Field reports on focussed elicitation of epistemics in Trans-Himalayan languages

Carl Bodnaruk

THE UNIVERSITY OF
SYDNEY

Introduction

- Report from fieldwork conducted on Chocangaca and Lhokpu
  - Both fairly brief, minimal opportunity for language learning
  - Both underdescribed, very limited published material (Tournadre and Rigzin 2015, Grollmann and Gerber 2018)
- Focus on usage of elicitation activities to generate naturalistic data
  - Specifically, given the circumstances, feasibility when larger research project is not an option
- What should work in theory?
  - What actually worked in practise?
  - What had an impact on the effectiveness?
- Conclusions: What is useful, but also when and where
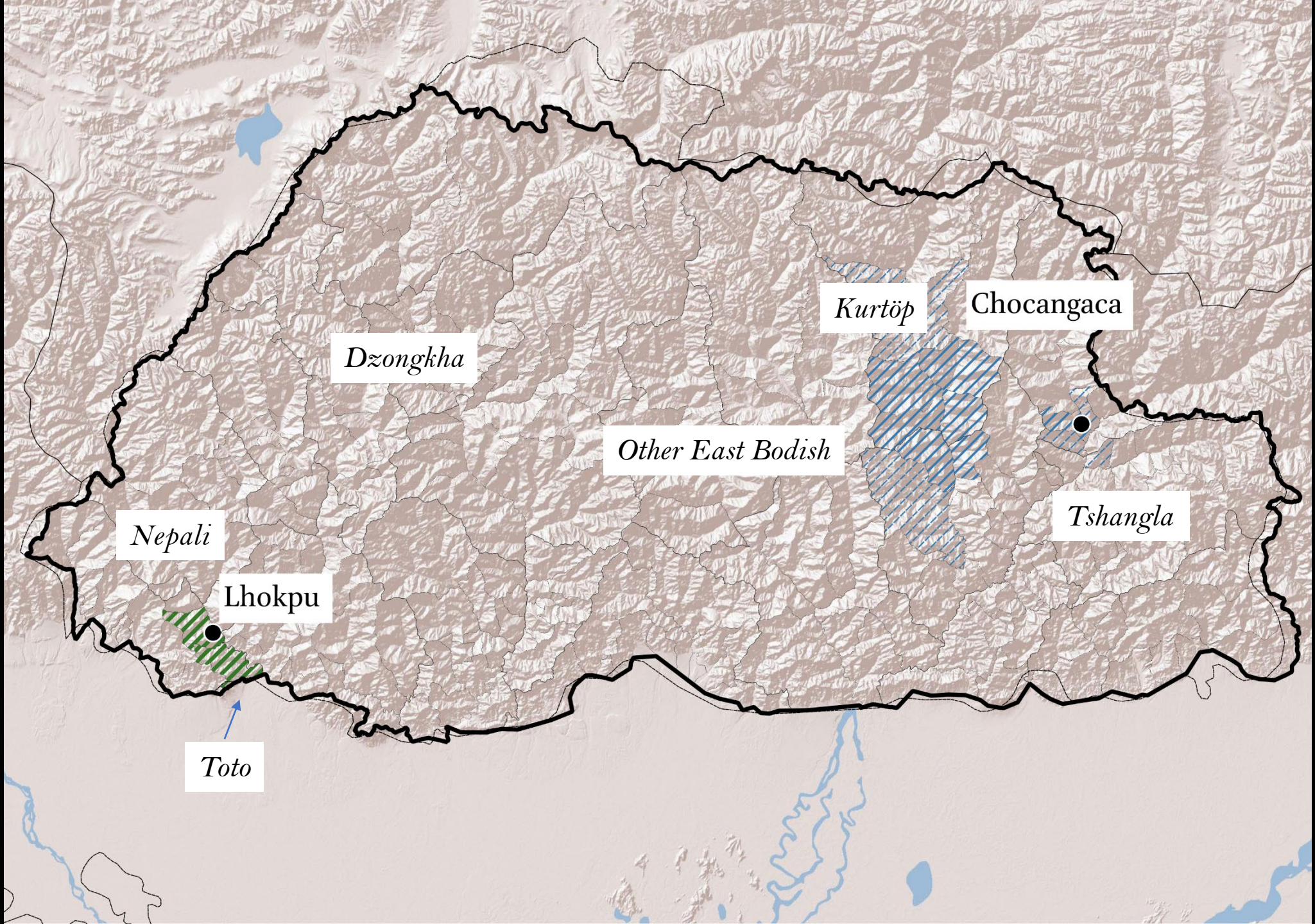
Chocangaca

