

# Les défis des grammaires et des dictionnaires des langues en voie d'extinction

The screenshot displays a linguistic analysis software interface. At the top, a timeline shows the audio file 'duoxu006.wav' with time markers from 00:00:48.000 to 00:00:57.000. Below the waveform, a transcription timeline is visible, with the text 'Katia Chirkova (CNRS-CRLAO)' overlaid. The transcription includes phonetic segments such as 'tʰe<sup>53</sup>pu<sup>21</sup> kʰa<sup>53</sup> la ja<sup>22</sup>ju<sup>33</sup>m...' and 'mi<sup>21</sup>va<sup>53</sup> ma<sup>53</sup> pʰa<sup>21</sup>'. The interface also shows a list of linguistic resources on the left, including 'A\_句子-txt-ers-CN [93]', 'A\_单词-txt-ers-CN [842]', 'Interlinear-文件名-e [0]', 'A\_翻译-gls-zh-CN [87]', 'A\_参与者注释-zh-C [93]', 'B\_句子-txt-ers-CN [110]', and 'B\_单词-txt-ers-CN'. At the bottom, the text 'Journées du Séminaire de Master LCT 17 et 18 mars 2021' is displayed, along with a small table containing the Chinese characters '有一天'.

duoxu006.wav 00:00:48.000 00:00:49.000 00:00:50.000 00:00:51.000 00:00:52.000 00:00:53.000 00:00:54.000 00:00:55.000 00:00:56.000 00:00:57.000

Katia Chirkova (CNRS-CRLAO)

A\_句子-txt-ers-CN [93]  
A\_单词-txt-ers-CN [842]  
Interlinear-文件名-e [0]  
A\_翻译-gls-zh-CN [87]  
A\_参与者注释-zh-C [93]  
B\_句子-txt-ers-CN [110]  
B\_单词-txt-ers-CN

有一天  
有一天

Journées du Séminaire de Master LCT  
17 et 18 mars 2021

# Linguistique descriptive

- l'objectif principal est d'aboutir à une description la plus exhaustive possible d'une langue auparavant non décrite
- trois résultats principaux (cf. Dixon 2007 : 19-20)
  1. une grammaire de référence complète
  2. un volume de textes
  3. un dictionnaire (ou un thésaurus)

# Langues en danger

- l'attrition linguistique et la disparition des langues
- le rythme auquel les langues disparaissent aujourd'hui est sans précédent
- depuis les années 1990 la crise de la diversité linguistique fait l'objet d'un intérêt grandissant des linguistes et du public (e.g. Dorian 1989 ; Dixon 1997 ; Grenoble et Whaley 1998, 2006 ; Crystal 2000 ; Evans 2010)

**Organisation des Nations Unis pour l'Éducation, la Science et la Culture**



## **VITALITE ET DISPARITION DES LANGUES**

**Groupe d'experts spécial de l'UNESCO sur les langues en danger**

# Critères UNESCO

- Facteur 1** : Transmission de la langue d'une génération à l'autre : proportion d'enfants parlant la langue
- Facteur 2** : Nombre absolu de locuteurs
- Facteur 3** : Taux de locuteurs sur l'ensemble de la population et âge des derniers locuteurs
- Facteur 4** : Utilisation de la langue dans les différents domaines publics et privés
- Facteur 5** : Réaction face aux nouveaux domaines et médias
- Facteur 6** : Matériels d'apprentissage et d'enseignement des langues
- Facteur 7** : Attitudes et politiques linguistiques au niveau du gouvernement et des institutions : usage et statut officiels
- Facteur 8** : Attitude des membres de la communauté vis-à-vis de leur propre langue
- Facteur 9** : Type et qualité de la documentation

# Critères UNESCO

Ces critères permettent de définir **six niveaux de vitalité de langues en danger** :

SURE (langues dominantes, nationales)	5
<hr/>	
STABLE mais MENACEE*	5-
PRECAIRE	4
EN DANGER	3
SERIEUSEMANT EN DANGER	2
MORIBONDE	1
MORTE	0

\* Classifiée comme "viable", c'est-à-dire qui a une chance de survivre jusqu'à la fin du 21ème siècle.



