# Constitution d'une base de données lexicales des langues sinitiques et contribution à l'étude de leur phylogénie

YIN Yuanhao (22100358)

INALCO, SDL (M2)
Parcours : Linguistique : langues, terrains, variations, typologie



- Introduction
- 2 Stratification
- 3 Données
- 4 Phylogénie
- Conclusion



## La classification des langues sinitiques

#### Les critères traditionnels :

- Les initiales obstruantes sonores du chinois moyen
- Les syllabes à codas \*-p/t/k du chinois moyen

#### Questions:

- Phonétique/phonologique
- Rétention commune
- Absence de la stratification



## Exemple 1

R. 李. Lǐ (1987) et Xióng & Zhāng (2012):

- Mandarin 官话
- ② Jin 晋语 (séparé du mandarin)
- <sup>3</sup> Wu 吴语
- Min 闽语
- Hakka 客家话
- Yue 粤语
- Xiang 湘语
- ❸ Gan 赣语
- ❷ Hui 徽语
- Pinghua et Patois 平话和土话



## Exemple 1

#### Mandarin:

- Nord-est 东北
- ② Beijing 北京
- ③ Jilu 冀鲁
- Jiaoliao 胶辽
- ⑥ Lanyin 兰银
- ☑ Jianghuai 江淮
- Sud-ouest 西南



## Exemple 1

#### R. 李. Lǐ (1987: A2) et Shěn (1994: 5):

Initiales du MC	西南	中原	冀鲁	兰银	北京/东北	胶辽	江淮	Jin (Taiyuan)
Sourde	2	1	1	4	1234	3	D	D1
Sonante	2	1	4	4	4	4	D	D2
Sonore	2	2	2	2	2	2	D	D2

Les tonèmes des syllabes à codas \*-p/t/k du chinois moyen dans les variantes du mandarin



## Exemple 2

#### X. 李. Lǐ (2012):

- Mandarin 官话
- ② Wu 吴方言
- ③ Xiang 湘方言
- 4 Min 闽方言
- ∮ Yue 粤方言
- **⑤** Gan 赣方言
- Hakka 客家方言 (distinct du Gan pour des raisons culturelles)
- ❷ Pinghua 平话
- Hui 徽州话 (de nature mélangée et transition nelle)

十余字 национален 文化

## Exemple 2

#### X. 李. Lǐ (2012:511):

Tonème du MC	A	В	C	D	
Mandarin	[p <sup>h</sup> ]			[p]	
Yue	[p <sup>h</sup> ]   [	[p <sup>h</sup> ]/[p	]	[p]	
Xiang	[b] (m	ajorité)	/[p]	[ph] (majorité)/[b]	
Wu	[b]				
Hui (Jixi)		[p <sup>h</sup>	] (majo	rité)/[p]	
Gan	[p <sup>h</sup> ]				
Hakka			[p <sup>1</sup>	1]	
Pinghua	[p]				
Min	[p] (majorité)/[pʰ]				

Les initiale obstruantes sonores du MC dans les langues sinitiques (prenons l'initiale

[\*b-] comme exemple) sans la stratification



## Exemple 2

Tonème du MC	A   B   C   D		
Mandarin	[p <sup>h</sup> ] [p]		
Taiyuan	[p]		
Rongcheng	[p]		
Yue	[p <sup>h</sup> ] [p]		
Xiang	[b]   [p <sup>h</sup> ]		
Wu	[b]		
Hui (Jixi)	[p <sup>h</sup> ]		
Gan	[p <sup>h</sup> ]		
Hakka	[p <sup>h</sup> ]		
Pinghua	[p]		
Min	[p]		

Les initiale obstruantes sonores du MC dans les langues sinitiques (prenons l'initiale

[\*b-] comme exemple) pour les couches héritées

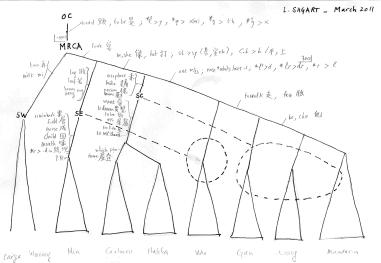


# La classification des langues sinitiques

Les critères pour la phylogénie linguistique, une classification tant synchronique que diachronique :