- Tibetic, not Dzongkha variety or descendant
  - Phonologically more conservative, also strong influence other non-Tibetic TH languages in Bhutan (East Bodish)
- Spoken in a primary area in Central Bhutan, as well as a smaller second area further east, split possibly occurred post-17th C (Bodt 2012)
  - Speakers here originate from Eastern area
- Data collected throughout 13-week Field Methods unit at Sydney Uni

Lhokpu

- Trans-Himalayan, exact subgrouping unclear
  - Potentially more closely related to Dhimal and Toto (Grollmann and Gerber 2018)
- Spoken in a few villages in South-Western Bhutan
- Data collected from speakers in Jigme village and over WhatsApp

Kurtöp

Chocangaca

Dzongkha

Other East Bodish

Nepali

Tshangla

Lhokpu

Toto
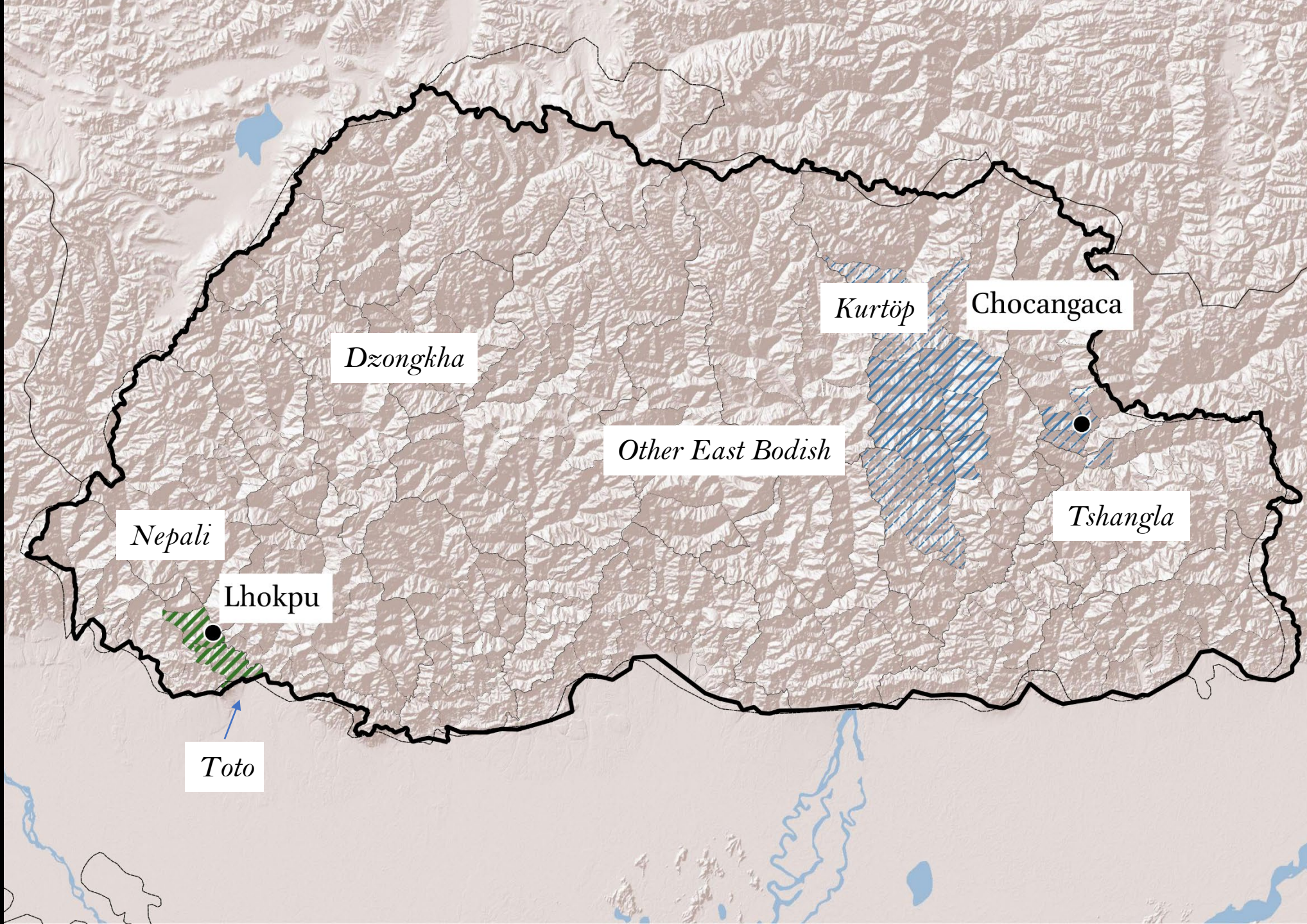
Theoretical Foundations

- Epistemics
  - Related to knowledge and its representation
  - Here: Strategies of communicating meta-propositional information about knowledge
    - What do I know, or how well do I know it?
    - How do I know it? When?
    - How do I feel about this?
    - Same for the addressee
  - Establishing shared ground between SAPs
    - Or: disputing shared ground?
- Deictic, dependent on context
- Where present in questions, tend to take addressee-perspective

Theoretical Foundations

- Fundamental goals with our data collection:
  - Accurate – must not be mistranscribed, mistranslated
    - Is the linguist's understanding of the data correct?
  - Naturalistic – must represent how people actually speak in a day-to-day context
    - Unless otherwise specified to a marked context
    - Is the data correct within the language itself?
