Sample Field Report: Assamese Language Documentation

By Divya Kadav

Note on the Data

The material presented in this report is a sample of the wider fieldwork, transcription, and annotation work carried out during my dissertation project on Assamese (Ujoni Variety).

The complete project involved:
- **Fieldwork sessions** with the information, using elicitation techniques i.e. English prompts with Assamese responses.
- **Phonetic transcription** of the data using the IPA Chart.
- **Morphological segmentation and interlinear glossing**, following standard convention.
- **Annotation and labelling** with the help of multiple datasets.
- **Documentation and analysis** of Grammatical Features.
- **Dissertation writing and presentation** of the data, where these methods and analyses were systematically applied and archived.

What is included here represents **only a portion** of the collected data. These examples serve to demonstrate the **methodology, documentation style, and analytical approach** applied throughout the dissertation. The full dataset is more extensive and has been systematically archived.

This sample illustrates my ability to:
1. Collect and manage primary linguistic data.
2. Transcribe and annotate unfamiliar language structures.
3. Apply linguistic tools (ELAN, IPA, LaTeX, MS Word/ Excel) for systematic documentation.
4. Analyse features of Assamese grammar and usage with reference to broader linguistic frameworks.

Field Report: Assamese Language Documentation

Researcher Information

- **Researcher Name**: Divya Kadav

- **Institution/Affiliation**: University of Mumbai

- **Location of Fieldwork**: Indo-Canadian Room, Ranade Bhavan

- **Language Studied**: Assamese

Informant Information

- **Informant Name:** Priyakhi Khanikar

- **Age:** 28

- **Hometown:** Duliajan, Dibrugarh District, Assam

- **Dialect/Variety:** Ujoni (Upper) Assamese

- **Educational Background:**

  - Schooling in Assam

  - Master's and M.Phil. at Delhi University

  - Primary medium of education: English

Language Background

- **Mother Tongue**: Assamese

- **Other Languages Spoken at Home**: Assamese, Bengali, Hindi, English

	Assamese	Bengali	Hindi	English
Read	✔	✔	✔	✔
Write	✔	✔	✔	✔
Understand	✔	✔	✔	✔

Background

The informant's **early education** began with three years at the basic level in Assamese. Over time, she developed familiarity with Bengali, since its script is highly similar to Assamese. Exposure to Bengali films and literature further influenced her, and she formally began studying the language in the **8th Standard**.

She attended **Vivekananda Kendriya Vidyalaya**, where she studied **English, Hindi, and Sanskrit** alongside Assamese. As a result, her linguistic repertoire expanded early, allowing her to engage with multiple languages across different domains.

The informant actively uses **all four languages (Assamese, Bengali, Hindi, and English)** for reading media and academic articles. In daily life, Assamese remains the **primary language** spoken with family and friends, especially in her hometown of **Duliajan, Dibrugarh District**. She reads newspapers in Assamese and predominantly uses it in informal contexts.

However, her language use also demonstrates **code-switching and multilingual flexibility**:

- With migrants, she adapts to **Bengali, Hindi, or Assamese**, depending on the interlocutor's preference.

- Assamese is spoken within her community, while Hindi and Bengali are used more often in urban and migratory interactions.

Culturally and geographically, the **Brahmaputra River divides Assam** into Upper (*Ujoni*) and Lower (*Namoni*) Assamese varieties. The informant identifies with **Upper Assamese** and notes that the script is consistent across both regional varieties.

Equipment Used

- **Notebook** → for field note-taking

- **ELAN software** → for transcription and annotation

- **IPA Chart** → for phonetic transcription reference

- **Microsoft Word & Excel** → for documentation and data organisation

- **LaTeX** → for dissertation writing and formatting

Methodology

The data collection process for each session followed a structured elicitation-based approach:-

1. The **researcher provided sentences in Englis**h to the informant as stimuli. These prompts were carefully selected to cover a range of syntactic and semantic structures.
2. The **informant produced the equivalent sentences in Assamese**. This ensured that the elicited data reflected Natural language usage while remaining aligned with the researcher's target structures.
3. The Assamese responses were transcribed using the **International Phonetic Alphabet (IPA)**, with the IPA chart serving as a reference to ensure accuracy of segmental representation.
4. Each sentence was segmented into morphemes. **Standard interlinear glossing conventions** were applied, including labels for grammatical categories. (e.g., LOC, N, SG, ADJ, 3PL).
5. The glossed sentences were rendered into idiomatic English to provide a clear semantic equivalent of the Assamese data. We did this by doing free translation initially, then matching it with the English Sentence.

