

凱達格蘭 · 巴賽語

Ketagalan · Basay — 從記憶到再生

[首頁](#) [文法](#) [教育推進](#) [研究成果](#) [研究筆記](#) [語音合成](#) [辭典](#)[Home](#) / [Research](#) / [This article](#)[中文](#) [日本語](#) [English](#)

The Syllable Inventory of the Basay Yilan Dialect (source=T+M)

*Evidence of Phonological Contact with Kavalan***Author:** Tsai, Yung-kuei (蔡永桂)**Date:** June 23, 2026**Type:** Original research (language contact / quantitative phonology)**License:** CC BY 4.0 **Citation ID:** basay.tw/research/2026-06-basay-syllable-TM/

Abstract

This paper analyzes the syllable inventory of the Yilan dialect data (source=T: Trobiawan contextual collection; source=M: Trobiawan vocabulary-only collection; 1,129 entries combined) of the Basay lexical database and compares it with the native vocabulary (source=B). The T+M data yield 315 syllable types across 38 onset categories — substantially larger than source=B (266 types, 22 onsets). Six onset categories exclusive to T+M — /q/, /z/, /ʒ/ (z'), /l/ (l'), /m/ (ml'), and /v/ (vl') — correspond directly to phonemes of Kavalan (噶瑪蘭語), a Formosan language spoken in the Yilan Plain adjacent to the Basay Yilan dialect area. The marked increase in CVVC (26 vs. 7 types) and onset-

cluster structures (22 vs. 6 types) in T+M further parallels Kavalan phonotactics. These findings support a hypothesis of phonological borrowing from Kavalan into the Basay Yilan dialect through prolonged contact.

Keywords: Basay, Trobiawan (T+M), Yilan dialect, Kavalan, language contact, phonological borrowing, syllable inventory

Cite this article

APA:

```
Tsai, Y.-k. (2026). The syllable inventory of the Basay Yilan dialect (source=T+M): Evidence of phonological contact with Kavalan. basay.tw. https://basay.tw/research/2026-06-basay-syllable-TM/en/
```

BibTeX:

```
@misc{tsai2026syllableTM_en,  
  author = {Tsai, Yung-kuei},  
  title = {The Syllable Inventory of the {Basay} {Yilan} Dialect (source=T+M)},  
  year   = {2026},  
  month  = {6},  
  url    = {https://basay.tw/research/2026-06-basay-syllable-TM/en/}  
}
```

1. Introduction

1.1 The T and M Source Types

Both source=T and source=M in the Basay lexical database represent Trobiawan vocabulary data. Source=T contains entries collected from connected discourse (contextual collection), while source=M consists of isolated vocabulary items

without sentence context. T and M share closely similar phonological profiles; they are treated as a single Trobiawan dataset (T+M) in this paper. Source S (113 entries), which shows extensive admixture of Kavalan vocabulary, is excluded.

1.2 Kavalan as a Contact Language

The Kavalan people (噶瑪蘭族) were the principal indigenous inhabitants of the Yilan Plain. Kavalan (噶瑪蘭語) is a Formosan Austronesian language documented to have uvular /q/, voiced fricative /z/, retroflex lateral /ɭ/, and voiced lateral fricative /ɮ/ (Li 2000) — precisely the phonemes that appear in T+M but are absent from source=B.

2. Method

Analysis followed the same procedure as the source=B paper. Source=S, =V, and PAN reconstructions were excluded.

Table 1. Orthography–IPA correspondence for source=T+M

Orthography	IPA	Description
n'	ŋ	Velar nasal
ɭ'	ɭ	Retroflex lateral
z'	ɮ	Voiced alveolar lateral fricative
o'	ə	Mid central vowel (schwa)
' (coda)	ʔ	Glottal stop (syllable-final coda)
q	q	Uvular/pharyngeal stop

Orthography	IPA	Description
ts	ts	Alveolar affricate
v	v	Voiced labiodental fricative
z	z	Voiced alveolar fricative
j	j ~ dʒ	Approximant or affricate

Note: source=T+M contains no occurrences of the onset h, /ʃ/ (s'), or /tʃ/ (ts') — all of which are present in source=B.

