Linguistics* 16.1, pp. 1–22.
- Bradley, D. (2021). Typological profile of Burmic languages. In: *The Languages and Linguistics of Mainland Southeast Asia: A comprehensive guide*. Ed. by P. Sidwell and M. Jenny. Berlin/Boston: De Gruyter Mouton, pp. 299–336.

- Bradley, D. and P. Mohanty (2023). Sociolinguistics of South Asia Tibeto-Burman, Austroasiatic and other languages. In: *The Routledge Handbook of Sociolinguistics Around the World*. Second Edition. London/New York: Routledge, pp. 184–197.
- Breugel, S. van (2014). *A Grammar of Atong*. Leiden: Brill.
- Brinton, L. J. (1988). *The development of English aspectual systems: Aspectualizers and post-verbal particles*. Cambridge University Press.
- Brown, C. H. (2020). Ethnobiology and the Hunter-Gatherer/Food Producer Divide. In: *The Language of Hunter-Gatherers*. Ed. by T. Güldemann, P. McConvell, and R. A. Rhodes. Cambridge University Press, pp. 76–88.
- Bufan, H. and D. Qingxia (1992). *Zangmiananyuzu yuyan cihui [A Tibeto-Burman Lexicon]*. Beijing: Central Institute of Minorities. Accessed via STEDT database.
- Burling, R. (2003). *The language of the Modhupur Mandi, Garo : Vol. I : Grammar*. The Scholarly Publishing Office, The University of Michigan, University Library.
- Butler, B. (2015). Approaching a phonological understanding of the sesquisyllable with phonetic evidence from Khmer and Bunong. In: *Languages of Mainland Southeast Asia: The State of the Art*. Ed. by N. J. Enfield and B. Comrie. Berlin/München/Boston: De Gruyter Mouton, pp. 443–499.
- Button, C. (2011). *Proto Northern Chin*. University of California.
- Bybee, J. L. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.
- Campbell, L. (1998). *Historical linguistics*. Edinburgh University Press.
- (2000). What’s wrong with grammaticalization? In: *Language sciences* 23.2-3, pp. 113–161.
- Castañó, G. G. (2019). On the existence of egophoricity across clause types in Totoró Namtrik. In: *Evidentiality, egophoricity and engagement*. Ed. by H. Bergqvist and S. Kittilä. Berlin: Language Science Press, pp. 165–196.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: *Subject and Topic*. Ed. by C. N. Li. New York: Academic Press, pp. 25–56.
- Chafe, W. (1994). *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press.

- Chafe, W. and J. Nichols (1986). Introduction. In: *Evidentiality: The Linguistic Coding of Epistemology*. Ed. by J. N. Wallace Chafe. Norwood, New Jersey: Ablex Publishing, pp. vii–xi.
- Chamoreau, C. (2021). Restrictive relative constructions in Pesh: A predominantly internally-headed relative clause language. In: *Relative Clause Structure in Mesoamerican Languages*. Ed. by E. L. Palancar, R. Z. Maldonado, and C. Chamoreau. Leiden: Brill, pp. 291–319.
- Chao, Y. R. (2011[1968]). *A Grammar of Spoken Chinese*. Beijing: Commercial Press.
- Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge University Press.
- (1982). Grammatical relations in Huichol. In: *Syntax and semantics vol 15: Studies in Transitivity*. Ed. by P. J. Hopper and S. A. Thompson, pp. 95–116.
 - (1989). *Language Universals and Linguistic Typology: Syntax and Morphology*. Second Edition. The University of Chicago Press.
 - (2013). *Alignment of Case Marking of Full Noun Phrases*. Ed. by M. S. Dryer and M. Haspelmath. WALS Online. URL: <https://wals.info/chapter/98>.
- Comrie, B., D. Forker, and Z. Khalilova (2012). Adverbial clauses in the Tsezic languages. In: *Clause Linkage in Cross-Linguistic Perspective: Data-Driven Approaches to Cross-Clausal Syntax*. Ed. by V. Gast and H. Diessel. Berlin/Boston: De Gruyter Mouton, pp. 157–190.
- Corbett, G. G. (2006). *Agreement*. Cambridge University Press.
- Coupe, A. R. (2008). *A Grammar of Mongsen Ao*. Berlin: Mouton De Gruyter.
- (2017). On the diachronic origins of converbs in Tibeto-Burman and beyond. In: *Socio-historical Linguistics in Southeast Asia: New Horizons for Tibeto-Burman Studies in honor of David Bradley*. Leiden: Brill, pp. 211–237.
 - (2018). South Asian perspectives on relative-correlative constructions. In: *The 51st International Conference on Sino-Tibetan Languages and Linguistics*, pp. 1–13.
 - (2020). Northern Sangtam phonetics, phonology and word list. In: *Linguistics of the Tibeto-Burman Area* 43(1), pp. 147–189.
- Craig, C. G. (1992). Classifiers in a Functional Perspective. In: *Layered Structure and Reference in a Functional Perspective: Papers from the Functional Grammar Conference, Copen-*

- hagen. Ed. by M. Fortescue, P. Harder, and L. Kristoffersen. Amsterdam/Philadelphia: John Benjamins Publishing Co., pp. 277–301.
- Croft, W. (2023). Word Classes in Radical Construction Grammar. In: *The Oxford Handbook of Word Classes*. Ed. by E. van Lier. Oxford University Press, pp. 213–230.
- Culhane, K., N. Peck, and U. Reinöhl (n.d.). The mystery of word-initial consonant loss in Kera'a. Unpublished manuscript.
- Van Dam, K. P. (2019). The tone system of Tangsa-Nocte and related Northern Naga varieties. PhD thesis. La Trobe University.
- Daudey, G. (2014). A Grammar of Wadu Pumi. PhD thesis. La Trobe University.
- David, A. B. (2015). *Descriptive Grammar of Bangla*. Berlin/Boston: Mouton De Gruyter.
- Davies, W. D. (1986). *Choctaw verb agreement and universal grammar*. Dordrecht/Boston/Lancaster/Tokyo: D. Reidel.
- De Swart, H. E. (2012). Verbal aspect across languages. In: *The Oxford Handbook of tense and Aspect*. Ed. by R. I. Binnick. Oxford University Press, pp. 752–780.
- DeLancey, S. (1986a). Evidentiality and volitionality in Tibetan. In: *Evidentiality: The Linguistic Coding of Epistemology*. Ed. by W. Chafe and J. Nichols. Norwood, New Jersey: Ablex, pp. 203–213.
- (1986b). Toward a history of Tai classifier systems. In: *Noun classes and categorization*. Ed. by C. G. Craig. Amsterdam/Philadelphia: John Benjamins, pp. 437–452.
 - (1990). Ergativity and the cognitive model of event structure in Lhasa Tibetan. In: *Cognitive Linguistics* 1.3, pp. 289–322.
 - (1991). The origins of verb serialization in Modern Tibetan. In: *Studies in Language* 15(1), pp. 1–23.
 - (1997). Grammaticalization and the Gradience of Categories. In: *Essays on Language Function and Language Type* 51, pp. 51–69.
 - (2002). Relativization and Nominalization in Bodic. In: *Proceedings of the Twenty-Eighth Annual Meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics*, pp. 55–72.
- DeLancey, S. (2011). Notes on verb agreement prefixes in Tibeto-Burman. In: *Himalayan Linguistics* 10.1, pp. 1–29.