# Sino-Tibetan

## Tibeto-Burman

## Chinese

**Kamarupan**  
NE India,  
W Burma

**Baic**  
SW China (Yunnan)

**Northern  
Mandarin**  
**SW Mandarin**

**Karenic**  
Burma, Thailand

**Himalayish**  
Tibet, Nepal,  
Bhutan, Sikkim

**Lolo-Burmese-Naxi**  
SW China, Burma,  
Thailand, Laos, Vietnam

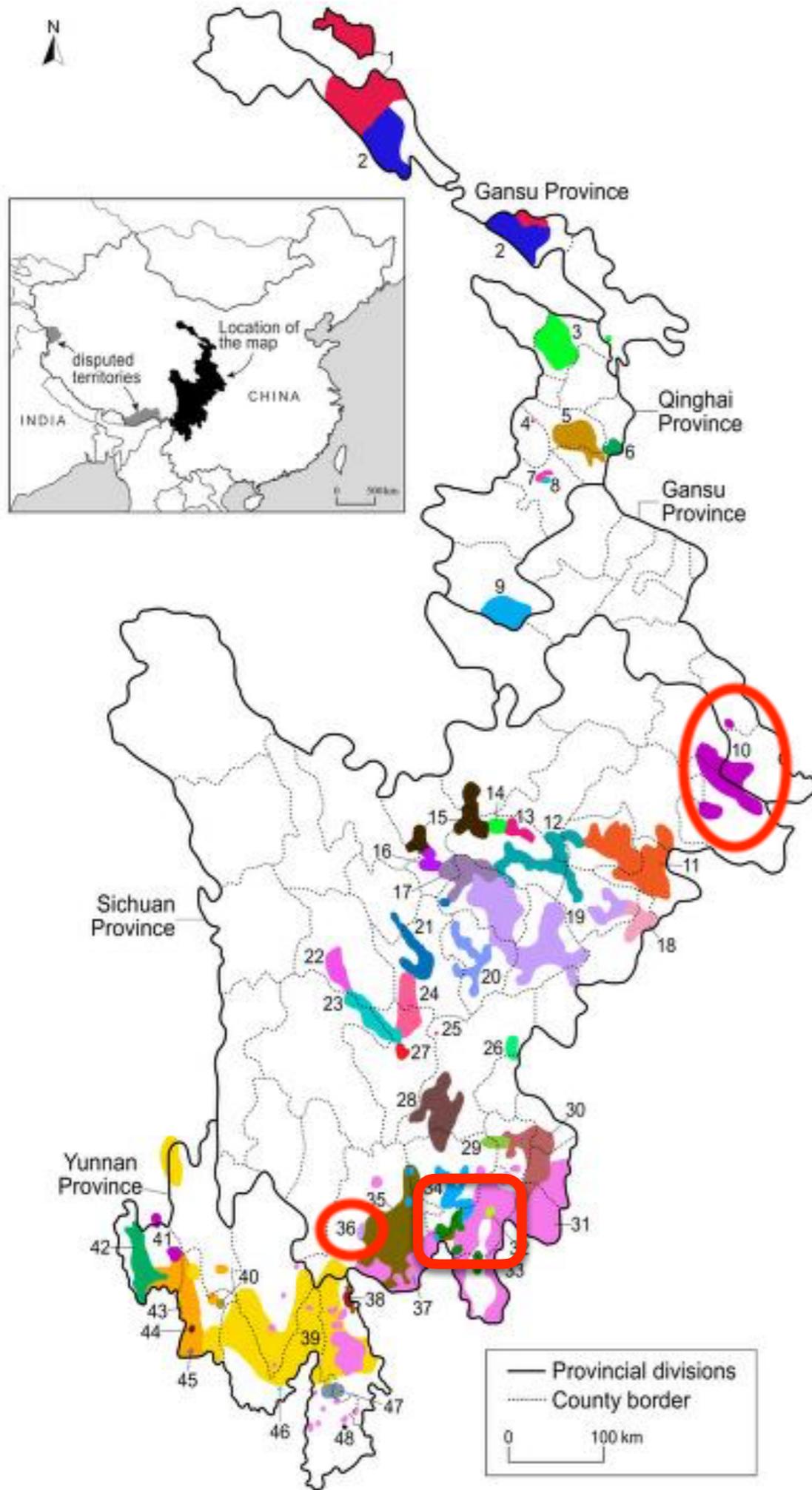
**Central**  
Wu, Xiang,  
Gan, Hakka

**Qiangic**  
SW China  
(Sichuan, Yunnan)

**Jingpho-Nungish-Luish**  
N Burma, SW China (Yunnan)

**Southern**  
Yue, Min

Based on: Matisoff, James A. 2003. *The Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press, 5.



le baima  
langue tibétique

le shuhi  
le lizu  
le ersu  
le duoxu  
langues qianguiques

par Chandra Jayasuriya, données linguistiques par Gerald Roche et Hiroyuki Suzuki, <https://archive.org/details/MinorityLanguagesOfTheEasternTibetosphere>

# Baima 白马 pé





Shuhi

书亨

ᠰᠤᠬᠢ

# Ersu 尔苏 ḙsḙ



Lizu  
里汝

lízú



Duoxu

多续

do<sup>33</sup> ɕu<sup>33</sup>



Langue  
Facteur

Baima

Shuhi

Lizu

Ersu

Duoxu

Facteur 1:  
Transmission  
de la langue  
d'une  
génération à  
l'autre

Facteur 2:  
Nombre absolu  
de locuteurs

Facteur 9:  
Type et qualité  
de la  
documentation



# Linguistique descriptive

- collecter des données primaires et fournir une analyse descriptive de ces données
- buts: décrire une langue sous tous ses aspects, présenter l'analyse sous la forme d'une grammaire et d'un dictionnaire
- public: linguistes descriptifs et comparatifs

# A Grammar and a Dictionary of Duoxu



# Problèmes

- décrire une langue sous tous ses aspects
  - pour un seul auteur / un seul ouvrage, il est impossible de produire une analyse qui sera également détaillée par rapport à tous les sous-systèmes linguistiques
  - il n'existe pas de procédure automatique et infaillible pour produire l'analyse descriptive à partir d'un corpus de données primaires
  - toute collection de données primaires permet divers types d'analyses même dans le cadre de la linguistique descriptive
  - il subsiste toujours une possibilité d'erreur

# Problèmes

- public: linguistes descriptifs et comparatifs
  - les données primaires peuvent intéresser diverses autres (sous-) disciplines (sociolinguistique, anthropologie, analyse du discours, histoire orale)
  - la communauté

# Solutions possibles

- l'accent sur la collecte des données primaires et la présentation de ces données sous forme ouverte
  - enregistrements audio/vidéo de bonne qualité avec annotation
    - textes interlinéaires
    - listes de mots, règles grammaticales etc.
      - la présentation de l'analyse sous une autre forme

# Solutions possibles

- projet collectif
  - participation d'autres chercheurs
  - des membres de la communauté

# Défis

- Comment recueillir un corpus représentatif?
- Comment couvrir un maximum d'aspects de la langue documentée?
- Comment rendre les données ouvertes et accessibles?
- Comment intégrer le travail linguistique aux besoins de la communauté?

