

# 浅井恵倫アーカイブ録音からのバサイ語母音音響特性の再構築

— Praatフォルマント測定とeSpeak-NG合成による消滅言語音声の再現 —

著者：蔡永桂 (Tsai Yung-kuei)

日付：2026年4月

種別：原著研究 (音響音韻論 / 言語記録)

ライセンス：CC BY 4.0 引用識別子：basay.tw/research/2026-04-basay-acoustic/

## 要旨

本稿は、台湾北東部の消滅したオーストロネシア語族フォルモサン諸語に属するバサイ語 (Basay/Basai) の母音体系について、フォルマント測定に基づく初めての音響再構築を提示する。6母音のフォルマント周波数 (F1・F2・F3) は、日本人言語学者・浅井恵倫が1936年頃にドロミエン (哆囉美遠、Trobiawan) 方言話者から録音したアーカイブ音声から、Praat音響分析ソフトウェアを用いて直接測定した。録音は国立台湾大学人類学博物館に所蔵されており、東京外国語大学言語アーカイブ (TUFS Language Archive) を通じて入手した。測定されたフォルマント値は、eSpeak-NG 1.52のSPECTSQ2バイナリ形式をリバースエンジニアリングし、Pythonツールを開発することで、動作する音声合成器として実装した。従来の転写記号類似性に基づく音素ファイル近似と比較すると、差異は顕著であり——特に/o/ (F1 : +80.5%) および/a/ (F2 : +30.5%) ——転写ベースの再合成が元の録音に記録された音響特性を信頼性をもって再現できないことが示さ

れた。合成母音はスペクトログラム分析により検証され、本研究の方法論は音響証拠に基づく消滅言語再構築の汎用的な枠組みとして位置づけられる。

キーワード：バサイ語、フォルモサン語族、音響音韻論、フォルマント測定、Praat、eSpeak-NG、アーカイブ録音、消滅言語記録

## □本稿の引用 / Cite this article

### APA :

蔡永桂 (2026). 浅井恵倫アーカイブ録音からのバサイ語母音音響特性の再構築：Praatフォルマント測定とeSpeak-NG合成による消滅言語音声の再現. basay.tw. <https://basay.tw/research/2026-04-basay-acoustic/>

### BibTeX :

```
@misc{tsai2026basay,  
  author = {蔡永桂 and Tsai, Yung-kuei},  
  title   = {浅井恵倫アーカイブ録音からの{バサイ}語母音音響特性の再構築},  
  year    = {2026},  
  month   = {4},  
  url     = {https://basay.tw/research/2026-04-basay-acoustic/},  
  note    = {Praatフォルマント測定とeSpeak-NG合成による消滅言語音声の再現}  
}
```

## 一、はじめに

バサイ語 (Basay、Basaiとも表記) は、台湾北東部沿岸——現在の台北・基隆・宜蘭平野——に歴史的に分布していた消滅したオーストロネシア語族の言語である。言語的には、バサイ語はオーストロネシア語族フォルモサ

ン語派に属し、カバラン語 (Kavalan) と最も近い関係にある。両言語は北東フォルモサン語支の唯一の構成員である。バサイ語の最後の流暢な母語話者は1930年代に日本人言語学者・浅井恵倫によって記録されており (Asai 1937)、その野帳および蓄音機録音は、この言語の一次的かつ代替不可能な記録資料となっている。

浅井の記録以来、後続の研究——李壬癸 (Li 1993, 1999, 2001) の基礎的研究、および曾立洋 (Tseng 2022) の比較再構築研究を含む——は主として音韻転写・文法分析・語彙再構築に焦点を当ててきた。これらの成果は極めて重要であるが、いずれも浅井のアーカイブ録音に内在する最も根本的な資源——音響信号そのもの——を十分に活用していない。管見の限り、これらの録音から直接フォルマント測定を行った先行研究は存在せず、バサイ語母音の音響特性は数値化されないまま残されていた。

本研究はこの空白を埋めるものである。Praat音響分析ソフトウェアを用いて、ドロミエン (Trobiawan) 方言話者の録音から直接F1・F2・F3フォルマント周波数を測定した。これらの実証的な数値は、eSpeak-NG 1.52エンジンのSPECTSQ2バイナリ形式をリバースエンジニアリングし、Pythonツールを開発することで、動作するテキスト音声合成器として実装した。その成果は、バサイ語母音の初めての音声学的根拠に基づく音響モデルである。

