

# Appendix for *Beyond Hill and Valley: A Sociogeographic Analysis of Argument Coding Complexity in the Eastern Himalayan Region*

## Appendix A

This appendix describes the sample, data collection and coding procedure in further detail.

### 1. Languages Sampled

Table 1 below details the sample. The sample was initially collected proportionally to size of subgroup, but after excluding languages for lack of adequate data sources, languages were included by convenience. However, this still results in a fairly genealogically diverse sample, as can be seen by the variety in subgroup and node below subgroup, given in the table below. The sample numbers to 32 languages. Data for K'man (glottocode:[gema1234](#)) and Asakian (glottocode:[chak1270](#)) were collected, but excluded for lack of adequate linguistic and ethnographic information to sufficiently code all variables with confidence.

Table 1: List of references for linguistic information

Language	Glottocode	Subgroup (node below top-level)	Node below subgroup	Citation
Rabha	<a href="#">rabh1238</a>	Kochic	Bodo-Garo	Joseph 2007
Bodo-Mech	<a href="#">bodo1269</a>	Bodo-Mech-Kachari	Bodo-Garo	Brahma 2015, Haokip 2018
Garو	<a href="#">garo1247</a>	Bodo-Garo	Bodo-Garo	Burling 2003
Tiwa	<a href="#">tiwa1253</a>	Sal (Brahmaputran)	Bodo-Garo	Dawson 2020a, 2020b
Turung	<a href="#">sing1264</a>	Sal (Brahmaputran)	Jingpho-Asakian	Morey, 2010
Khamniungan	<a href="#">khia1236</a>	Sal (Brahmaputran)	Patkaian	Van Dam and Tham, p.c.
Hakhun Tangsa	<a href="#">noct1238</a>	Sal (Brahmaputran)	Patkaian	Boro 2017
Hills Karbi	<a href="#">karb1241</a>	Kuki-Chin-Naga	Karbic	Konnerth, 2014
Manipuri	<a href="#">mani1292</a>	Kuki-Chin-Naga	Manipuri	Chelliah, 1997
Ao Naga	<a href="#">mong1332</a>	Kuki-Chin-Naga	Angami-Ao	Coupe, 2007
Angami	<a href="#">anga1288</a>	Kuki-Chin-Naga	Angami-Ao	Kevichusa-Ezung, 2014
Suansu	<a href="#">suan1234</a>	Kuki-Chin-Naga	Tangkhu1-Maring	Ivani, 2023
Zeme	<a href="#">zeme1240</a>	Kuki-Chin-Naga	Zemeic	Chanu, 2017
Mizo	<a href="#">lush1249</a>	Kuki-Chin-Naga	Kuki-Chin	Chhange, 1993
Thado Chin	<a href="#">thad1238</a>	Kuki-Chin-Naga	Kuki-Chin	Haokip, 2014
Lamkang	<a href="#">lamk1238</a>	Kuki-Chin-Naga	Kuki-Chin	Thounaojam and Chelliah, 2007,

				Chelliah et al., 2019
Daai Chin	<a href="#">daai1236</a>	Kuki-Chin-Naga	Kuki-Chin	So-Hartmann, 2009
Bori-Karko	<a href="#">bori1243</a>	Macro-Tani	Eastern Tani	Lalrempuii, 2005
Milang	<a href="#">mila1245</a>	Macro-Tani	Koro-Holon	Modi, 2017
Mising	<a href="#">miny1240</a>	Macro-Tani	Mising-Padam-Miri-Minyong	Prasad, 1991
Galo	<a href="#">galo1242</a>	Macro-Tani	Subansiri	Post, 2007
Kera'a	<a href="#">idum1241</a>	Digarish	Digarish	Peck and Reinöhl, p.c.
Hrusso Aka	<a href="#">hrus1242</a>	Hruso	Unclear	D'Souza, 2021
Tibetan	<a href="#">tibe1272</a>	Bodic	Central Tibetan	DeLancey, 2003
Brokpa/Brokpake	<a href="#">brok1248</a>	Bodic	Southern Tibetan	Wangdi, 2021
Bjokapakha	<a href="#">bjok1234</a>	Bodic	Tshanglic	Grollmann, 2020
Tamangic	<a href="#">tama1367</a>	Bodic	Ghale-Tamangic	Owen-Smith, 2015
Limbu	<a href="#">limb1266</a>	Himalayish	Kiranti	van Driem, 1987
Lepcha	<a href="#">lepc1244</a>	Himalayish	Himalayish	Plaisier, 2007
Puroik	<a href="#">puro1234</a>	Kho-Bwa	Kho-Bwa	Lierberherr, 2017
Miji	<a href="#">miji1239</a>	Miji	Miji	Weedall, 2021
Trung/Drung	<a href="#">drun1238</a>	Nungish	Gunong	Perlin, 2020

## 2. Coding of sociocultural variables

The following table links ethnographic and other literature used to code sociocultural variables to the languages. Sometimes, linguistic sources (such as grammars) were re-used if they were the best available source on sociocultural information. Oftentimes, linguistic sources were used in conjunction with other sources to fully code sociocultural variables.

