巴賽語固有語彙（source=B）の音節目録

——ソース別分析による音韻体系の再記述——

著者：蔡永桂（Yung-kuei Tsai）
日付：2026年6月23日
種別：原著論文（言語類型論 / 音韻計量分析）
ライセンス：CC BY 4.0 **引用識別**：basay.tw/research/2026-06-basay-syllable-B/

要旨 Abstract

本稿は、巴賽語（Basay）辞書データベースのうち固有語彙（source=B、1,117エントリ）のみを対象として音節目録を計量的に抽出し、その音韻体系を記述する。従来の混合分析（B・T・M・S・V を一括処理）を修正し、ソース別分析を適用した結果、固有語彙からは頻度2以上の音節266種・onset 22種が確認された。音節構造はCVC型（134種・54%）が最多であり、CV型（75種・25%）がこれに続く。onset として `h`、`s'`（ʃ）、`ts'`（tʃ）、`sj` が固有語彙に特徴的であり、他方言（T・M）に多く見られる `q`・`z`・`z'`（ʒ）・`l'`（ɮ）は固有語彙に出現しない。この結果は、従来の混合分析で「巴賽語の特徴」と記述されていた一部の音素が、実際には宜蘭方言（T・M）固有またはカバラン語接触に起因する可能性を示唆する。

キーワード：巴賽語・固有語彙・音節目録・ソース別分析・音韻体系・台湾原住民語

引用 / Cite this article

APA：

```
蔡永桂（2026）. 巴賽語固有語彙（source=B）の音節目録——ソース別分析による音韻体系の再記述. basay.tw. https://basay.tw/research/2026-06-basay-syllable-B/ja/
```

BibTeX：

```
@misc{tsai2026syllableB_ja,
  author = {蔡永桂 and Tsai, Yung-kuei},
  title  = {巴賽語固有語彙（source=B）の音節目録——ソース別分析による音韻体系の再記述},
  year   = {2026},
  month  = {6},
  url    = {https://basay.tw/research/2026-06-basay-syllable-B/ja/}
}
```

---

1. はじめに

巴賽語（Basay）は台湾北部の平埔族・巴賽族が使用した消滅南島語系言語である。17世紀のオランダ統治期文書に語彙が記録されており、中央研究院語言學研究所を中心に記録保存・語言復振の取り組みが行われている（李壬癸 1996；2000）。

先行研究（前稿・混合分析版）では辞書データの全エントリ（PAN再建形を除く2,364件）を一括して分析し、486種の音節目録を提示した。しかし、辞書データには複数のソースコードが付与されており、それぞれ異なる語彙層を代表している。これらを一括処理することは、異なる方言・言語接触の痕跡を混同することになり、音韻体系の実態を曇らせる危険がある。

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	カバラン語等の混入が疑われる資料
V	5	不明
PAN	960	祖先語再建形

本稿では source=B の1,117エントリのみを分析対象とし、他のソースはすべて除外した。

2. データと方法

音節抽出は以下の手順で行った：注記・括弧・代替形を除去し第一形式のみ使用；(C*)V(V?)(C?) テンプレートに基づく音節切分アルゴリズムを適用；`'`起始音節を除外；頻度1の音節を除外；音節構造をV・VC・VV・VVC・CV・CVC・CVV・CVVC・otherの9類型に分類。

表1　source=B 正書法・IPA 対照表

正書法	IPA	説明
`n'`	ŋ	軟口蓋鼻音
`s'`	ʃ	硬口蓋歯茎摩擦音
`o'`	ə	中舌中段母音（シュワー）

正書法	IPA	説明
`'`（コーダ）	ʔ	声門閉鎖音（前音節韻尾）
`ts`	ts	歯茎破擦音
`ts'`	tʃ	硬口蓋歯茎破擦音
`sj`	sj	硬口蓋歯茎摩擦音変種
`j`	j〜dʒ	接近音または破擦音