# Un corpus représentatif

- les objectifs de recueillir un corpus représentatif, de couvrir un maximum d'aspects de la langue documentée, de mettre à disposition l'ensemble des données recueillies : la documentation linguistique
- la documentation linguistique vise à collecter et à archiver à long terme des bases de données documentant les pratiques langagières d'une communauté
- la documentation linguistique a également pour but de recueillir des informations linguistiques et culturelles liées à cette communauté et à sa langue afin de créer « a lasting, multipurpose record of a language » (Himmelman 2006 : 1)

# Un corpus représentatif

- la représentativité dépend de l'échantillon, de sa constitution et de la sélection des participants
  - un éventail large de registres linguistiques à enregistrer (du spontané au plus soigné)
  - le travail avec plusieurs locuteurs
- ce dernier point soulève nécessairement le problème de la variation linguistique dans la communauté, d'autant plus important dans le cas des langues à tradition orale, que ces dernières n'ont par définition pas de normes participant de l'usage écrit

# Rendre les données ouvertes

- l'objectif d'archiver les données d'une façon pérenne et ouverte contribue à la transparence des données utilisées dans un but d'analyse, donnant ainsi un côté empirique et vérifiable aux recherches linguistiques
- la mise à disposition de l'ensemble des données primaires permet aux autres utilisateurs de vérifier l'analyse et les conclusions, et de reproduire les résultats, ce qu'impliquent les caractères définitoires du savoir scientifique que sont la falsifiabilité et la réfutation par l'expérience (cf. Popper 1961, 1969)

# Archivage

- assurer la pérennité et la disponibilité des données collectées sur le terrain
- la mise à disposition de l'ensemble des données recueillies entraîne une économie des ressources déployées pour la recherche, tels que le financement des travaux de terrain, en rendant les données de terrain accessibles à:
  - d'autres linguistes
  - aux membres de la communauté
  - aux chercheurs d'autres domaines, voire toutes autres parties intéressées

# Communauté

- l'objectif d'intégrer le travail linguistique aux besoins de la communauté aide à impliquer les locuteurs dans le processus de la collecte de données, augmente pour les locuteurs le prestige de la langue et de la culture (souvent très limité dans le cas des langues menacées) et aide les communautés à savoir la manière dont elles aimeraient être aidées (par exemple, Grenoble et Whaley 1996: 220-221, Hinton 2001, Grenoble 2013)
- avec un accès généralisé à des appareils d'enregistrement audio et vidéo de bonne qualité, et une augmentation du nombre de plateformes de partage de données numériques, les locuteurs commencent à jouer un rôle plus important dans le processus de collecte et de partage des données

