

Can Oral Hokkien Language be Written? How Can Wikidata help?

slide: <https://hackmd.io/@wikidata-tw/Syy3-LJ5>
Wikimedia Taiwan Tân Sĩ-lĩm

Who am I?

- Supaplex
- OpenStreetMap  Wikidata
- Wikimedia Taiwan 

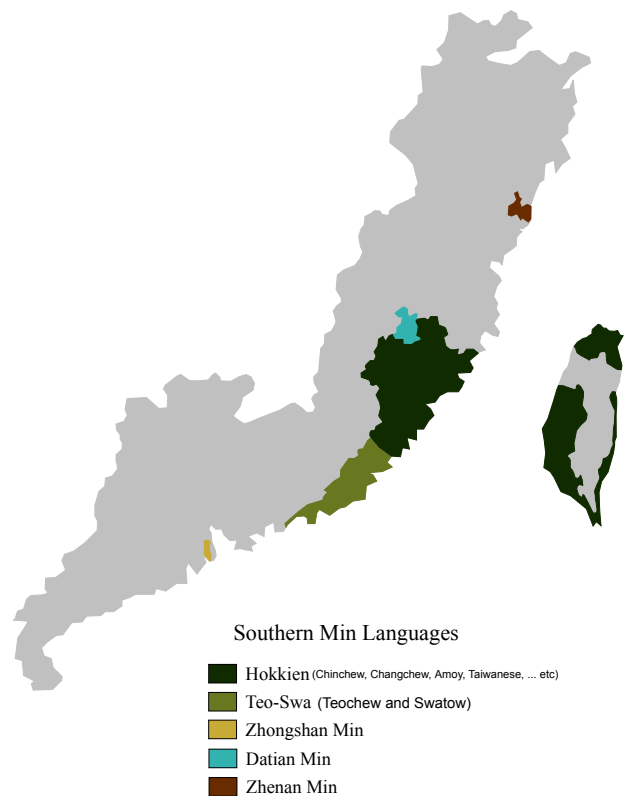
Highlight of Previous Sp

- The Numbers of Written Form Article
Taigi is quite weak, but it is improving
years
- Emphasis on you can write in Hokk
Wikipedia too
- Main namespace on nan Wikipedia
POJ(Latinization form), also Possib
Hanji Characters
- **Manpower needed**

Wikimania 2023

- Talk about Taigi on Wikimania 2023 ([record](#))
- The talk is on EASAP track, not many participants, and not many locals Singapore
- Some Chinese people attended, not sure about their attitudes

Introduction of Min Nan lâm/Hokkien Language





Introduction of Taigi/M Bân-lâm/Hokkien Lan

- Total Speaking Population: 48 Million
- One of the Main Languages: Taiwan
Fujian(Hokkien)
- Minority Languages: [The Phillipines](#),
[Singapore](#), [Malaysia](#), [Indonesia](#),
Thailand(Teochew)

The Past of Taiwanese

- One Language Policy
- Mandarin in Education System
- Entertainment: Taigi Songs,
- Market language

The Recent Development of Taiwanese Taigi


- End of the One Language Policy in 1993 which Limited the Broadcasting time of Taigi or Hakka
- Taigi Language Examination: Ministry of Education, National Chengchong
- Became National Language in 2017
- Online Dictionaries
 - Taigi Dictionary
 - Chinese Dictionary
 - Taigi Dictionary

PTS Taigi Station

台語新詞辭庫

台語台收集生活中的新興詞彙
定期邀請專家學者開會討論
建議適當的台語講法
建立台語新詞的詞庫
開放予民眾查詢使用

窯烘麩

音讀： iô-hang-pháng 

對應華語： 窯烤麵包

生炭地

音讀： senn-thuà

sinn-thuà

The Possible Split of

- Appeal for a (<https://iso639-3.sil.org/lists/2021-045>) ISO 639-3 Code by ITH (意傳科技) and other NGOs
- Taiwan Taigi: oan
- nan split into Eleven Codes: Luezhon, Hainanese, Haklau, Datian Min, Zha, Quanzhou, Zhenan Min, Teo-Swa, Min etc.

