

Challenges and network building in the creation and use of the Akha legacy collection

Authors:

Jose Benavides, Søren Borch, Mary Burke, Shobhana Chelliah

3rd Discussion on Legacy Materials Conference
(DiLegMa-3)

Paris Institute for Advanced Studies

16-17 April 2026

Introduction

- Legacy collection: Inga-Lill Hansson's Akha collection is used as an example of Documentary linguistic data.
- Challenges: Practical, technical, ethical, and political.
- Networking: For solving the challenges.

“Documentary linguistic data may be acquired not only firsthand, but by consulting materials that were produced by scholars, missionaries, speakers, and others in the past.” (Dobrin and Schwartz 2021:1).

We'll use the Akha Death Ritual as a case study to discuss common challenges found in legacy collections.

Presentation outline

1. The Akha language
2. The Akha legacy collection
3. Technical challenges
4. Network building
5. Conclusions

1. The Akha language

- Tibeto-Burman (Lolo-Burmese)
- Spoken by ≈620,000 people
- Across Myanmar, Laos, Thailand, Viet Nam & China

- Facts: Eberhard et al., 2026
- Regional names: *Aini*, *Kaw*, *A ka daw*
- In China, included in Hani ethnic group
- ISO 639-3 [ahk](#); Glottocode [akha1245](#)



2. The Akha legacy collection, origin

Collected and developed by Prof. Inga-Lill Hansson
Copenhagen University Denmark and Lund University Sweden.

Akha language and culture:

- Rituals, festivals, traditional music
- Narratives, songs, chants, conversations

Documentation:

- Field work 1970-1999, Chiang Rai Province, Thailand
- Analysis & research 1974-2022



Content archived as per now

- **Photo** (81)
 - Many more ready for upload.
- **Audio** (190)
 - Cassette tapes (80 digitized, ~100 hours).
- **Text** (224)
 - Notebooks, 150+
 - Transcriptions, translations, analyses
 - Publications and presentations.

Has been archived in the CoRSAL archive:
<https://digital.library.unt.edu/explore/collections/AHKLR/>



More coming up in the future.

The Akha Death Ritual as our case study

The Akha Death Ritual reflects the Akha beliefs about:

- Ancestors
- Life cycles
- Transition from life to death
- Cosmology

Hansson's recorded Death Ritual gives us:

- Audio recordings from three days reciting plus follow-up interviews
- Specialized ritual vocabulary used in chants and narratives
- A vocabulary that differs from everyday language.
- Preserved vocabulary that might otherwise be lost.

In this presentation we'll use Inga-Lill Hansson's 1984 version of the Akha Death Ritual:

- 22 chapters, 418 pages,
- Akha text and free translation in English.



The buffalo being prepared to sacrifice at the Death Ritual



The phima and the buffalo killer preparing for the Death Ritual

3. Technical challenges

Taking the Akha Death Ritual as a case study, we identify the following challenges:

- 3.1. Irregular/non-standard transcriptions
- 3.2. Incomplete/non-standard glossing
- 3.3. Migration and data retrieval

3.1. Irregular/non-standard transcriptions

- During the early transcriptions, Hansson developed a very sophisticated system to transcribe faithfully the Akha language, sounds and tones, (i.e., aspiration, vowels, etc.)
- This was before the international standard IPA.
- Some of the letters and tone diacritics were typed using a special font of an IBM ballpoint typewriter.

mò djiq djiq ə thó ə Akha

mò djiq djiq ə thó ə English

1. sjhí' sjhó tshà lánj àxhánj

1. Having died on earth being bent

2. tshà lánj à, lánj doq àxhɛ

2. the door where the sun goes up

Hansson, 1984: Akha Death Ritual, Ch1. p.5 and p.6

Different transcription systems

Hansson's system deviated greatly from previous writing styles like the one developed by Lewis in the 50s and 60s.

For instance, consider the two transcriptions for "*the sun (and) the moon*":

náŋ-ma ba-la

(Hansson, 1984: Ch1. line 5, p.5)

nah^v ma ba la

(Lewis, 1968: 218)

Some consonants, vowels, aspirations, and tones are marked in different ways in the two systems:

Akha tones (Hansson, 1987:1)

(Lewis, 1968: xii)

Non-laryngealized: Ex.