*Table 2: List of references for sociocultural, agricultural and geographic information.*

Language	Glottocode	Subgroup (node below top-level)	Node below subgroup	Citation
Rabha	<a href="#">rabh1238</a>	Kochic	Bodo-Garo	Goswamy et al, 2024
Bodo-Mech	<a href="#">bodo1269</a>	Bodo-Mech-Kachari	Bodo-Garo	Boro, 2024
Garó	<a href="#">garo1247</a>	Bodo-Garo	Bodo-Garo	Burling, 1963
Tiwa	<a href="#">tiwa1253</a>	Sal (Brahmaputran)	Bodo-Garo	Doloi et al, 2024
Turung	<a href="#">sing1264</a>	Sal (Brahmaputran)	Jingpho-Asakian	Morey, 2010
Khiamniungan	<a href="#">khia1236</a>	Sal (Brahmaputran)	Patkaian	Van Dam and Tham, p.c.
Hakhun Tangsa	<a href="#">noct1238</a>	Sal (Brahmaputran)	Patkaian	Boro 2017
Hills Karbi	<a href="#">karb1241</a>	Kuki-Chin-Naga	Karbic	Hansepi and Laisram, 2022
Manipuri	<a href="#">mani1292</a>	Kuki-Chin-Naga	Manipuri	Meitei and Sharma, 2023
Ao Naga	<a href="#">mong1332</a>	Kuki-Chin-Naga	Angami-Ao	Mills, 1926
Angami	<a href="#">anga1288</a>	Kuki-Chin-Naga	Angami-Ao	Hutton, 1921
Suansu	<a href="#">suan1234</a>	Kuki-Chin-Naga	Tangkhul-Maring	Singh and Choudhary, 2015
Zeme	<a href="#">zeme1240</a>	Kuki-Chin-Naga	Zemeic	Longkumer, 2007

Mizo	<a href="#">lush1249</a>	Kuki-Chin-Naga	Kuki-Chin	Chhangte, 1993
Thado Chin	<a href="#">thad1238</a>	Kuki-Chin-Naga	Kuki-Chin	Lalthansangsa, 2018
Lamkang	<a href="#">lamk1238</a>	Kuki-Chin-Naga	Kuki-Chin	Thounaojam and Chelliah, 2007
Daai Chin	<a href="#">daai1236</a>	Kuki-Chin-Naga	Kuki-Chin	Chan and Takeda, 2016
Bori-Karko	<a href="#">bori1243</a>	Macro-Tani	Eastern Tani	Raj, 2010
Milang	<a href="#">mila1245</a>	Macro-Tani	Koro-Holon	Modi, 2017
Mising	<a href="#">miny1240</a>	Macro-Tani	Mising-Padam-Miri-Minyong	Chutia, 2020
Galo	<a href="#">galo1242</a>	Macro-Tani	Subansiri	Post, 2007
Kera'a	<a href="#">idum1241</a>	Digarish	Digarish	Peck and Reinöhl, p.c.
Hrusso Aka	<a href="#">hrus1242</a>	Hruso	Unclear	Nimachow et al. 2011
Tibetan	<a href="#">tibe1272</a>	Bodic	Central Tibetan	DeLancey, 2003
Brokpa/Brokpake	<a href="#">brok1248</a>	Bodic	Southern Tibetic	Wangdi, 2021
Bjokapakha	<a href="#">bjok1234</a>	Bodic	Tshanglic	Akamatsu, 2015
Tamangic	<a href="#">tama1367</a>	Bodic	Ghale-Tamangic	Owen-Smith, 2015
Limbu	<a href="#">limb1266</a>	Himalayish	Kiranti	Limbu, 2013
Lepcha	<a href="#">lepc1244</a>	Himalayish	Himalayish	Thapa and Allay, 2021
Puroik	<a href="#">puro1234</a>	Kho-Bwa	Kho-Bwa	Ramjuk, 2018
Miji	<a href="#">miji1239</a>	Miji	Miji	Weedall, 2021
Trung/Drung	<a href="#">drun1238</a>	Nungish	Gunong	Perlin, 2020

Full bibliographic references for all data sources can be found at the end of this document. For languages where the citation says ‘p.c.’, the researcher reached out to academics known to be working on the language. They were asked to fill out a questionnaire, given below. The questionnaire also reflects my own methodology when coding information from a grammar or ethnographic data source.

### 3. Questionnaire

Lang: (glottocode), filled out by: (name)

1. Do markers on the noun distinguish core grammatical relations? (Are there case markers?)  
**If no, ignore Q2 -Q8**
2. Is the only argument of a one-place (intransitive) predicate marked? If so, by what forms? (affixes or adpositions)
3. Is an agentive argument of a one-place (intransitive) predicate marked? If so, by what forms? (affixes or adpositions)
4. Is a patientive argument of a one-place (intransitive) predicate marked? If so, by what forms? (affixes or adpositions)
5. Is the agentive argument of a two-place (transitive) predicate marked? If so, by what forms? (affixes or adpositions)
6. Is the patientive argument of a two-place (transitive) predicate marked? If so, by what forms? (affixes or adpositions)
7. Is the marker for the agentive argument of a transitive predicate optional? If so, under what conditions? (if known)
8. Is the marker for the patientive argument of a transitive predicate optional? If so, under what conditions? (if known)
9. Are there verbal indexes for core grammatical relations in affirmative declarative clauses? If yes, which macro-roles (of *intransitive subject*, *agent*, and *undergoer*) are indexed? If yes

and different in different TAME categories, answer for category with most distinctions. **If no, ignore Q10-Q18.**

10. Is the only argument of a one-place (intransitive) predicate indexed on the verb? If so, what persons and numbers can be indexed on the verb?
11. Is the agentive argument of a two-place (transitive) predicate indexed on the verb? If so, what persons and numbers?
12. Is the patientive argument of a two-place (transitive) predicate indexed on the verb? If so, what persons and numbers?
13. Are the indexes for intransitive argument the same as the ones for the agentive transitive argument?
14. Are the indexes for intransitive argument the same as the ones for the patientive transitive argument?
15. What is the maximal paradigm of agent and undergoer combinations for the transitive clause that cause some morphological marking on the verb?
16. Is there a hierarchy that determines the expression of verbal indexes?
17. Is there an inverse marker on the verb?
18. Are there any TAME requirements for the verbal indexes?
19. What is the pragmatically unmarked constituent order in the transitive clause?
20. Are other constituent orders in the transitive clause possible?
21. Any outstanding comments?