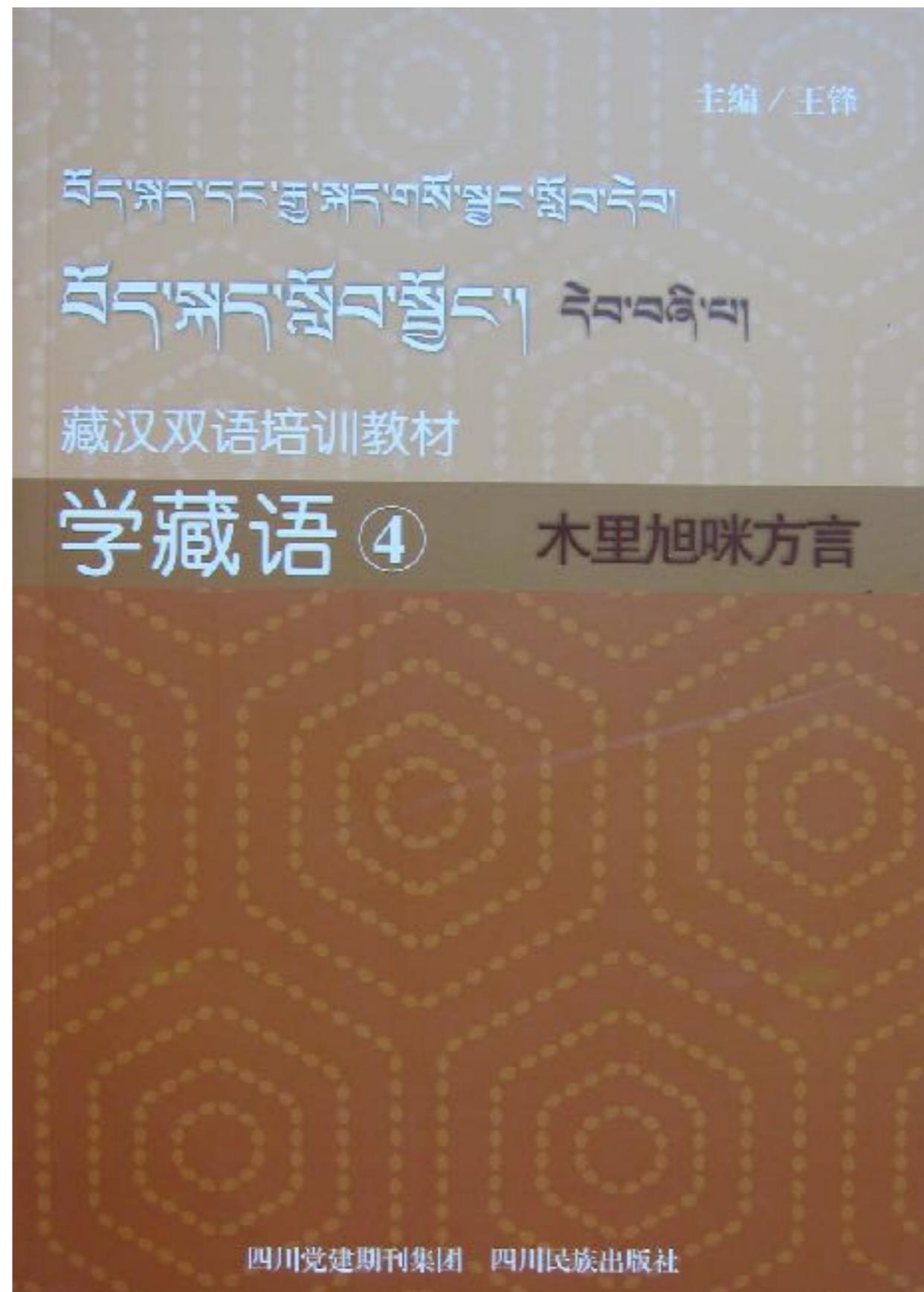
# Implications

- un linguiste travaillant sur le terrain peut être amené à mesurer la pertinence descriptive et documentaire des données collectées par la possibilité pour un apprenant d'extraire suffisamment d'informations à partir de ces données pour parler couramment la langue documentée (Grenoble 2013 : 53)
- un linguiste travaillant sur le terrain peut être amené à accepter un travail dans le cadre d'une équipe de revitalisation linguistique (comprenant rédaction de supports pédagogiques, aide à la formation des enseignants et enseignement de la langue) (Gerdt 2010)

# Lizu



un DVD de chansons traditionnelles de la communauté lizu du comté de Muli (2011)



Chirkova, Katia et Duoding. 2013. *Muli Xumi Fangyan* [La langue xumi du comté de Muli]. Collection Xue Zangyu [Learn Tibetan] Series, Vol. 4. Chengdu: Minzu Chubanshe

## 三十四、反义词

དོན་ཚྭ་གཞི་ཚེག

མང་པོ་དང་ཉུང་ཉུང་།

忙波/档/扭扭。

棒/泥。

bē<sup>33</sup> ne<sup>55</sup>

1、大小。

这里有很多人。

ཆེན་པོ་དང་ཆུང་ཆུང་།

འདི་ན་མི་མང་པོ་འདུག།

钦波/档/邛邛。

迪纳/迷/忙波/杜。

堆/支。

哈弄/亨/巴哈易。

du<sup>33</sup> tʂɿ<sup>55</sup>

xe<sup>33</sup> no<sup>55</sup> hī<sup>55</sup> be<sup>33</sup> be<sup>55</sup> ji<sup>35</sup>.

这是一棵大树。

那里人很少。

ཤིང་ལུང་འདི་ཆེན་པོ་ཞིག་རེད།

དེ་སར་མི་ཉུང་ཉུང་ཉུང་རེད།

薪朋/迪/钦波/细热。

得萨/迷/哈江/扭扭/热。

哈日/森棒/嘛堆/叭棒。

堤弄/亨/喀尼易。

xe<sup>33</sup> zɿ<sup>55</sup> sī<sup>55</sup> pɔ̃<sup>31</sup> m<sup>33</sup> du<sup>33</sup> dʒi<sup>55</sup>

tʰi<sup>33</sup> no<sup>55</sup> hī<sup>55</sup> kʰe<sup>55</sup> ne<sup>55</sup> ji<sup>35</sup>.

pɔ̃<sup>55</sup>.

# Ersu

表1 尔苏语 38 个单辅音

ERS	IPA	ERS	IPA	ERS	IPA	ERS	IPA
b	p	sh	ʃ	z	ts		
p	p <sup>h</sup>	ssh	ʒ	c	ts <sup>h</sup>	dr	tɬ
bb	b	l	l	zz	dz	tr	tɬ <sup>h</sup>
m	m	lh	ɬ	n	n	ddr	ɬɬ

ERS	IPA	ERS	IPA	ERS	IPA	ERS	IPA
w	w	j	tɕ	s	s	r	ɽ
f	f	q	tɕ <sup>h</sup>	ss	z	g	k
v	v	jj	dʒ	zh	tʃ	k	k <sup>h</sup>
d	t	x	ɕ	ch	tʃ <sup>h</sup>	gg	g
t	t <sup>h</sup>	xx	ʒ	zzh	dʒ	h	x
dd	d	y	j			ng	ŋ

表5 前置鼻音 14 个复辅音例举

ERS	IPA	ERS	IPA	ERS	IPA	ERS	IPA
nbb	mb	nc	nts <sup>h</sup>	mgg	ŋg	ntr	ɳt <sup>h</sup>
np	mp <sup>h</sup>	nzzh	nɕʒ	mk	ŋk <sup>h</sup>	njj	ndʒ
ndd	nd	nch	ntʃ <sup>h</sup>	nz	ndz	nq	ntɕ <sup>h</sup>
nt	nt <sup>h</sup>	nddr	ɳdɽ				