- (2015). The origins of postverbal negation in Kuki-Chin. In: *North East Indian Linguistics* 7, pp. 203–212.
 - (2019). Non-singular pronouns in Tibeto-Burman (Trans-Himalayan). In: *Linguistics of the Tibeto-Burman Area* 42.1, pp. 110–136.
- Denny, J. P. (1979). The ‘Extendedness’ Variable in Classifier Semantics: Universal Features and Cultural Variation. In: *Ethnolinguistics: Boas, Sapir and Whorf Revisited*. Ed. by M. Mathiot. Berlin/Boston: De Gruyter Mouton, pp. 97–119.
- Deutscher, G. (2009). Nominalization and the origin of subordination. In: *Syntactic Complexity: Diachrony, acquisition, neuro-cognition, evolution*. Ed. by T. Givón and M. Shibatani. Amsterdam: John Benjamins, pp. 199–214.
- Diessel, H. (1999). *Demonstratives: Form, Function and Grammaticalization*. Amsterdam/Philadelphia: John Benjamins.
- (2015). Usage-based construction grammar. In: *Handbook of Cognitive Linguistics*, pp. 295–321.
- Diffloth, G. (1979). The Wa languages. In: *Linguistics of the Tibeto-Burman Area* 5.2, pp. 1–182.
- (2005). The contribution of linguistic palaeontology to the homeland of Austroasiatic. In: *The Peopling of East Asia: Putting Together the Archaeology, Linguistics and Genetics*. Ed. by R. B. L. Sagart and A. Sanchez-Mazas. London: Routledge Curzon, pp. 77–80.
 - (2008). Shafer’s ‘parallels’ between Khasi and Sino-Tibetan. In: *North East Indian Linguistics*. Ed. by S. Morey and M. Post. Foundation Books, pp. 93–104.
- Dixon, R. M. W. (1982). *Where have all the adjectives gone?: and other essays in semantics and syntax*. Vol. 107. Berlin/New York/Amsterdam: Mouton Publishers.
- (1986). Noun classes and noun classification in typological perspective. In: *Noun classes and categorization*. Ed. by C. G. Craig, pp. 105–112.
 - (2000). A typology of causatives: form, syntax and meaning. In: *Changing Valency: Case Studies in Transitivity*. Ed. by R. M. W. Dixon and A. Y. Aikhenvald. Cambridge University Press, pp. 30–83.
- Dixon, R. M. W. (2004). Adjective Classes in Typological Perspective. In: *Adjective Classes: A Cross-Linguistic Typology*. Oxford University Press.

- (2006a). Complement Clauses and Complementation Strategies in Typological Perspective. In: *Complementation: A Cross-Linguistic Typology*. Ed. by Robert M. W. Dixon and Alexandra Y. Aikhenvald. Oxford University Press, pp. 1–48.
 - (2006b). Complement clauses and complementation strategies in typological perspective. In: *Complementation*. Ed. by R. M. W. Dixon and A. Y. Aikhenvald. Oxford University Press, pp. 1–48.
 - (2009). The Semantics of Clause Linking in Typological Perspective. In: *The Semantics of Clause Linking A Cross-Linguistic Typology*. Ed. by A. Y. Aikhenvald and R. M. W. Dixon. Oxford University Press, pp. 1–55.
 - (2022). *A New Grammar of Dyrbal*. Oxford University Press.
- Doley, S. K. and M. W. Post (2012). Classifiers in Mising. In: *North East Indian Linguistics*. Ed. by G. Hyslop, S. Morey, and M. W. Post. New Delhi: Cambridge University Press India, pp. 243–268.
- Doornenbal, M. (2009). A Grammar of Bantawa: Grammar, paradigm tables, glossary and texts of a Rai language of Eastern Nepal. PhD thesis. Rijksuniversiteit te Leiden.
- Dryer, M. S. (2003). Word order in Sino-Tibetan languages from a typological and geographical perspective. In: *The Sino-Tibetan languages*. Ed. by G. Thurgood and R. J. LaPolla. London: Routledge, pp. 43–55.
- (2007a). Clause types. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 1. Cambridge University Press, pp. 224–275.
 - (2007b). Word order. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. Vol. 1. Cambridge University Press, pp. 61–131.
 - (2013). *Order of Subject, Object and Verb*. Ed. by M. S. Dryer and M. Haspelmath. WALS Online. URL: <https://wals.info/chapter/81>.
- Dutta, D. K. (2006). The marriage system of the Miju-Mishmis. In: *Marriage and culture: Reflections from Tribal Societies of Arunachal Pradesh*. Ed. by T. Mibang and M. C. Behera. Vol. 1. New Delhi: Mittal, pp. 235–244.
- Elford, L. W. and M. Elford (1998). *Dene (Chipewyan) Dictionary*. Prince Albert, Saskatchewan: Northern Canada Mission Distributors.
- Ellen, R. (2016). Is There a Role for Ontologies in Understanding Plant Knowledge Systems? In: *Journal of Ethnobiology* 36.1, pp. 10–28.

- (2017). Categorizing Natural Objects: Some Issues Arising from Recent Work in Cognitive Anthropology and Ethnobiological Classification. In: *Classification from Antiquity to Modern Times: Sources, Methods, and Theories from an Interdisciplinary Perspective*. Ed. by T. Pommerening and W. Bisang. Berlin/Boston: De Gruyter, pp. 263–278.
 - (2020). *The Nuaulu world of plants: Ethnobotanical cognition, knowledge and practice among a people of Seram, eastern Indonesia*. Canon Pyon: Sean Kingston.
- Enfield, N. J. (2005). Areal linguistics and Mainland Southeast Asia. In: *Annual Review of Anthropology* 34, pp. 181–206.
- (2006). Heterosemy and the grammar-lexicon trade-off. In: *Catching Language: The Standing Challenge of Grammar Writing*. Ed. by F. K. Ameka, A. Dench, and N. Evans. Berlin/ New York: De Gruyter Mouton, pp. 297–320.
 - (2007). *A Grammar of Lao*. Berlin/New York: De Gruyter Mouton.
 - (2009). Case relations in Lao, a radically isolating language. In: *The Oxford Handbook of Case*. Ed. by A. Malchukov A.; Spencer. Oxford University Press, pp. 808–820.
 - (2021). *The Languages of Mainland Southeast Asia*. Cambridge Language Surveys. Cambridge University Press.
- Enfield, N. J. and B. Comrie (2015). Mainland Southeast Asian languages: State of the art and new directions. In: *Languages of Mainland Southeast Asia: The State of the Art*. Ed. by N. J. Enfield and B. Comrie. Berlin/München/Boston: De Gruyter Mouton, pp. 1–28.
- Eriksen, T. H. (2015). *Small Places, Large Issues: An Introduction to Social and Cultural Anthropology*. Sidmouth, England: PLUTO.
- Evans, J. P. and J. Manyu (2021). The sounds of Tawrã (Digaru-Mishmi), a Tibeto-Burman language. In: *Linguistics of the Tibeto-Burman Area* 44.1, pp. 1–26.
- (2022). *Tawrã Dictionary*. URL: <https://www.webonary.org/tawra/?lang=en>.
- Evans, J. P. (2017). Common tone sandhi processes across Sino-Tibetan languages. In: *Tonal Change and Neutralization*. Ed. by H. Kubozono and M. Giriko. Berlin: De Gruyter Mouton, pp. 223–248.
- Evans, J. P. (2022). Classifiers before numerals: Origin of an atypical morpheme order. In: *Linguistics of the Tibeto-Burman Area* 45.1, pp. 1–21.
- Evans, N. (2007). Insubordination and its uses. In: *Finiteness: Theoretical and Empirical Foundations*. Ed. by I. Nikolaeva. Oxford Universit Press, pp. 366–431.

- Foley, W. A. (2007). A typology of information packaging in the clause. In: *Language typology and syntactic description*. Ed. by T. Shopen. 2nd ed. Vol. 1. Cambridge University Press, pp. 362–446.
- (2010). Clause linkage and nexus in Papuan languages. In: *Clause Linking and Clause Hierarchy : Syntax and Pragmatics*. Ed. by I. Bril. Amsterdam: John Benjamins, pp. 27–50.
 - (2023). Word-Class systems and other grammatical properties. In: *The Oxford Handbook of Word Classes*. Ed. by E. van Lier. Oxford University Press, pp. 109–134.
- Forker, D. (2019). Grammatical relations in Sanzhi Dargwa. In: *Argument Selectors: A new perspective on grammatical relations*. Amsterdam/Philadelphia: John Benjamins, pp. 69–106.
- Gast, V. and H. Diessel (2012). The typology of clause linkage: status quo, challenges, prospects. In: *Clause Linkage in Cross-Linguistic Perspective: Data-Driven Approaches to Cross-Clausal Syntax*. Ed. by V. Gast and H. Diessel. Berlin/Boston: De Gruyter Mouton, pp. 1–36.
- Genetti, C. (2005). The participial construction of Dolakhā Newar: Syntactic implications of an Asian converb. In: *Studies in Language* 29, pp. 35–87.
- (2006). Complement Clause Types and Complementation Strategy in Dolakha Newar. In: *Complementation: A Cross-Linguistic Typology*. Oxford University Press.
 - (2011). Nominalization in Tibeto-Burman languages of the Himalayan area: A typological perspective. In: *Nominalization in Asian languages: Diachronic and typological perspectives*. Amsterdam & Philadelphia: John Benjamins, pp. 163–193.
- Genetti, C. et al. (2021). Direction and associated motion in Tibeto-Burman. In: *Linguistic Typology* 25.2, pp. 345–388.
- Gerlach, B. and J. Grijzenhout (2000). Clitics from different perspectives. In: *Clitics in Phonology, Morphology and Syntax*. Ed. by B. Gerlach and J. Grijzenhout. Amsterdam/Philadelphia: John Benjamins, pp. 1–29.
- Gil, D. (2005). From repetition to reduplication in Riau Indonesian. In: *Studies on Reduplication*. Ed. by B. Hurch. Berlin/Boston: De Gruyter Mouton, pp. 31–64.
- Givón, T. (1971). Historical syntax and synchronic morphology: An archaeologist's field trip. In: *Chicago linguistic society*. Vol. 7. 1, pp. 394–415.