本研究はまた、従来のアプローチ——転写記号類似性によるeSpeak音素ファイル選択——の不十分さを、それらのファイルのフォルマント値と本研究の実測目標値とを比較することによって実証する。差異は音響的に顕著であり、転写ベースの近似は測定ベースの再構築の代替とはなり得ないことを示す。

## 二、背景と先行研究

### 2.1 浅井恵倫の録音

本研究の一次音響資料は、浅井恵倫が台湾でのフィールドワーク（1936年頃）中に録音した蓄音機録音群である。録音には宜蘭地域のバサイ語ドロミエン（哆囉美遠、Trobiawan）方言話者の音声記録されている。原音盤は国立台湾大学人類学博物館に所蔵されており、本研究では東京外国語大学言語アーカイブ（TUFS Language Archive）を通じて複製を入手した。録音内容には「ガジの話（3）」「Mutravaiの話（1）」などの叙述テキストが含まれる。

ドロミエン方言はバサイ語の記録された二方言区のうちの一つを代表する。浅井（1937）によれば、主要なドロミエン方言インフォーマントのMutravaiはカバラン語も話しており、若干の音声転移の可能性は排除できない。この言語接触状況は音響データの代表性に関する限界として認識されるが、これらの録音がバサイ語音声の現存する唯一の音響記録であるという重要性を減じるものではない。

### 2.2 先行言語記録

浅井恵倫の1937年野帳は、バサイ語音韻論の一次転写記録である。李壬癸（1999）は現在最も包括的な音韻分析を提供しており、約15の子音音素と4〜6母音の母音体系を確認している。曾立洋（2022）はバサイ語とカバラン語の比較に基づき、4母音（\*i・\*u・\*ə・\*a）と4重母音を有する原北東フォルモサン語の再構築を提示している。

これらの研究はいずれも母音フォルマンの音響測定を試みていない。本研究は転写からの推論に依存せず、アーカイブ録音から直接こうした測定値を提供する。

## 2.3 eSpeak-NGとフォルマント合成

eSpeak-NGはKlattモデルに基づくフォルマント合成器を用いたオープンソースのテキスト音声合成エンジンである。母音音質はSPECTSQ2形式のバイナリファイルで指定され、フォルマント周波数・帯域幅・振幅パラメータを含むスペクトルフレームが符号化される。合成エンジンはこれらのバイナリファイルから直接フォルマント値を読み取り（`peaks[1]`・`peaks[2]`・`peaks[3]`がそれぞれF1・F2・F3に対応）、Hz単位の精度でフォルマント指定が可能である。本研究では、測定されたフォルマント値からSPECTSQ2形式ファイルを直接生成するPythonスクリプトを開発し、再現可能でプラットフォームに依存しない合成パイプラインを実現した。

## 三、方法

### 3.1 音響測定

フォルマント周波数はPraat (Boersma & Weenink) を用いて浅井のアーカイブ録音から測定した。分析対象はドロミエン方言録音のうち叙述テキスト「ガジの話(3)」および「Mutravaiの話(1)」である。スペクトログラム表示で母音トークンの定常部分を同定し、各トークンの中間点においてPraatのフォルマント追跡器でF1・F2・F3値を抽出した。得られた測定値は浅井が1936年頃に録音したドロミエン方言話者によるバサイ語母音の音響実現を表す。第3の録音「Mutravaiの話(4)」は自動母音分類手続きの方法論開発に別途使用したが、合成器に実装されたフォルマント値への寄与はない。

フォルマント測定の信頼性を評価し録音雑音の潜在的影響を検討するため、追加の第4録音「Saturai(1)」(歌唱録音)を分析し、9,991の分析フレームを抽出した。実装に使用した2つの散文叙述と歌唱録音の3つのソースにわたって、F2の変動係数(CV)は全6母音で10%未満(範囲:1.9~6.0%)、F1のCVは4母音(/e/・/a/・/o/・/ə/)で5%未満であった。F1のCVが高かった2母音(/i/:16.3%、/u/:12.2%)は散文録音でのトークン数が少なく(それぞれN=40およびN=174)、音響測定誤差ではなく統計的サンプリン

グ効果によるものと考えられる。これらの結果は、フォルマント測定が異なる録音セッションおよび発話スタイルにわたって頑健であり、1930年代の録音雑音が測定値に実質的な影響を与えないことを示す。

帯域幅推定値（B1・B2・B3）は、測定されたフォルマント配置と整合する標準的な弛緩発声規範に従って設定した。アーカイブ録音の音質からは帯域幅情報を信頼性をもって抽出できないためである。測定フォルマント値と帯域幅推定値を表1に示す。

