



# Language documentation and vocal technologies for minority languages in Vietnam

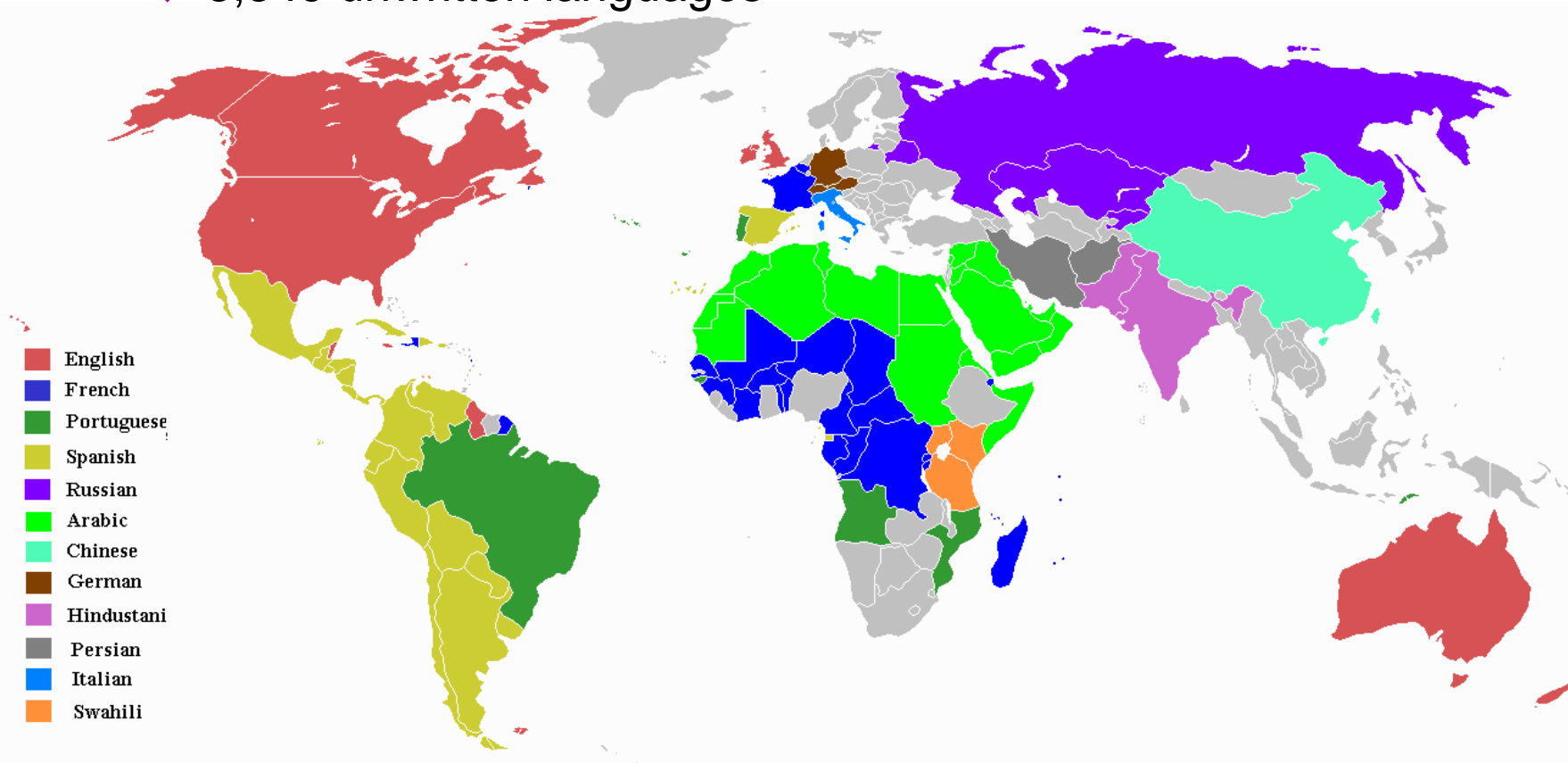
*MAC Dang-Khoa*  
*Speech Communication department*

**International Research Institute MICA**  
Multimedia, Information, Communication & Applications  
UMI 2954

Hanoi University of Science and Technology  
1 Dai Co Viet - Hanoi - Vietnam

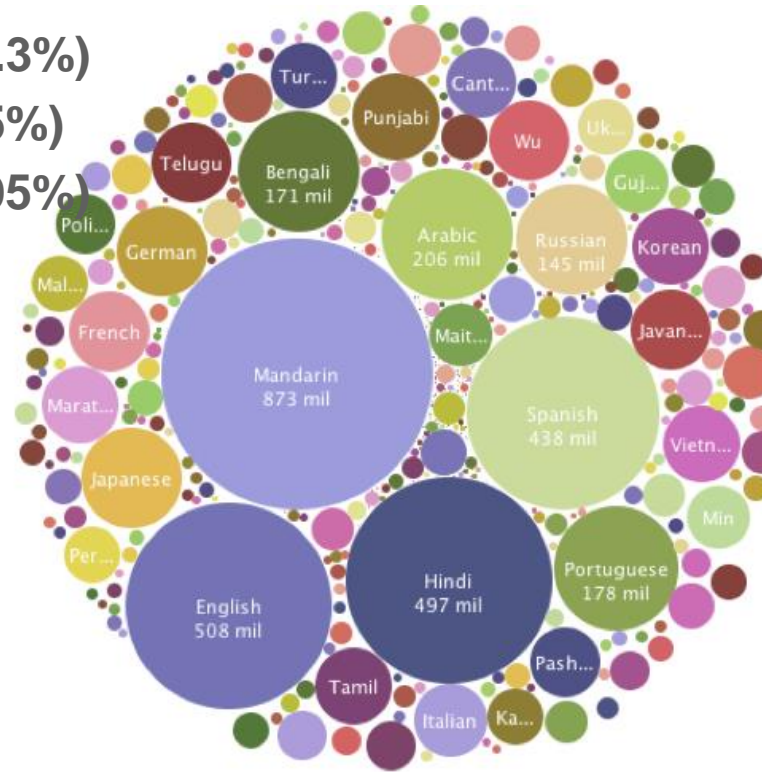
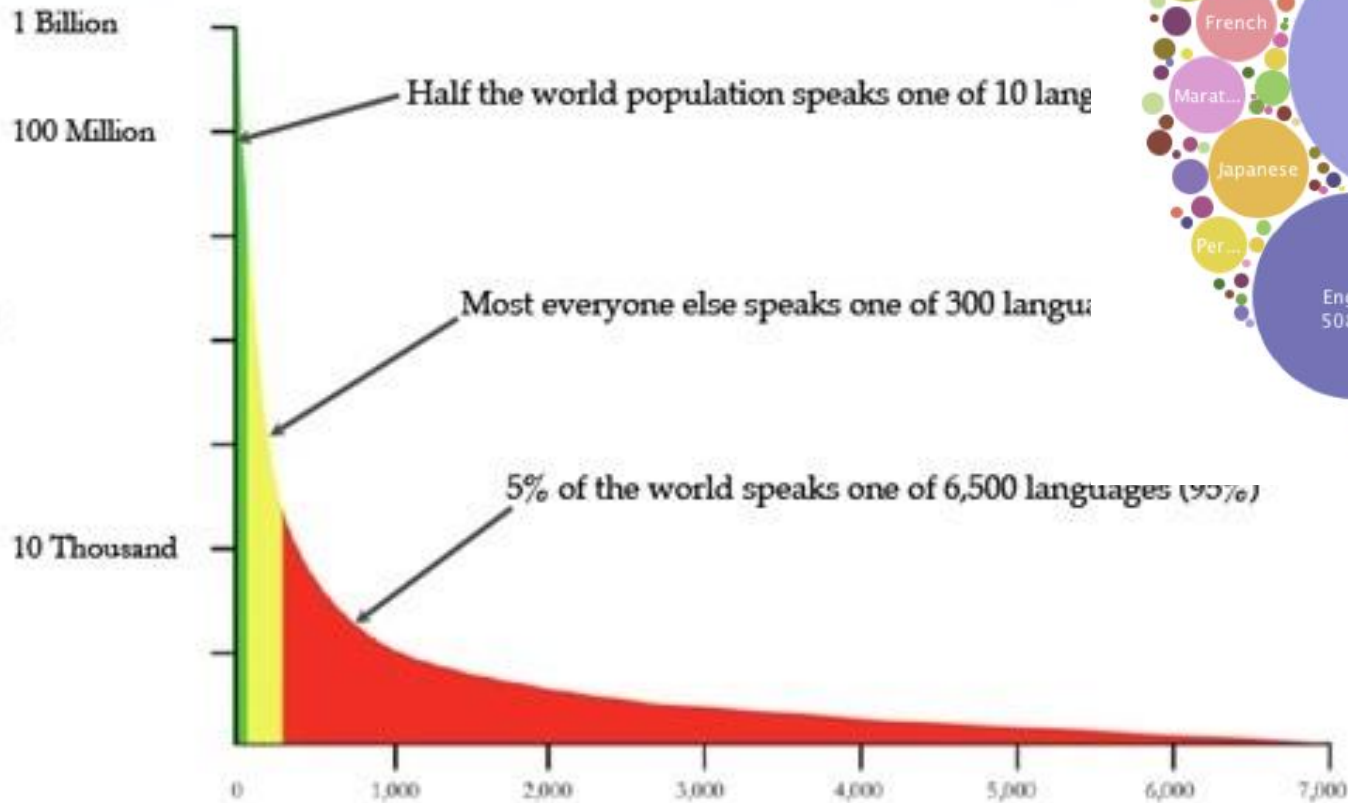
# Languages in the world