- (1985). *Syntax: A functional-typological introduction*. Vol. 1. Amsterdam/Philadelphia: John Benjamins.
 - (1986). Prototypes: Between Plato and Wittgenstein. In: *Noun classes and categorization: Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983*. Ed. by C. G. Craig. Amsterdam/Philadelphia: John Benjamins, pp. 77–102.
 - (1997). Grammatical relations: An introduction. In: *Grammatical Relations: A functionalist perspective*. Ed. by T. Givón. Amsterdam/Philadelphia: John Benjamins, pp. 1–84.
 - (2001a). *Syntax: An Introduction*. Vol. 1. Amsterdam/Philadelphia: John Benjamins.
 - (2001b). *Syntax: An Introduction*. Vol. 2. Amsterdam/Philadelphia: John Benjamins.
 - (2018). *On Understanding Grammar*. Revised edition. Amsterdam/Philadelphia: John Benjamins.
- Gordon, M. K. (2016). *Phonological Typology*. Oxford University Press.
- Goswami, G. and J. Tamuli (2004). Asamiya. In: *The Indo-Aryan Languages*. Ed. by D. J. George Cardona. New York & Londong: Routledge, pp. 429–484.
- Green, A. D. (2005). Word, foot and syllable structure in Burmese Syntax. In: *Studies in Burmese linguistics*. Ed. by J. Watkins. Canberra: Pacific Linguistics, pp. 1–25.
- Greenberg, J. (1954). A quantitative approach to the morphological typology of language. In: *IJAL* 26, pp. 178–194.
- (2013[1972]). Numeral classifiers and substantival number: Problems in the genesis of a linguistic type. In: *Eight Decades of General Linguistics*. Ed. by F. Kiefer and P. van Sterkenburg. Leiden/Boston: Brill, pp. 305–333.
- Greenberg, J. H. (1974). Studies in numerical system, 1: Double numeral system. In: *Working Papers on Language Universals*. Vol. 14. Language universals project, Stanford University, pp. 75–89.
- Grierson, G. A. (1909). *Linguistic Survey of India: Tibeto-Burman Family*. Vol. 3.1: *General Introduction, Specimens from the Tibetan Dialects, the Himalayan Dialects and North Assam Group*. Calcutta: Superintendent of Government Printing, India.
- Grinevald, C. (2000). A morphosyntactic typology of classifiers. In: *Systems of nominal classification*. Ed. by G. Senft. Cambridge University Press, pp. 50–92.

- (2003). Classifiers systems in the context of a typology of nominal classification. In: *Perspectives on Classifier Constructions in Sign Languages*. Ed. by K. Emmorey. New York/London: Psychology Press, pp. 91–109.
- Guillaume, A. and H. Koch (2021). Introduction: associated motion as a grammatical category in linguistic typology. In: *Associated Motion*. Ed. by A. Guillaume and H. Koch. Berlin/Boston: de Gruyter, pp. 3–30.
- Gurdon, P. R. T. (1914). *The Khasis*. London: MacMillan and Co.
- Haiman, J. (1978). Conditionals are topics. In: *Language* 54.3, pp. 564–589.
- (1981). On some origins of switch-reference marking. In: *Switch Reference and Universal Grammar: Proceedings of a symposium on switch reference and universal grammar*, pp. 105–128.
- Hale, A. (1980). Person markers: Finite conjunct and disjunct verb forms in Newari. In: *Papers in South-East Asian Linguistics* 7, pp. 95–106.
- Hampe, B. and C. Lehmann (2013). Partial coreference. In: *Languages Across Boundaries: Studies in Memory of Anna Siewierska*. Ed. by D. Bakker and M. Haspelmath. Berlin/Boston: De Gruyter Mouton, pp. 159–196.
- Handel, Z. J. (2003). A concise introduction to Old Chinese phonology. In: *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction*. Ed. by J. A. Matisoff. University of California Press, pp. 543–575.
- Hanks, W. F. (2018[1996]). *Language and Communicative Practices*. New York/London: Routledge.
- Hargreaves, D. (2005). Agency and intentional action in Kathmandu Newar. In: *Himalayan Linguistics* 5, pp. 1–48.
- So-Hartmann, H. (2009). *A Descriptive Grammar of Daai Chin*. University of California.
- Haspelmath, M. (1993). Inchoative/causative verb alternations. In: *Causatives and Transitivity*. Ed. by B. Comrie and M. Polinsky. Amsterdam: John Benjamins, pp. 87–120.
- Haspelmath, M. (1995). The converb as a cross-linguistically valid category. In: *Converbs in Cross-Linguistic Perspective: Structure and Meaning of Adverbial Verb Forms - Adverbial Participles, Gerunds*. Ed. by M. Haspelmath and E. König. Berlin: Walter de Gruyter, pp. 1–56.

- (2007). Coordination. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 2. Cambridge University Press, pp. 1–51.
 - (2009). Framework-Free Grammatical Theory. In: *The Oxford Handbook of Linguistic Analysis*. Ed. by B. Heine and H. Narrog. Oxford University Press, pp. 341–366.
 - (2011). On S, A, P, T, and R as comparative concepts for alignment typology. In: *Linguistic Typology* 15.3, pp. 535–567.
- Hayes, B. (1985). Iambic and trochaic rhythm in stress rules. In: *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, pp. 429–446.
- (1995). *Metrical stress theory: principles and case studies*. University of Chicago Press.
 - (2009). *Introductory Phonology*. Hoboken, New Jersey: Wiley-Blackwell.
- Heine, B. (2000). On the rise of new-event markers in Kxoe. In: *Mehr als nur Worte...: Afrikanistische Beiträge zum 65. Geburtstag von Franz Rottland*. Ed. by A. M. R. Vossen and A. Meissner. Cologne: Rüdiger Köppe, pp. 269–284.
- Hellwig, B. (2011). *A Grammar of Goemai*. Berlin/Boston: De Gruyter Mouton.
- (2019). *A Grammar of Qaqet*. Berlin/Boston: De Gruyter Mouton.
- Hengeveld, K. J. R. and A. Siewierska (2004). Parts-of-speech systems and word order. In: *Journal of Linguistics* 40, 527 – 570.
- Hildebrandt, K. (2004). A grammar and glossary of the Manange language. In: *Tibeto-Burman Languages of Nepal: Manange and Sherpa [Pacific Linguistics 557]*. Ed. by C. Genetti. Canberra: Pacific Linguistics, pp. 3–189.
- Himmelman, N. P. (1996). Demonstratives in narrative discourse a taxonomy of universal uses. In: *Studies in Anaphora*. Ed. by B. A. Fox. Amsterdam/Philadelphia: John Benjamins, pp. 205–254.
- Hock, H. H. (1991). *Principles of Historical Linguistics*. Berlin/New York: De Gruyter Mouton.
- Hock, H. H. (2016a). Contact and convergence. In: *The Languages and Linguistics of South Asia: A Comprehensive Guide*. Ed. by H. H. Hock and E. Bashir. Berlin/Boston: De Gruyter Mouton, pp. 241–374.
- (2016b). Finite and nonfinite subordination. In: *The Languages and Linguistics of South Asia: A Comprehensive Guide*. Ed. by H. H. Hock and E. Bashir. Berlin/Boston: De Gruyter Mouton, pp. 567–584.