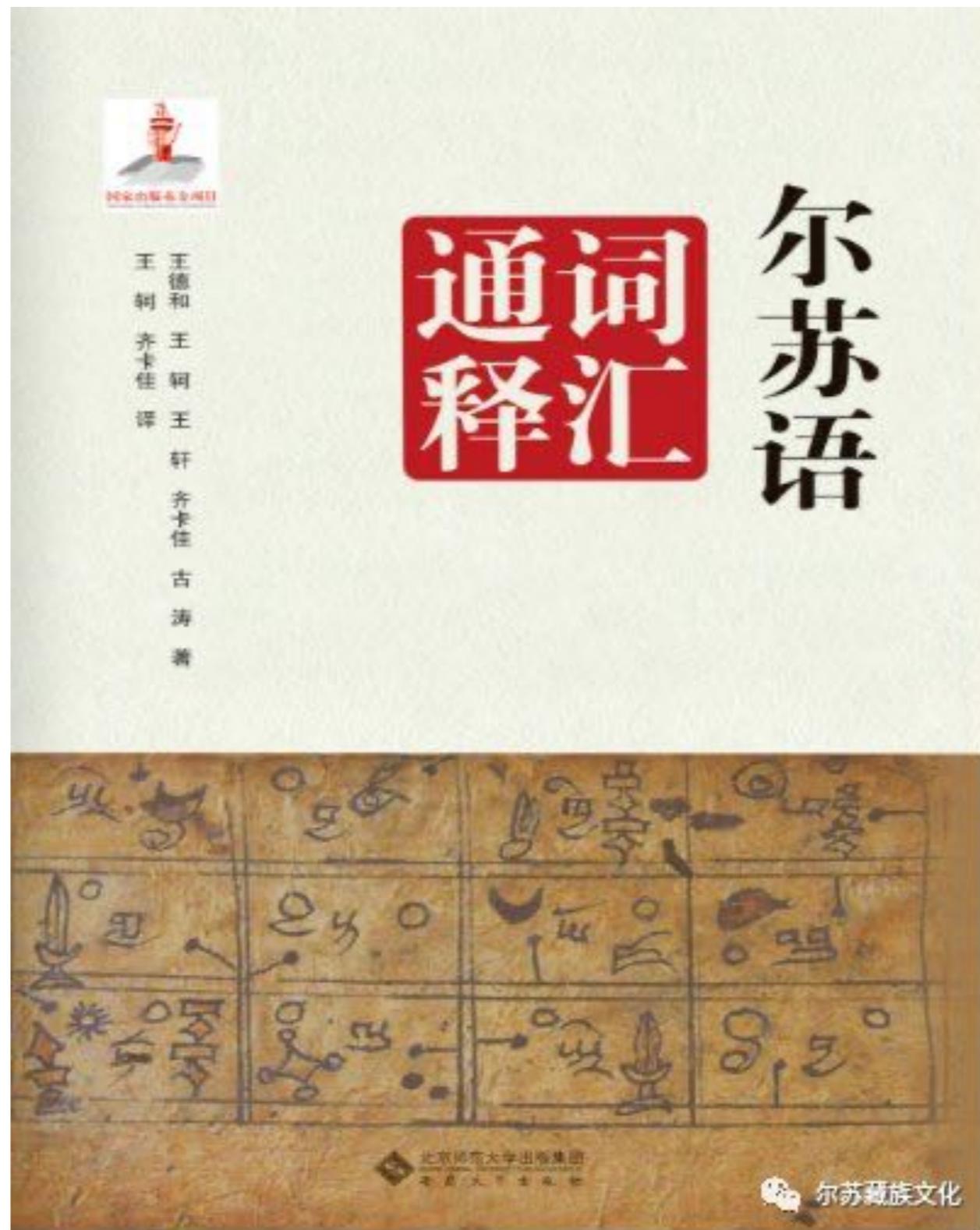
表9 尔苏语元音例表

ERS	IPA	ERS	IPA	ERS	IPA	ERS	IPA
i	i	e	ɛ	a	a	o	o
yi	ʒ	u	ʏ	er	ə	ar	ɑ

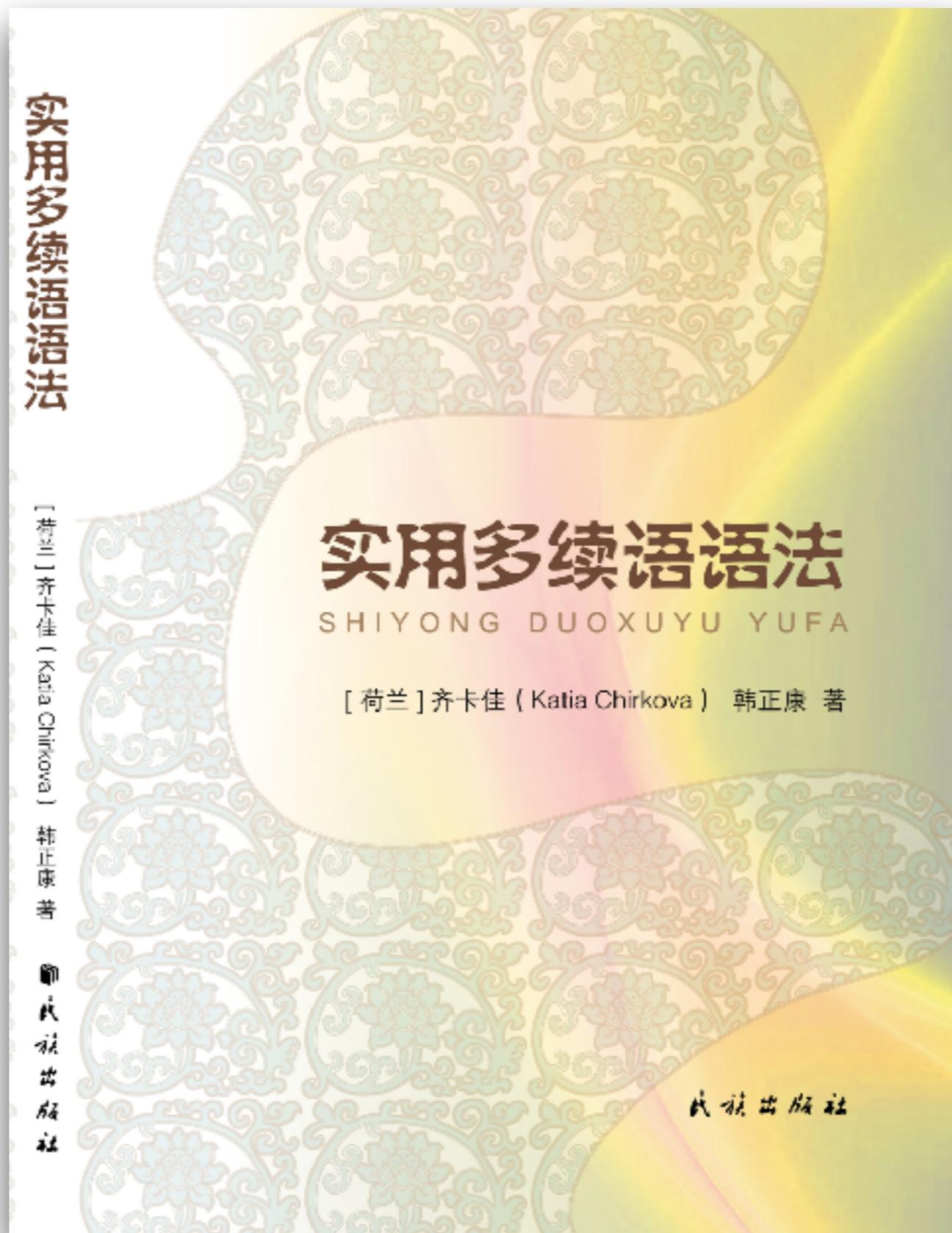
尔苏语元音组词例举如下(表 10):

