

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

[ホーム](#) / [研究成果](#) / 本稿

[中文](#) [日本語](#) [English](#)

巴賽語音節目録の再分析

——ソース別分析による混合分析の訂正と言語接触の証拠——

著者：蔡永桂 (Yung-kuei Tsai)

日付：2026年6月23日

種別：方法論的訂正論文 (音韻分析 / 言語接触)

ライセンス：CC BY 4.0 引用識別：basay.tw/research/2026-06-basay-syllable-revised/

要旨 Abstract

本稿は、巴賽語 (Basay) の音節目録に関する先行の混合分析 (B・T・M を一括処理した486種の目録) を訂正し、ソース別分析の結果を比較考察する。固有語彙 (source=B、1,117件) の単独分析では266種・22 onset が確認され、CVC型 (54%) が最多構造型であった。宜蘭方言 (source=T+M、1,129件) の単独分析では315種・38 onset が確認され、B に存在しない **q**・**z**・**z'** (**ɟ**)・**l'** (**l**) などの音素が固有語彙の2倍近い規模で出現する。これらの「余剰」音素は宜蘭平野において隣接するカバラン語 (噶瑪蘭語) の音韻体系と一致し、言語接触による音韻借用の証拠と解釈される。混合分析でこれらが「巴賽語の特徴」として記述されたことは方法論的誤謬であり、本稿はその訂正を行うとともに、ソース別分析が消滅言語の音韻記述において不可欠であることを論じる。

キーワード：巴賽語・音節目録・ソース別分析・言語接触・カバラン語・方法論的訂正・台湾原住民語

引用 / Cite this article

APA :

蔡永桂 (2026). 巴賽語音節目録の再分析——ソース別分析による混合分析の訂正と言語接触の証拠. basay.tw. <https://basay.tw/research/2026-06-basay-syllable-revised/ja/>

BibTeX :

```
@misc{tsai2026syllableRevised_ja,
  author = {蔡永桂 and Tsai, Yung-kuei},
  title = {巴賽語音節目録の再分析},
  year = {2026},
  month = {6},
  url = {https://basay.tw/research/2026-06-basay-syllable-revised/ja/}
}
```

1. はじめに：混合分析の問題

1.1 先行分析の経緯

先行稿（混合分析版）では、巴賽語辞書データベース（`basay_dict.jsonl`）のうち PAN 再建形を除く2,364エントリを一括処理し、頻度2以上の音節486種を抽出して音節目録を提示した。そこでは `q`・`z`・`z'` (`ʃ`)・`ɹ'` (`l`)・複雑クラスター onsetなどを巴賽語の音韻的特徴として記述した。

しかし、この分析には重大な方法論的問題がある。辞書データには複数のソースコードが付与されており（表1）、それぞれ異なる語彙層を代表している。固有語彙（B）と宜蘭方

言 (T・M)、さらにカバラン語混入が疑われる資料 (S) を一括処理することは、異なる音韻体系を同一の「巴賽語」として分析することになる。

ソース	件数	内容
B	1,117	巴賽語固有語彙
T	588	Trobiawan (文脈付き採集)
M	541	Trobiawan (語彙のみ採集)
S	113	カバラン語等混入が疑われる資料 (本稿除外)
V	5	不明 (除外)
PAN	960	祖先語再建形 (除外)

本稿では S・V・PAN を除外した上で、B と T+M を分離して分析し、結果を比較することで混合分析の問題を修正する。

2. 方法

音節抽出手順は前稿と同一とし、以下を追加修正した：' 起始音節を前音節韻尾の転写アーティファクトとして除外；音節構造分類において onset の有無を反映（前稿では VC・VV が誤って CVC・CVV に分類されていた点を修正）；S ソース（113件）はカバラン語語彙の大量混入が確認されたため除外。

3. 結果：B と T+M の比較

3.1 規模の比較

項目	B (固有語彙)	T+M (宜蘭方言)
エントリ数	1,117	1,129
音節種数 (頻度2以上)	266	315
onset 種数	22	38
共通音節	128種 (B の48%)	128種 (T+M の41%)
固有音節	138種	187種