- World population: 7,294,154,164
- 7,097 **live** languages
  - ◆ 3,349 unwritten languages



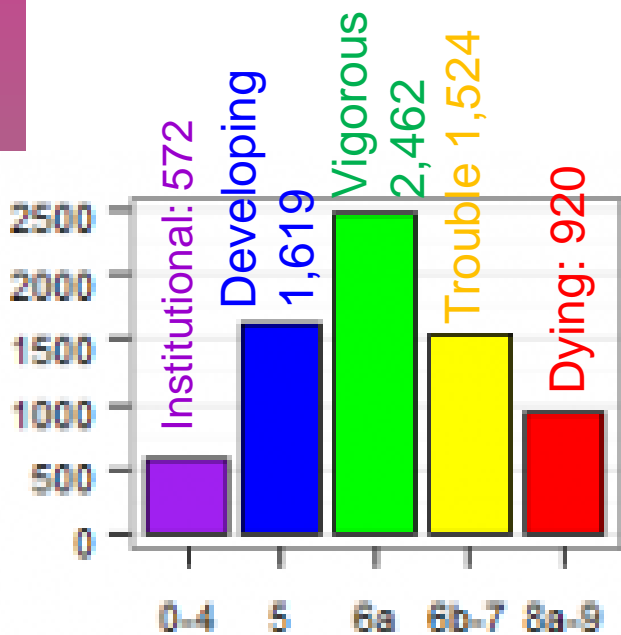
# Language speaking distribution

- 50 % population speak 10 languages (1.3%)
- 95% population speak 300 languages (5%)
- 5% population speak 6500 languages (95%)



# Language status

## Expanded Graded Intergenerational Disruption Scale (EGIDS)



## Status of world languages

Source: <https://www.ethnologue.com>

L	Label	Description
0	International	The language is widely used between nations in trade, knowledge exchange, and international policy.
1	National	The language is used in education, work, mass media, and government at the national level.
2	Provincial	The language is used in education, work, mass media, and government within major administrative subdivisions of a nation.
3	Wider Communication	The language is used in work and mass media without official status to transcend language differences across a region.
4	Educational	The language is in vigorous use, with standardization and literature being sustained through a widespread system of institutionally supported education.
5	Developing	The language is in vigorous use, with literature in a standardized form being used by some though this is not yet widespread or sustainable.
6a	Vigorous	The language is used for face-to-face communication by all generations and the situation is sustainable.
6b	Threatened	The language is used for face-to-face communication within all generations, but it is losing users.
7	Shifting	The child-bearing generation can use the language among themselves, but it is not being transmitted to children.
8a	Moribund	The only remaining active users of the language are members of the grandparent generation and older.
8b	Nearly Extinct	The only remaining users of the language are members of the grandparent generation or older who have little opportunity to use the language.
9	Dormant	The language serves as a reminder of heritage identity for an ethnic community, but no one has more than symbolic proficiency.
10	Extinct	The language is no longer used and no one retains a sense of ethnic identity associated with the language.

# Many languages are dying !

## ■ Endangered languages

EGIDS	Living languages			Number of speakers				
	Count	Percent	Cumulative	Total	Percent	Cumulative	Mean	Median
0	6	0.1	0.1%	1,910,305,410	29.3610	29.3610%	318,384,235	255,399,910
1	96	1.4	1.4%	1,944,722,578	29.8900	59.2511%	20,257,527	5,634,000
2	68	1.0	2.4%	627,218,790	9.6402	68.8913%	9,223,806	1,236,470
3	166	2.3	4.7%	584,974,240	8.9909	77.8822%	3,523,941	743,750
4	236	3.3	8.1%	252,622,839	3.8828	81.7650%	1,070,436	114,465
5	1,619	22.8	30.9%	677,113,264	10.4071	92.1721%	418,229	25,100
6a	2,462	34.7	65.6%	448,099,167	6.8872	99.0593%	182,006	10,000
6b	1,061	14.9	80.5%	50,831,670	0.7813	99.8406%	47,909	3,000
7	463	6.5	87.0%	9,535,479	0.1466	99.9872%	20,595	1,000
8a								250
8b								15
9								0
Totals	7,097	100.0		6,506,259,160	100.0000			

Languages saving => Documentation

## ■ Language disappearing <sup>1</sup>

<https://www.ethnologue.com>

- ◆ From 1950, **421** languages were disappeared
- ◆ “*1/3 of the world’s languages are in danger of disappearing in the next few decades*”
- ◆ One language dies every 14 days
- ◆ Language changing, mixing

<sup>1</sup> <http://www.sil.org>

# Languages documentation

## ■ What?

- ◆ “the methods, tools, and theoretical underpinnings for compiling a representative and lasting multipurpose record of a natural language or one of its varieties” (Himmelman 1998)