表1. バサイ語母音フォルマント周波数（ドロミエン方言）。F1-F3は浅井アーカイブ録音（1936年頃）からPraatを用いて測定。帯域幅（B1-B3）は標準的弛緩発声推定値。

母音 (IPA)	F1 (Hz)	F2 (Hz)	F3 (Hz)	B1 (Hz)	B2 (Hz)	B3 (Hz)	出典
/i/	269	1696	2128	60	80	120	実測値
/e/	639	1701	2245	70	90	120	実測値
/ə/ (中舌母音)	858	1703	2180	100	100	120	実測値
/a/	1064	1671	2185	120	110	130	実測値
/o/	805	1191	1949	80	100	130	実測値

母音 (IPA)	F1 (Hz)	F2 (Hz)	F3 (Hz)	B1 (Hz)	B2 (Hz)	B3 (Hz)	出典
/u/	381	1219	1910	60	90	130	実測値

### 3.2 SPECTSQ2形式のリバースエンジニアリング

SPECTSQ2バイナリ形式は、eSpeak-NGソースコード ( `spect.cpp` ・ `spect.h` ・ `synthesize.h` ) の解析と既存母音ファイルの16進ダンプとの照合によりリバースエンジニアリングした。主要な知見は以下の通りである：

- ファイルヘッダはマジックシグネチャ `SPECTSQ2` ( `0x53504543 + 0x54535132` ) を使用し、リトルエンディアンの長さプレフィックス付き名前文字列が続く。
- 浮動小数点値 ( 時間・F0・長さ・dx ) は標準的なIEEE 754 64ビット倍精度ではなく、**80**ビット拡張精度 ( 各10バイト ) で格納される。
- F1・F2・F3はそれぞれ `peaks[1]` ・ `peaks[2]` ・ `peaks[3]` に符号化され、`peaks[0]` はF0領域を符号化する。
- `KLATT_Kopen` ( `klatt_param[]` インデックス5 ) は50に設定しなければ、摩擦雑音のない有声弛緩発声合成が得られない。

### 3.3 バイナリファイルの生成

6つのバサイ語母音のSPECTSQ2ファイルは、既存のeSpeak-NG母音ファイルをテンプレートとしてパッチするPythonスクリプト

( `gen_basay_vowels_v8.py` ) により生成した。パッチ手順は：(1) `peaks[1-3].pkfreq` および `formants[1-3].freq` を測定F1-F3値に置換；(2) `peaks[0].pkheight` および `peaks[4-8].pkheight` をゼロにして非フォルマントスペクトル成分を抑制；(3) 調波スペクトル配列

( `spect[]` ) をゼロにしてテンプレート母音から継承されるスペクトル着色効果を除去。 `KLATT_Kopen` 値およびその他の合成パラメータはテンプレートから保持し、クリーンな有声出力を確保する。

## 四、結果

### 4.1 従来の転写近似値との比較

今回の音響測定以前、eSpeak中のバサイ語母音は転写記号類似性により選択された既存音素ファイルで近似されていた(例: /i/ → `i_fnt`、/o/ → `o_mid`)。表2はこれらの近似ファイルのフォルマント値と浅井録音から測定した値を比較する。

表2. 従来の転写近似値と浅井アーカイブ録音実測値の比較。\* `i_fnt` は旧SPECTSEQ形式を使用しており、キーフレーム高さデータを持たない。

母音	パラメータ	従来近似値 (Hz)	実測値 (Hz)	差 (Hz)	差 (%)
/i/	F1	~0 *	269	—	—
	F2	~0 *	1696	—	—
	F3	~0 *	2128	—	—
/e/	F1	584	639	+55	+9.4%
	F2	1820	1701	-119	-6.5%
	F3	2560	2245	-315	-12.3%
/ə/	F1	581	858	+277	+47.7%

母音	パラメータ	従来近似値 (Hz)	実測値 (Hz)	差 (Hz)	差 (%)
	F2	1653	1703	+50	+3.0%
	F3	2477	2180	-297	-12.0%
/a/	F1	872	1064	+192	+22.0%
	F2	1280	1671	+391	+30.5%
	F3	2660	2185	-475	-17.9%
/o/	F1	446	805	+359	+80.5%
	F2	883	1191	+308	+34.9%
	F3	2485	1949	-536	-21.6%
/u/	F1	371	381	+10	+2.7%
	F2	1276	1219	-57	-4.5%
	F3	2308	1910	-398	-17.2%