The Different Names and of Hokkien

- Taiwan: Tâi-uân Bân-lâm-gú/Tâi-uân-gí/Tâi-uân Bân-lâm-gír
- Historic Name: Tâi-oân-ōe, Tâi-gí
- New Official Name: Taiwanese Taigi
- South East Asia: Hokkian/Hokkien/
- The Phillipine: Lan-lang-oe/咱人話
- Hakka's called Hoklo(鶴老話)

Taiwanese Taigi on Wik

- [Min Nan Wikipedia](#)
- [Wikidata](#) nan/nan-hani labels, o
- Wikidata [Lexeme](#) (Wikidata Lex
- [Wiktionary](#)

nan Community

- the original Nan Wikipedians
- Hó-pêng-iú Wikimedians User Group
 - Tâi-uân ê kong-biō-bûn-huà v
- Wikidata Taiwan community

Bân-lâm-gú(nan) Wik

Tâi-tang-koān (Taitong County) J

W D C
WIKI DATA CON



WIKIPEDIA
Chū-iū ê Pek-kho-choân-su

[Main Page](#)

[Bûn-chiu" bók-chhù](#)

[Random page](#)

[Current events](#)

[Donate](#)

[Pian-chip](#)

[Help](#)

[Community portal](#)

[Thó-lūn](#)

[Recent changes](#)

[Tools](#)

[What links here](#)

[Related changes](#)

[Upload file](#)

[Special pages](#)

[Permanent link](#)

[Page information](#)

[Cite this page](#)

[Wikidata item](#)

[Print/export](#)

[Download as PDF](#)

[Printable version](#)

[In other projects](#)

Page [Discussion](#)

[Read](#)

[Edit](#)

[Edit source](#)

[View history](#)



[More](#)

Tâi-tang-koān

Tâi-tang-koān (**Hàn-jī**: 臺東縣) sī **Tâi-oân tang-pō'ê 1-ê koān**.

Contents [hide]

- [Hêng-chèng tan-ūi](#)
 - [Chhī](#)
 - [Tìn](#)
 - [Hiong](#)

Hêng-chèng tan-ūi [edit | edit source]

Chhī [edit | edit source]

- [Tâi-tang-chhī](#)

Tìn [edit | edit source]

- [Koan-san-tìn](#)
- [Sêng-kong-tìn](#)

Hiong [edit | edit source]

- [Hái-toa"-hiong](#)
- [lân-pêng-hiong](#)
- [Kim-hong-hiong](#)
- [Lân-sū-hiong](#)
- [Lèk-tó-hiong](#)
- [Lòk-ia̍-hiong](#)

The Problem of nan Wi

- More Favor of POJ, lack of Hanji Content
- Too much Administrative Articles, not so many Scinece Articles
- Lack o Comm
- Many A linking

Problem of nan Wikipedia Spam

垃圾ê頁面4 [kái goân-sí-bé]

[Fodboldtrøjer Udsalg 90677](#)、[Kajian Platform Dagangan Pocket Option Untuk Malaysia: Panc](#)
[Remarkable Web Site - Binary Options Will Help You Get There](#)，其中Fodboldtrøjer Udsalg 9
單純拍廣告做SEO--ê。標示管理員來處理 @A-lú-mih、@Kaihsu、@Taigiholic、@Yoxem，希
章的頁面，勞力！2023-nî 7-goèh 8-jit (Pài 6) 15:56 (UTC)~ [Supaplex \(thó-lūn\)](#) 2023-nî 7-goèh
(UTC) [[回覆](#)]

Problem Case 1: Dupli Poland Place

Legarda (Pho-lân)

Redirect page

↳ [Legarda](#)