## ■ To

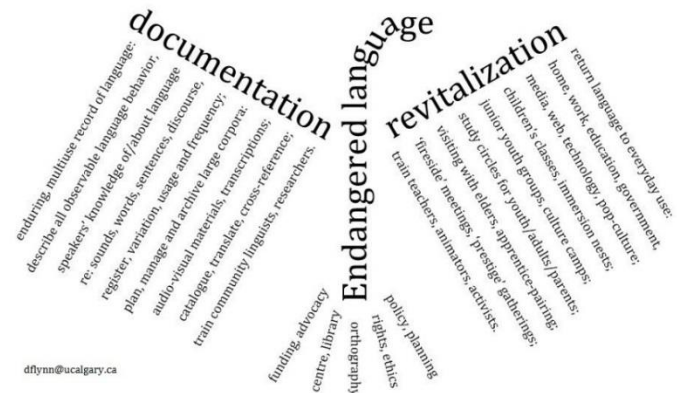
- ◆ Preserve language (endangered languages)
- ◆ Material for language study
- ◆ Material for Natural language and Speech processing

## ■ Tasks

- ◆ Collecting : recording, taking pictures, gathering written documents, ...
- ◆ Processing : analyzing, systematizing, transcribing, translating, ...
- ◆ Archiving: storing, publishing

## ■ Among 7097 languages

- ◆ < 600 well documentation languages
- ◆ 3,349 unwritten languages

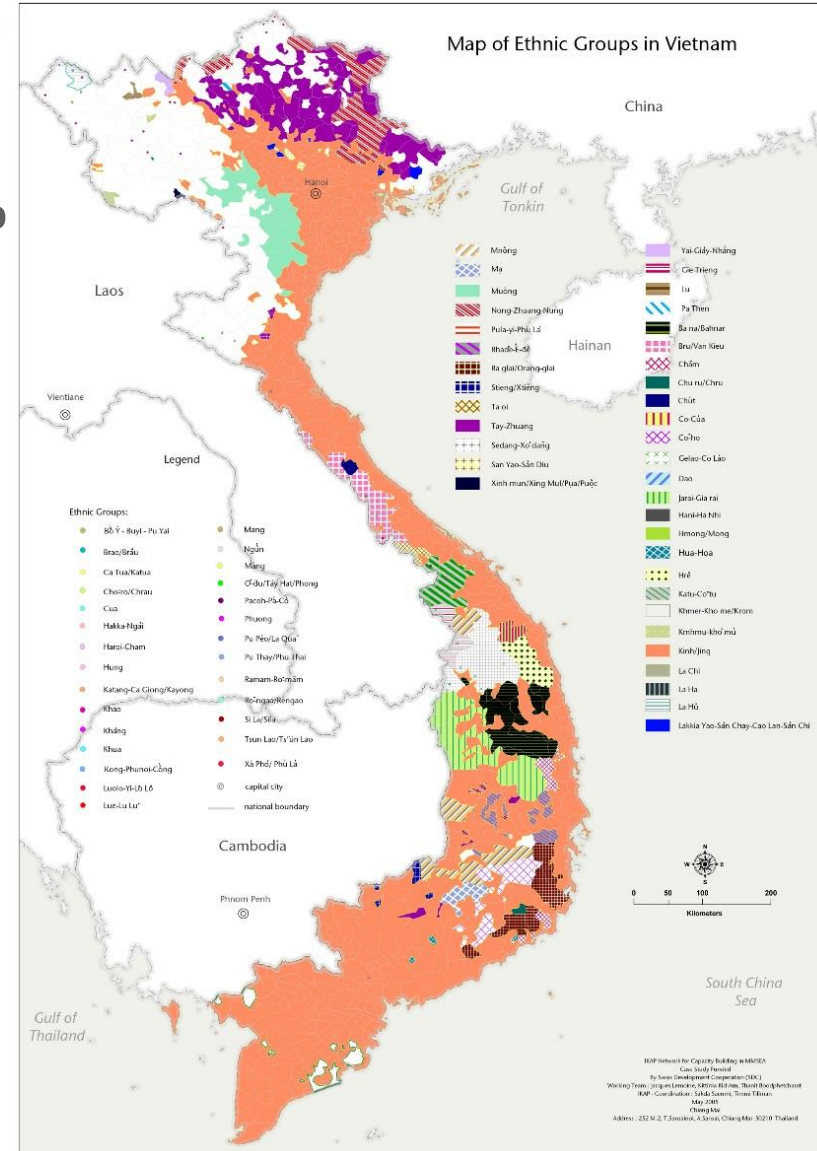


# Community

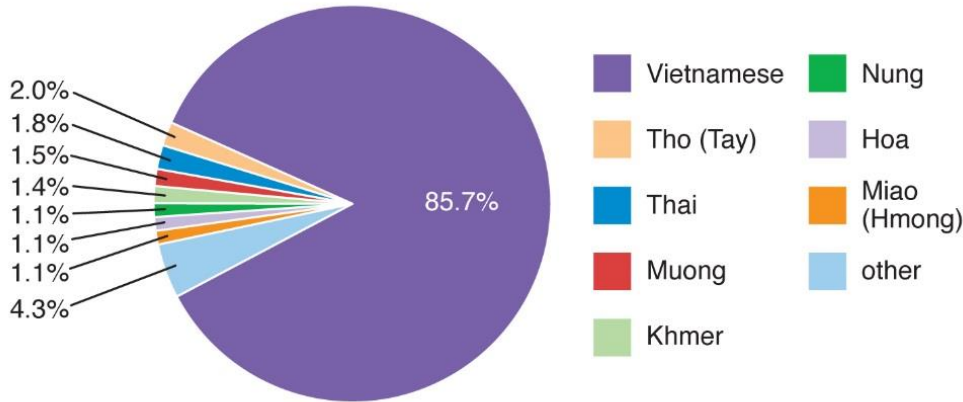


# Vietnamese minority languages

- 54 ethnic groups
- Vietnamese (Kinh): > 85%
- 6 **dying** (level 8a-9)
- 1 **extinct**



Ethnic composition (2003)

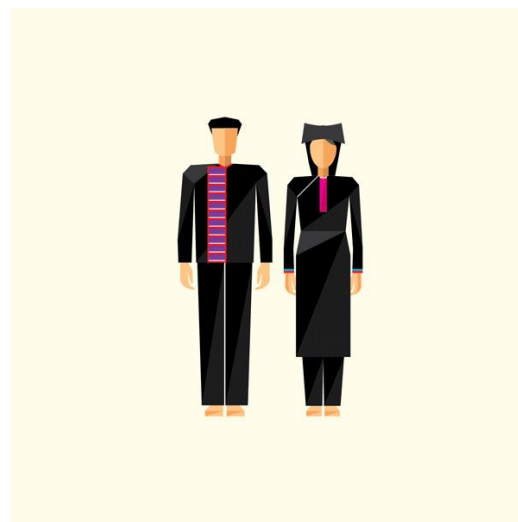




# MICA's AuCo collection



- **ÂuCo**: **Audio Copora**
- From 2007
- Language documentation:
  - ◆ Minorities
  - ◆ Vietnam an neighbors
- **Tasks**
  - ◆ Collection
  - ◆ Digitalization
  - ◆ Documentation: annotation, transcription
  - ◆ Analysis: Phonology, phonetic
  - ◆ Archiving: Online access