Some geographic/cultural questions, if possible:

1. What is the largest settlement where this language is primarily spoken?
2. How many speakers are there?
3. Is the language spoken as an L2 by any other speech communities?
4. What is the type of agriculture traditionally practiced? (intensive/irrigated, extensive/shifting/slash-and-burn/*jhum*, horticulture/gathering, or casual) (based on <https://d-place.org/parameters/EA028#1/30/152>)
5. What are/were the jurisdictional levels beyond the local community? (None/autonomous bands or villages, petty chiefdom, larger chiefdom, state, complex state) (based on <https://d-place.org/parameters/EA033#1/30/153>)
6. Any outstanding comments?

## Appendix B

This appendix describes the statistical analyses in further detail.

### 1. Introduction

All code and results for the statistical analyses can be found in the R notebook *synthesis.rmd*, at <https://github.com/patrickdas/ehr-argument-coding>.

Table 3 below details the analyses conducted in this study. Some of these analyses which were not fully detailed in the paper for reasons of space are given in full below.

*Table 3: Details of all statistical models used in this study.*

Model type	Response variable	Predictor variable	Random effect	N
Logistic regression	Hill/Valley	Elevation	None	31
Logistic regression	Hill/Valley	Terrain roughness	None	31
Logistic regression	Hill/Valley	Elevation,	None	31

		Terrain roughness		
Mixed effects	Complexity	Hill/Valley	Subfamily	31
Mixed effects	Complexity	Agricultural intensity	Subfamily	31
Mixed effects	Complexity	L2 Status	Subfamily	31
Mixed effects	Complexity	Population Size	Subfamily	31
Mixed effects	Complexity	Political Organization	Subfamily	31

## 2. Hill/Valley classification as a predictor of complexity

The primary analysis of this study is whether Hill/Valley classification can effectively predict linguistic complexity. To this end, I ran six linear mixed effects models. Linear mixed-effects models were fit with hill\_valley as a fixed effect (categorical: Hill, Split, Valley with Hill as reference) and a random effect for subfamily. The tables below present the estimated coefficients, standard errors, degrees of freedom, t-values, p-values, and R<sup>2</sup> values for each model with Limbu included. Models with Limbu excluded are available in the paper S4.2. Hill was set as the reference level for all models.

*Table 4: Mixed-effects model predicting cell complexity from Hill/Valley classification with Limbu included. Model formula: scaled\_cell\_complexity ~ hill\_valley + (1 | subfamily). N = 32 languages; 10 subfamilies*

Term	Estimate (β)	Std. Error	df	t	p
Intercept (Hill)	0.14	0.25	9.79	0.58	0.58
Split	−0.55	0.42	25.32	−1.30	0.20
Valley	−0.10	0.471	28.79	−0.21	0.83

R<sup>2</sup>: Marginal = 0.051; Conditional = 0.106

*Table 5: Mixed-effects model predicting form complexity from Hill/Valley classification with Limbu included. Model formula: scaled\_form\_complexity ~ hill\_valley + (1 | subfamily). N = 32 languages; 10 subfamilies*

Term	Estimate (β)	Std. Error	df	t	p
Intercept (Hill)	0.050	0.254	12.26	0.196	0.848
Split	−0.329	0.426	26.30	−0.772	0.447
Valley	0.056	0.478	28.86	0.117	0.907

R<sup>2</sup>: Marginal = 0.022; Conditional = 0.083

*Table 6: Mixed-effects model predicting case complexity from Hill/Valley classification with Limbu included. Model formula: scaled\_case\_marking\_complexity ~ hill\_valley + (1 | subfamily). N = 32 languages; 10 subfamilies*

Term	Estimate (β)	Std. Error	df	t	p
Intercept (Hill)	−0.264	0.239	12.09	−1.104	0.291
Split	0.542	0.414	26.57	1.309	0.202
Valley	0.712	0.462	28.85	1.541	0.134

R<sup>2</sup>: Marginal = 0.093; Conditional = 0.122

### 3. Ecological analyses

Three logistic regression models were fitted to test whether ecological variables predict Hill/Valley societal classification, using the sensitivity dataset (N = 31, excluding Limbu). The dependent variable was a binary classification where Hill = 1 and Valley/Split = 0.