Problem Case 2: Thousands of Non Wikidata-linked Articles



1,328	SOHO-chòk	che
1,329	Lùi-piát talk:Bàng-gà	che
1,330	Tiuⁿ Oán-chheng	che
1,331	Tâi-tòk tóng-kong	che
1,332	XMM-Newton	che
1,333	Tiong-kok Jîn-bîn Chèng-tī Hiáp-siong Hōe-gī	che
1,334	Sió-hô-khu	che
1,335	Âng-mńg	che
1,336	Lú-sèng-khe	che
1,337	Tom Arnold	che
1,338	Hûi-luh-mu	che
1,339	Sī-chhiòng	che
1,340	How To Advance Psychic Ability	che

Wikidata Labels & Descriptions

nan, nan-hani

諾魯文	Republik Tsiene	尚未定義描述
Nāhuatl	Taihuan	尚未定義描述
Min Nan (Hanji)	中華民國	東亞的國家
巴克摩挪威文	Taiwan	nasjon i Asia
低地德文	Republiek China	尚未定義描述

POJ and Hanji: Permanent Duplicated Items

- Hakka, Bang-ue both have Hanji and Romanization articles
- Linkage of Same language but different system by Wikidata

Permanent Duplicate Item(P2959)

*this item duplicates another item in Wikidata
and the two can't be merged. This property is used when
one Wikimedia project includes content from another
project, e.g. in different script or different
languages (only applies to some Wikimedia
wikis, e.g.: cdowniki, gomwikin, guoyanwikin,
nanwiki). Use P31=Q1736292 for a specific
other wikis.*

Example of Permanent Duplicated Item

Taiwan Permanent Duplicated



陳述

性質



維基媒體永久重複頁面



0 條

讀音 (音訊檔案)



1 條

[LL-Q34239 \(kok\)-Fredericknoronha-Taiwan.wav](#)

1.2秒；115 KB

作品或名稱語言

孔卡尼語 繁體中文 (已轉換拼寫)

永久重複項



中華民國



0 條

[显示派生陈述](#)

維基百科 (4 條目)



[\[移動\]](#)

[cdo](#) [中華民國](#)

[gom](#) [Taiwan](#)

[hak](#) [中華民國](#)

[zh_min_nan](#) [Pang-chân:中華民國](#)

維基教科書 (0 條目)



維基新聞 (0 條目)



維基語錄 (0 條目)



Relation between Perm Duplicated Item



	Main Item	Permanent Duplica
Wikidata	Q865	Q22337
zh-min-nan	Tiong-hôa-bîn-kok	中華民國
hak	Chûng-fà Mìn-koet	中華民國
cdo	Dǔng-huà Mìng-guók	中華民國

Merits of Writing Hokk Wikidata

- Using few Words of short Sentence Description
- An already Existing Structure
 - ** Use English as Reference to Write
 - ** No need to rely on Mandarin
- Writing form: POJ, nan-hani, or Mix
- Linked Data

Other Sinitic Languages

Written System

- Chinese characters:
Cantonese(yue),
Wu(wuu)
- Romanization:
Hakka(hak), **Bàng-**
uâ(cdo),
Hokkien(nan)

Merits

- Big City
Wu(Shi)
Nan(Tai)
- Music In
Yue(Ho)
Nan(Tai)
- Educati
Yue(Sc)
Nan(Sc)

Fun thing on Wikid

Taigi Hanji

Peh-ōe-jī

ata.org/wiki/Q113584516?uselang=nan-hani

has part(s) <i>English</i>	<ul style="list-style-type: none">對台南永康到台北永康，破碎的台北夢 ... series ordinal <i>English</i> 1 ...新北市的新貧階級，予人看見幸福的盡頭 ... series ordinal <i>English</i> 2 ...討債戀愛過程告別青春，佇西門町享受速食愛情 ... series ordinal <i>English</i> 3 ...權力關係的砌砗枋，用自信妝扮東區的查某人 ... series ordinal <i>English</i> 4 ...仁愛路圓環的修羅場，第三者的遊戲規則 ... series ordinal <i>English</i> 5 ...
----------------------------	--