6. Using multiple datasets, the transcribed sentences were cross-referenced for consistency in annotation. This process enabled the batch of students to collaboratively establish labelling standards and refine glossing accuracy.

7. Notes were recorded in a field notebook, with further transcription and annotation completed in ELAN, Microsoft Word, and Excel.

Glossary of Abbreviations

- **1SG** – first person singular

- **1PL** – first person plural

- **3** – third person

- **3PL** – third person plural

- **AUX** – auxiliary verb

- **ADJ** – adjective

- **ALL** – allative case (movement towards)

- **CLF** – classifier

- **DAT** – dative case (marking recipient or goal)

- **ERG** – ergative case (agent of a transitive verb)

- **GEN** – genitive case (possession)

- **INF** – infinitive

- **INST** – instrumental case (means/by/with)

- **LOC** – locative case (place, location)

- **N.PL** – plural noun

- **N.SG** – singular noun

- **PST** – past tense

- **PROG** – progressive aspect (ongoing action)

- **QUANT** – quantifier

- **SG** – singular

Session 1

Topic: Animals
Date: 01/08/2024
Elicitors: Avinash and Renuka

1.dʒɔntu - animal

habi=t	bɔhut	bɔnoɾ-ija	dʒɔntu	as-e
jungle=LOC	many	wild-ADJ	animal	be-3
There are many wild animal in the jungle.				

habi-t	bɔhut	bɔnor-ija	dʒɔntu	tʰak-e
jungle=LOC	many	wild-ADJ	animal	be-3
There live many wild animal in the jungle.				

habi=t	ɛ-ta	dʒɔntu	ase
jungle=LOC	one-CLF	animal	be-3
There is an animal in the jungle.			

2. goɹu - cow

ama=ɹ	ɛ-ta	goɹu	ase
1PL=GEN	one-CLF	cow	AUX
We have a cow.			

mʊ=l	ɛ-ta	mɔh	ase
1SG=GEN	one-CLF	buffalo	AUX
I have a buffalo.			

3.kuːkˈuɹ - Dog

ama=ɹ	du-ta	kuːkuɹ	ase
1PL=GEN	two-CLF	dog	AUX
We have two dogs.			

ama=ɹ	ɛ-ta	kuːkuɹ	ase
1PL=GEN	one-CLF	dog.SG	AUX
We have a dog.			

kuːkuɹ	gɔɾosia	dʒɔntu
dog	domestic	animal
Dogs are domestic animals.		

ama=ɹ	kuːkuɹ	ase
1PL=GEN	dog	AUX
We have dogs.		

ama=ɹ	bɔhut	kuːkuɹ	ase
1PL=GEN	many	dog	AUX
We have many dogs.			

4.gadʰəv - donkey

gadʰəv-bʊl=e	gɔdʰul	bɔstu	koɹija-bo	paɹ-e
donkey-PL=ERG	heavy	luggage	carry-INF	can-3
Donkey can carry heavy luggage.				

gadʰəv-bʊl=e	gɔdʰul	bɔstu	koɹi-je
donkey-PL=ERG	heavy	luggage	carry-3
Donkeys carry heavy luggage.			

5.haṯi - Elephant

hati-bʊl	xoŋal	dʒɔntu	hoi
Elephant-PL	ferocious	animal	be
Elephants are ferocious animal.			

6.sagoli- goat

sagoli-ɛ-ta	bɔn=ot	soɹ-i	as-e
goat-one-CLF	garden=LOC	graze-PROG	be-3
A goat is grazing in the yard.			

sagoli-bʊl	bɔno=t	soɹ-i	as-e
goat-PL	garden=LOC	graze-PROG	be-3
Goats are grazing in the yard.			

7. Lion - xiŋˈho

xiŋˈho-bʊl	xoŋal	dʒɔntu
lion-PL	ferrocious	animal
Lions are ferocious animals.		

8.bandɔɹ - Monkey

kali	mɔi	ɛ-ta	bandɔɹ	dekʰi-l-u
yesterday	1SG	one-CLF	monkey	see-PST-1
I saw a monkey yesterday.				

kali	mɔi	bandɔɹ	ɛ-ta	dekʰi-l-u
yesterday	1SG	monkey	one-CLF	see-PST-1
I saw a monkey yesterday.				