# DoReMiFa project

- **Données des Recherches linguistiques de Michel Ferlus en Asie du sud-est**

- ◆ Digitizing the collections of Michel Ferlus
- ◆ 2014 – 2015

- **Data**

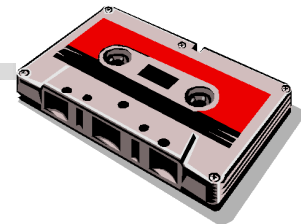
- ◆ >40 ethnic languages
- ◆ > 200 cassette tapes + manuscript, recording from 1963 – 2003

- **Project groups**

- ◆ Promoter: Alexis Michaud
- ◆ 6 Linguists + IT experts
- ◆ Linguistic students



# Data collection



- Recorded data
- Fieldwork recording



# Data collection

## Wordlist

- ◆ Common wordlist for fieldwork
  - ★ Ecole Française d'Extrême-Orient (1938)
- ◆ >2800 words
- ◆ Available : [HAL](#)



1	1	I	Ciel	Sky	天	trời	megh	me:c	mè:k
2	1a	III	Clair, ciel clair	Clear sky	晴天	trời trong	-	-	
3	1b	III	Sombre, ciel sombre	Dark sky	阴天	(trời) râm	-	-	
4	2	I	Soleil	Sun	太阳	mặt trời	thài	t'haj	t'haj ?ej
5	2a	III	Briller, le soleil brille	Shine (the sun shines)	闪耀 (太阳闪耀 ) , 有阳光 太阳光, 阳光 (	(năng) chiếu	-	-	
6	2b	III	Rayon de soleil	Sun beam					
7	2c	III	Voilé, soleil voilé	Dim (sunlight); cloudy (sky); hazy					
8	2d	III	Eclipse de soleil	Solar eclipse					
9	2.1	II	Ombre	Shadow					
10	2.1a	III	Ombre de maison	Shadow of the house					
11	2.1b	III	Ombre d'arbre	Shadow of the tree					
12	3	I	Lune	Moon					
13	3.1	III	Briller, la lune brille	Bright moon					



EFE0-CNRS-SOAS Word List  
for Linguistic Fieldwork in Southeast Asia

*Version 1.0*

Lexique EFE0-CNRS-SOAS  
pour les enquêtes linguistiques en Asie du Sud-Est

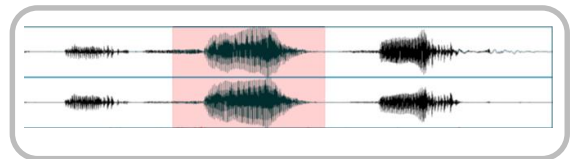
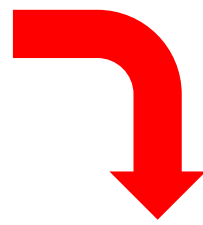
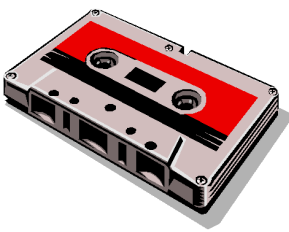
*Version 1.0*

法國遠東學院、法國國家科學研究院與倫敦亞非研究學院 東南亞語言調查詞彙表  
1.0 版

Bảng từ EFE0-CNRS-SOAS  
Dùng cho nghiên cứu điền dã ngôn ngữ học ở Đông Nam Á

*bản 1.0*

# Digitalization

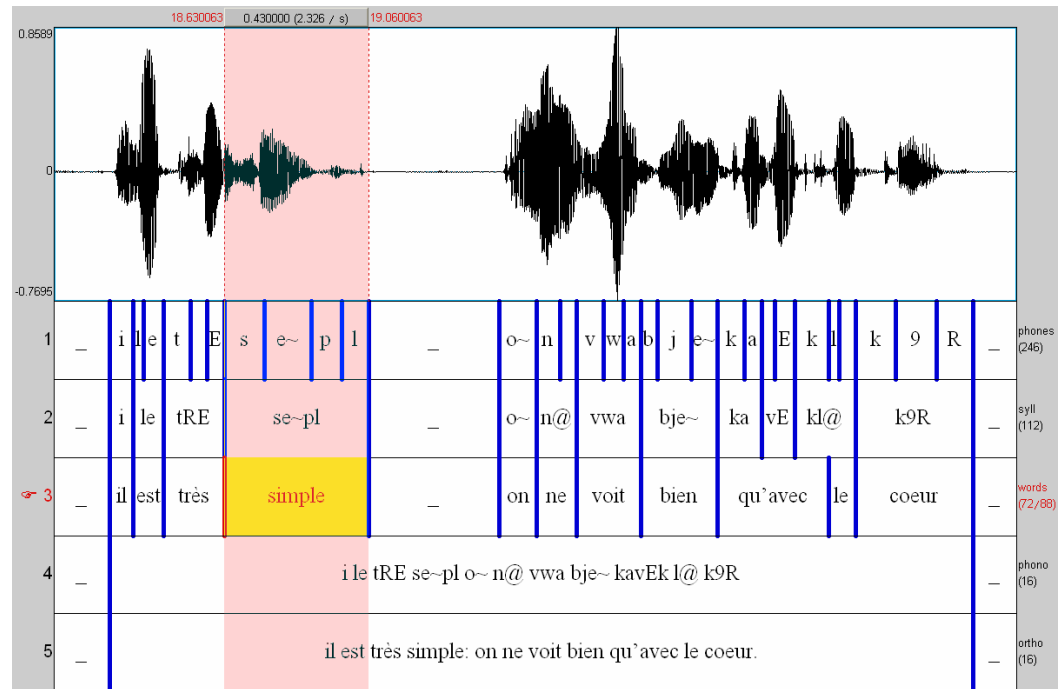


Signal digital (WAV, 24-bit, 48,000 Hz)