High ' ló

Mid zero lo

Low ` lò

Laryngealized:

High ´q lóq

Mid q loq

Low `q lòq

Tones: Oral vowels occur on three level tones:

/^v/ a high tone (a^v ba^v)

// a mid tone (no tone mark) (ba la)

/_v/ a low tone (ba_v yu_v)

Laryngealized vowels occur on two level tones:

/[^]/ a mid tone (ba[^] k'ot)

/_^/ a low tone (ba_^-eu)

Common Akha Orthography (CAO)–2009

- Lewis' and Hansson's systems differ substantially from the official CAO with respect to creaky voice and tone marking.
- Compare the Lewis system against the CAO:
 Lewis: Nga[˘] A_v san_vbo_˘ dzaw. "I study Akha language"
 CAO: Ngar Aqkaq sanqbovq dzaw. "I study Akha language"

Tone	CAO 2009	CAO examples
High tone	-r	<i>aqnyir</i> 'younger'
Mid tone	-	<i>ba la</i> 'moon'
Low tone	-q	<i>zaq</i> 'children'
High creaky	-vr	<i>tivr</i> 'little'
Mid creaky	-v	<i>gav</i> 'cold'
Low creaky	-vq	<i>ar byievq</i> 'bamboo root'

3.2. Incomplete/non-standard glossing

mò djiq djiq ə thó ə

Akha

1. s jhí' s jhó tshà lán̄ àxhán̄
2. tshà lán̄ à, lán̄ doq àxhe
3. s jhí' s jhó bi dèq nòq lán̄ la doq àxhe
4. nán̄ lə dzò phí' àda à thí' γà zá zí' thó lə dèq le miá
5. nò le ló la àmɛ' à, nán̄-ma ba-la nò
6. gán̄ lèq ló la àmɛ' à, nán̄-ma ba-la gán̄
7. tshó gán̄ dzò mò phu shu tjhö án̄ ný
8. ý gán̄ à, la sjaq thm̄ γòq liq án̄ jǒ ə nǵá
9. nó ló à, mà shí dɛ γa ànm̄ mà nǵé
10. shí shó tshó xaq làq-bȳ

Hansson's free and rich translation

mò djiq djiq ə thó ə

English

1. Having died on earth being bent
2. the door where the sun goes up
3. you die, others are let to live, the door where the moon comes out
4. the phí'-ma recites for a lifetime
5. the phí'-ma must think long ahead, long like the river floating upstream,
think clearly like the sun and the moon
6. must be clever like the clear water upstream, like the clear sun and moon
7. let old people live in the village being clever
8. the clear water flowing in the bending river
9. you are not a young one going to the dancing place
10. you are an adult who knows a lot, like the Chinese

Challenge using the free translations

- The division of sentences sometimes do not correspond to free translations.
- Some of the free translations are inconsistent with the numbering of the sentences.
- The direct meaning of a word is sometimes hard to see in the free translations.

mò djiq djiq ə thó ə

1. s jhí s jhó tshà lán̄ àxhán̄
2. tshà lán̄ à, lán̄ doq àxhə
3. s jhí s jhó bi dèq nòq lán̄ la doq àxhə
4. nán̄ lə dzə phí-à à thî yà zá zí thó lə dèq le miá
5. nò le ló la àmɛ̀ à, nán̄-ma ba-la nò

mò djiq djiq ə thó ə

1. Having died on earth being bent
2. the door where the sun goes up
3. you die, others are let to live, the door where the moon comes out
4. the phí-ma recites for a lifetime
5. the phí-ma must think long ahead, long like the river floating upstream,
think clearly like the sun and the moon

Hansson, 1984: Akha Death Ritual, Ch1. p.5 and p.6

Glossings from her publications

We have a large number of ritual text lines that have been glossed by Hansson, and published. We expect to find many more in computer files that we have not yet been able to open.

They were done, while Hansson still knew the glosses (thus not needing translation in the notebooks).

Ex. 6

Ritual: Chmq bi gar sar ngaq neh maq cawq chivq-eu mir-a
Shan road level I NP not jump able VP SP

Modern: Miqchmq-eu garma yawsar-angr ngar maq chang nya-eu mir-a
Shan NP road level NP I not walk able VP SP

I can 't walk on the level Shan road

All these glossings may be transferred into a modern FLEEx-like system, and thus, may be used for human learning and for AI machine learning.