- Lexical
- Innovation commune
- Stratification

#### Sagart (2011):







## Critères phonétiques de la stratification

Sagart & Xu (2001), Chén (2003) et Chén (2005):

- Correpondances phonétiques systématiques entre les doublets/triplets/quartets... sur une syllabe dans différents mots ("wen-bai")
- Connaissances de l'évolution phonologique du chinois (OC, MC, etc.)
- Critère (étendu) de cohérence



## Exemple

```
"目" du Wu de Wenzhou:
```

- couche I: [ne], "三四 ~" (Yóu & Yáng 1998: 158), "trois ou quatre jours"
- couche II: [niai], "生~" (Yóu & Yáng 1998: 1471), "anniversaire" (couche héritée)
- couche III: [zai], "~ 记" (Pān 1998: 54), "journal intime"

## Situations exceptionnelles

- Cáo (2022):
   Les mots grammaticaux peuvent subir le changement phonétique exceptionnel ("fossilisation", "lénition", etc.).
   e.g. "去" ("aller", verbe directionnel), "渠" (prénom 3SG"), "家" (suffixe), etc.
- Lǐ (1985):
   Les toponymes peuvent garder les prononciations archaïques.
  - e.g. "富"[pʰu] de "大富车" de la province du Hebei (Wáng 2017: 36–37)

Une base de données déjà établie par Wu & List (2023) Traitement des données :

- Filtration des morphèmes non saillants, e.g. les suffixe "子", "头", etc.
- Annotation des couches : 1 v.s. 2/3 pour Min et 2 v.s. 1/3/4 pour non Min
- Correction de l'étymologie, e.g. "拨" > "畀" ("donner") du Wu de Suzhou
- Concepts à exclure (mal traduit, trop de variation, étymologie non déterminée, etc.)
- COGIDs partielles (des syllabes) > COGIDs pleines (des mots) مرفية على المحاصلة ا

# Méthodes phylogénétiques

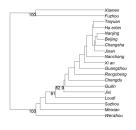
Hall (2018 : 74–75) et Pellard & Ryder & Jacques (2023 : 15–16) :

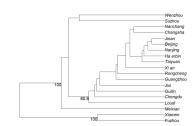
10 10).			
		Avantage	Inconvénient
Méthodes de diastances	UPGMA NJ NN	rapide	présupposition non pertinente
Méthodes de caractères	MP ML BI	"rule of thumb" populaire	"long branch attraction" chronophage

On applique les méthodes d'UPGMA, d'NJ et de MP aux données suivant le tutoriel de M. Pellard<sup>1</sup>.