  - Manageable – must be recorded in a way that is usable in analysis down the line
    - Can the linguist effectively use the data in practical terms?

Methodology

- Speakers are not typically consciously aware of the distinctions
  - Forms can be hard to elicit in isolation (Grzech et al. 2020)
- Ideal method may be recording naturalistic data in situ and creating a full corpus
  - Not feasible in a shorter project, also can be unwieldy
  - As observer, key context regarding knowledge not necessarily available
- Middle ground: elicitation activities that allow naturalistic conversation within a certain topic, with established contexts
  - Contrastive contexts within activities or across multiple activities could highlight contrasts in epistemic marking

Methodology - Activities

- Family Problems Picture Task (San Roque et al 2012; Barth and Evans 2017)

- Man and Tree Picture Sets (Levinson et al 1992)

- Also ran: 20 Questions Game (Gawne 2020)

Methodology - Activities

- Family Problems Picture Task (San Roque et al 2012; Barth and Evans 2017)
  - Four stages
  - Participants describe of a set of illustrations depicting various scenes out of order
  - Participants confer and put the illustrations in an order that tells a story
  - Participants tell the story in third person
  - Participants tell the story in first person, taking on the role of one of the characters
- Covers:
  - Visual Evidence (description)
  - Inferential Evidence (description and discussion)
  - Non-origo or equal authority (description, both participants are seeing the images for the first time together)
  - Participatory evidence (first person telling)