*Table 7: Logistic regression model predicting Hill/Valley classification from elevation and terrain roughness. Model formula: hill\_binary ~ scaled\_altitude + scaled\_stdev\_slope, family = binomial. N = 31 languages*

Parameter	Estimate	Std. Error	z-value	p-value
Intercept	0.341	0.369	0.922	0.356
Elevation (scaled)	0.262	0.451	0.581	0.561
Terrain roughness (scaled)	0.076	0.422	0.181	0.856

**Model fit:** AIC = 47.516, Null deviance = 42.165 (df = 30), Residual deviance = 41.516 (df = 28), N = 31

*Table 8: Logistic regression model predicting Hill/Valley classification from elevation only. Model formula: hill\_binary ~ scaled\_altitude, family = binomial. N = 31 languages*

Parameter	Estimate	Std. Error	z-value	p-value
Intercept	0.341	0.369	0.924	0.355
Elevation (scaled)	0.301	0.396	0.760	0.448

**Model fit:** AIC = 45.548, Null deviance = 42.165 (df = 30), Residual deviance = 41.548 (df = 29), N = 31

*Table 9: Logistic regression model predicting Hill/Valley classification from terrain roughness only. Model formula: hill\_binary ~ scaled\_stdev\_slope, family = binomial. N = 31 languages*

Parameter	Estimate	Std. Error	z-value	p-value
Intercept	0.329	0.366	0.898	0.369
Terrain roughness (scaled)	0.200	0.368	0.544	0.587

**Model fit:** AIC = 45.866, Null deviance = 42.165 (df = 30), Residual deviance = 41.866 (df = 29), N = 31

#### 3.1 Does coding Split as either Hill/Valley affect prediction by ecological variables?

In an alternate analysis, I checked whether coding Split languages as Hill instead affects the prediction by ecological variables. The results show that this reclassification does not improve ecological prediction. While the models show better overall fit (lower AIC values), the ecological predictors themselves remain non-significant across all models. Elevation and terrain roughness still fail to predict societal classification (all  $p > 0.4$ ), reinforcing the conclusion that the ecological foundations of the Hill/Valley framework are not supported in this dataset.

Table 10: Logistic regression model predicting Hill/Valley classification from elevation and terrain roughness with alternative coding. Model formula:  $\text{hill\_valley\_binary} \sim \text{scaled\_altitude} + \text{scaled\_stdev\_slope}$ , family = binomial.  $N = 31$  languages.

Parameter	Estimate	Std. Error	z-value	p-value
Intercept	1.172	0.452	2.591	0.010**
Elevation (scaled)	0.327	0.610	0.536	0.592
Terrain roughness (scaled)	0.495	0.493	1.005	0.315

**Model fit:** AIC = 38.753, Null deviance = 35.403 (df = 30), Residual deviance = 32.753 (df = 28),  $N = 31$

Table 11: Logistic regression model predicting Hill/Valley classification from elevation only with alternative coding. Model formula:  $\text{hill\_valley\_binary} \sim \text{scaled\_altitude}$ , family = binomial.  $N = 31$  languages

Parameter	Estimate	Std. Error	z-value	p-value
Intercept	1.157	0.450	2.573	0.010*
Elevation (scaled)	0.617	0.551	1.120	0.263

**Model fit:** AIC = 37.847, Null deviance = 35.403 (df = 30), Residual deviance = 33.847 (df = 29),  $N = 31$

Table 12: Logistic regression model predicting Hill/Valley classification from terrain roughness only with alternative coding. Model formula:  $\text{hill\_valley\_binary} \sim \text{scaled\_stdev\_slope}$ , family = binomial.  $N = 31$  languages

Parameter	Estimate	Std. Error	z-value	p-value
Intercept	1.141	0.439	2.597	0.009**
Terrain roughness (scaled)	0.637	0.436	1.462	0.144

**Model fit:** AIC = 37.058, Null deviance = 35.403 (df = 30), Residual deviance = 33.058 (df = 29),  $N = 31$

## 4. Sociocultural analyses

### 4.1 Predicting complexity via L2\_status

Linear mixed-effects models were fit with L2\_status as a fixed effect (binary: 1 = present, 0 = not) and a random effect for subfamily. The tables below present the estimated coefficients, standard errors, degrees of freedom, t-values, p-values, and  $R^2$  values for each model.

Table 13: Mixed-effects model predicting cell complexity from L2 status. Model formula:  $\text{scaled\_cell\_complexity} \sim \text{L2} + (1 \mid \text{subfamily})$   $N = 31$  languages; 10 subfamilies

Term	Estimate ( $\beta$ )	Std. Error	df	t	p
Intercept	0.110	0.198	13.30	0.554	0.589
L2	-0.618	0.274	28.22	-2.257	<b>0.032*</b>

**$R^2$ :** Marginal = 0.139; Conditional = 0.222

Table 14: Mixed-effects model predicting form complexity from L2 status. Model formula:  $scaled\_form\_complexity \sim L2 + (1 | subfamily)$   $N = 31$  languages; 10 subfamilies

Term	Estimate ( $\beta$ )	Std. Error	df	t	p
Intercept	-0.008	0.193	16.30	-0.043	0.966
L2	-0.434	0.261	28.40	-1.663	0.107

R<sup>2</sup>: Marginal = 0.079; Conditional = 0.189

Table 15: Mixed-effects model predicting case-marking complexity from L2 status. Model formula:  $scaled\_case\_marking\_complexity \sim L2 + (1 | subfamily)$   $N = 31$  languages; 10 subfamilies