<sup>1</sup> https://tpellard.github.io/phylolinguistique/

#### **UPGMA**



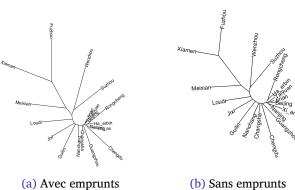


(a) Avec emprunts

(b) Sans emprunts

FIG. : Arbres d'UPGMA testés avec la technique de BS a [ C O

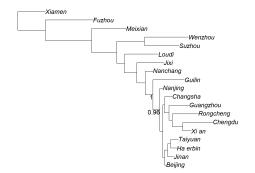
## NJ



its (b) Sans emprunt

FIG. : Arbres générés avec la méthode d'NJ | n a l c o

## MP avec Ratchet

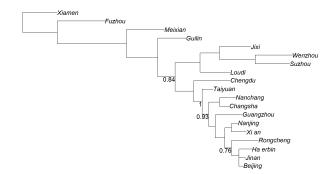




19/33

FIG. : Arbres MPR testés avec la technique de BS : avec emprunts

## MP avec Ratchet





20/33

FIG. : Arbres MPR testés avec la technique de BS : sans emprunts

### MP avec BaB

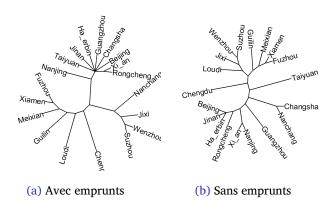


Fig. : Arbres consensus strict des arbres MP | n a l c c

#### Observation

- Les Min de Fuwhou et Xiamen constituent un clade.
- Les Wu de Suzhou et Wenzhou constituent un clade.
- Le Hakka se sépare après le Min et ne constituent jamais un clade avec le Gan de Nanchang.
- Les Xiang de Changsha et Loudi ne constituent jamais un clade.



#### Discussion

Les statuts anormaux du Yue de Guangzhou et du mandarin de Chengdu : Deux types de biais résultant de l'écriture dans les enquêtes sur le terrain

	Avec emprunts	Sans emprunts
Beijing	76%	76%
Jinan	74%	75%
Guangzhou	70%	69%
Rongcheng	69%	68%
Changsha	62%	59%
Nanchang	62%	59%
Jixi	61%	61%
Loudi	58%	58%
Chengdu	56%	55%
Suzhou	54%	53%
Meixian	52%	51%
Wenzhou	50%	51%
Xiamen	40%	36%
Fuzhou	22%	32%

e.g. "给" v.s. "畀" ("donner") du Yue de Guangzhou



#### Discussion

- La nativisation étymologique des emprunts (Aikio 2007)
   e.g. "太阳" v.s. "日头/热头" ("soleil") du Yue de Guangzhou (Zhān 2002 : 389) et du Wu de Suzhou (Yè 1988 : 346)
- Situation opposée: les emprunts détectables partagent la même forme que les mots remplacés
  e.g. "四" (\*sijH) [sq] ("quatre") v.s. "死" (\*sijX) [si]
  ("mourir") du Wu de Suzhou (Liú & Wáng & Băi 2007: 96)



### Conclusion

- L'exclusion des emprunts influencerait dans différentes mesures les résultats phylogénétiques.
- Les résultats phylogénétiques confirment dans une mesure ceux de Sagart (2011).
- Il faut bien veiller à la méthodologie des enquêtes sur le terrain pour obtenir les matériaux pertinents.
- Il faut chercher des solutions aux emprunts indétectables et faussement exclus.



## Bibliographie I

- Aikio, Ante. 2007. Etymological nativization of loanwords. A case study of Saami and Finnish. Teoksessa Ida Toivonen–Diane Nelson (toim.), Saami Linguistics. 17-52.
- Cáo, Zhìyún 曹志耘. 2022. Zhèjiāng jīnhuá tāng xī fāngyán de yǔ liú yīn biàn 浙江金华汤溪方言的语流音变[La variation contextualle dans le dialecte de Tangxi de la ville de Jinhua, province de Zhejiang]. Fāngyán 方言[Dialect] 44(2). 8.



# Bibliographie II

Chén, Zhōngmǐn 陈忠敏. 2003. Wúyǔ jí línjìn fāngyán yúyùn de dúyīn céngcì—jiān lùn "jīnlíng qièyùn" yúyùn de yīnzhí 吴语及邻近方言鱼韵的读音层次—兼论"金陵切韵"鱼韵的音值[Une discussion sur les différences de prononciation de la catégorie Yu (鱼) dans le wu et les dialectes voisins: avec quelques réflexions supplémentaires sur sa reconstruction au sein de Qieyun]. Yǔyánxué lùncóng 语言学论丛[Essays on Linguistics] 27. 11-55.