差異が最も顕著なのは/o/ ( F1 : +80.5%、 F2 : +34.9% ) であり、次いで/ə/ ( F1 : +47.7% ) および/a/ ( F2 : +30.5% ) である。/o/のF1差359 Hzは音響的に重要な意味を持つ：近似ファイルはこの母音を中母音域 ( 446 Hz ) に置いていたが、実測値 ( 805 Hz ) はより高い中後舌母音として識別する。後舌母音 ( /o/・/u/ ) のF3値は従来近似で400~536 Hz系統的に過大評価されており、前舌母音テンプレートを誤用した結果である。

/i/の場合は最も極端である：従来使用されていた近似ファイル ( `i_fnt` ) は旧SPECTSEQ形式を採用しており、キーフレーム高さデータが含まれず、今回の再構築以前はこの母音が実質的に音響的未定義の状態にあった。バサイ

語合成器における/i/の現在の実装は、浅井録音の実測値 (F1 : 269 Hz、F2 : 1696 Hz、F3 : 2128 Hz) に完全に基づいており、比較すべき有効な先行近似値は存在しない。この事実は転写近似アプローチの根本的な不十分さをさらに浮き彫りにする：類型論的に最も典型的で通言語的に最も安定した高前舌母音である/i/に対してすら、従来の手法は使用可能な音響出力を生成できなかった。

## 4.2 スペクトログラムによる検証

合成母音はSoXを用いたスペクトログラム分析により検証した。スペクトログラムはF1の順序が期待される類型論的階層に従うことを確認する：F1(/a/) > F1(/o/) > F1(/e/) > F1(/u/) > F1(/i/)、これは普遍的な母音高さ次元と一致する。前後対立も正しく表現されており、/i/・/e/・/ə/のF2は1700 Hz付近に集中し、/o/・/u/のF2は1200 Hz付近に集中しており、実測値および/o/・/u/に期待される後舌母音特性と一致する。

合成母音はすべて4 kHz以上に残留広帯域雑音を示すが、これはフォルマント指定の問題ではなくKlatt合成エンジンの有声化モデルの特性であり、F1-F3実装の音響的妥当性には影響しない。

## 五、考察

本研究の中核的貢献は方法論的なものである：アーカイブ録音からの音響測定——たとえ音質の限られた録音であっても——が、転写ベース近似とは実質的かつ系統的に異なるフォルマント値をもたらすことを実証した。表2に記録された差異はランダムな変動ではなく、近似アプローチの特定の構造的欠陥を反映している：/i/に対して `i_fnt` を使用すると有効なフォルマントデータのないファイルが生成される；/o/に対して `o_mid` を使用するとF1が359 Hz過小評価され母音高さが誤表現される；/o/と/u/に前舌母音テンプレートを使用するとF3が400～536 Hz系統的に過大評価される。

これらの欠陥は元の録音に立ち返ることなしには発見できなかったものである。本研究は、浅井のアーカイブ録音が音響音韻論的目的のために収集されたものではないにもかかわらず、Praatによるフォルマント抽出に十分な信号品質を有することを実証する。この発見はバサイ語を超えた意義を持つ：他の消滅フォルモサン語族言語のアーカイブ録音が存在する場合、それらも同様に音響測定に適している可能性を示唆する。

本研究の測定にはいくつかの限界がある。第一に、録音された話者（ドロミエン方言）はカバラン語も話しており、音声転移の可能性を完全に排除できない。第二に、正確な録音日付（1936年頃）は不確かである。第三に、帯域幅値は録音から測定されたのではなく、標準的な弛緩発声規範に基づいて推定された。今後の研究では帯域幅の直接測定を試み、適切な録音が特定できる場合にはドロミエン方言のフォルマント値を新社（Xinzhe）方言と比較すべきである。

測定信頼性を評価するため、3つの独立した録音を分析した：実装に使用した2つの散文叙述（ガジの話（3）とMutravaiの話（1））および追加の歌唱録音（Saturai（1）、N=9,991フレーム）。3つのソースにわたる変動係数は全6母音のF2で10%未満（範囲：1.9～6.0%）、4母音（/e/・/a/・/o/・/ə/）のF1で5%未満であった。F1変動が高かった2母音（/i/：16.3%、/u/：12.2%）は散文録音でのトークン数が少なく（それぞれN=40およびN=174）、録音雑音ではなくサンプリング効果によるものと考えられる。これらの結果はフォルマント測定が異なる録音環境および発話スタイルにわたって頑健であることを示す。