*** Example 6 was taken from the Conference handout from 35th ICSTLL, November 2002, p.5, Hansson, "Translation of the Ritual Akha Language into Modern Akha", digital.library.unt.edu/ark:/67531/metadc2404318/*

3.3. Migration and data retrieval

The STEDT system

Hansson used the STEDT-system as a rich transcription system.

- The Sino-Tibetan Etymological Dictionary and Thesaurus

STEDT remains an invaluable resource for Sino-Tibetan languages, but interoperability is limited by specialized encoding conventions:


- non-standardized symbols
- project-specific transcription practices.

The Sino-Tibetan Etymological Dictionary and Thesaurus, STEDT:

Large-scale compilation of comparative lexical and etymological data in Sino-Tibetan languages, begun in 1987 at UC Berkeley, directed by James A. Matisoff and team. Later digitized into a custom database system (~2011)

Transferring STEDT-files

As expected, we have managed to locate Hansson's original Death Ritual transcription files, but we have not yet been able to transfer properly. We have encountered challenges opening/copying the data:

Name ↑	Modified	Size
 Death 1-3 Stedt	Dec 18, 1999	48.23 KB

(Hansson's original STEDT file Ch1-3. retrieved from her laptop)

Ch	Line	Text
1	1	s h i s h ø tsh\ a la≥ \ axh a≥
1	2	tsh\ a l a≥ l a≥ doq \ axh%
1	3	s h i s h ø bi d\%q ≥\oq l a≥ la doq \ axh%
1	4	n a≥ lø dz\\ö ph i \ ada \ a th\ i ©\ a z a z i th o l% d\%q le m i a
1	5	n\\ö le l ø la \ am% n a≥ ma ba la n\\ö
1	6	g a≥ l\eq l ø la \ am % n a≥ ma ba la n\\ö
1	7	tsh ø g a≥ dz\\ö m\ø phu shu tjhö a≥ n y
1	8	y g a≥ la sjag th\\m ©\oq liq a≥ j ö ' ≥ aa
1	9	n ø l ø \ a m\ a s h\ i d% ©a \ an m m\ a ≥
1	10	sh\ i sh ø tsh ø xaq l\ aq b\ y

(Screenshot of the Death 1-3 STEDT file, first 10 lines Ch 1. [displayed in Notepad])

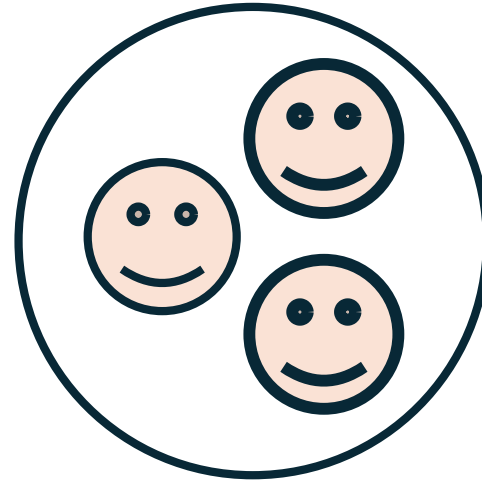
Challenges when transferring files

Overall, we have had three challenges when attempting to transfer data from outdated programs** into modern software (e.g., FLEx):

- Encoding mismatch
- Missing dependencies (e.g., font, keyboard mapping, encoding converter)
- Structure format layer (e.g., auxiliary configuration files).

** We believe Hansson may have used the early database tool **Shoebox** (1987) created by SIL. But we have not yet found the fonts.