重要：source=B には `l'`（ɮ）・`z'`（ɮʒ）・`q`・`z` が出現しない。

3. 結果

3.1 全体統計

項目	値
分析エントリ数	1,117件
音節種数（頻度2以上）	266種
onset 種数	22種
最高頻度音節	la（148回）
高頻度（≥50回）	12種
中頻度（10〜49回）	53種

項目	値
低頻度（2〜9回）	201種

3.2 音節構造別分布

注：`o'` は単一母音 /ə/ のため `no'`→CV、`ko'`→CV 等コーダなしとして分類。

構造	種数	割合	説明
V	4	1.5%	母音のみ（a, i, o, u）
VC	1	0.4%	母音＋コーダ（at）
VV	2	0.8%	二重母音（ai, au）
VVC	1	0.4%	二重母音＋コーダ（oat）
CV	75	24.8%	基本音節型
CVC	**134**	**53.8%**	**最多型**
CVV	36	13.5%	二重母音核
CVVC	7	2.6%	二重母音核＋コーダ
other	6	2.3%	クラスター等
合計	**266**	**100%**	

3.3 onset 別分布

onset	IPA	音節種数	計（回）	代表音節
∅	—	8	41	a, i, o, u

onset	IPA	音節種数	計（回）	代表音節
b	b	19	207	ba, be, bu
h	h	20	138	ha, hi, he
j	j～dʒ	6	44	ja, jen, jan
k	k	23	228	ka, ke, ku
l	l	32	434	la, li, lu
m	m	21	215	ma, man, mu
n	n	20	211	na, nan, nu
n'	ŋ	3	17	n'a, n'o
p	p	22	243	pa, pu, pi
r	r	10	65	ra, ri, ru
s	s	30	476	se, sa, su
s'	ʃ	3	16	s'i, s'a
sj	sj	3	13	sja, sje
t	t	26	328	te, ta, ti
ts	ts	8	60	tsa, tse
ts'	tʃ	2	9	ts'i, ts'a
v	v	2	5	va, ve
w	w	5	52	wa, wan

4. 考察

4.1 CVC 優勢という新発見

固有語彙（source=B）を単独分析すると CVC 型が 54% を占める。これは、巴賽語固有語彙が韻尾子音（コーダ）を積極的に利用する音節体系を持つことを示す。フォルモサ語群（台湾南島語）は語末子音を保持するものが多く、ポリネシア系言語のような CV 優勢とは対照的である（Blust 1999）。source=B の CVC 優勢はこの類型論的傾向と整合する。

4.2 B 固有の onset 音素

`h`・`s'`（ʃ）・`ts'`（tʃ）・`sj` は source=B に出現するが、T+M には出現しない。特に `h` onset は 20 種・138 回と B で大きな存在感を持ち、基本語彙の多くを担う。
`s'`（ʃ）・`ts'`（tʃ）・`sj` の存在は、巴賽語固有語彙に硬口蓋性（palatality）の対立があったことを示唆する。

4.3 不在音素の意味

/q/、/z/、/ɣ/（z'）、/l/（l'）は source=B に出現せず、これらの音素が巴賽語固有の音韻目録に属さないことを示す。混合分析でこれらが現れたのは宜蘭方言（T+M）のエントリ混入によるものであり、カバラン語接触仮説と関連付けることができる（姉妹論文参照）。

---

5. おわりに

本稿は巴賽語辞書データベースの source=B（固有語彙）1,117エントリのみを対象とした音節目録分析を行い、266種・22 onset を確認した。CVC型（134種・50%）が最多構造型であり、混合分析で見えていた「CV優勢」は他ソース混入による錯覚であった。`h`・

`s'`（ʃ）・`ts'`（tʃ）・`sj`が固有語彙に特徴的な onset 音素である。`q`・`z`・`z'`（ʒ）・`l'`（ɮ）は固有語彙に出現せず、これらの音素の帰属については再検討が必要である。

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.
- 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

「すべての音節は、ひとつの記憶。」

© 2026 basay.tw · CC BY 4.0 ｜ ホームへ戻る