Carroll et al 2009

Lhokpu speakers arranging the illustrations in a story

Methodology - Activities

- Man and Tree Picture Sets (Levinson et al 1992)
  - A set of images depicting plastic figures of a man, bushes, and pigs, in various numbers and arrangements
    - Originally intended for the documentation of spatial reference
  - Run as a guessing game/matcher-director task
    - One participant has the full set of images in front of them, and is the guesser
    - The other has the set face down, and draws one image card to describe it to the guesser (the describer)
    - The guesser, asking questions of the describer, guesses which card the describer has drawn from the set
  - Here, used initially for elicitation of epistemic forms
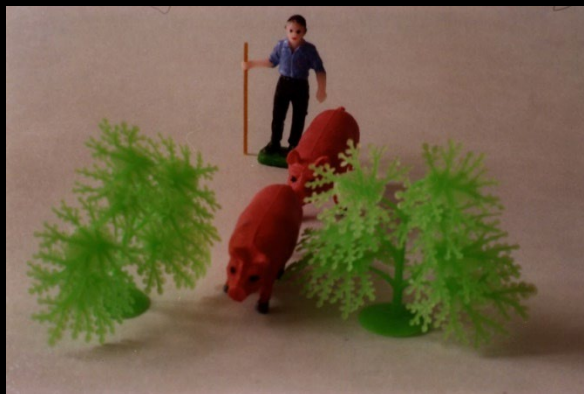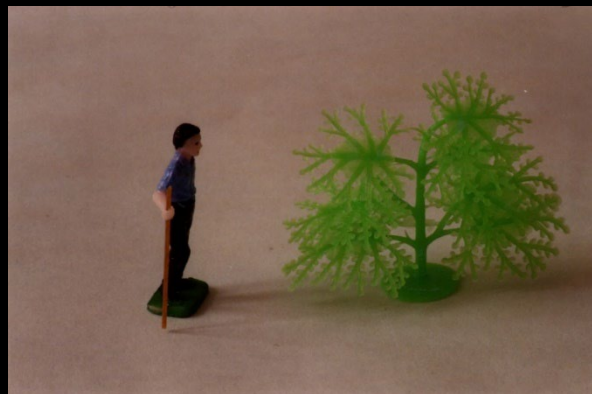
Methodology - Activities

- Man and Tree Picture Sets (Levinson et al 1992) cont.
- Covers:
  - Unequal epistemic authority, speaker- and non-speaker-authority (describer vs guesser)
  - Visual evidence
  - Questions polling visual evidence, and differences in authority and confidence

Field Report

- More or less used these activities in isolation
  - Lhokpu – Wordlist and some basic grammar elicited
  - Chocangaca – Much more elicitation throughout field methods course
  - Very different levels of general understanding of the languages
  - This turned out to be a challenge
- Lhokpu – ran activities multiple time with different pairs, each pair had at least one participant with some English proficiency (younger male) and one other
- Chocangaca – ran each activity once only with both participants also proficient in English and Dzongkha
- As a result, have demographic variation in the Lhokpu data but not really in the Chocangaca data

Descriptive Conclusions

- Chocangaca Equative copulas
  - Confirmation of evidential contrasts in verbal morphology
- Lhokpu some confirmation of evidential contrasts in existential copulas
  - Possible evidential contrasts in verbal morphology, but remains uncertain
  - Various other areas, e.g., topographical deixis (as per Levinson et al's (1992) original intentions

Chocangaca

- Equative copulas
  - Two forms identified, *yincok* used throughout the Family Task, *yin* used in confirmation questions in the Man and Tree game
  - *Yincok* for visual evidence, equal/non-origo authority
  - *Yin* used with reference to origo-authority, specifically when confirming cards *Origo-authority?*

(1) *tsikpa*   *re*   *za-te*   *tamo…*   *thong-di-ya*
     angry   even   become-PAST.ALLO   look   see-PRES.PROG?-INT
     'See, they were getting angry… You see?'

     yincok
     COP.EVD
     'Yes'

(2) *ula*   *yin-nam*
     DEM.PROX   COP-INT
     'Is it this one?'

     ula   *yin*
     DEM.PROX   COP
     'It is.'

Note: listener-perspective in questions ("you have authority")

Chocangaca

- Notably, *yin* <u>only</u> regularly occurs in confirmation statements in the Man and Tree sets
- Contrasts in context
  - *Yin:* speaker authority, visual evidence?, previous knowledge
  - *Yincok:* equal authority/non-speaker authority, visual evidence, new knowledge
  - Remaining agnostic as to the exact analysis of these forms here
    - In-depth presentation on just these forms next week at the Himalayan Languages Symposium in Paris
- Some other forms noted, not going to discuss here

Chocangaca

- Confirmation of forms from elicitation
  - Epistemic marking in perfective/past paradigm

Participatory/Ego	-la	1st person story
Factual/Authoritative	-te	1st person story
Evidential/Mirative	-teme	3rd person story, descriptions

(3) *delen   nga dang   nga moisa   la zha-te   duk-la*
      then   1.SG and   1.SG wife   [work]-NF   PROG-**PST.EGO**
      'Then me and my wife were working.' (Family Problems task 1st Person Story)

(4) *kho-i   moisa   jab zhak-te   namesame*
      3.SG-GEN   wife   [beat]-PST.**NON_EGO**   much
      'His wife was beaten a lot.'