3.2 音節構造の比較

構造	B	T+M	増減	解釈
V	4	4	0	同等
VC	1	2	+1	微差
VV	2	2	0	同等
VVC	1	0	-1	微差
CV	75	73	-2	ほぼ同等
CVC	134	159	+25	T+M に多い
CVV	36	27	-9	B に多い
CVVC	7	26	+19	T+M に顕著に多い
other (クラスター)	6	22	+16	T+M に顕著に多い

3.3 onset の比較

区分	onset
Bのみ	h, s' (ʃ) , ts' (tʃ) , sj, sw, n'ts
T+Mのみ	q, z, z' (ʒ) , l' (l̥) , ml' (ml̥) , vl' (vl̥) , y, km, kn, kt, ml, mn, mr, ms, mt, pr, pts, sm, tl, tm, tn, tr, zm
共通	∅, b, j, k, l, m, n, n' (ŋ) , p, r, s, t, ts, v, w

共通 onset は15種のみであり、B と T+M の音韻体系は onset レベルで大きく異なる。

4. 考察

4.1 混合分析の誤謬と訂正

混合分析での記述	実態
q (口蓋垂音) が存在する	T+M のみ。B には出現しない
z (有声摩擦音) が存在する	T+M のみ。B には出現しない
z' (ʒ) が存在する	T+M のみ。B には出現しない
l' (l̥) が存在する	T+M のみ。B には出現しない
CV 型が優勢	混合の錯覚。B 単独では CVC 型が54%で最多
音節種数486種	B : 266種・T+M : 315種 (重複128種を除く)

4.2 T+M 余剰音素のカバラン語起源仮説

T+M に固有の6 onset (q, z, z', l', ml', vl') はすべてカバラン語の基本音素と一致する。宜蘭方言 (T・M) 話者とカバラン族は宜蘭平野において地理的に隣接しており、長期的な言語接触が存在したと推定される。Thomason & Kaufman (1988) の言語接触理論によれば、T+M における **q** (128回・18種) ・ **z** (98回・21種) の高頻度出現は、音素としての体系的統合を示唆する。

4.3 B 固有音素の再評価

B に固有で T+M に不在の **h** ・ **s'** (ʃ) ・ **ts'** (tʃ) ・ **sj** は、巴賽語の音韻的個性を担う音素として再評価される必要がある。特に **h** onset (20種・138回) は B 固有語彙の基本音節 (ha, hi, he, ho, hu 等) を形成しており、巴賽語音韻体系の核心的要素である。

4.4 方法論的教訓

本稿の事例は以下の教訓を示す：

1. ソース混合は音節目録を過大評価させる (266種・315種 → 混合486種)
2. ソース混合は音節構造の実態を歪める (CVC 優勢 → CV 優勢に見える)
3. ソース混合は接触起源の音素を固有音素として誤分類させる
4. S ソースのような信頼性の低いデータは早期に除外すべきである

5. おわりに

本稿は先行の混合分析を訂正し、固有語彙 (B) 266種・22 onset と宜蘭方言 (T+M) 315種・38 onset という分離された音節目録を提示した。主要な知見は以下のとおりである：

1. 巴賽語固有語彙はCVC型優勢 (54%) であり、混合分析で見えたCV型優勢は誤りであった
2. T+M の余剰音素 (q, z, z', l' 等) はカバラン語接触に起源する可能性が高い
3. B 固有の h ・ s' ・ ts' こそが巴賽語音韻体系の個性を担う

4. ソース別分析は消滅言語の音韻記述において不可欠の手続きである

参考文献

- 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.
- Fenk-Oczlon, G., & Fenk, A. (2021). Linguistic complexity. *Frontiers in Communication*, 6, 626032.
- 李壬癸 (1996). 《宜蘭縣南島民族與語言》. 宜蘭縣：宜蘭縣政府.
- 李壬癸 (2000). 《台灣南島語言的語音符號系統》. 台北：文鶴出版.
- Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
- 中央研究院語言學研究所（編）. 《巴賽語辞書データベース》（`basay_dict.jsonl`）. 台北：中央研究院.

[← 研究成果一覧に戻る](#)

「訂正こそが、科学誠実性の証明。」

© 2026 basay.tw · CC BY 4.0 | [ホームへ戻る](#)