実装されていないパラメータの一つに語彙アクセントがある。バサイ語のアクセント位置は、現存する録音から十分な確信をもって決定できない：アーカイブ音声では、F0上昇・持続時間・振幅などの音響的相関によって強調音節を信頼性をもって特定できない。予備的観察は2音節語での末尾アクセント・3音節語での前末尾アクセントのパターンを示唆し、オーストロネシア語族によく見られるアクセント型と整合するが、これは確認された規則ではなく仮説にとどまる。現在の実装はeSpeak-NGのデフォルトアクセント付与を使用している。ソースコード検査により、eSpeak-NG 1.52の言語設定ファ

イルでは `stressrule` パラメータが実装されていないことが確認されており、`lang`ファイルのみではアクセント付与を制御できない。より信頼できる音響証拠が得られた時点——理想的にはアーカイブ録音中の同一語彙の複数トークンにわたる持続時間とF0の系統的測定によって——このパラメータは改訂されるべきである。

本稿で述べたeSpeak-NG実装は完全に再現可能である。Python生成ツール・SPECTSQ2形式ドキュメント・フォルマント目標値は研究コミュニティに公開されており、新たな音響証拠が得られた際に合成器を更新できる。

特筆すべき知見の一つは、実測フォルマント値および合成出力における/a/と/a/・/o/の近融合である。6母音のF2値は狭い後方域(1191~1703 Hz)に系統的に圧縮されており、/a/・/o/・/ə/の3母音のF2は特に近接している。この圧縮現象は、相互に補強する少なくとも3つの要因と整合する。第一に、高齢話者における音韻的ドリフトは十分に記録されており、前舌・中舌母音の後退・咽頭腔の拡大・口唇収縮力の低下などが含まれ、これらはすべて母音空間のF2幅を縮小する。第二に、録音された話者はカバラン語も話していたが、カバラン語の再構築母音目録(\*i・\*u・\*ə・\*a)は独立した/o/・/e/を含まないため、この話者において/a/と隣接母音の音韻対立が弱まっていた可能性がある。第三に、曾立洋(2022)によるバサイ語再構築は4母音(\*i・\*u・\*ə・\*a)のみを含んでおり、バサイ語の/e/・/o/が後期拡張または接触借用を表し、音韻的定着度が相応に弱いことを示唆する。したがって、合成母音における/a/と/a/・/o/の知覚的近融合は、合成手法の限界ではなく、元の録音において音韻対立が実際に弱まっていた状況を忠実に反映している可能性がある。この観察はそれ自体が音韻論的知見を構成する：これらのアーカイブ録音は母音空間収縮過程の音響証拠を保存しており、他の消滅言語および危機言語の末期話者の音声学において記録されてきた現象と一致する。

## 六、結論

---

本研究は、浅井恵倫のアーカイブ録音（1936年頃、ドロミエン方言、東京外国語大学を通じて入手）から直接フォルマント測定を行うことに基づく、バサイ語母音体系の初めての音響再構築を提示した。6母音のフォルマント周波数 $F1 \cdot F2 \cdot F3$ をPraatで抽出し、リバーズエンジニアリングされたバイナリ形式とPython合成ツールを通じてeSpeak-NG 1.52に実装した。従来の転写近似との比較は、元の録音に立ち返ることなしには検出できなかった実質的な差異を明らかにした。得られた合成器はバサイ語の初めての音声学的根拠に基づく音響モデルを構成し、本研究の方法論は現存するアーカイブ録音を持つ消滅言語の証拠に基づく音響再構築のための再現可能な雛形を提供する。

## 参考文献

---

- 浅井恵倫 (Asai, Erin) . 1937. 《バサイ》野帳. [録音1936年頃、ドロミエン方言.] 原本：国立台湾大学人類学博物館；複製：東京外国語大学言語アーカイブ.
- Boersma, Paul & David Weenink. Praat: doing phonetics by computer. [www.praat.org](http://www.praat.org)
- 李壬癸 (Li, Paul Jen-kuei) . 1993. 三種台湾消滅言語の新資料. 《中央研究院歴史語言研究所集刊》63(2): 301–323.
- 李壬癸 (Li, Paul Jen-kuei) . 1999. バサイ語のいくつかの問題. E. Zeitoun & P.J.K. Li (編)、《第8回オーストロネシア語学国際会議選集》、635–664頁. 台北：中央研究院.
- 李壬癸 (Li, Paul Jen-kuei) . 2001. バサイ語の言語的地位. 《言語と言語学》2(2): 155–171.
- 曾立洋 (Tseng, Li-yang) . 2022. 北東フォルモサン語の再構築. 修士論文.
- Dunn, Reece H. et al. eSpeak-NG. [github.com/espeak-ng/espeak-ng](https://github.com/espeak-ng/espeak-ng)

[← 研究成果一覧に戻る](#)