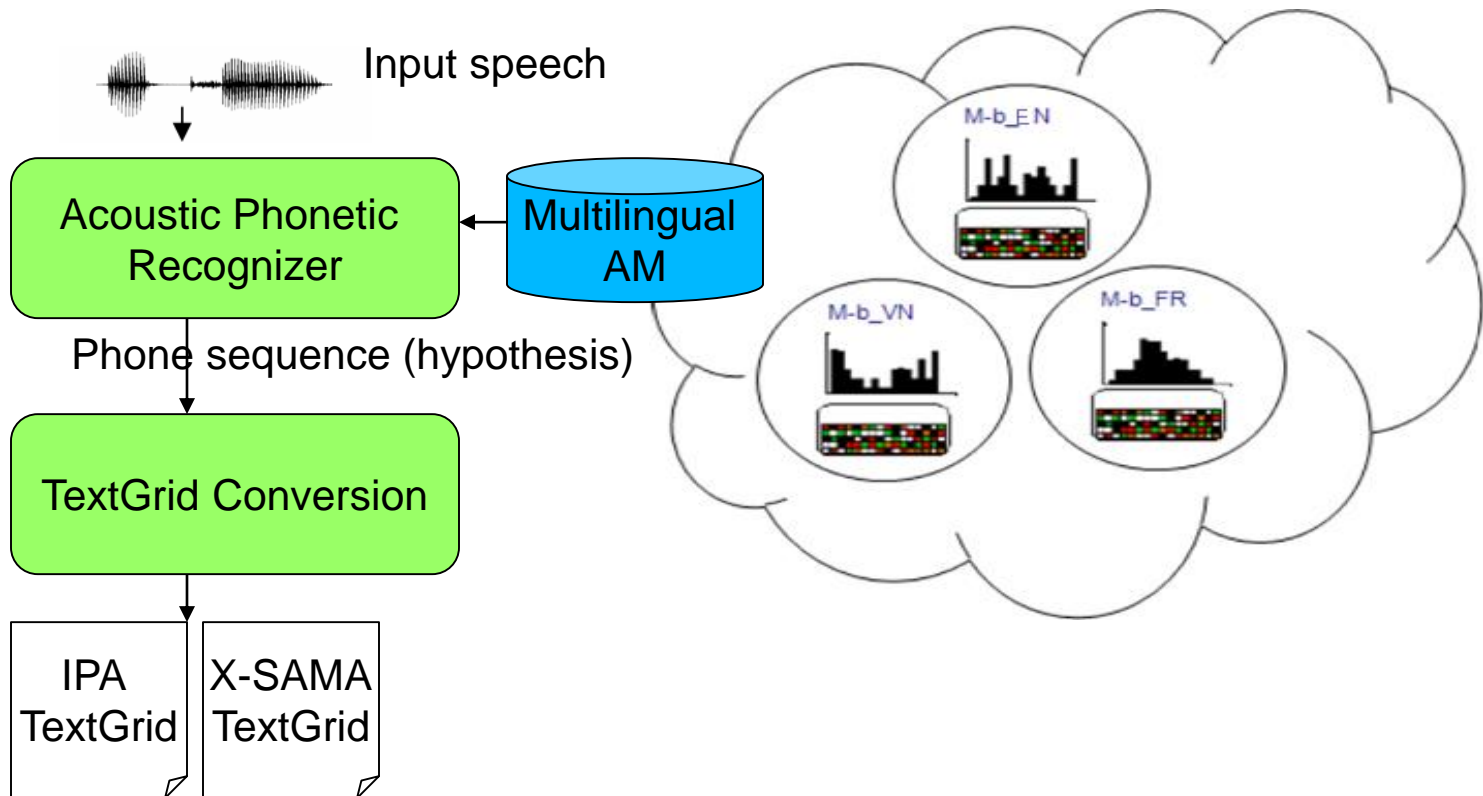
# Processing

- **Transcription - Manually**
  - ◆ Linguistics /Phonetician
  - ◆ Time-consuming
    - ★ > 10h working for 1 hour of signal (Word level)
    - ★ > 100h working for 1 hour of signal (Phone level)



# Processing (2)

- **Transcription: Semi-automatically**
  - ◆ Multilingual speech recognition



# Processing (3)

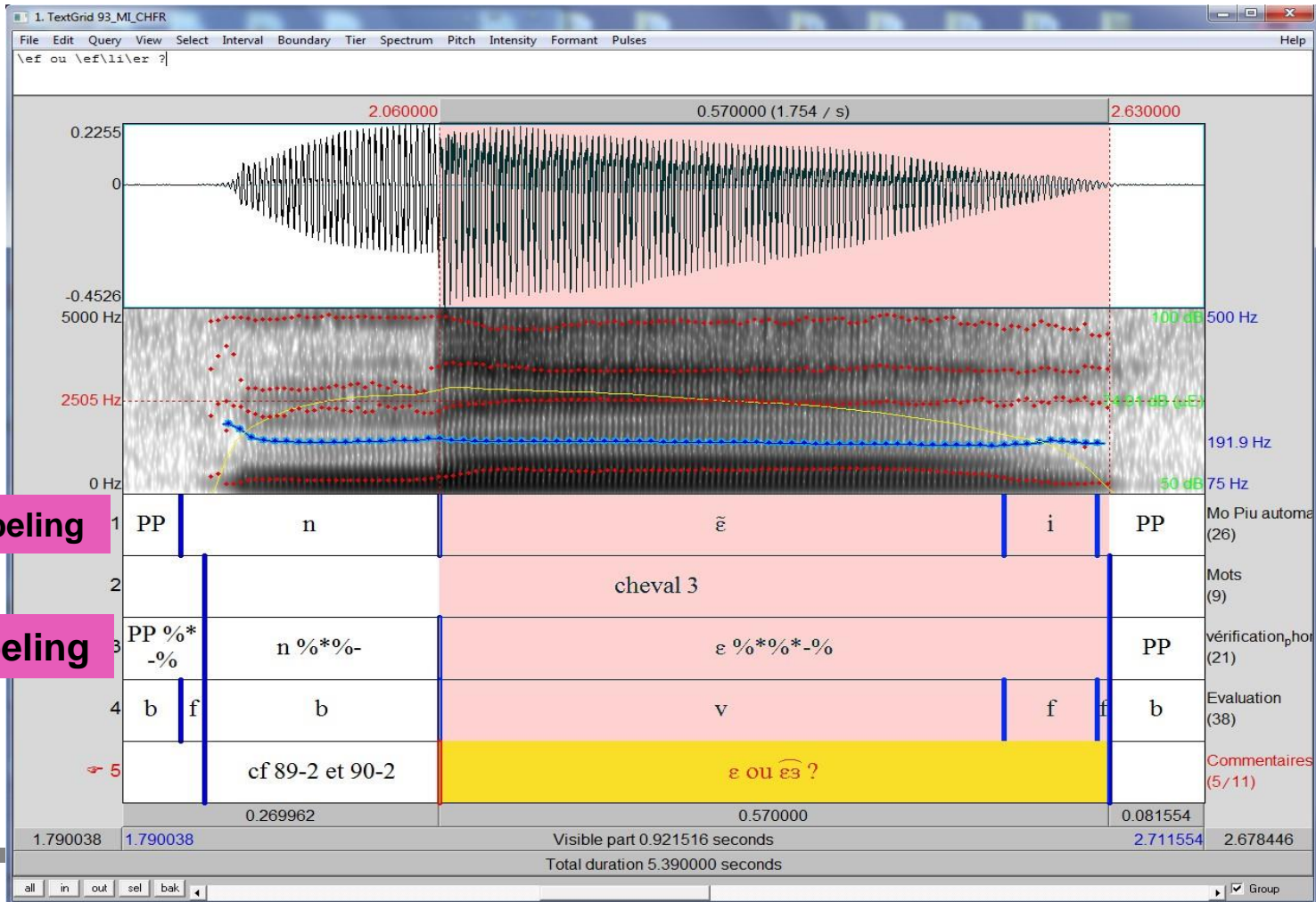
- **Transcription: Semi-automatically**
  - ◆ Experiments of phone level transcription on
    - ★ Green Mong (Mo Piu) language [1]
      - <500 speakers, unwritten languages, lack of linguistic study
      - Methods: 5 languages supply acoustic models : Vietnamese (VN), Mandarin (CH), Khmer (KH), French (FR), English (EN), each one trained on big corpora
    - ★ Na language [2]
      - Acoustic model from 5 languages: 40 English phones, 43 French phones, 34 Mandarin phones, 41 Vietnamese phones, and 36 Khmer phones