表 10 尔苏语元音构成音节的词汇例举

ERS	IPA	汉译	ERS	IPA	汉译	ERS	IPA	汉译	ERS	IPA	汉译
issa	ízá	儿子	zi	tsí	头发	i	ji	屋	zyi	tsʒ	喂
ela	élá	公羊	ze	tsé	云彩	u	u	鸣	zu	tsý	灶、豆
a	á	我	za	tsá	捆、拴	ar	à	咬			
ojja	ódzá	梨	zo	tsó	堆积	er	é	年	erbbe	óbé	法螺



Wang Dehe, Wang Ke, Wang Xuan, Katia Chirkova, Gu Tao. 2020. *Ersuyu Cihui Tongshi* [Dictionnaire de la langue ersu]. Hefei: Anhui Chubanshe.



Chirkova, Katia et Han Zhengkang. 2016. *Shiyong Duoxuyu Yufa* [Une grammaire pratique de la langue duoxu]. Pékin : Minzu Chubanshe.



Han Zhengkang, Katia Chirkova, Yuan Xiaowen. 2020. *Sichuan Mianning Duoxuhua* [La langue duoxu du comté de Mianning de la province du Sichuan]. Pékin : Shangwu Yinshuguan.

# Collections en ligne

- Duoxu : <http://elararchive.org/duoxu-chirkova-0318>
- Ersu : <http://elararchive.org/ersu-chirkova-0316>
- Lizu : <http://elararchive.org/lizu-chirkova-0317>
- Shuhi/Xumi : <http://elararchive.org/xumi-chirkova-0189>
- Ersu and Xumi: Comparative and Cross-Varietal Documentation of Highly Endangered Languages of South-West China (491 enregistrements) (<https://pangloss.cnrs.fr/>; <https://cocoon.humanum.fr/>)

## duoxu-chirkova-0318

Object Type: Folder  
In root of archive

Browse Archive > duoxu-chirkova-0318

Sort by Type



### duoxu-chirkova-0318

duoxu-chirkova-0318

#### Collection

Collection Handle: <http://hdl.handle.net/>

Title: Duoxu: Documentation of a Cri...

#### Location

Continent: Asia

Country: China

Region: Asia

Address: Mianning County, Sichuan ...

Latitude: 28.549809

Longitude: 102.177082

#### Languages

##### Language

Name: Duoxu

##### Project

ID: MDP0257

Name:

Description:

#### Project Information

Deposit Status: Collection online

Funding Body: ELDP

#### Actor

Full Name: Katia Chirkova

Role: Depositor

Affiliation: CNRS-CRLAO

#### Actor

Full Name: 多续语 正康 韩

Role: Depositor

Affiliation: 教师

Show more ▼

**0318-00000000 - "Old ox" festival**

0318-00000000 - "Old ox" festival

**Session Information**

Session Handle: http://hdl.handle.n..  
 Title: "Old ox" festival  
 Description: Wu Decai (吴德才) spea..  
 Date created: 2013-04-12

**Location**

Continent: Asia  
 Country: China  
 Region: Asia  
 Address: Heshang village, Mianning..