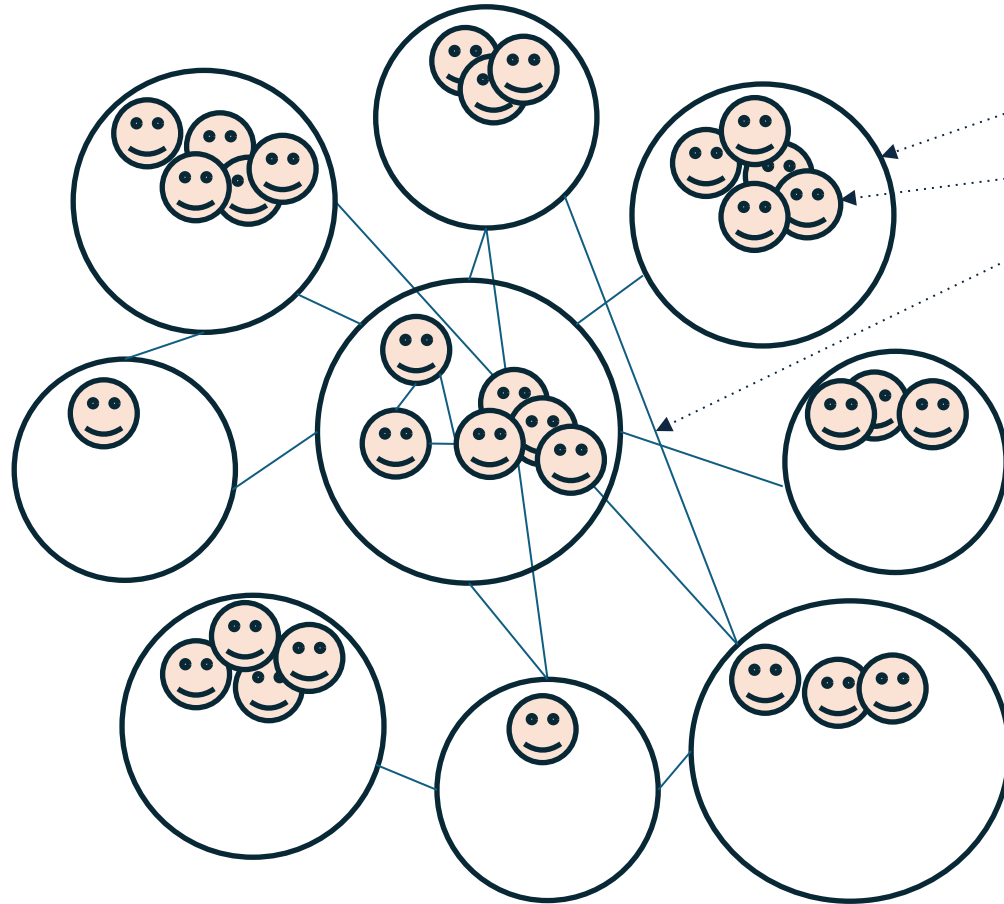
4. Networking



4.1 Expanding the network

4.2 Expanding accessibility to the legacy

4.1. Expanding the Network



Nodes:

- Groups of actors

Relations:

- Between nodes

Results:

- Solutions to challenges

Node 1: Hansson's Friends, 1970's...



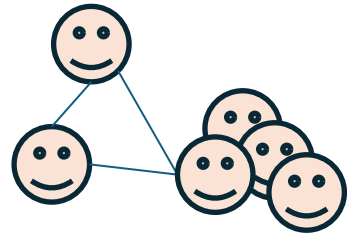
1977 Søren visits Inga-Lill in Maeby, Thailand



1980-1992 Inga-Lill lives with the Borch-family in Copenhagen.



Søren taught Inga-Lill to use a spreadsheet for concordance.



Common friends:

- Inga-Lill's annual Christmas lunch
- Lise reads aloud the Death Ritual
- Typing and translation into Danish

The Akha-fund:

- A friend's donation
- Accelerated digitization



www.akha.dk

Inga-Lill's and Søren's strategy:

- Vision: Archive becomes accessible.
- Value: Open sharing.
- Mission: Project management

Personal relations matter in networking

Node 2: Angels from CoRSAL, 2022

Visit in Malmö November 2022. CoRSAL team: Prof. Shobhana Chelliah and Mary Burke.



Shobhana interviewing Inga-Lill
youtu.be/EkCrg2bMKw0



Shobhana interviewing Søren
youtu.be/UErGAXg_I9M



The CoRSAL fund

Tasks during visit:

- Getting context, metadata etc.
- Packing material for CoRSAL
- Planning for future digitization



Mary Burke



Inga-Lill's will:
“My material is to be shared with Akha communities and researchers.”