1. Caelen-Haumont G, Sam S, Castelli E (2011) *Automatic Labeling and Phonetic Assessment for an Unknown Asian Language: The Case of the “ Mo Piu ” North Vietnamese Minority (early results)*. In: Asian Language Processing (IALP), 2011 International Conference on. IEEE, pp 260–263
2. Thi-Ngoc-Diep DO, Alexis M, Eric C (2015) *Towards the Automatic Processing of Yongning Na (Sino-Tibetan): Developing a “Light” Acoustic Model of the Target Language and Testing “Heavyweight” Models from Five National Languages*. In: The 4th International Workshop on Spoken Language Technologies for Under-resourced Languages (SLTU’14. St Petersburg, Russia



# Processing (4)

- **Transcription: Semi-automatically**
  - ◆ **Phone transcription result for Mo Piu**

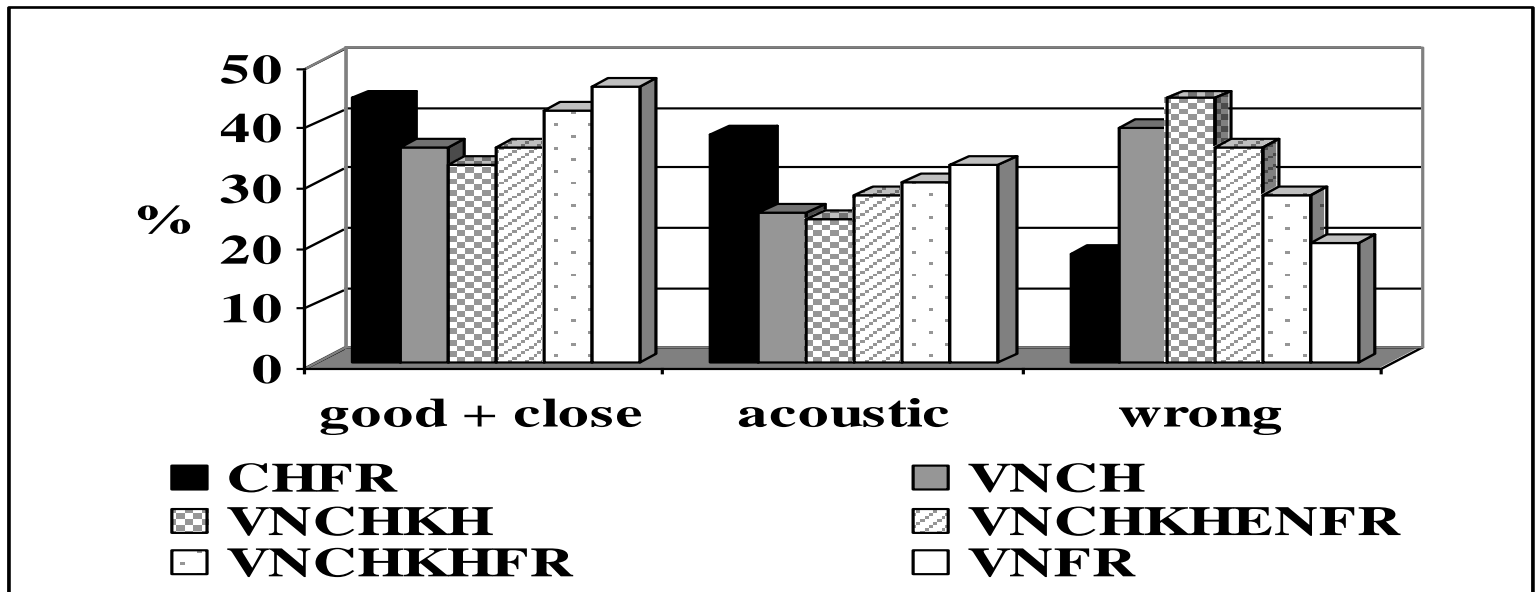


Automatic labeling

Manual labeling

# Processing (5)

- **Transcription: Semi-automatically**
  - ◆ Phone transcription result for Mo Piu



# Archiving (1)

## ■ Open Language Archives Community (OLAC)

Firefox | Open Language Archives Community | www.language-archives.org | language archive

### OLAC: Open Language Archives Community

[HOME](#) | [DOCUMENTS](#) | [ABOUT](#) | [ARCHIVES](#)  
[NEWS](#) | [ORGANIZATION](#) | [TOOLS](#) | [SERVICES](#)

#### OLAC Mission

OLAC, the Open Language Archives Community, is an international partnership of institutions and individuals who are creating a worldwide virtual library of language resources by: (i) developing consensus on best current practice for the digital archiving of language resources, and (ii) developing a network of interoperating repositories and services for housing and accessing such resources.

#### News

**New OLAC Search Service:** In December 2010, to mark our 10th anniversary, OLAC announces a new search service... [More...](#)

**Language Archives Tutorial at LSA:** In January 2009, an OLAC-endorsed tutorial on language archives was held at the winter meeting of the Linguistic Society of America... [More...](#)

**OLAC Presented at three Pacific Conferences:** In 2008 and 2009, OLAC is being presented at conferences in Australia, the Philippines, and Hawaii... [More...](#)

**OLAC Presents at Driver Summit:** In January 2008, the DRIVER project convened a Summit [re] an operational confederation of digital repositories... [More...](#)

#### Find Language Resources

Try OLAC's new search engine at: <http://search.language-archives>

Search

Archive: -- all archives --

Region: [Africa](#) [Americas](#) [Asia](#) [Europe](#) [Pacific](#)

#### OLAC Coverage

OLAC Archives contain over 100,000 records, covering resources

# Archiving (2)

## ■ Participating Archives

Firefox OLAC Metrics +

www.language-archives.org/metrics/compare language archive

OLAC Archive Metrics Comparative Archive Metrics

(Click column headers to sort)

Archive	Overall Rating	Number of Resources	Number of Resources Online	Distinct Languages	Distinct Linguistic Subfields	Distinct Linguistic Types	Distinct DCMI Types	Average Elements Per Record	Average Encoding Schemes Per Record	Average Metadata Quality Score	Last Updated	Known Integrity Problems
Collections de Corpus Oraux Numeriques (CoCoON ex-CRDO)	★★★★★	4784	4783	110	17	3	4	30.3	21.4	9.9	2013-09-04	0
Multimodal Learning and teaching Corpora Exchange	★★★★★	49	49	4	5	1	2	60.9	41.1	9.9	2013-08-29	<a href="#">6</a>
POLLEX-Online	★★★★★	65	65	60	0	1	1	13.0	9.0	9.8	2013-09-06	0
The LDC Corpus Catalog	★★★★★	560	0	70	0	2	4	23.5	7.0	9.5	2013-08-19	<a href="#">3</a>
TransNewGuinea.org	★★★★★	315	315	275	0	1	1	11.0	7.0	9.5	2013-09-05	<a href="#">2</a>
WALS Online	★★★★★	2678	2678	2479	0	1	1	11.0	7.0	9.5	2013-08-31	0
TALKBANK Data repository	★★★★★	216	216	30	7	1	3	16.6	9.1	9.3	2013-04-25	0
Glottolog 2.0	★★★★★	7651	7651	7636	0	1	1	10.0	7.0	9.3	2013-09-06	<a href="#">4</a>
The Sociolinguistic Archive and Analysis Project (SLAAP)	★★★★★	33	0	3	1	1	1	14.0	9.0	9.3	2013-05-22	0