3. Results

3.1 Comparison with Source=B

Parameter	T+M	B	Difference
Entries	1,129	1,117	—
Syllable types (freq. ≥ 2)	315	266	+49
Onset categories	38	22	+16
Shared syllables	128	128	—
Exclusive syllables	187	138	—

3.2 Syllable Structure Comparison

Structure	T+M	B	Δ	Note
V	4	4	0	Equal
VC	2	1	+1	Minor
VV	2	2	0	Equal
VVC	0	1	-1	Minor
CV	73	66	-3	Near equal
CVC	159	134	+25	T+M higher
CVV	27	36	-9	B higher
CVVC	26	7	+19	Markedly T+M higher
other (clusters)	22	6	+16	Markedly T+M higher

3.3 T+M-Exclusive Onsets

Onset	IPA	Types	Tokens	Kavalan correspondence
q	q	18	128	Kavalan has uvular /q/ ✓
z	z	21	98	Kavalan has voiced /z/ ✓
ɭ'	ɭ	5	23	Kavalan has retroflex /ɭ/ ✓
z'	ʒ	5	16	Kavalan has lateral fricative /ʒ/ ✓
mɭ'	mɭ	3	12	Kavalan-type cluster ✓
vɭ'	vɭ	2	4	Kavalan-type cluster ✓

Onset	IPA	Types	Tokens	Kavalan correspondence
y	j	2	19	—
Other clusters	—	16	~41	—

4. Discussion

4.1 The Contact Hypothesis

The six Kavalan-corresponding onset types in T+M (q, z, z', l', ml', vl') collectively account for approximately 176 occurrences across 34 syllable types. Their complete absence from source=B is the critical distributional fact. Under Thomason & Kaufman's (1988) framework, the high frequency and systemic integration of these phonemes in T+M suggests phonological borrowing at the system level, not mere lexical diffusion.

4.2 CVVC and Long-Vowel Structures

The 26 CVVC types in T+M (vs. 7 in B) include forms such as maan, laan, zaay, zian, and z'ian. Kavalan distinguishes phonological vowel length and allows CVVC structures (Li 2000). The parallel increase in CVVC types in T+M — particularly involving the z and z' onsets which are themselves Kavalan-derived — supports the contact hypothesis at the level of syllable structure.

4.3 Source=B Onsets Absent from T+M

The absence of h, /ʃ/ (s'), and /tʃ/ (ts') from T+M deserves attention. Kavalan lacks an onset /h/ phoneme; the loss or non-transfer of /h/ in T+M may reflect contact-

induced attrition. The absence of the palatal series /j/ and /tʃ/ may similarly reflect convergence toward the Kavalan phonological type, which lacks these segments.

5. Conclusion

The T+M Yilan dialect data yield 315 syllable types across 38 onsets. Six onset categories exclusive to T+M correspond to Kavalan phonemes, constituting evidence of phonological contact with Kavalan in the Yilan Plain. The CVVC and cluster expansions reinforce this conclusion at the level of syllable structure. Conversely, the native-vocabulary phonemes /h/, /ʃ/, and /tʃ/ are absent from T+M, suggesting contact-induced attrition. These findings reframe T+M not as a straightforward dialect of Basay but as a contact variety in which Kavalan phonological influence is substantial.

References

- Blust, R. (1999). Subgrouping, circularity and extinction. In E. Zeitoun & P. J.-K. Li (Eds.), *Selected papers from the Eighth International Conference on Austronesian Linguistics* (pp. 31–94). Academia Sinica.
- Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Blackwell.
- Li, Paul Jen-kuei. (1996). *The Formosan Tribes and Languages in I-Lan*. Yilan: Yilan County Government.
- Li, Paul Jen-kuei. (2000). *The Phonological System of Taiwan Austronesian Languages*. Taipei: Crane Publishing.
- Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
- Institute of Linguistics, Academia Sinica (Ed.). *Basay Lexical Database* (`basay_dict.jsonl`). Taipei: Academia Sinica.

[← Back to Research](#)

"Contact leaves its mark on every syllable."

© 2026 basay.tw · CC BY 4.0 | [Return to Home](#)