Batch-1 uploaded Autumn 2022

<https://digital.library.unt.edu/explore/collections/AHKLRL/>

More batches

- Batch 2: 2023, summer
- Batch 3: 2023, autumn

Node 3: Excavating Inga-Lill's Archive, 2024

Jose Benavides, Linguist from CoRSAL, in Malmö & Copenhagen, May - June 2024



Finding:

- Publications
- Work papers
- Files on computers



Organizing:

- Physically on shelves
- Digitally in computers



Analyzing:

- Identifying challenges
- Solving problems

A screenshot of a spreadsheet or database interface. The table contains multiple rows of data, with columns for various fields. Some cells are highlighted in blue, and there are some red and yellow cells, possibly indicating errors or specific data points. The text is small and difficult to read, but it appears to be a structured list of information.

Batch-4:

Jose helped:

- Finding files for CoRSAL
- Metadating

Node 4: Multiple Relations, 2025

Yuan Mingze in Malmö & Copenhagen, January & August 2025. Søren in Heidelberg March 2025.



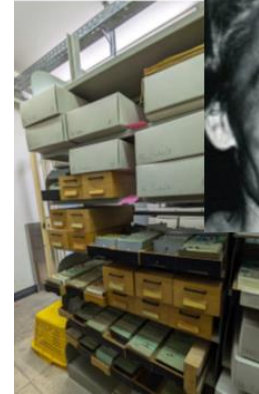
Yuan Mingze:

- Anthropologist
- Han Chinese
- Beijing & Heidelberg



Yuan's research:

- Field work in Yunnan
- Akha ritual and belief
- Uses Hansson's ritual texts
- Compares rituals +50 years
- Akha communities contacts



Scholz' Akha legacy:

- Heidelberg University
- Akha anthropological
- Yuan "excavates"
- Our 2 legacies cooperate
- [Scholz' Archive](#)

Inga-Lill's legacy:

Yuan helped:

- Improving metadata
- Adding more photos
- Using insight in Akha life
- Contact to Akha speakers

Node 5: Efficiency with Artificial Intelligence 2026

Giulio Ongaro, Hong Kong University. Visits to Malmö 2020-2022. AI project 2026.



Giulio Ongaro:

- Anthropologist
- Akha field studies, Laos
- Italian
- Hong Kong University

	<u>mò djiq djiq ə thó ə</u>
1.	sji' sji' tshà lán àxháŋ
2.	tshà lán à, lán doq àxhe
3.	sji' sji' bi dèq nòq lán la doq àxhe
4.	nán lo dzò phi' àda à thi' γà zá zí thó le dèq le míá
5.	nò le ló la àmɛ à, nán-ma ba-la nò
6.	gán lèq ló la àmɛ à, nán-ma ba-la gán
7.	tshó gán dzò mò phu shu tji' hō áŋ ný
8.	ý gán à, la sjaq thèn γòq liq áŋ jō ə nǎa
9.	nó ló à, mà shí de γa ànín mà nǎ
10.	shí shó tshó xaq làq-bý

Giulio's work with Inga-Lill:

- Glossing parts of the Death Ritual
- Using it also for his own research
- Mutual benefit



AI project of Giulios group

Material:

- Akha ritual language
- Four collections (Inga-Lill + 3 more)

Goals:

- Audio → Text
- IPA-like text → Akha-CAO text
- Translation, Akha → English/Chinese

First results:

- Challenges identified for full project.

Node 6: Subcontractors



Subcontractors:

- IT-consultants
- Linguists
- Digitization services
- Worldwide

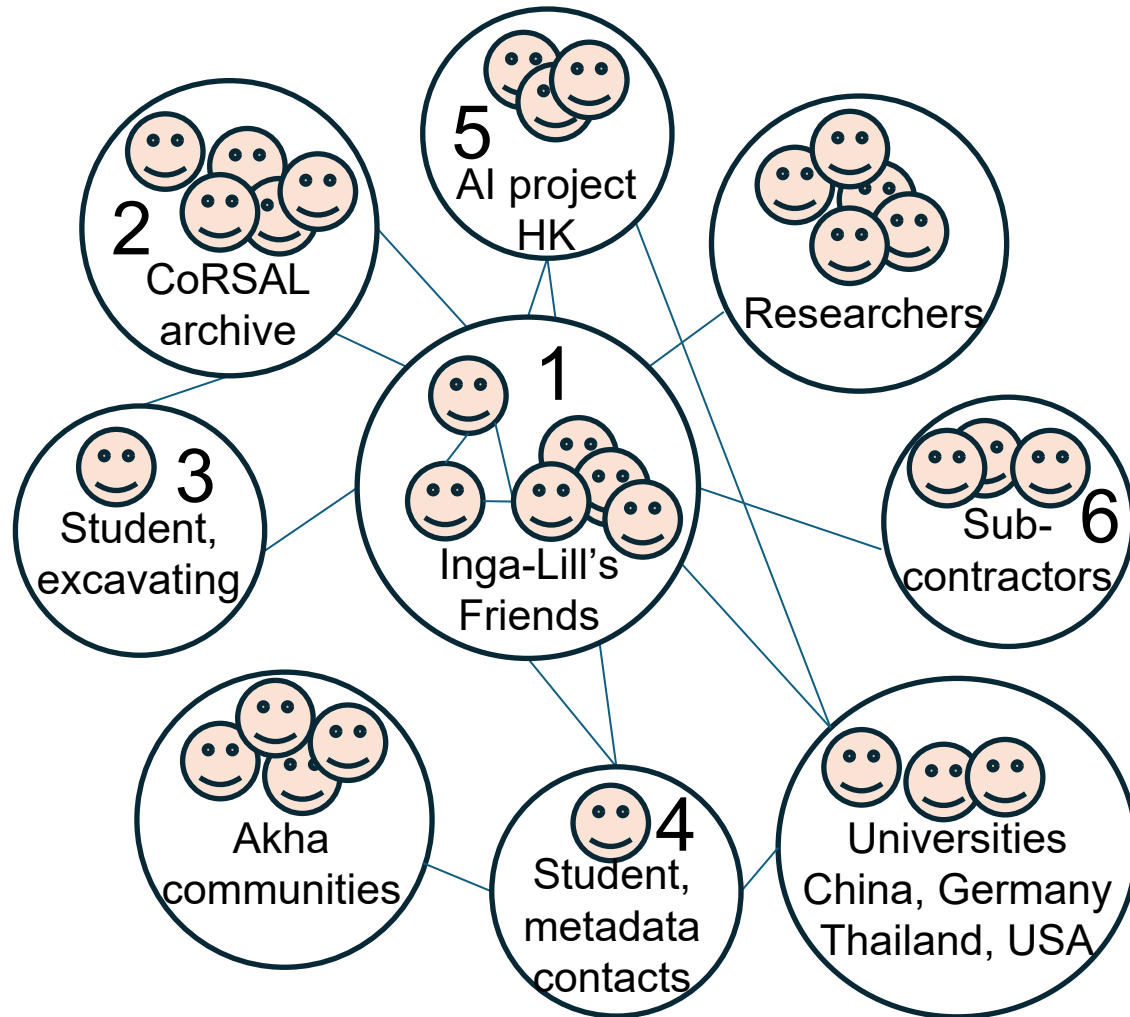
Goals:

- Typing and transferring paper/pdf to text file
- Digitization
- Transcription
- Glossing
- Migration and data retrieval

Methods:

- Payed individuals and commercial services
- Invitation through the internet and through contacts
- Financial funding is needed.

The Growing Network



Network:

- Nodes and actors (interests)
- Relations (win-win)

Networking results:

- Keeping the Akha legacy alive
- Involving and motivating actors
- Solving challenges

Methods:

- Social and professional relations
- Contact through contacts
- Motivation and patience

4.2 Expanding Accessibility to the Legacy

User-accessibility through:

- CoRSAL's website
- Hansson's Friends' website



CoRSAL's Akha archive

Website:

- Advanced tools for search, deep links, etc.

Content:

- Files and metadata

Users:

- Primary: Academic researchers (linguists, anthropologists, ...)
- Others: Akha communities, music researchers, general public, ...

Plans for augmented accessibility:

- Personal outreach to raise awareness
- Internet search engine optimization?
- Customized user-interfaces for selected user-groups?



[digital.library.unt.edu
/explore/collections/
AHKLR/](https://digital.library.unt.edu/explore/collections/AHKLR/)



Hansson's Friends' Akha Website

Website:

- Basic tools

Content:

- Example materials (ritual texts, stories, music)
- Videos from interviews and symposiums
- Deep links to CoRSAL
- Links to numerous Akha related sites

Users:

- Academics: Linguist, anthropologists, ...
- Non-academics: Akhas, friends, the public, ...



www.akha.dk



Webmaster
Søren

Plans for augmented accessibility:

- Personal outreach to raise awareness
- Redesign, also for mobil devices
- Customized user-interfaces for selected user-groups
- Supplementary, professional interface with files in Dropbox and metadata in Excel.