- (2016c). Old and Middle Indo-Aryan. In: *The Languages and Linguistics of South Asia: A Comprehensive Guide*. Ed. by H. H. Hock and E. Bashir. Berlin/Boston: De Gruyter Mouton, pp. 18–35.
 - (2016d). Syntax and semantics. In: *The Languages and Linguistics of South Asia: A Comprehensive Guide*. Ed. by H. H. Hock and E. Bashir. Berlin/Boston: Mouton De Gruyter, pp. 501–630.
- Hockett, C. F. (1967). The Quantification of functional load. In: *WORD* 23.1-3, pp. 300–320.
- Holden, J. (2013). *Benasní-I Remember: Dene Sų́líné Oral Histories with Morphological Analysis*. Leiden/Boston: Brill.
- Hopper, P. J. and E. C. Traugott (2003). *Grammaticalization*. Cambridge University Press.
- Hopper, P. J. (1986). Some Discourse Functions of Classifiers in Malay. In: *Noun Classes and Categorization: Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983*. Ed. by C. G. Craig. Amsterdam/Philadelphia: John Benjamins, pp. 309–326.
- (1991). On some principles of grammaticization. In: *Approaches to grammaticalization*. Ed. by E. C. Traugott and B. Heine. Amsterdam/Philadelphia: John Benjamins, pp. 17–35.
- Hotz, R. (forthcoming 2025). Classifiers in Tawrā Mishmi. In: *Linguistics of the Tibeto-Burman Area* 48.2.
- Do-Hurinville, D. T. and H. L. Dao (2019). Vietnamese. In: *The Mainland Southeast Asia Linguistic Area*. Ed. by A. Vittrant and J. Watkins. Berlin/Boston: De Gruyter Mouton, pp. 384–431.
- Hyman, L. M. (1985). *A theory of phonological weight*. Dordrecht/Cinnaminson: Foris Publications.
- Hyman, L. M. (2006). Word-prosodic typology. In: *Phonology* 23.2, pp. 225–257.
- (2007). Kuki-thaadow: An african tone system in Southeast Asia. In: *UC Berkeley Phon-Lab Annual Report* 3.3.
 - (2009). How (not) to do phonological typology: the case of pitch-accent. In: *Language sciences* 31.2-3, pp. 213–238.

- Hyslop, G. (2013). The Kurtöp -si construction: Converbs, clause-chains and verb serialization. In: *Functional-Historical Approaches to Explanation: In honor of Scott DeLancey*. Ed. by T. Thornes et al. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 155–178.
- (2017). *A Grammar of Kurtöp*. Leiden: Brill.
 - (2021). Between stress and tone: Acoustic evidence of word prominence in Kurtöp. In: *Language Documentation and Conservation* 15, pp. 551–575.
- Iwasaki, S. and I. P. Horie (2005). *A Reference Grammar of Thai*. Cambridge University Press.
- Jacques, G. (2016). Tangut, Gyalrongic, Kiranti and the nature of person indexation in Sino-Tibetan/Trans-Himalayan. In: *Linguistics Vanguard* 2.1, pp. 1–13.
- (2021). *A Grammar of Japhug*. Berlin: Language Science Press.
- Jacques, G., A. Lahaussais, and S. Zhang (2021). Associated motion in Sino-Tibetan, with a focus on Gyalrongic and Kiranti. In: *Associated Motion*. Ed. by A. Guillaume and H. Koch. Berlin/Boston: de Gruyter, pp. 819–854.
- Jacques, G. and C. Stevens (2024). Linguistic, archaeological and genetic evidence suggests multiple agriculture-driven migrations of Sino-Tibetan speakers from Northern China to the Indian subcontinent. In: *Quaternary International* 711, pp. 1–20.
- Jacquesson, F. (2015). Persons and grammar in Meyor. In: *Language and Culture in Northeast India and Beyond. In Honour of Robbins Burling. Asia-Pacific Linguistics*, pp. 100–113.
- Jaker, A. and P. J. Howson (2022). An acoustic study of Tetsôt'iné stress: Iambic stress in a quantity-sensitive tone language. In: *Phonology* 39.1, pp. 1–39.
- Jenny, M. and S. S. Hnin Tun (2017). *Burmese: A Comprehensive Grammar*. London: Routledge.
- Jenny, M. (2000). The aspect system of Thai. In: *Aktionsart and Aspectotemporality in Non-European Languages: Proceedings from a Workshop held at the University of Zurich*, pp. 97–140.
- (2011). Burmese in Mon syntax: external influence and internal development. In: *Austroasiatic Studies: papers from ICAAL 4*. Ed. by S. Srichampa, P. Sidwell, and K.

- Gregerson. Mon-Khmer Studies Journal Special Issue 3. SIL International. Mahidol University, Pacific Linguistics, pp. 48–64.
- (2021). Pragmatics and syntax in the languages of MSEA. In: *The Languages and Linguistics of Mainland Southeast Asia: A comprehensive guide*. Ed. by P. Sidwell and M. Jenny. Berlin/Boston: De Gruyter Mouton, pp. 825–854.
- Jenny, M., T. Weber, and R. Weymuth (2014). The Austroasiatic languages: a typological overview. In: *The Handbook of Austroasiatic Languages. Vol. 1*. Ed. by M. Jenny and P. Sidwell. Leiden/Boston: Brill, pp. 13–143.
- Jiang, D., D. Li, and H. Sun (2013). *Darangyu yanjiu (Darang language research)*. Beijing: Minzu Press.
- Jing, G. and M. Liu (2023). The multifunctionality of Mandarin diào: A cross-linguistical perspective. In: *Chinese Language and Discourse* 14(1), pp. 18–49.
- Jørgen, R. (1995). *Minor Mlabri: A Hunter-Gatherer Language of Northern Indochina*. Museum Tusculanum Press, University of Copenhagen.
- Joseph, U. V. (2006). *Rabha*. Leiden: Brill.
- (2009). The numeral ‘one’ in Khasi and Karbi. In: *North East Indian Linguistics*. Ed. by S. Morey and M. Post. New Delhi: Foundation / Cambridge University Press India, pp. 149–161.
- Kager, R. (2007). Representation. In: *The Cambridge Handbook of Phonology*. Ed. by P. d. Lacy. Cambridge Handbooks in Language and Linguistics. Cambridge University Press, pp. 195–228.
- Kari, J. (1990). *Ahtna Athabaskan Dictionary*. Alaska Native Language Center, University of Alaska Fairbanks.
- Kato, A. (2021). Typological profile of Karenic languages. In: *The Languages and Linguistics of Mainland Southeast Asia: A comprehensive guide*. Ed. by P. Sidwell and M. Jenny. Berlin/Boston: De Gruyter Mouton, pp. 337–368.
- Kemmerer, D. (2017). Categories of object concepts across languages and brains: the relevance of nominal classification systems to cognitive neuroscience. In: *Language, Cognition and Neuroscience* 32.4, pp. 401–424.
- King, J. T. (2009). *A Grammar of Dhimal*. Leiden, Brill.

- Kirby, J. and M. Brunelle (2017). Southeast Asian tone in areal perspective. In: Cambridge University Press, pp. 703–731.
- Kittilä, S. and F. Zúñiga (2010). Introduction: Benefaction and malefaction from a cross-linguistic perspective. In: *Benefactives and malefactives: typological perspectives and case studies*. Ed. by F. Zúñiga and S. Kittilä. Amsterdam/Philadelphia: John Benjamins, pp. 1–28.
- Kobayashi, M. and B. Tirkey (2017). *The Kurux Language: Grammar, Texts and Lexicon*. Leiden: Brill.
- König, E. and P. Siemund (2007). Speech act distinctions in grammar. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. Vol. 1. Cambridge University Press, pp. 276–324.
- Konnerth, L. and A. Teo (2014). Acoustic-statistical and perceptual investigations of Karbi tones: A peculiar case of incomplete neutralization of F0. In: *North East Indian Linguistics* 6, pp. 13–37.
- Konnerth, L. A. (2014). A Grammar of Karbi. PhD thesis. University of Oregon.
- Kuteva, T. et al. (2019). *World Lexicon of Grammaticalization*. Cambridge University Press.
- Ladefoged, P. and K. Johnson (2006). *A Course in Phonetics*. Thomson Wadsworth Boston, MA.
- Lambrecht, K. (1994). *Information Structure and Sentence Form: Topics, Focus, and the Mental Representations of Discourse Referents*. Cambridge University Press.
- Langacker, R. W. (2002). *The Cognitive Basis of Grammar*. Berlin/New York: De Gruyter Mouton.
- (2008). *Cognitive Grammar: A Basic Introduction*. Oxford University Press.
- LaPolla, R. J. (2001). The Role of migration and language contact in the development of the Sino-Tibetan language family. In: *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Ed. by R. M. W. Dixon and A. Y. Aikhenvald. Oxford University Press, pp. 294–327.
- LaPolla, R. J. (2003). Overview of Sino-Tibetan morphosyntax. In: *The Sino-Tibetan Languages*. Ed. by G. Thurgood and R. J. LaPolla. 2nd ed. London: Routledge, pp. 22–42.