(5) *kho-ki   moisa   dang   busa   jap-teme,   dele meme   chi-ki   thong-teme*
      3.SG-GEN   wife   and   son   beat-**PST.VIS**,   then grandfather   one-ERG   see-**PST.VIS**
      '..he beat his wife and his son, and then a grandfather saw them...'

Lhokpu

- Language's phonology been a challenge
- Partly able to confirm forms from elicitation
  - Evidential distinction in existential copulas *mi* vs *mihã*
    - Appears similar in contrast to Chocangaca data – maybe because of similar contexts?

(6)   *nosam rang-ka   [ganmo   mi]*
             mind   PRON-GEN   [wife   COP.EXIST]
             'In his mind, "My wife is there".' (Family Problems)

(7)   *ka-lok   dze   niʰ-pu mi*
             1.SG-OBL   dog   two-CLF COP.EXIST
             'I have two dogs.' (Elicited)

(8)   *kona   i-du   meʰ   mihã*
             then   PROX-LOC   fire   COP.EXIST.EVD
             'then here there is fire' (Man and Tree sets)

(9)   *kanka   it-dra   mihã*
             old.man   one-CLF   COP.EXIST.EVD
             'There is one old man.' (Family Problems)

Lhokpu

- Elicited vs Activity
  - Equative copula given in elicitation as [baʰ], does not appear at all in activity data
  - Instead, copula appears to be is [haʰ]
  - Functional difference unclear, maybe epistemic (mirative/authority?)

(10) *ka   Australia-na   baʰ*
     1.SG   Australia-ABL   COP
     'I am from Australia.' (Elicited)

(11) *i   table   ek-ka   baʰ*
     DEM.PROX   table   red-ADJ   COP
     'This table is red.' (Elicited)

(12) *ja-ka   ha-yi*
     attach-ADJ COP-INT
     'Are they attached?'

     ja-ka haʰ
     attach-ADJ COP
     'They are attached.' (Man and Tree task)

Lhokpu challenges

- Lack of foundational understanding of full verbal morphology
- Uncertainty surrounding phonology
  - Status of coda position [h]
  - [ka-ha miŋ kaːl ba] parsed as "My name is Carl-ba", [ka-ha miŋ kaːl baʰ] as "My name is Carl"
- Verb endings in activity data appears to include [-a], [-aʰ], [-aha] among others
  - Working theory [-aʰ] evidential marker, appears throughout activity data but without any clear non-evidential meanings on alternative forms
  - Also, nasal vowels (cf. mihã vs -aʰ) complicate

Lhokpu challenges

(13)   *ŋan   dokm̥eŋ-su   dzoŋ-do-aʰ*
person walking.stick-COM   stand-PROG-**EVD?**
'The person is standing with the stick.'

Final h dropped medially?

(14)   *siŋ-hõ   hut-a   dzoŋ-do-a   le~le*
tree-TOWARDS   look-**?** ASP-PROG-**EVD?**   DOWNHILL~ADV
'Looking downhill towards a tree.'

Seems to match *–a(l)* given in Grollmann and Gerber (2018) in function but not in form, unless we assume phonological variation where /l/ is [h] in Jigme village.