## Bibliographie III

- Chén, Zhōngmǐn 陈忠敏. 2005. Lún yǔyīn céngcì de shíjiān xiānhòu 论语音层次的时间先后[Quelques notes sur les étapes chronologiques des strates sonores]. Yǔyán yánjiū jíkān 语言研究集刊[Bulletin of Linguistic Studies] (1). 123-132.
- Hall, Barry G. 2018. *Phylogenetic trees made easy*. Fifth Edition. New York: Oxford University Press.
- Lǐ, Róng 李荣 (éd.). 1987. Zhōngguó yǔyán dìtú jí 中国语言 地图集[Atlas des langues de la Chine]. Hongkong: Longman.
- Lǐ, Rúlóng 李如龙. 1985. Dìmíng zhōng de gǔyīn 地名中的 古音[Les prononciations archaïques dans les toponymes]. Yǔwén yánjiū 语文研究[Linguistic Research] (1). 30-30. a

## Bibliographie IV

- Lǐ, Xiǎofán 李小凡. 2012. Pínghuà de guīshǔ hé hànyǔ fāngván fēnlèi 平话的归属和汉语方言分类[L'affiliation de Pinghua et la classification des dialectes chinois]. Yŭyán kēxué 语言科学[Linguistic Sciences] 11(5). 509-517.
- 🖥 Liú, Lìlǐ 刘俐李 & Hóngzhōng 王洪钟 Wáng & Yíng 柏 莹 Bǎi. 2007. Xiàndài hànyǔ fāngyán héxīn cí tèzhēng cí jí 现 代汉语方言核心词·特征词集「Collection de mots de base et de mots caractéristiques dans les dialectes chinois modernes].

Nanjing: Phoenix Publishing & Media, Inc.



# Bibliographie V

- Pān, Wùyún 潘悟云. 1998. Wēnzhōu huà yīndàng 温州话音档[Enregistrement du dialecte de Wenzhou]. In Hóu, Jīngyī侯精一 (éd.), Xiàndài hànyǔ fāngyán yīnkù 现代汉语方言音库[Phonobase des dialectes chinois modernes. Shanghai: Shanghai Educational Publishing House.
- Pellard, Thomas & Robin J. Ryder & Guillaume Jacques. 2023. The Family Tree model. In Ledgeway, Adam et al. (éds.), *The Wiley Blackwell companion to diachronic linguistics*. preprint 2023-05-11. https:

```
//www.academia.edu/101656989/The_Family_Tree_model.
```

## Bibliographie VI

Sagart, Laurent. 2011. Classifying Chinese dialects/Sinitic languages on shared innovations.

PaperpresentedattheSéminaireSino-TibétainduCRLAO (2011-03-28). Downloadable at https://www.academia.edu/19534510 (accessed February 4, 2018).

- Sagart, Laurent & Shixuan Xu. 2001. History through loanwords: the loan correspondences between Hani and Chinese. *Cahiers de linguistique Asie orientale* 30(1), 3-54.
- Shěn, Míng 沈明. 1994. Tàiyuán fāngyán cídiǎn 太原方言詞典 [Dictionnaire du dialecte de Taiyuan]. Nanjing: Jiangsu Education Publishing House.

## Bibliographie VII

- Wáng, Jiājiā 王佳佳. 2017. Héběi yì xiàn dìmíng de yǔyán xué kǎochá 河北易县地名的语言学考察[Une étude linguistique des toponymes dans le district de Yi, province du Hebei]. Bohai University. ([mém. de mast.]).
- Wu, Mei-Shin & Johann-Mattis List. 2023. Annotating cognates in phylogenetic studies of Southeast Asian languages. *Language Dynamics and Change* 1(aop). 1-37.

# Bibliographie VIII

- I Yè, Xiánglíng 葉祥苓. 1988. Sūzhōu fāngyán zhì 苏州方言志[Chorographie dialectal de Suzhou]. Nanjing: Jiangsu Education Publishing House.
- You, Rǔjié 游汝杰 & Qiánmíng 楊乾明 Yáng. 1998. Wēnzhōu fāngyán cídiǎn 溫州方言詞典 [Dictionnaire du dialecte de Wenzhou]. Nanjing: Jiangsu Education Publishing House.
- In Italian, Bóhuì 詹伯慧. 2002. Guǎngdōng yuè fāngyán gàiyào 广东粤方言概要[Un aperçu des dialectes Yue dans le Guangdong]. Guangzhou: Jinan University Press.