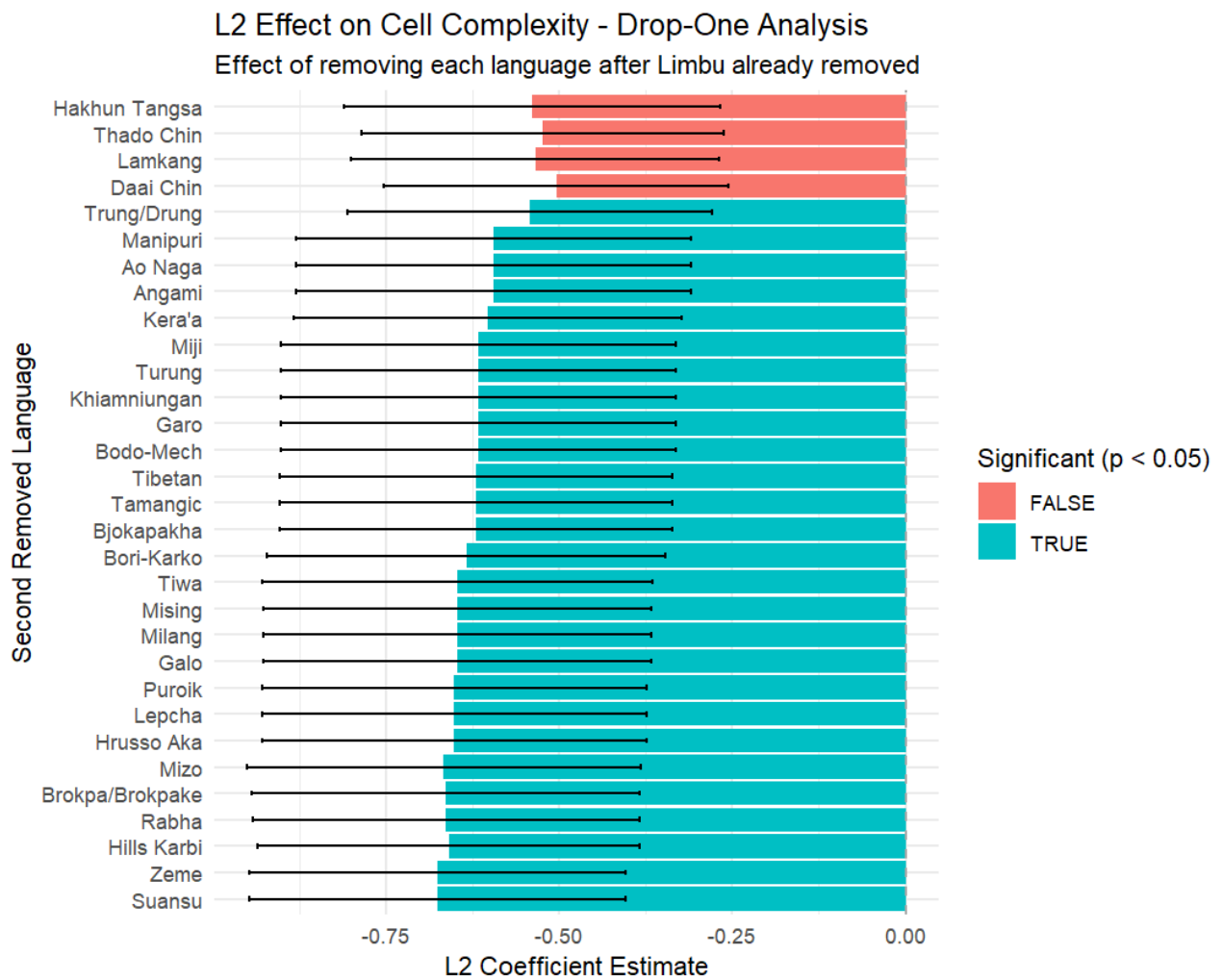
Term	Estimate ( $\beta$ )	Std. Error	df	t	p
Intercept	-0.233	0.251	8.54	-0.929	0.378
L2	0.556	0.360	27.81	1.546	0.133

R<sup>2</sup>: Marginal = 0.072; Conditional = 0.130

## 4.2 Testing robustness of effect of L2\_status on cell complexity

To test whether the significant relationship found between L2\_status and cell complexity was truly robust, I wrote a function to loop through the revised dataset (excluding Limbu) to exclude one language at a time, to see whether the relationship maintained. The code can be found in the associated R-markdown document (*synthesis-code*) on the github, while I provide a graphical representation of the results of this process here.





The median p-value is 0.038, while the coefficient estimates range from -0.676 to -0.504. The result is significant in 27 out of 31 iterations.

### 4.3 Predicting complexity via agricultural intensity

Linear mixed-effects models were fit with **agricultural intensity** as a fixed effect (binary: 1 = Intensive/Irrigated, 0 = Other) and a random effect for subfamily. The tables below present the estimated coefficients, standard errors, degrees of freedom, *t*-values, *p*-values, and *R*<sup>2</sup> values for each model.

*Table 16: Mixed-effects model predicting cell complexity from agricultural intensity. Model formula: scaled\_cell\_complexity ~ agriculture\_binary + (1 | subfamily) N = 31 languages; 10 subfamilies*

Term	Estimate ( $\beta$ )	Std. Error	df	<i>t</i>	<i>p</i>
Intercept	-0.054	0.185	8.09	-0.292	0.777
Agricultural intensity	-0.533	0.381	27.24	-1.398	0.173

*R*<sup>2</sup>: Marginal = 0.057; Conditional = 0.158

Table 17: Mixed-effects model predicting form complexity from agricultural intensity. Model formula: *scaled\_form\_complexity* ~ *agriculture\_binary* + (1 | *subfamily*) *N* = 31 languages; 10 subfamilies

Term	Estimate ( $\beta$ )	Std. Error	df	t	p
Intercept	-0.104	0.177	10.94	-0.588	0.568
Agricultural intensity	-0.506	0.350	27.72	-1.446	0.159