**Keywords**

Keyword: Wu Decai  
 Keyword: Old Ox Festival  
 Keyword: Retelling  
 Keyword: Translation  
 Keyword: Discussion  
 Keyword: Traditional Festival

**Content**

Genre: Narrative

**Languages**

**Language**  
 Name: Duoxu

**Language**

Name: Chinese, Mandarin

**Actor**

Full Name: Han Zhengkang (韩正康)

**Actor**

Full Name: Jin Wanxiang

**Actor**

Full Name: Katia Chirkova

**Actor**

Full Name: Ma Wenming

**Actor**

Full Name: Ma Wenxiu

**Actor**

Full Name: Song Deyou

**Actor**

Full Name: Wang Dehe (王德和)

**Actor**

Full Name: Wu Decai (吴德才)

**Actor**

Full Name: Wu Dehui

**Actor**

Full Name: Wu Denglian

**Actor**

Full Name: Wu Rongfu (伍荣富)

**Actor**

Full Name: Ma Wenming

**Actor**

Full Name: Wu Rongfu (伍荣富)

**Actor**

Full Name: Wu Denglian

**Actor**

Full Name: Ma Wenxiu

**Actor**

Full Name: Wu Decai (吴德才)

**Actor**

Full Name: Wu Dehui

**Actor**

Full Name: Wang Dehe (王德和)

**Actor**

Full Name: Jin Wanxiang

**Actor**

Full Name: Song Deyou

**Actor**

Full Name: Han Zhengkang (韩正康)

**Actor**

Full Name: Katia Chirkova

**Media Files**

Filename: wdc-2013-04-12-05.wav

File handle: http://hdl.handle.net/2..

File type: Audio

Access: R

Filename: wdc-2013-04-12-06.wav

File handle: http://hdl.handle.net/2..

File type: Audio

Access: R

Filename: 2013-04-16-01.mts

File handle: http://hdl.handle.net/2..

File type: Video

Access: R

Filename: 2013-04-16-01.wav

File handle: http://hdl.handle.net/2..

File type: Audio

Access: R

Filename: 2013-04-16-02.mts

File handle: http://hdl.handle.net/2..

File type: Video

Access: R

Filename: 2013-04-16-02.wav

File handle: http://hdl.handle.net/2..

File type: Audio

Access: R

Recording Equipment: Fostex FR-2..

Show more ▼

VIEW GALLERY

Search within

FOLDER Q

REFINE YOUR SELECTION

No refinements available for these results



2013-04-16-01  
mpeg-video

U\_UserAccess

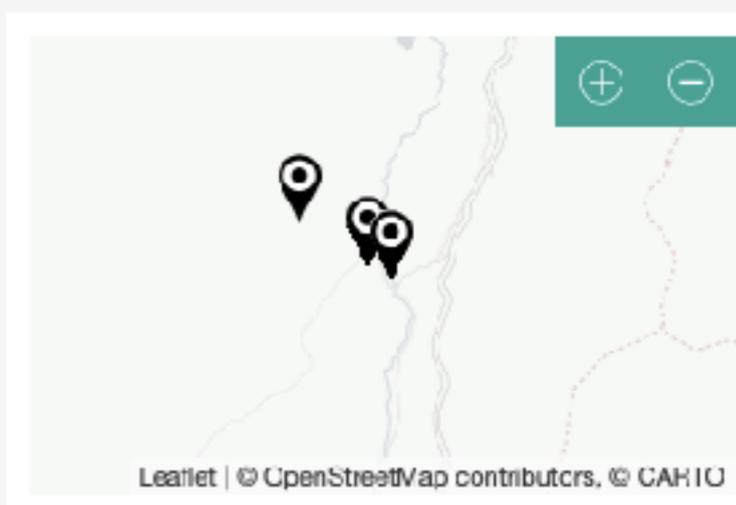


2013-04-16-01  
wave

U\_UserAccess

## Duoxu

La langue duoxu (/do<sup>33</sup>-ɕu<sup>33</sup>-na<sup>31</sup>/, Duōxù 多须 ou Duōxū 多须 en chinois mandarin) est la langue traditionnelle des Tibétains duoxu du comté de Miǎnníng (冕宁县), qui est situé dans dans la Préfecture autonome de Liángshān Yí (凉山彝族自治州) dans la province du Sichuan (四川省) en République populaire de Chine. Le duoxu fait partie d'un groupe de trois langues étroitement apparentées (le duoxu, le lizu, le ersu) avec une séparation récente d'un ancêtre commun (Sūn 1983; Wáng 2010). Les trois langues sont actuellement classées comme membres du sous-groupe qianguique de la famille des langues tibéto-birmanes (e.g. Sūn 1983, 2001; Bradley 1997: 36-37). Le lizu est parlé dans le Comté autonome tibétain de Mùlǐ (木里藏族自治县, *smi li rang skyong rdzong* en tibétain écrit, ci-après WT) et dans les comtés de Jiǔlóng (九龙县, WT *bryad zur*, dans la préfecture autonome tibétaine de Gānzī 甘孜州, WT *dkar mdzes*) et de Miǎnníng (冕宁县). Le ersu est parlé dans les comtés de (i) Gānlùò (甘洛县), (ii) Yuèxī (越西县) (les deux comtés se trouvent dans la Préfecture autonome de Liángshān Yí 凉山彝族自治州), (iii) Shímáncǎn (石棉县) et (iv) Hànyuáncǎn (汉源县) (dans la municipalité de Yǎ'ān 雅安), voir carte 1.



### Chercheurs

Chirkova, Katia

## Recherche

Mot-clé: Duoxu x



## Affichage

10 entrées

← « 1 2 3 4 5 6 7 8 9 10 » →

### Children Stories 01

 annotations  audio

"Children Stories 01" 2015. Duoxu; Chinese, Mandarin. Wu Rongfu (speaker); Han Zhengkang (researcher); Chirkova, Katia (depositor). Editeur(s): Centre de recherches linguistiques sur l'Asie orientale.

[notice](#) 

### Children Stories 02

 annotations  audio

"Children Stories 02" 2015. Duoxu; Chinese, Mandarin. Wu Rongfu (speaker); Han Zhengkang (researcher); Chirkova, Katia (depositor). Editeur(s): Centre de recherches linguistiques sur l'Asie orientale.