- (2006). On grammatical relations as constraints on referent identification. In: *Voice and Grammatical Relations: In Honor of Masayoshi Shibatani*. Ed. by T. Tsunoda and T. Kageyama. Amsterdam/Philadelphia: John Benjamins, pp. 139–151.
 - (2008a). Relative Clause Structures in the Rawang Language. In: *Language and Linguistics* 9.4, pp. 797–812.
 - (2008b). Nominalization in Rawang. In: *Linguistics of the Tibeto-Burman area* 31.2, pp. 45–65.
 - (2010). Copula constructions in Rawang. In: *Linguistic Studies in Chinese and Neighboring Languages*, pp. 1059–1076.
 - (2016). On categorization: Stick to the facts of the languages. In: *Linguistic Typology* 20.2, pp. 365–375.
 - (2017). Noun-modifying clause constructions in Sino-Tibetan languages: Rethinking theoretical and geographical boundaries. In: *Noun-Modifying Clause Constructions in Languages of Eurasia*. Ed. by Y. Matsumoto, B. Comrie, and P. Sells, pp. 91–103.
 - (2023). Grammatical Relations. In: *The Cambridge Handbook of Role and Reference Grammar*. Ed. by D. Bentley et al. Cambridge University Press, pp. 269–291.
- LaPolla, R. J. and C. Huang (2003). *A Grammar of Qiang: with annotated texts and glossary*. Berlin/New York: De Gruyter Mouton.
- LaPolla, R. J. and D. Poa (2006). On describing word order. In: *Catching Language: The Standing Challenge of Grammar Writing*. Ed. by F. K. Ameka, A. Dench, and N. Evans. Berlin/New York: De Gruyter Mouton, pp. 269–296.
- Law, A. (2002). Cantonese sentence-final particles and the CP domain. In: *UCL Working Papers in Linguistics* 14.
- Lehmann, C. (1984). *Der Relativsatz: Typologie seiner Strukturen; Theorie seiner Funktionen; Kompendium seiner Grammatik*. Tübingen: Gunter Narr.
- (1995[1982]). *Thoughts on Grammaticalization*. München/Newcastle: Lincom Europa.
- Lehmann, C., Y.-M. Shin, and E. Verhoeven (2004). *Direkte und indirekte Partizipation: zur Typologie der sprachlichen Repräsentation konzeptueller Relationen*. Munich: Lincom Europa.
- Li, C. N. and S. A. Thompson (1976). Subject and topic: A new typology of language. In: *Subject and Topic*. Ed. by C. N. Li. New York: Academic Press, pp. 457–489.

- Li, X. and W. Bisang (2012). Classifiers in Sinitic languages: From individuation to definiteness-marking. In: *Lingua* 122.4, pp. 335–355.
- Lieberherr, I. (2017). A Grammar of Bulu Puroik. PhD thesis. University of Bern.
- Liljegren, H. (2016). *A Grammar of Palula*. Berlin: Language Science Press.
- Lin, Y. (2012). By no means marginal: Privative tone in Zhuokeji Rgyalrong. In: *Language and Linguistics*, 13(4), pp. 625–662.
- Löbner, S. (2011). Concept Types and Determination. In: *Journal of Semantics* 28.3, pp. 279–333.
- Longacre, R. E. (2007). Sentences as combinations of clauses. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 2. Cambridge University Press, pp. 372–420.
- Lowe, E. J. (2012). Individuation, Reference, and Sortal Terms. In: *Perception, Realism, and the Problem of Reference*. Ed. by A. Raftopoulos and P. Machamer. Cambridge University Press, pp. 123–141.
- Lustig, A. (2010). *A Grammar and Dictionary of Zaiwa: Vol. 1*. Leiden: Brill.
- Luthra, P. N. (1971). North-East Frontier Agency Tribes: Impact of Ahom and British Policy. In: *Economic and Political Weekly* 6.23, pp. 1143–1149.
- Lyons, J. (1977). *Semantics*. Cambridge University Press.
- Malchukov, A. (2015). Valency classes and alternations: Parameters of variation. In: *Valency Classes in the World's Languages: Introducing the Framework, and Case Studies from Africa and Eurasia*. Ed. by A. Malchukov and B. Comrie. Berlin/Boston: De Gruyter Mouton, pp. 73–130.
- Manova, S. and W. U. Dressler (2005). The morphological technique of conversion in the inflecting-fusional type. In: *Approaches to conversion/zero-derivation*. Ed. by L. Bauer and S. Valera Hernández. Münster/New York: Waxman, pp. 67–102.
- Masica, C. P. (1993). *The Indo-Aryan languages*. Cambridge University Press.
- Matisoff, J. A. (1973a). *The Grammar of Lahu*. University of California Press.
- Matisoff, J. A. (1973b). Tonogenesis in Southeast Asia. In: *Consonant types and tone (Southern California Occasional Papers in Linguistics 1)*. Los Angeles: University of Southern California 34. Ed. by L. M. Hyman, pp. 71–96.

- (1991). Areal and universal dimensions of grammaticalization in Lahu. In: *Approaches to Grammaticalization: Volume II. Types of grammatical markers*. Ed. by E. C. Traugott and B. Heine. Amsterdam/Philadelphia: John Benjamins, pp. 383–454.
 - (2000). Tibeto-Burman tonology in an areal context. In: *The Fifth International Symposium on Languages and Linguistics*. Vietnam National University, Ho Chi Minh City University of Social Sciences and Humanities, pp. 85–151.
 - (2003). *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction*. University of California Press.
 - (2010). Toward a Eurasian bestiary: (I) OTTER in Tibeto Burman and MonKhmer: (II) JACKAL in Sino-Tibetan and Indo-European. In: *Bulletin of National Museum of Ethnology* 34(3).
 - (2018). Rethinking the Proto-Tibeto-Burman *a-prefix: Glottal and nasal complications. In: *Journal of Asian and African Studies* 96, pp. 29–69.
- Matthews, P. H. (1991). *Morphology*. 2nd ed. Cambridge University Press.
- Mattiola, S. (2019). *Typology of Pluractional Constructions in the Languages of the World*. Amsterdam/Philadelphia: John Benjamins.
- Mayer, M. (1969). *Frog, Where Are You?* New York: Puffin.
- McCarthy, J. and A. Prince (1986). *Prosodic Morphology*. University of Massachusetts, Amherst and Brandeis University. Unpublished manuscript.
- Mene, T. and S. K. Chaudhuri (2019). The Idu Mishmi tribe of frontier state of Arunachal Pradesh. In: *Change and Continuity among Tribes*. Ed. by T. Mene and S. K. Chaudhuri. Bhopal: Indira Gandhi Rashtriya Manav Sangrahalaya, pp. 7–14.
- Mithun, M. (1984). The evolution of noun incorporation. In: *Language* 60.4, pp. 847–894.
- (2001). *The Languages of Native North America*. Cambridge University Press.
- Modi, Y. (2006). The complexity and emergence of Hindi as Lingua Franca in Arunachal Pradesh. In: *Vital Voices: Endangered Language and Multilingualism*. Ed. by R. Elangaiyan et al. CIIL, Mysore, India: Foundation for Endangered Languages, pp. 161–165.
- Modi, Y. (2015). Towards an understanding of language distribution in the Tani area: Social organization, expansion and migration. In: *Language and Culture in Northeast India and Beyond: In Honor of Robbins Burling*. Ed. by M. W. Post, S. Morey, and S. DeLancey, pp. 245–264.