# Archiving (3)

## ■ Pangloss collection:

Accueil > Accueil Pangloss > Les langues du Caucase > Corpus oubykh

### Oubykh

Fiche technique sur l'oubykh

La langue oubykh a fait l'objet de nombreuses études par le caucasologue Georges Dumézil (1898-1986). Elle représentait, pour les médias de la fin du 20e siècle, le type même de langue en danger, puis de langue éteinte sous nos yeux avec la disparition du dernier locuteur en 1992. La richesse de son inventaire de consonnes est restée légendaire dans le monde de la linguistique.

Les textes présentés ici proviennent d'une série de dix bandes enregistrées par G. Dumézil à Paris en 1968 avec le dernier locuteur, M. Tevfik Esenç (ou Tevfik Sanıç). Nous présentons ici une série de sept enregistrements de récits tirés de la bande n° 3.

Les transcriptions et gloses mot-à-mot en français de quatre de ces histoires, de l'écriture de Dumézil lui-même, ont été retrouvées et sont disponibles ici sous forme numérisée en format pdf. L'une de ces histoires, "La chair de poisson rend intelligent", a été retrouvée également dans une deuxième transcription, comprenant des gloses de morphème en plus des gloses de mot, préparée par Dumézil pour une linguiste non-spécialiste, Mme Christine Leroy. La meilleure lisibilité de cette deuxième annotation nous a encouragé à l'informatiser (sans les gloses de morphème) et à la synchroniser avec l'enregistrement sonore, avec l'aide de Mme Dina Dabjen-Bailly, locutrice de tcherkesse.

<a href="#">Corpus abzakh</a>
<a href="#">Corpus bjedough</a>
<a href="#">Corpus chapsough</a>
<a href="#">Corpus laze</a>
<a href="#">Corpus oubykh</a>
<a href="#">vers Les langues du Caucase</a>

# Archiving (5)

## ■ Data format standard:

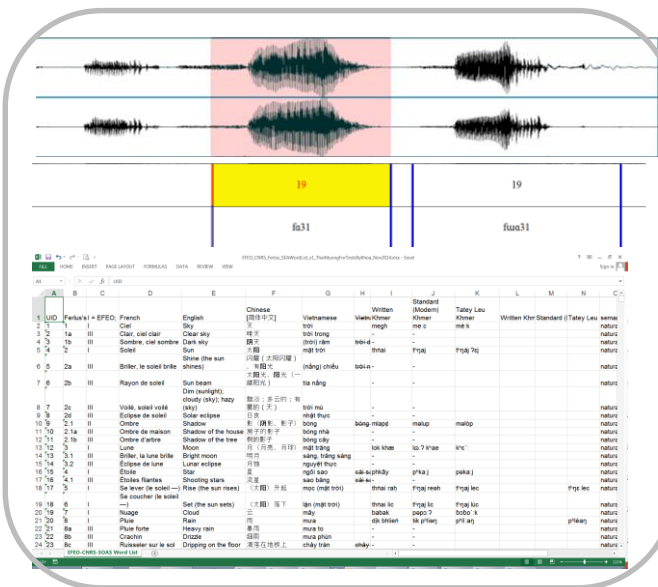
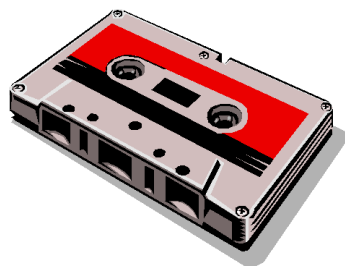
```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/xsl" href="view_text.xsl"?>
<WORDLIST id="EFEO_CNRS_Ferlus_SEAWordList_v1_ThaiMuongForTestsByKhoa_Nov2014" xml:lang="mtg">
  <HEADER>
    <TITLE xml:lang="eng">Vocabulary list for Tai Muong, part 1</TITLE>
    <SOUNDFILE href=" ../audio/TaiMuong13_voc1_items1a82_1.wav"/>
  </HEADER>
  <NOTE xml:lang="eng" message="This is an...>
  <W id="W1">
    <FORM kindOf="phono">phā</FORM>
    <FORM kindOf="ortho">fa:53</FORM>
    <NOTE xml:lang="eng" message="UID = 1">
    <NOTE xml:lang="eng" message="Ferlus!">
    <NOTE xml:lang="eng" message="Layer i">
    <NOTE xml:lang="eng" message="Categor">
    <NOTE xml:lang="fra" message="Catégor">
    <TRANSL xml:lang="vie">trời</TRANSL>
    <TRANSL xml:lang="eng">sky</TRANSL>
    <TRANSL xml:lang="cmn">天</TRANSL>
    <TRANSL xml:lang="fra">Ciel</TRANSL>
    <TRANSL xml:lang="khm-written">meḡh</TRANSL>
    <TRANSL xml:lang="khm-standard-moderr">
    <TRANSL xml:lang="khm-tatevleu">me:k</TRANSL>
    <AUDIO start="0.722" end="1.787"/>
  </W>
```

Phono  Traduction (ELL)

	Phono	Ortho	Traduction (ELL)	Other
W1	<input type="checkbox"/> <input type="checkbox"/>	ut <sup>h</sup> ij <sup>2</sup> toh	(rắn) cắn bite (snake)	(蛇) 咬 mordre (serpent)
W2	<input type="checkbox"/> <input type="checkbox"/>	ut <sup>h</sup> ij <sup>2</sup> toh	(rắn) cắn bite (snake)	(蛇) 咬 mordre (serpent)
W3	<input type="checkbox"/> <input type="checkbox"/>	ut <sup>h</sup> ij <sup>2</sup> tla:l	bò crawl (snake)	(蛇) 爬 ramper
W4	<input type="checkbox"/> <input type="checkbox"/>	ut <sup>h</sup> ij <sup>2</sup> tla:l	bò crawl (snake)	(蛇) 爬 ramper
W5	<input type="checkbox"/> <input type="checkbox"/>	ut <sup>h</sup> ij <sup>2</sup> tah kuso:i <sup>2</sup>	rắn lột xác, rắn thay da sloughed skin (of snake)	(蛇) 的) 蜕皮 mue de serpent
W6	<input type="checkbox"/> <input type="checkbox"/>	ut <sup>h</sup> ij <sup>2</sup> tah kuso:i <sup>2</sup>	rắn lột xác, rắn thay da sloughed skin (of snake)	(蛇) 的) 蜕皮 mue de serpent