R<sup>2</sup>: Marginal = 0.059; Conditional = 0.185

Table 18: Mixed-effects model predicting case-marking complexity from agricultural intensity. Model formula: *scaled\_case\_marking\_complexity* ~ *agriculture\_binary* + (1 | *subfamily*) *N* = 31 languages; 10 subfamilies

Term	Estimate ( $\beta$ )	Std. Error	df	t	p
Intercept	-0.046	0.207	5.39	-0.222	0.832
Agricultural intensity	+0.247	0.501	27.21	+0.494	0.626

R<sup>2</sup>: Marginal = 0.008; Conditional = 0.022

#### 4.4 Predicting complexity via speaker population

Linear mixed-effects models were fit with **scaled\_population** as a fixed effect (binary: 1 = Intensive/Irrigated, 0 = Other) and a random effect for subfamily. The tables below present the estimated coefficients, standard errors, degrees of freedom, *t*-values, *p*-values, and R<sup>2</sup> values for each model.

Table 19: Mixed-effects model predicting cell complexity from speaker population size. Model formula: *scaled\_cell\_complexity* ~ *scaled\_population* + (1 | *subfamily*) *N* = 31 languages; 10 subfamilies

Term	Estimate ( $\beta$ )	Std. Error	df	t	p
Intercept	-0.054	0.185	8.09	-0.292	0.777
Agricultural intensity	-0.533	0.381	27.24	-1.398	0.173

R<sup>2</sup>: Marginal = 0.027; Conditional = 0.134

Table 20: Mixed-effects model predicting form complexity from speaker population size. Model formula: *scaled\_form\_complexity* ~ *scaled\_population* + (1 | *subfamily*) *N* = 31 languages; 10 subfamilies

Term	Estimate ( $\beta$ )	Std. Error	df	t	p
Intercept	-0.104	0.177	10.94	-0.588	0.568
Agricultural intensity	-0.506	0.350	27.72	-1.446	0.159

R<sup>2</sup>: Marginal = 0.009; Conditional = 0.114

Table 21: Mixed-effects model predicting case-marking complexity from speaker population size. Model formula: *scaled\_case\_marking\_complexity* ~ *scaled\_population* + (1 | *subfamily*) *N* = 31 languages; 10 subfamilies

Term	Estimate ( $\beta$ )	Std. Error	df	t	p
Intercept	-0.046	0.207	5.39	-0.222	0.832
Agricultural intensity	0.247	0.501	27.21	0.494	0.626

R<sup>2</sup>: Marginal = 0.014; Conditional = 0.042

### 4.5 Predicting complexity via political organization

Linear mixed-effects models were fit with **political\_organization** as a fixed effect (binary: 1 = State, 0 = Acephalous/other) and a random effect for subfamily. The tables below present the estimated coefficients, standard errors, degrees of freedom, *t*-values, *p*-values, and *R*<sup>2</sup> values for each model.

Table 22: Mixed-effects model predicting cell complexity from political organization. Model formula: *scaled\_cell\_complexity* ~ *political\_organization* + (1 | *subfamily*). *N* = 31 languages; 10 subfamilies

Term	Estimate (β)	Std. Error	df	t	p
Intercept	−0.075	0.187	7.69	−0.403	0.698
Political organization	−0.492	0.428	28.47	−1.149	0.260
R <sup>2</sup> : Marginal = 0.045; Conditional = 0.176					

Table 23: Mixed-effects model predicting form complexity from political organization. Model formula: *scaled\_form\_complexity* ~ *political\_organization* + (1 | *subfamily*)*N* = 31 languages; 10 subfamilies

Term	Estimate (β)	Std. Error	df	t	p
Intercept	−0.122	0.178	10.46	−0.685	0.508
Political organization	−0.482	0.393	28.57	−1.226	0.230
R <sup>2</sup> : Marginal = 0.045; Conditional = 0.176					

Table 24: Mixed-effects model predicting case-marking complexity from political organization. Model formula: *scaled\_case\_marking\_complexity* ~ *political\_organization* + (1 | *subfamily*)*N* = 31 languages; 10 subfamilies

Term	Estimate (β)	Std. Error	df	t	p
Intercept	−0.013	0.204	4.38	−0.062	0.953
Political organization	0.049	0.553	28.90	0.088	0.930
R <sup>2</sup> : Marginal = 0.000; Conditional = 0.031					

### Bibliographic References

Akamatsu, Yoshio. 2015. *Study on Change of Agricultural Livelihood and Its Contemporary Problems in Eastern Bhutan*. Ehime University (愛媛大学) PhD Dissertation. <https://cir.nii.ac.jp/crid/1910020910687471360>. (7 March, 2025).

Boro, Krishna. *A Grammar of Hakhun Tangsa*. United States -- Oregon: University of Oregon PhD Dissertation. <https://www.proquest.com/docview/1949396335/abstract/A397B9623F0545FFPQ/1>. (2 April, 2024).

Boro, Purnima. 2024. Planting Dreams and Harvesting Smiles: Traditional Paddy Practices among the Bodos. *International Journal of Engineering Development and Research*. IJEDR (www.ijedr.org) 12(4). 177–184.

Burling, Robbins. 2003. *Language of the Modhupur Mandi (Garó), Volume 1*. Ann Arbor: Michigan Publishing, University of Michigan Library. <https://doi.org/10.3998/spobooks.bbv9808.0001.001>.