- (2017). *The Milang Language: Grammar and texts*. PhD thesis. University of Bern.
- Modi, Y. and M. W. Post (2022). Applicatives in Macro-Tani languages (Trans-Himalayan, Eastern Himalaya): Forms, functions and historical origins. In: *Applicative Morphology: Neglected Syntactic and Non-syntactic Functions*. Ed. by S. Pacchiarotti and F. Zúñiga. Berlin, Boston: De Gruyter Mouton, pp. 299–328.
- Morey, S. (2012). *Turung: A variety of Singpho language spoken in Assam*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- (2014). Studying tones in North East India: Tai, Singpho and Tangsa. In: *Language Documentation and Conservation* 8, pp. 637–671.
- (2020). Syllable duration in Tai Phake: The interaction between vowel length and tone length. In: *Himalayan Linguistics* 19.2, pp. 69–79.
- Mulder, M. (2018). Documenting a fuzzy tone system: The case of Muklom Tangsa. In: *Journal of the Southeast Asian Linguistics Society, Special Publication No. 2*, pp. 113–123.
- Mya-Tu, U. et al. (1967). *The Tarons in Burma: The results of a scientific expedition by the Burma Medical Research Society*. Rangoon: Central Press.
- NEFA (nd). *A Dictionary of the Taraon Language - For the Use of Officers in the North-East Frontier Agency Administration, Taraon-English-English-Taraon*. Shillong: Philological Section, Research Dept.
- Newman, J. (2009). A cross-linguistic overview of ‘eat’ and ‘drink’. In: *The Linguistics of Eating and Drinking*. Ed. by J. Newman. Amsterdam: John Benjamins, pp. 1–26.
- Ngai, S. S. (2021). *A Grammar of Shaowu: A Sinitic Language of Northwestern Fujian*. Berlin/Boston: De Gruyter Mouton.
- Noonan, M. (1997). Versatile nominalizations. In: *Essays on Language Function and Language Type: Dedicated to T. Givón*. Ed. by J. H. Joan L. Bybee and S. A. Thompson. Amsterdam/Philadelphia: John Benjamins, pp. 373–410.
- Noonan, M. (1999). Converb constructions in Chantyal. In: *Topics in Nepalese Linguistics* 15(1). Ed. by Y. P. Yadava, pp. 1–23.
- (2003). Recent language contact in the Nepal Himalaya. In: *Language variation: papers on variation and change in the Sinosphere and in the Indosphere in honour of James A.*

- Matisoff*. Ed. by B. M. David Bradley Randy LaPolla and G. Thurgood. Canberra: Pacific Linguistics, pp. 65–88.
- (2007). Complementation. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 2. Cambridge University Press, pp. 52–150.
- Norman, J. and T.-l. Mei (1976). The Austroasiatic in Ancient South China: Some lexical evidence. In: *Monumenta Serica* 32, pp. 274–301.
- Oh, Y. M. et al. (2015). Bridging phonological system and lexicon: Insights from a corpus study of functional load. In: *Journal of Phonetics* 53. On the cognitive nature of speech sound systems, pp. 153–176.
- Olson, M. L. (1981). Barai clause junctures: Toward a functional theory of inter-clausal relations. PhD thesis. Australian National University.
- Olsson, B. (2013). *Iamitives: Perfects in Southeast Asia and Beyond*. MA thesis.
- Opgenort, J. R. (2005). *A Grammar of Jero: With a Historical Comparative Study of the Kiranti Languages*. Leiden: Brill.
- Overall, S. E. (2017). A typology of frustrative marking in Amazonian languages. In: *The Cambridge Handbook of Linguistic Typology*. Ed. by A. Y. Aikhenvald and R. M. W. Dixon. Cambridge University Press, pp. 477–512.
- Pacchiarotti, S. and F. Zúñiga (2022). Introduction. In: *Applicative Morphology: Neglected Syntactic and Non-syntactic Functions*. Ed. by S. Pacchiarotti and F. Zúñiga. Berlin, Boston: De Gruyter Mouton, pp. 1–18.
- Parker, S. G. (2002). Quantifying the sonority hierarchy. PhD thesis. University of Massachusetts Amherst.
- Payne, T. E. (1997). *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge University Press.
- Perlin, R. (2020). *A Grammar of Trung*. Himalayan Linguistics: University of California at Santa Barbara.
- Peterson, J. (2010). *A Grammar of Kharia: A South Munda language*. Leiden/Boston: Brill.
- Pichetpan, N. and M. W. Post (2021). Bare classifier phrases in Thai and other Mainland Asian Languages: implications for classifier theory and typology. In: *Linguistic Typology* 25.3, pp. 461–506.

- Pittayaporn, P. (2015). Typologizing sesquisyllabicity: The role of structural analysis in the study of linguistic diversity in Mainland Southeast Asia. In: *Languages of Mainland Southeast Asia: The State of the Art*. Ed. by N. J. Enfield and B. Comrie. Berlin, München, Boston: De Gruyter Mouton, pp. 500–528.
- Plaisier, H. (2006). *A Grammar of Lepcha*. Leiden/Boston: Brill.
- Post, M. W. and T. Kanno (2013). Apatani phonology and lexicon, with a special focus on tone. In: *Himalayan Linguistics* 12(1), pp. 17–75.
- Post, M. (2006). Compounding and the structure of the Tani lexicon. In: *Linguistics of the Tibeto-Burman Area* 29(1), pp. 41–60.
- Post, M. W. (2007). *A Grammar of Galo*. PhD thesis. Melbourne: La Trobe University.
- (2008). Verbs of posture, existence, location and possession and their grammaticalization pathways in the Tani languages. In: *North East Indian Linguistics*. Ed. by S. Morey and M. W. Post. New Delhi: Cambridge University Press India, pp. 127–152.
 - (2009). The semantics of clause linking in Galo. In: *The Semantics of Clause Linking A Cross-Linguistic Typology*. Ed. by A. Y. Aikhenvald and R. M. W. Dixon. Oxford University Press, pp. 74–95.
 - (2010). Predicate derivations in the Tani languages: Root, suffix, both or neither. In: *North East Indian Linguistics* 2, pp. 175–197.
 - (2011a). Nominalization and nominalization-based constructions in Galo. In: *Nominalization in Asian languages: Diachronic and typological perspectives*. Amsterdam & Philadelphia: John Benjamins, pp. 255–287.
 - (2011b). Nominalization-based constructions in Tibeto-Burman languages: Typology and evolution. Paper presented at the Association for Linguistic Typology 9th Biennial Conference, University of Hong Kong, July 22.
 - (2011c). Topographical deixis and the Tani languages of North East India. In: *North east Indian linguistics* 3, pp. 137–154.
 - (2012). Oral literature in the Eastern Himalaya: Review of ‘The Sun Rises: A Shaman’s Chant, Ritual Exchange and Fertility in the Apatani Valley’, by Stuart Blackburn. Leiden/Boston: Brill: 2010; ISBN 9789004175785; 401 pages. In: *Linguistics of the Tibeto-Burman Area* 35.1, pp. 113–124.

- Post, M. W. (2013). Person-sensitive TAME marking in Galo: Historical origins and functional motivation. In: *Functional-historical approaches to explanation*. Ed. by T. Thornes et al. Amsterdam/Philadelphia: John Benjamins, pp. 107–130.
- (2015a). Tones in Northeast Indian languages, with a focus on Tani: A fieldworker’s guide. In: *Language and Culture in Northeast India and Beyond: In Honour of Robbins Burling*, pp. 182–210.
 - (2015b). Sino-Tibetan negation and the case of Galo: explaining a distributional oddity in diachronic terms. In: *Language and Linguistics* 16.3, pp. 431–464.
 - (2017). *The Tangam Language: Grammar, Lexicon and Texts*. Leiden: Brill.
 - (2019). Topographical deixis in Trans-Himalayan (Sino-Tibetan) languages. In: *Transactions of the Philological Society* 117.2, pp. 234–255.
 - (2020). The Eastern Himalayan ethno-linguistic diversity hotspot: Where is it, why is it significant, why is it endangered, and what should “we” do about it? In: *Aligarh Journal of Linguistics* 9(1-2), pp. 29–56.
 - (2022). Classifiers in a language with articles: Recent evolution of a typologically unusual Asian classifier system in the Tani languages of northeast India. In: *Asian Languages and Linguistics* 3.2, pp. 239–267.
- Post, M. W. and R. Burling (2017). The Tibeto-Burman languages of Northeast India. In: *The Sino-Tibetan languages*. Ed. by G. Thurgood and R. J. LaPolla. 2nd ed. London: Routledge, pp. 213–242.
- Post, M. W. and Y. Modi (2024). Clause chaining and related structures in Macro-Tani languages: Synchronic and diachronic perspectives. In: ed. by H. S. Sarvasy and A. Y. Aikhenvald. Oxford University Press, pp. 547–575.
- Post, M. W., S. Morey, and T. Huber (2022). Ethnolinguistic Prehistory of the Eastern Himalaya: Diversity and Its Sources. In: *Ethnolinguistic Prehistory of the Eastern Himalaya*. Ed. by M. W. Post, S. Morey, and T. Huber. Leiden: Brill, pp. 1–21.
- Post, M. W. and J. T.-S. Sun (2017). Tani languages. In: *The Sino-Tibetan Languages [Second Edition]*. Ed. by G. Thurgood and R. J. LaPolla. London/New York: Routledge, pp. 322–337.
- Post, M. W. and T. J. Sun (in prep.). *A reconstruction of Proto-Tani*.