5. Conclusion and reflections

*Handout conference
paper: <https://akha.dk/>*

Hansson's legacy collection:

- Hansson's material is rich and dates two generations back. The Akha culture is changing.
- It is the first scientific, linguistic research material on the Akha language. Contains historic ritual vocabulary.
- Hansson's collection has been saved thanks to the network. CoRSAL has played an essential role.

Challenges:

1. Transcription: Even the best system is difficult to transfer to modern standards after 50 years.
2. Glossing: Rich free translation, handwritten notes, and examples of scientific glossing do not suffice without an Akha ritual language connoisseur.
3. Data retrieval: IT-files from old SW-systems are difficult to use because of non-standard symbols, outdated file formats, missing fonts etc.

Networking:

- Challenge solving: Engaging motivated people in the network is a way to handle our challenges.
- Network knowledge: Necessary to engage organizations and professional actors of many abilities.
Networking method: Personal motivation and relations matter to ensure the "energy". Win-win is important.
- Accessibility: The collection must be easy to find on the internet. User-friendly interfaces are momentous.

Contact information

For: Questions. Cooperation for solving challenges. Joining the network. Invitations.

Authors

Jose Benavides, Ph.D. candidate in linguistics, CL minor, Indiana University, Linguistics department, josbenav@iu.edu, linguistics.indiana.edu/about/graduate-students/index.html

Søren Borch, Ph.D., retired associate professor, Copenhagen Business School, consultant, smb@condiv.dk, www.linkedin.com/in/sorenborch/

Mary Burke, Assistant Professor, Digital Archivist, University of North Texas, Department of Information Science, mary.burke@unt.edu, www.linkedin.com/in/mary-burke-096333167/

Shobhana Chelliah, Professor, Linguistics, Indiana University, Linguistics department, schellia@iu.edu, linguistics.indiana.edu/about/faculty/chelliah-shobhana.html

Some other actors in the network

Inga-Lill Hansson, retired associate professor at Lund University, Sweden, M.A. and postgraduate from Copenhagen University, Denmark. Lives in Malmö, Sweden. Bibliography: digital.library.unt.edu/ark:/67531/metadc2404244/. CV: <https://digital.library.unt.edu/ark:/67531/metadc2404243/>. Email c/o smb@condiv.dk.

Yuan Mingze, Anthropologist, Ph.D. student Akha rituals, Beijing University & Heidelberg University, yuanmz1995@163.com, www.linkedin.com/in/%E9%93%AD%E6%B3%BD-%E5%8E%9F-19b14132b/

Giulio Ongaro, Dr. Anthropology, Assistant Professor, Hong Kong University of Science and Technology, giulio.ongaro@ust.hk, www.linkedin.com/in/giulio-ongaro-5b59b160/

References

- Austin, P. K. (2017). Language documentation and legacy text materials. *Asian and African Languages and Linguistics*, 11.
- Austin, P. K. (2013). Language documentation and meta-documentation. In M. Jones & S. Ogilvie (Eds.), *Keeping Languages Alive: Documentation, Pedagogy and Revitalization* (pp. 3–15). Cambridge: Cambridge University Press.
- Dobrin, L., & Schwartz, S. (Eds.) (2021). The social lives of linguistic legacy materials. *Language Documentation and Description*, 21.
- Eberhard, David M., Gary F. Simons, and Alison J. Robinson (eds.). 2026. *Ethnologue: Languages of the World*. Twenty-ninth edition. Dallas, Texas: SIL Global. Online version: <https://www-ethnologue-com.proxyiub.uits.iu.edu/language/ahk/>
- Lahaussois, A. (2021). Glossing in the linguistic survey of India. Some insights into early 20th century glossing practices. *Historiographia Linguistica*. <https://doi.org/10.1075/hl.00081.lah>.
- Lehmann, C. (2004). Interlinear morphemic glossing. In G. Booij, C. Lehmann, J. Mugdan & S. Skopeteas (Eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung*. (pp. 1834–1857). W. de Gruyter.
- Lupke, F. (2011). Orthography development. In P. K. Austin & J. Sallabank (Eds.), *Handbook of Endangered Languages*. Cambridge University Press.



Thank you!

Q&A