9. nigʰoni - Mouse

hɛi	gʰɔɹ-tu=t	bɔhut	nigʰoni	as-e
that	house-CLF=LOC	many	mouse	be-3
There are a lot of mice in that house				

10.Snake -xap

mʊ=l	xap=oloi	bʰɔi	lage
1SG=GEN	snake=ALL	scare	feel
I am scared of snakes			

xap	ɛ-dal	kʊmɔl	dʒibɔl
snake	one-CLF	soft	creature

Snake is a soft animal

11. claw - tʰeŋ

sɔɹai-dʒoni=ɹ	gʰɔɹ-tu-t	kʰub	dʰun-ija
bird-CLF=GEN	claw-PL-CLF	very	pretty-ADJ
The bird's claws are very pretty.			

sɔɹai-dʒuni=ɹ	tʰen-tu	kʰub	dʰun-ija
bird-CLF=GEN	claw-CLF	very	pretty-ADJ
This bird's claws are very pretty.			

sɔɹai-dʒuni=ɹ	tʰen-tu	bɔhut	danoɹ
bird-CLF=GEN	claw-CLF	very	big
This bird's claw is very big.			

xiŋˈho-tu=je	ˈhɔlin-tu=k	tʰeŋ=ele	maɹi-j-e
tiger-CLF=ERG	deer-CLF=DAT	claw=INST	hit-PST-3
The tiger hit the deer with it's claw.			

Session 2

Topic: Birds
Date: 05/08/2024
Elicitors: Avinash and Renuka

sɔɹai-ɛ-dʒoni	gos-zupa=t	boh-i	as-e
bird-one-CLF	tree-CLF=LOC	sit-PROG	be-3
A bird is sitting on the tree.			

sɔɹai-ɛ-dʒoni	uɹ-i	as-e
bird-one-CLF	fly-PROG	be-3
A bird is flying.		

sɔɹai-dʒoni	uɹ-i	as-e
bird-CLF	fly-PROG	be-3
The bird is flying.		

sɔɹai-bʊl	uɹ-i	as-e
bird-PL	fly-PROG	be-3
The birds are flying.		

sɔɹai	uɹ-i	ase
bird	fly-PROG	be-3
Birds are flying.		

sɔɹai-bʊl	gɔs=ot	boh-i	as-e
bird-PL	tree=LOC	sit-PROG	be-3
Birds are sitting on the tree.			

do-dʒoni-sɔɹai	uɹ-i	as-e
two-CLF-bird	fly-PROG	be-3
Two birds are flying		

mɔi	ɛ-dʒoni	gʰɔnsilika	aɹu	bɔhut	kauɹi	dekʰi-l-u
1SG	one-CLF	sparrow	and	many	crow	see-PST-1
I saw a sparrow and many crows.						

kaoɹi-bʊl	kola/kula
crow-PL	black
The crows are black	

hɛi	kaoɹi-dʒoni	kola
that	crow-CLF	black
That crow is black.		

mɔɹa	ama=ɹ	dʒatija	sɔɹai
peacock	1PL=GEN	national	bird
Peacock are our national bird.			

moɹa=ɹ	pakʰi	kʊmol
peacock=GEN	feather	soft
Peacocks feather are soft		

ama=ɹ	gʰoɹ=oɹ	usoɾ-ot	bɔhut	gʰɔnsilika	ase
1PL=GEN	house=GEN	near=LOC	many	sparrow	be
There are many sparrow near my house.					

ama=ɹ	gʰɔɹ=oɹ	gʰɔnsilika	ase
1PL=GEN	house=GEN	sparrow	be
There are sparrows near our house.			

mɔi	ɛ-dʒoni	gʰɔnsilika	aɹu	bɔhut	kauɹi	dekʰi-l-u
1SG	one-CLF	sparrow	and	many	cɾow	see-PST-1
I saw a sparrow and many crows.						

Session 3

Topic: Colours
Date: 08/08/2024
Elicitors: Palakshi, Divya, Rishi, Tithi

tai=ɹ	sula-tu=l	nisina	kola
3FSG=GEN	top-CLF=GEN	like	black
black like heɾ top			

xi	ɹɔŋbiɹɔŋ=oɹ	kapuɹ	pindʰ-ise
3MSG	colourful=GEN	cloth	wear-PROG
He is wearing colourful clothes			

1	dark red	datʰ ɹoŋa
2	red	ɹɔŋa
3	orange	komola (colloquial) geɾua (formal)
4	dark yellow	datʰ haɾodʰija
5	light yellow	patəl haɾodʰija
6	yellow	haɾodʰija
7	light green	bʰato heudʒija
8	green	heudʒija
9	dark green	datʰ heudʒija
10	light blue	axakʰi nila
11	blue	nila
12	dark blue	datʰ nila
13	violet	benunija