- Pu, M.-M. (1997). Zero anaphora and grammatical relations in Mandarin. In: *Grammatical Relations: A functionalist perspective*. Ed. by G. T. Amsterdam/Philadelphia: John Benjamins, pp. 281–322.
- Pulu, J. (1991). *A Phrase Book on Taraon Language*. Guwahati, Assam: Purbadesh Mudran.
- Rai, I. P. et al. (2009). Mundum: a case study of Chintang ritual language. In: *Locality, History, Memory: The Making of the Citizen in South Asia*, pp. 20–33.
- Rai, N. K. and W. Winter (1997). Triplicated verbal adjuncts in Bantawa. In: *Papers in Southeast Asian Linguistics No. 14*: ed. by D. Bradley. Pacific Linguistics, pp. 119–134.
- Rai, N. K. et al. (2005). Triplication and ideophones in Chintang. In: *Contemporary issues in Nepalese linguistics*, pp. 205–209.
- Rai, N. K. et al. (2011). *Chintaṅ śabdakoś tatha vyakaran*. [Chintang dictionary and grammar]. Kathmandu: Chintang Language Research Program.
- Regmi, D. R. and A. Regmi (2018). *A Grammar of Western Tamang*. Munich: Lincom GmbH.
- Reinöhl, U. (2022). Locating Kera’a (Idu Mishmi) in its Linguistic Neighbourhood: Evidence from Dialectology. In: *Ethnolinguistic Prehistory of the Eastern Himalaya*. Leiden/Boston: Brill, pp. 232–263.
- Rice, K. and S. Hargus (2005). Introduction. In: *Athabaskan Prosody*. Ed. by S. Hargus and K. Rice. Amsterdam/Philadelphia: John Benjamins, pp. 1–45.
- Rijkhoff, J. (1991). Nominal aspect. In: *Journal of Semantics* 8.4, pp. 291–309.
- Ring, H. (2015). A Grammar of Pnar. PhD thesis. Nanyang Technological University.
- Robinson, W. (1855). Notes on the languages spoken by the Mi-Shmis. In: *Journal of the Asiatic society of Bengal* 24.4, pp. 307–324.
- Rodriguez, B. (2008). *Der Hühnerdieb*. Wuppertal: Hammer.
- Rondo, J. (2006). Marriage system of the Idus: a profile. In: *Marriage and culture: Reflections from Tribal Societies of Arunachal Pradesh*. Ed. by T. Mibang and M. C. Behera. Vol. 1. New Delhi: Mittal, pp. 141–145.
- Rose, F. and A. Van linden (2023). The derivational use of classifiers in Western Amazonia. In: *Binominal Lexemes in Cross-Linguistic Perspective: Towards a Typology of Complex Lexemes*. Ed. by S. Pepper, F. Masini, and S. Mattioli. Berlin/Boston: De Gruyter Mouton, pp. 237–276.

- RStudio Team (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc. Boston, MA.
- Sagart, L. et al. (2019). Dated language phylogenies shed light on the ancestry of Sino-Tibetan. In: *Proceedings of the National Academy of Sciences of the United States of America* 116.21, pp. 10317–10322.
- San Roque, L. (2018). Egophoric patterns in Duna verbal morphology. In: *Egophoricity*. Ed. by S. Floyd, L. San Roque, and E. Norcliffe. Amsterdam/Philadelphia: John Benjamins, pp. 406–436.
- San Roque, L., S. Floyd, and E. Norcliffe (2018). Egophoricity: an introduction. In: *Egophoricity*. Ed. by S. Floyd, L. San Roque, and E. Norcliffe. Amsterdam/Philadelphia: John Benjamins, pp. 1–78.
- Sarvasy, H. (2015). Breaking the clause chains: Non-canonical medial clauses in Nungon. In: *Studies in Language* 39:3, pp. 664–697.
- Sasse, H.-J. (2015). Syntactic categories and subcategories. In: *Syntax: Theory and Analysis. Volume 1*. Ed. by T. Kiss and A. Alexiadou. Berlin/München/Boston: De Gruyter Mouton, pp. 158–217.
- Sastry, G. D. P. (1984). *Mishmi grammar*. Mysore: Central Institute of Indian Languages.
- (1991). *Mishmi-English-Hindi Dictionary*. Mysore: Central Institute of Indian Languages.
- Saxena, A. (1988). On syntactic convergence: The case of the verb ‘say’ in Tibeto-Burman. In: *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, pp. 375–388.
- Schachter, P. and T. Shopen (2007). Parts-of-speech systems. In: *Language typology and syntactic description*. Ed. by T. Shopen. 2nd ed. Vol. 1. Cambridge University Press, pp. 3–61.
- Schackow, D. (2015). *A Grammar of Yakkha*. Berlin: Language Science Press.
- (2019). Grammatical relations in Yakkha (Kiranti). In: *Argument Selectors: A new perspective on grammatical relations*. Amsterdam/Philadelphia: John Benjamins, pp. 469–510.
- Schiffman, H. F. and F. Harold (1999). *A Reference Grammar of Spoken Tamil*. Cambridge University Press.
- Schikowski, R. (in prep.). *A Grammar of Chintang*.

- Schikowski, R., B. Bickel, and N. P. Paudyal (2015). Flexible valency in Chintang. In: *Valency Classes in the World's Languages*. Berlin: Mouton De Gruyter, pp. 669–708.
- Schultze-Berndt, E. (2000). Simple and complex verbs in Jaminjung. A study of event categorisation in an Australian language. PhD thesis. Nijmegen: Max Planck Institute.
- Seifart, F. (2005). The structure and use of shape-based noun classes in Miraña (North West Amazon). PhD thesis. Radboud University Nijmegen.
- Shafer, R. (1955). Classification of the Sino-Tibetan Languages. In: *WORD* 11:1, pp. 94–111.
- Sharma, M. (2014). Phonological changes in the Hindi lexicon: a case of Meghalaya Hindi. In: *North East India linguistic society* 6. Ed. by G. Hyslop et al., pp. 193–212.
- Shimelman, A. (2017). *A Grammar of Yauyos Quechua*. Berlin: Language Science Press.
- Shorto, H. (2006). *A Mon-Khmer Comparative Dictionary*. Canberra: Pacific Linguistics.
- Sidwell, P. (2014). Austroasiatic classification. In: *The Handbook of Austroasiatic Languages*. Ed. by M. Jenny and P. Sidwell. Leiden: Brill, pp. 144–220.
- (2015). *The Palaungic Languages: Classification, Reconstruction and Comparative Lexicon*. München: Lincom Europa.
- (2018). *The Khasian Languages: Classification, Reconstruction and Comparative Lexicon*. München: Lincom Europa.
- Sidwell, P. and R. Blench (2011). The Austroasiatic Urheimat: The Southeastern Riverine Hypothesis. In: *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*. Ed. by N. J. Enfield. Canberra: Pacific Linguistics, pp. 315–343.
- Sidwell, P. and F. Rau (2014). Austroasiatic Comparative-Historical Reconstruction: An Overview. In: *The Handbook of Austroasiatic Languages*. Ed. by M. Jenny and P. Sidwell. Leiden: Brill, pp. 221–363.
- Singh Grewal, D. (1997). *Tribes of Arunachal Pradesh: Identity, Culture and Languages*. Vol. 2. Delhi: South Asia Publications.
- Slade, B. (2016). Compound verbs in Indo-Aryan. In: *The Languages and Linguistics of South Asia: A Comprehensive Guide*. Ed. by H. H. Hock and E. Bashir. Berlin/Boston: De Gruyter Mouton, pp. 559–567.
- Steriade, D. (2003). Syllables in phonology. In: *International Encyclopedia of Linguistics*. Ed. by W. J. Frawley. 2nd Edition. Oxford University Press. Available at: <https://www.>

oxfordreference.com/display/10.1093/acref/9780195139778.001.0001/acref-9780195139778-e-1062?rskey=44hYmk&result=1061#acref-9780195139778-eSub-0209/.