Burling, Robbins. 2017. *Rengsanggri: Family and Kinship in a Garo Village*. Philadelphia: University of Pennsylvania Press.

Chan, Nyein & Shinya Takeda. 2016. The Transition Away From Swidden Agriculture and Trends in Biomass Accumulation in Fallow Forests: Case Studies in the Southern Chin Hills of Myanmar. *Mountain Research and Development*. International Mountain Society 36(3). 320–331. <https://doi.org/10.1659/MRD-JOURNAL-D-14-00083.1>.

Chanu, Sapam Sangita. 2017. *A Descriptive Grammar Of Zeme*. Silchar: Assam University PhD Dissertation. <https://shodhganga.inflibnet.ac.in:8443/jspui/handle/10603/293382>. (10 December, 2024).

Chelliah, Shobhana Lakshmi. 1997. *A Grammar of Meithei*. Berlin: Walter de Gruyter.

Chelliah, Shobhana, David Peterson, Tyler Utt, Evaline Blair & Sumshot Khular. 2019. Lamkang verb conjugation. *Himalayan Linguistics* 18(1). <https://doi.org/10.5070/H918143199>.

Chhangte, Lalnunthangi. 1993. *Mizo syntax*. United States -- Oregon: University of Oregon PhD Dissertation. <https://www.proquest.com/docview/304094920/abstract/AEDAE49C4598483FPQ/1>. (27 March, 2025).

Chutia, Humen. 2020. Traditional Agricultural Practices of the Mising Community of Assam. *RESEARCH REVIEW International Journal of Multidisciplinary* 05(06). 52–55. <https://doi.org/10.31305/rrijm.2020.v05.i06.010>.

Coupe, A. R. 2008. A Grammar of Mongsen Ao: In *A Grammar of Mongsen Ao*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110198522>.

Dawson, Virginia. 2020. Person marking in Tiwa. *Linguistics of the Tibeto-Burman Area*. John Benjamins Publishing Company 36(1). 91–99. <https://doi.org/10.3316/informit.323731798654843>.

Dawson, Virginia Ellen. 2020. *Existential Quantification in Tiwa: Disjunction and Indefinites*. United States -- California: University of California, Berkeley PhD Dissertation. <https://www.proquest.com/docview/2462470108/abstract/2337A142CB424FB7PQ/1>. (2 September, 2024).

DeLancey, Scott. 2003. Lhasa Tibetan. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages* (Routledge Language Family Series), 270–288. 1. publ. in paperback. London: Routledge.

Doloi, Gargi, Bijoy Timung & Kongkon Bordoloi. 2024. The Tiwa Community of Assam: A Historical and Cultural Study. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. <https://doi.org/10.2139/ssrn.4935389>.

Driem, George van. 2011. *A Grammar of Limbu*. Walter de Gruyter.

D'Souza, V. A. 2021. *Aspects of Hrusso Aka phonology and morphology*. Oxford: University of Oxford PhD Dissertation. <https://ora.ox.ac.uk/objects/uuid:1f47a3ab-a444-4166-a4d7-3a26c6532610>. (10 December, 2024).

Ezung, Mimi Kevichüsa. 2014. Case and Tone Systems in Tenyidie. *Interdisciplinary Journal of Linguistics* 7. 51–58.

Goswamy, Nabanita, Joyraj Kalita & Sayantani Banerjee. 2024. Forest-Dwelling Rabha Community in Alipurduar, West Bengal: An Ethnographic Study. *International Journal of Recent Advances in Multidisciplinary Topics*. IJRAMT 5. <https://doi.org/10.5281/ZENODO.13366828>.

Grollmann, Selin. 2020. *A Grammar of Bjokapakha*. Leiden: Brill.

Hansepi, Laxmi & Rena Laisram. 2022. Karbi Women and Environmental Conservation. *Grassroots Journal of Natural Resources* 05(03). 1–17. <https://doi.org/10.33002/nr2581.6853.050301>.

Haokip, Marykim. 2014. *Grammar of Thadou-Kuki: A Descriptive Study*. New Delhi: Jawaharlal Nehru University PhD Dissertation.

Haokip, Pauthang & Daimalu Brahma. 2018. Differential Case Marking in Bodo. *Journal of the Southeast Asian Linguistics Society* 11(1). 1–15.

Ivani, Jessica K. 2023. Suansu language from northeastern India: A field report. *Linguistics of the Tibeto-Burman Area*. John Benjamins 46(1). 138–163. <https://doi.org/10.1075/ltba.22005.iva>.

J H Hutton M A. 1921. *The Angami Nagas*. London: Macmillan and Company, limited. <http://archive.org/details/in.ernet.dli.2015.81454>. (23 June, 2025).

Joseph, U. V. 2007. *Rabha* (Brill's Tibetan Studies Library, Languages of the Greater Himalayan Region v. 5/1). Leiden ; Boston: Brill.

Konnerth, Linda Anna. 2014. *A grammar of Karbi*. University of Oregon PhD Dissertation. <https://www.proquest.com/docview/1535090744/abstract/EB991E2644FB4E91PQ/1>. (12 May, 2025).

Lalrempuii, C. 2005. *Morphology of the Adi language of Arunachal Pradesh*. MA thesis. North Eastern Hill University <https://shodhganga.inflibnet.ac.in/handle/10603/60955>. (5 February, 2025).

Lalthansanga, Chawngthu. 2018. Mizo Values In Agricultural Sustenance: Lessons From Mizoram. In, 188–196. Atlantis Press. <https://doi.org/10.2991/aapa-18.2018.18>.