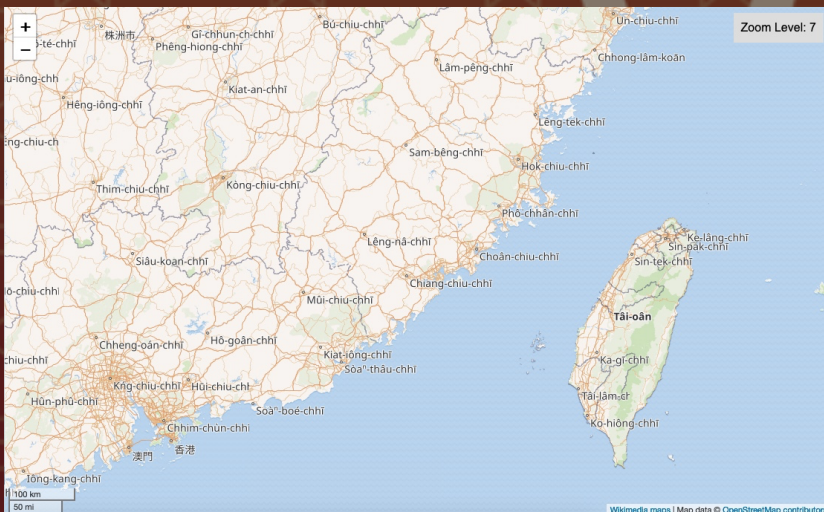
ata.org/wiki/Q113584516?uselang=nan

網路社群追蹤者 繁體中文 使用語言

可分為 繁體中文

Tù Tâi-lâm Tâi-pak-bân	系列序號
Sin-pak-chh chīn-thâu	系列序號
Thó-chè loá híong-siú s	系列序號
Khoán-lék k Tang-khu	

Wikimedia Map by Wikimedia Foundation



Wikimedia Map la
OpenStreetMap d
Taibun hanji, Tâi-l
maps.wikimedia.o
TL#16/24.9725/1

Q258647

郵差 (Q258647) ...

1994年邁克爾·拉德福導演的義大利電影
郵差 (1994年電影)

▼ 更多其他語言

語言	標籤	描述	類
中文（臺灣）	郵差	1994年邁克爾·拉德福導演的義大利電影	型
德文	Der Postmann	Film von Michael Radford (1994)	
英文	Il Postino: The Postman	1994 film by Michael Radford	F I T
客家話	尚未定義標籤	尚未定義描述	

Plan for nan Wikipe

- Let's Start Contribution: POJ or Hanji
- Taiwan MOE [Recommend Hanji](#) or POJ Characters
- Hanji-POJ linked by Permanent Duplicate Items on Wikidata

Plan for nan Wikipedia

- POJ (Wikidata Main item) - Hanji A
(Wikidata Permanent Duplicated item)
- Administrative units
- Science

Please Join the Writing Campaign



- ~~Stop the agreement~~
- Ignored those claimed that violate the study
- Need Code of Code to Stop the Ext Unproduce Speach
- The Possible split of nan ISO-639- might imclubated new Wikipedia S
- Edit nan/nan-hani Labels or Descri Wikidata
- Wiktonary
- Wikidata Lexeme

OSM x Wikidata Monthly Meetup

- 06/03 19:30 @MozSpace
- [Telegram Wikimedia nan](#)
- [Facebook Wikidata Taiwan](#)

To-siā! sùn-mùng-ń!
 you!



Q&A:

Communication Channels

- [GitHub](#) 、 [Wikidata](#) 、 [OpenStreetMap](#)
- [email](#)
- Facebook Group: [Wikidata Taiwan](#)
[OpenStreetMap Taiwan](#)