- Stolz, T. et al. (2008). Split possession. In: *Studies in Language Companion Series* 101.
- Sun, H. (1991). *Zangmianyu yuyin he cihui [Tibeto-Burman phonology and lexicon]*. Beijing: Chinese Social Sciences Press. Accessed via STEDT database.
- Sun, H. et al. (1980). *Menba, Luoba, Dengren de yuyan [The languages of the Menba, Luoba and Deng peoples]*. Ed. by B. S. S. Press. Beijing: Chinese Social Sciences Press. Accessed via STEDT database <<http://stedt.berkeley.edu/search/>> on 2023-02-16.
- Sun, T. J. (1993). A Historical-Comparative Study of the Tani Branch of Tibeto-Burman. PhD thesis. University of California, Berkeley.
- Taid, T. R. (2010). *Mising Gompir Kumsung: A Dictionary of the Mising Language. Mising to Mising, Assamese, English*. Assam: Anundoram Borooah Institute of Language, Art and Culture.
- Talmy, L. (2007). Lexical typologies. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Cambridge University Press, pp. 66–168.
- Tawsik, S. (2014). *Kaman-Tawrah*. 2nd ed. Self-published.
- Thompson, S. A., R. E. Longacre, and S. J. J. Hwang (2007). Adverbial clauses. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 3. Cambridge University Press, pp. 237–300.
- Thurgood, G. (1999). *From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change: With an Appendix of Chamic Reconstructions and Loanwords*. Oceanic Linguistics Special Publications, No. 28. Honolulu: University of Hawai'i Press.
- Timberlake, A. (2007). Aspect, tense, mood. In: *Language Typology and Syntactic Description*. Ed. by T. Shopen. 2nd ed. Vol. 3. Cambridge University Press, pp. 280–333.
- Turin, M. (2007). *Linguistic diversity and the preservation of endangered languages: A case study from Nepal*. International Centre for Integrated Mountain Development (ICI-MOD).
- van Driem, G. (1987). *A Grammar of Limbu*. Berlin: Mouton De Gruyter.
- (2001). *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region*. Vol. 1-2. Leiden/Boston/Köln: Brill.

- (2011). Tibeto-Burman subgroups and historical grammar. In: *Himalayan Linguistics* 10, pp. 31–39.
 - (2012). Glimpses of the ethnolinguistic prehistory of Northeastern India. In: ed. by T. Huber and S. Blackburn. Leiden/Boston: Brill, pp. 187–211.
 - (2014). Trans-Himalayan. In: *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area*. Ed. by T. Owen-Smith and N. Hill. Berlin/Boston: De Gruyter Mouton, pp. 11–40.
- Van Valin, R. D. and R. J. LaPolla (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.
- Västi, K. and S. Kittilä (2014). Semantic roles and verbless constructions: A Finnish challenge for verb-centered approaches. In: *Studies in Language* 38, pp. 512–542.
- Vittrant, A. (2019). Burmese. In: *The Mainland Southeast Asia Linguistic Area*. Ed. by A. Vittrant and J. Watkins. Berlin/Boston: De Gruyter Mouton, pp. 56–130.
- Vittrant, A. and M. Allasonnière-Tang (2021). Classifiers in Southeast Asian languages. In: *The Languages and Linguistics of Mainland Southeast Asia: A comprehensive guide*. Ed. by P. Sidwell and M. Jenny. Berlin/Boston: De Gruyter Mouton, pp. 733–772.
- Vries, M. de (2018). Relative Clauses in Syntax. In: Oxford University Press.
- Watkins, J. (2002). *The phonetics of Wa: Experimental phonetics, phonology, orthography and sociolinguistics*. Pacific linguistics, Research School of Pacific and Asian Studies. The Australian National University.
- Watters, D. E. (2008). Nominalization in the Kiranti and Central Himalayish languages of Nepal. In: *Linguistics of the Tibeto-Burman Area* 31.2, pp. 1–43.
- (2009). *A Grammar of Kham*. Cambridge University Press.
- Widmer, M. (2014). A descriptive Grammar of Bunan. PhD thesis. University of Bern.
- Widmer, M. (2015). The transformation of verb agreement into epistemic marking: evidence from Tibeto-Burman. In: *Agreement from a Diachronic Perspective*. Ed. by J. Fleischer, E. Rieken, and P. Widmer. Berlin/München/Boston: De Gruyter Mouton, pp. 53–74.
- (2020). Same same but different: On the relationship between egophoricity and evidentiality. In: *Evidentiality, egophoricity, and engagement*. Ed. by H. Bergqvist and S. Kittilä. Berlin: Language Science Press, pp. 263–287.

- Widmer, M. and M. Zemp (2017). The epistemization of person markers in reported speech. In: *Studies in Language* 41.4, pp. 33–75.
- Widmer, M. and F. Zúñiga (2017). Egophoricity, involvement, and semantic roles in Tibeto-Burman languages. In: *Open linguistics* 3.1, pp. 419–441.
- Widmer, M. et al. (2021). Morphological structure can escape reduction effects from mass admixture of second language speakers. In: *Studies in Language* 45(4), pp. 707–752.
- Wierzbicka, A. (1985). *Lexicography and Conceptual Analysis*. Ann Arbor: Karoma.
- Witzlack-Makarevich, A. (2011). *Typological Variations in Grammatical Relations*. PhD thesis. University of Leipzig.
- (2019). Argument selectors: A new perspective on grammatical relations: An introduction. In: *Argument Selectors: A new perspective on grammatical relations*. Ed. by A. Witzlack-Makarevich and B. Bickel. Amsterdam/Philadelphia: John Benjamins, pp. 1–38.
- Wolfenden, S. N. (1929). *Outlines of Tibeto-Burman Linguistic Morphology: With Special Reference to the Prefixes, Infixes and Suffixes of Classical Tibetan and the Languages of the Kichin, Bodo, Năgă, Kuki-Chin and Burma Groups*. Vol. 12. Royal Asiatic Society.
- Wu, M.-S., T. Bodt, and T. Tresoldi (2022). Bayesian phylogenetics illuminate shallower relationships among Trans-Himalayan languages in the Tibet-Arunachal area. In: *Linguistics of the Tibeto-Burman Area* 45.2, pp. 171–210.
- Yap, F. H., K. Grunow-Hårsta, and J. Wrona (2011). Introduction: Nominalization strategies in Asian languages. In: *Nominalization in Asian Languages: Diachronic and typological perspectives*. Amsterdam/Philadelphia: John Benjamins, pp. 1–59.
- Yip, M. (2002). *Tone*. Cambridge University Press.
- Zeisler, B. (2024). Facts and attitudes: on the so-called ‘factual’ markers of the modern Tibetic languages. In: *[HL ARCHIVE 14]. Himalayan Linguistics*.
- Zeite, L. (2022). *Dictionary of Meyor. Comprising Meyor to English and English to Meyor*. Itanagar: Centre for Endangered Languages, AITS.
- Zemp, M. (2018). *A Grammar of Purik Tibetan*. Leiden: Brill.
- Zoller, C. P. (2023). *Indo-Aryan and the Linguistic History and Prehistory of North India*. Vol. 1-2. Wiesbaden: Harrassowitz.

- Zúñiga, F. and D. Creissels (2024). Applicative constructions: An introductory overview.
In: *Applicative Constructions in the World's Languages*. Ed. by F. Zúñiga and D. Creissels.
Berlin/Boston: De Gruyter Mouton, pp. 1–56.
- Zúñiga, F. and S. Kittilä (2019). *Grammatical voice*. Cambridge University Press.
- Zwicky, A. (1985). Clitics and particles. In: *Language* 61, pp. 283–305.