14	dark pink	datʰ gulopija
15	pink	gulopija
16	black	kola
17	white	boga (common term) kʰu'bʰro
18	light brown	matiɾomija
19	golden	xun ɾomija
20	colour	ɹɔŋa
21	colours	ɹɔŋ-bʊl

dark	datʰ
orange (fruit)	komola
light	patəl
mustard flower	halodʰija
parrot	bʰato
sky	axak
like sky	axaki
brinjal	beŋena
gold	xun
colour = colour	ɹɔŋ = boɾonija
black like coal	koiːla kola
white hair	poka suli
gray like ash	sai boɹonija
off white	joha saul boɹonija saul (ɾice)
wheatish skin colour	mitʰa boɹonija
British white skin	dʰehla boga
skin	sal

white	boga
whiter than white	dʰehla
ripe banana leaf green colour (used for eating)	kopotija heudʒija
leaf	potija
leaf	patʰ
catfish colour (light brown skin colour)	maguɹ
colourful	ɹɔŋbiɹɔŋ
cloud	meg
cotton	kopa
soft like cotton	kopahi kumol
black cloud	kola meg
flower	pʰul

Analysis

Classifiers	
-kʰon	countable concrete object
-kʰini	mass noun
-tu	small, countable, animate
-ta	Numerals
-dʒoni	birds and a female human
-dʒon	male human
-dal	long, thin objects
-goruki	human male/female

Cases	
-e	nominative
-k	accusative
-ɹ / l	Genitive
-eɹe / ele / ile	instrumental
-t	Locative
-loi	Dative/allative
poɹa	from

Word	Components	Structure Type
Golpata	gol (neck) + pata (leaf)	Noun-Noun Compound
Nakpʰuli	nak (nose) + pʰuli (ring)	Noun-Noun Compound
Kanpʰuli	kan (ear) + pʰuli (ring)	Noun-Noun Compound
Kʰelpotʰal	kʰel (play) + potʰal (field/ground)	Noun-Noun Compound
Bʰeɹasagoli	bʰeɹa (sheep) + sagoli (goat)	Noun-Noun Compound
Loɹasuwali	loɹa (boy/child) + suwali (girl/child)	Noun-Noun Compound
Galigalaz	gali (abuse) + galaz (abuse)	Reduplication Compound

Stress Markers

Syllable Adjectives

	Stress on 1st Syllable	Stress on 2nd Syllable
Non-Derived Words	/ˈstjuː.pɪd/	/ɪnˈseɪn/
	/ˈkruː.əl/	/naɪˈiːv/
	/ˈprɪt.i/	/əˈweər/
	/ˈʌg.li/	/daɪˈrekt/
	/ˈɔː.fəl/	/suːˈpɜːb/
	/ˈsɪm.pəl/	/fəˈlɔːn/
	/ˈpɜː.fekt/	/səˈriːn/
		/əˈləʊn/
Derived Words	/ˈkʌn.ɪŋ/	/məˈriːn/
	/ˈhəʊp.fəl/	/əˈsliːp/
	/ˈpiːs.fəl/	/əˈlaɪv/
	/ˈfʌn.i/	/ɪkˈspəʊzd/
	/ˈlʌv.ɪŋ/	/prɪˈpeəd/
	/ˈtʃɑː.mɪŋ/	

o	Note Section	ENG WORD	ASSAMESE		SENTENCE			

| S2 | Session 2
Birds
05/08/2024
Elicitor - Avinash and Renuka | | | | | | | |

| 14 | sɔɹai gidʒoni - The bird
(Used for the specific kind of bird)
sɔɹai edʒoni - Any bird | Bird | sɔɹai | | | | | |

				sɔɹai-e-dʒoni	gos-ʐupa=t	boh-i	as-e
				bird-one-CLF	tree-CLF=LOC	sit-PROG	be-3
				A bird is sitting on the tree.			

				sɔɹai-ɛ-dʒoni	uɹ-i	as-e
				bird-one-CLF	fly-PROG	be-3
				A bird is flying.		

	a = ɛ-dʒoni the = dʒoni			sɔɹai-dʒoni	uɹ-i	as-e
				bird-CLF	fly-PROG	be-3
				The bird is flying.		

Figure: Sample from Session 2 (Birds) showing elicited Assamese data with IPA transcription, morpheme-by-morpheme glossing, and free English translation.