Lieberherr, Ismael. 2020. *A grammar of Bulu Puroik*. Lieberherr, Ismael (2017) *A grammar of Bulu Puroik thesis*. Bern: Universität Bern PhD Dissertation. <https://doi.org/10.24442/boristheses.2251>.

Limbu, Ramesh. 2013. Exploring Indigenous Knowledge System and Limbu Cultural Identity in Nepal. *Journal of University of Human Development* 2. 122–137.

Longkumer, Arkotong. 2007. Religious and Economic Reform: The Gaidinliu Movement and the Heraka in the North Cachar Hills. *South Asia: Journal of South Asian Studies* 30(3). 499–515. <https://doi.org/10.1080/00856400701714096>.

Meitei, Irom Luckychand & Gurumayum Jadumani Sharma. 2023. An Analytical Study on the Inseparable Relations Between Food, Population, Arable Land and Economy in Manipur from the Perspective of Sustainability. In Subhash Anand, Madhushree Das, Rituparna Bhattacharyya & R. B. Singh (eds.), *Sustainable Development Goals in Northeast India: Challenges and Achievements*, 539–551. Singapore: Springer Nature. [https://doi.org/10.1007/978-981-19-6478-7\\_28](https://doi.org/10.1007/978-981-19-6478-7_28).

- Mills, James Philip. 1926. *The Ao Nagas*. London: Macmillan and Company, limited.
- Modi, Yankee. 2017. *The Milang Language: Grammar and Texts*. Universität Bern PhD Dissertation.
- Morey, Stephen. 2010. *Turung: a variety of Singpho language spoken in Assam* (Pacific Linguistics). Vol. 614. Canberra: Research School of Pacific and Asian Studies, Australian National University.
- Nimachow, Gibji, R. C. Joshi & Oyi Dai. 2011. Role of indigenous knowledge system in conservation of forest resources—A case study of the Aka tribes of Arunachal Pradesh. *IJTK* Vol.10(2) [April 2011]. NISCAIR-CSIR, India. <http://nopr.niscpr.res.in/handle/123456789/11503>. (27 February, 2025).
- Owen-Smith, Thomas. 2015. *Grammatical relations in Tamang, a Tibeto-Burman language of Nepal*. London,: SOAS University of London phd. <https://doi.org/10.25501/SOAS.00023664>.
- Peck, Naomi & Uta Reinöhl. personal communication. Kera'a questionnaire.
- Perlin, Ross. 2020. A Grammar of Trung [HL Archive 8]. *Himalayan Linguistics* 0(0). <https://doi.org/10.5070/H918244579>.
- Plaisier, Heleen. 2007. *A Grammar of Lepcha* (Languages of the Greater Himalayan Region). Vol. 5. Leiden and Boston: Leiden: E.~J.~Brill. <https://stedt.berkeley.edu/stedt-cgi/rootcanal.pl/source/HP-GL>.
- Post, Mark William. 2007. *A Grammar of Galo*. La Trobe University PhD Dissertation.
- Prasad, Bal Ram. 1991. *Mising Grammar* (CIIL Grammar Series). Mysore: Central Institute of Indian Languages.
- Raj, Saravanan. 2010. Traditional knowledge, innovation systems and democracy for sustainable agriculture: A case study on Adi tribes of Eastern Himalayas of North-East India. In Emilie COUDEL, Hubert DEVAUTOUR, Christophe-Toussaint SOULARD & Bernard HUBERT (eds.), *Innovation and Sustainable Development in Agriculture and Food*, 10 p. Montpellier, France: Cirad-Inra-SupAgro. <https://hal.science/hal-00523309>. (27 February, 2025).
- Ramjuk, Tame. 2018. Understanding Tribal Life and Livelihood in Changing Times: A Study on the Puroiks of Arunachal Pradesh. *International Journal of Research and Analytical Reviews* 5(3). 604–616.
- Singh, L. Jitendro & Dhrupad Choudhury. 2015. Fallow-Management Practices Among The Tangkhuls Of Manipur: Safeguarding provisioning and regulatory services from shifting-cultivation fallows. In *Shifting Cultivation and Environmental Change*. Routledge.
- So-Hartmann, Helga. 2008. *A Descriptive Grammar of Daai Chin*. England: University of London, School of Oriental and African Studies (United Kingdom) Ph.D. <https://www.proquest.com/docview/2185733381/abstract/E058A5DAFEDC4DF8PQ/1>. (12 May, 2025).
- Tham, Keen & Kellen Parker van Dam. personal communication. Khamniungan questionnaire.

Thapa, Kishore & Sanjita Allay. 2021. Lepcha agriculture on Primitive Agro-system perspective. *Science of Eastern Himalaya Journal* 1(1). 77–82.

Thounaojam, Harimohon & Shobhana L. Chelliah. 2007. The Lamkang language: grammatical sketch, texts and lexicon. *Linguistics of the Tibeto-Burman Area* 30(1).  
<https://doi.org/10.32655/LTBA.30.1.01>.

Wangdi, Pema. 2021. *A Grammar of Brokpa: a Trans-Himalayan language of Bhutan*. James Cook University phd. <https://doi.org/10.25903/58yg-we65>.

Weedall, Christopher. 2021. A grammar of West Kameng Sajolang (Miji): (Upper Dzang and Khellong villages). <http://hdl.handle.net/1885/251256>. (4 February, 2025).