[notice](#) 

### Children Stories 03

 annotations  audio

"Children Stories 03" 2015. Duoxu; Chinese, Mandarin. Wu Rongfu (speaker); Han Zhengkang (researcher); Chirkova, Katia (depositor). Editeur(s): Centre de recherches linguistiques sur l'Asie orientale.

[notice](#) 

### Topological Relations Stimuli 01

 vidéo

"Topological Relations Stimuli 01" 2014. Duoxu; Chinese, Mandarin. Wu Rongfu (consultant); Han Zhengkang (researcher); Chirkova, Katia (depositor). Editeur(s): Centre de recherches linguistiques sur l'Asie orientale.

[notice](#) 

## Résultats

142 ressources trouvées

### Lieux d'enregistrement



China 142

### Langues



Ersu 112

Mandarin Chinese 90

Khams Tibetan 30

### Collections



Collection Pangloss 142

Ersu and Xumi: Comparative and Cross-Varietal Documentation of Highly Endangered Languages of South-West China 142

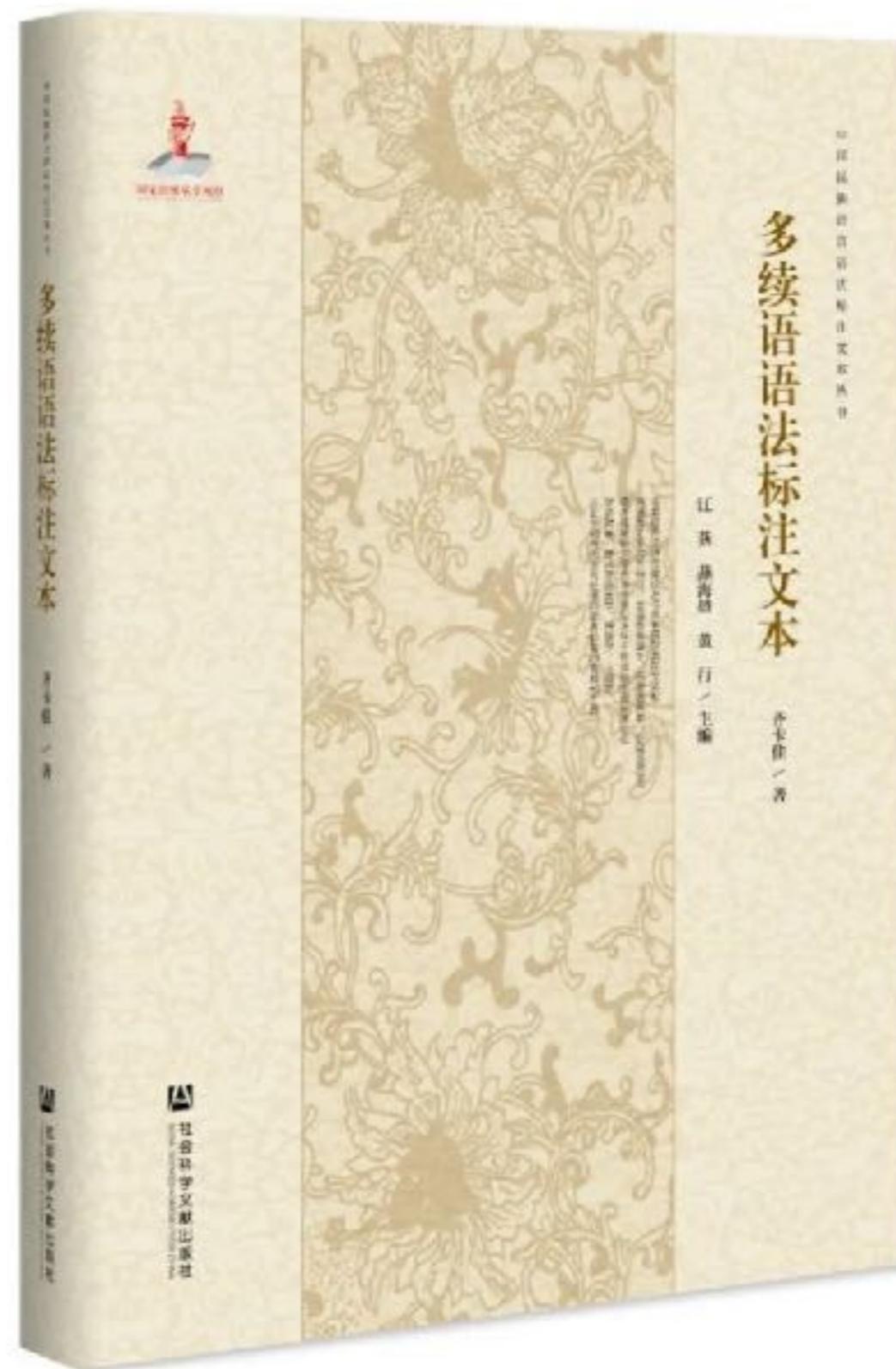
Fonds Katia Chirkova 142

Fonds Katia Chirkova: DUOXU 112

### Editeurs



Centre de recherches linguistiques sur l'Asie orientale 142



Chirkova, Katia. 2020. *Duoxuyu Yufa Biao-zhu Wenben* [Textes en duoxu avec une analyse grammaticale et une traduction en chinois]. Pékin: Minzu Chubanshe.

# Conclusion

- langues en voie d'extinction
- les données primaires
  - qualité d'enregistrement
  - représentativité
  - annotation
  - accessibilité
- grammaires, dictionnaires
  - un projet collaboratif, ouvert, interdisciplinaire

Merci de votre attention!

contact and info:

[katia.chirkova@gmail.com](mailto:katia.chirkova@gmail.com)

[www.katia-chirkova.info](http://www.katia-chirkova.info)