> "The allophoric gerund *-a(l)* functions both as a gerund, i.e. to derive deverbal adverbs, and as a finite allophoric verbal marker, marking a verb form as not being experienced personally by the speaker or as not belonging to the personal knowledge of the speaker." (20-21)

Other C -> [h] instances between my data and Grollmann and Gerber's (ŋa -> ha)

Lhokpu Challenges

- In the Man and Tree sets, use of confirmation questions in Chocangaca
- Not done in Lhokpu
  - Differences in how the task was explained?
  - Different relationship between participants and myself?
  - Maybe it was lucky to see it so clearly in CN?
- Gawne (2020) notes not all participants clicked with all tasks

(2) *ula   yin-nam*
DEM.PROX   COP-INT
'Is it this one?'

ula   yin
DEM.PROX   COP
'It is.'

Conclusions

Lhokpu	Little foundational analysis/elicitation supporting data from activities	Verbal morphology difficult to analyse from activity data alone
Chocangaca	Larger amount of direct elicitation and analysis prior to running and analysing activities	Verbal morphology, copulas much easier to analyse

- Limited prior analysis = limited conclusions
- More extensive prior analysis = better conclusions
- From wider research (Grzech et al 2020), yet more prior analysis and general data collection produces still clearer data

Conclusions

- Tends to agree with literature that a balanced corpus is ideal (back to Himmelmann (1998)?)
- Also allowed for conclusions in other domains
  - Man and Tree sets do also allow for some conclusions on spatial reference and deixis
  - Does definitely create an environment where relevant tokens are more likely to be used
    - Just need a wider foundation to be able to analyse them effectively
- Also relevant: Consultant training and experience

References

Barth, Danielle & Evans, Nicholas. 2017. SCOPIC Design and Overview. *Language Documentation and Conservation* (SP12).

Bodt, Timotheus A. 2012. *The New Lamp Clarifying the History, Peoples, Languages and Traditions of Eastern Bhutan and Eastern Mon*. Wageningen: Monpasang Publications.

Carroll, Alice & Evans, Nicholas & Hoenigman, Darja & San Roque, Lila. 2009. *The family problems picture task*. Designed for use by the Social Cognition and Language Project. A collaboration of The Australian National University, Griffith University, University of Melbourne and the Max Planck Institute for Psycholinguistics.

Gawne, Lauren. 2020. Looks like a duck, quacks like a hand: Tools for eliciting evidential and epistemic distinctions, with examples from Lamjung Yolmo (Tibetic, Nepal). *Folia Linguistica* 54(2).

Grollmann, Selin & Gerber, Pascal. 2018. Linguistic evidence for a closer relationship between Lhokpu and Dhimal. *Cahiers de Linguistique Asie Orientale*. Brill 47(1). 1–96. (doi:10.1163/19606028-04701004)

Grzech, Karolina & Schultze-Berndt, Eva & Bergqvist, Henrik. 2020. Knowing in interaction: An introduction. *Folia Linguistica* 54(2). 281–315. (doi:10.1515/flin-2020-2041)

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36. 161-195.

Levinson, S. C., Brown, P., Danzinger, E., De León, L., Haviland, J. B., Pederson, E., & Senft, G. (1992). Man and Tree & Space Games. In S. C. Levinson (Ed.), Space stimuli kit 1.2 (pp. 7-14). Nijmegen: Max Planck Institute for Psycholinguistics.

San Roque, Lila & Gawne, Lauren & Hoenigman, Diana & Miller, Julia Colleen & Rumsey, Alan & Spronck, Stef & Carroll, Alice & Evans, Nicholas. 2012. Getting the Story Straight: Language Fieldwork Using a Narrative Problem-Solving Task. *Language Documentation & Conservation* 6. 135–174.

Tournadre, Nicolas Laurent & Rigzin, Karma. 2015. Outline of Chocha-Ngacha. *Himalayan Linguistics*. 14(2).

Social factors

- Social relations, hierarchies can also affect epistemic usage
- E.g. usage of ego forms relating to family members rather than self
- Also: Right to claim authority over information may be socially conditioned in some cases

- Lhokpu project used pairs of varied demographics (younger vs older, male vs female, various levels of education and English proficiency)
  - Not yet clear if this worked, though arguably would need a higher level of insight into cultural expectations
  - Possible that challenges in running activities in some cases may be attributed to social expectations
    - Younger participants unwilling to correct older participants
    - Lack of confirmation questions giving epistemic authority to a younger speaker