# Data packaging

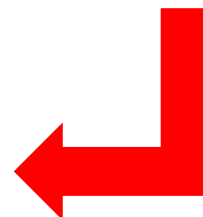


TextGrid

Wordlist

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/xsl" href="view_text.xsl"?>
<WORDLIST id="EFEO_CNRS_Ferlus_SEAWordList_v1_ThaiMuongForTestsByKhoa_Nov2014" xml:lang="mtq">
  <HEADER>
    <TITLE xml:lang="eng">Vocabulary list for Tai Muong, part 1</TITLE>
    <SOUNDFILE href="../audio/TaiMuong13_voc1_items1a82_1.wav"/>
  </HEADER>
  <NOTE xml:lang="eng" message="This is an example of DOREMIFA project"/>
  <W id="W1">
    <FORM kindOf="phone">phã</FORM>
    <FORM kindOf="ortho">fa:53</FORM>
    <NOTE xml:lang="eng" message="UID = 1"/>
    <NOTE xml:lang="eng" message="Ferlus's numbering = 1"/>
    <NOTE xml:lang="eng" message="Layer in word list: I (EFEO)"/>
    <NOTE xml:lang="eng" message="Category / Semantic field: natural environment"/>
    <NOTE xml:lang="fra" message="Catégorie / Champ sémantique: environnement naturel"/>
    <TRANSL xml:lang="vie">trời</TRANSL>
    <TRANSL xml:lang="eng">sky</TRANSL>
    <TRANSL xml:lang="cmn">天</TRANSL>
    <TRANSL xml:lang="fra">Ciel</TRANSL>
    <TRANSL xml:lang="khm-written">megh</TRANSL>
    <TRANSL xml:lang="khm-standard-modern">meic</TRANSL>
    <TRANSL xml:lang="khm-tateylau">meik</TRANSL>
  </W>
</WORDLIST>
```

Annotation



XML generation

# Publishing

AUR logo (A, U, R) with a pink double-headed arrow pointing to an audio waveform. Below the waveform is the text "cOCoOn" and "Collections de COrpus Oraux Numériques". At the bottom is the "Pangloss Collection" logo.

## Current

Map of Southeast Asia with location markers. Below the map is a list of statistics:

- 42 languages/dialects
- 120 hours of recordings
- 100 annotated documents

Pangloss Collection website header. Logos for CNRS, cOCoOn, Huma-Num, EFL, and ANR. Navigation menu: Lacito Home, The Oral Archives, Archive Access, Submit resources, Tools, Search.

Lacito Home > Archive Home > Asian Languages > Muong Corpus

### Mường corpus

**Muong (Mường)** is the closest relative of Vietnamese, the national language of Vietnam. Within the Vietic group of the Austroasiatic language family, the Northern Vietic subgroup (also known as "Viet-Muong") consists of Vietnamese and Muong. Among the language varieties which evolved historically out of Proto-Viet-Muong, Vietnamese gradually became the most widespread and influential; in Michel Ferlus's classification, all the other language varieties of this subgroup are referred to as "Muong". Linguistically speaking, the Nguồn language is also part of Muong.

#### Resources :

All available resources here

#### The archived documents: vocabulary, and some narratives

The recordings made available here were conducted by Michel Ferlus, in collaboration with colleagues in Vietnam and Laos. They mostly contain lexical data; there are also some narratives. As Ferlus's field notes have not been verified item by item to finalize the word list, it was decided to put the data online without a transcription, rather than run the risk of publishing data that contain mistaken transcriptions.

#### References

The EFEO-CNRS-SOAS Word List for Linguistic Fieldwork in Southeast Asia, used to elicit the vocabulary, is available [here](#).

About the Mường language, see Ferlus (2004) and references therein:

Ferlus, Michel. 2004. The Origin of Tones in Viet-Muong. In Somsongue Burusphat (ed.), *Papers from the Eleventh Annual Conference of the Southeast Asian Linguistics Society 2001*, 297–313. Tempe, Arizona: Arizona State University Programme for Southeast Asian Studies Monograph Series Press.

*Classification of the Vietic languages (a.k.a. Viet-Muong, Việt-Chứt) after Ferlus 1996, with links to the corresponding pages in the Pangloss Collection:*

1. Maleng	2. Arem	3. Chứt	4. Aheu	5. Hung	6. Thổ	7. Mường (this page)	8. Vietnamese
-----------	---------	---------	---------	---------	--------	----------------------	---------------

These materials were digitized by the *International Research Institute MICA (HUST – CNRS/UMI-2954 – Grenoble INP)* in partnership with *CNRS-LACITO* as part of the *DO-RE-MI-FA* project (Sept. 2014-Feb. 2016), funded by the *Digital Scientific Library*, a programme of the French Ministry for Higher Education and Research.

bsn BIBLIOTHÈQUE SCIENTIFIQUE NUMÉRIQUE DIGITAL SCIENTIFIC LIBRARY



AUR logo

Last updated 2016.

- Aheu C
- Arem C
- Bahing
- Burme:
- Bru Co
- Cao La
- Car-Nic Corpus
- Chang I
- Chong I
- Chong I
- Chong I
- Chứt C
- Cuối C
- Hayu C
- Hung C
- Japhug
- Khaling
- Kháng I
- Khmu C
- Kha Ph
- Koyi R
- Ksing M
- Kurdist



# Publishing (2)

- Examples: [http://lacito.vjf.cnrs.fr/archivage/index\\_en.htm](http://lacito.vjf.cnrs.fr/archivage/index_en.htm)

Vocabulary list for Hoa Binh dialect of the Mường language as spoken in the commune of Dan Chu [female speaker], part 1 of 2  
Language : Mường

Researcher(s) : Ferlus, Michel

Speaker(s) : Nguyễn Thị Hồng Vinh; Nguyen, Thi Hong Vinh; Nguyen, Thi Hong Vinh

00:22 / 47:23 Continuous playing:

Transcription

Translation

Notes

fr  en  cn  vn  khm-written  khm-standard-modern  khm-tateyleu

W1		trời	sky	天	ciel	megh	me:c	mè:k
W2		trời	sky	天	ciel	megh	me:c	mè:k
W3		mặt trời	sun	太阳	soleil	thñai	ʔñaj	ʔñaj ʔej
W4		mặt trời	sun	太阳	soleil	thñai	ʔñaj	ʔñaj ʔej
W5		bóng	shadow	影 ( 阴影 , 影子 )	ombre	mlap'	məɭup	məɭòp
W6		bóng	shadow	影 ( 阴影 , 影子 )	ombre	mlap'	məɭup	məɭòp
W7		mặt trăng	moon	月 ( 月亮 , 月球 )	lune	lok khæ	lɑ:ʔ kʰæ	kʰɛ:
W8		mặt trăng	moon	月 ( 月亮 , 月球 )	lune	lok khæ	lɑ:ʔ kʰæ	kʰɛ:
W9		ngôi sao	star	星	étoile	phkây	pʰka j	pəka j
W10		ngôi sao	star	星	étoile	phkây	pʰka j	pəka j
W11		mọc (mặt trời)	rise (the sun rises)	( 太阳 ) 升起	se lever (le soleil —)	thñai rah	ʔñaj reəh	ʔñaj lec
W12		mọc (mặt trời)	rise (the sun rises)	( 太阳 ) 升起	se lever (le soleil —)	thñai rah	ʔñaj reəh	ʔñaj lec

# Thank you for attention!

