

CHAPTER 1. THỐNG KÊ ÂM TIẾT BAHNARIC

1. Sơ lược âm tiết Bahnaric:

Cấu trúc âm tiết của tiếng Ba Na dạng tổng quát như sau: Cp1VCp2 Cm(1,2) Vm Cf. Trong cấu trúc này, Cp1VCp2 là tiền âm tiết, Cm(1,2) Vm Cf là âm tiết chính.

Cấu trúc âm tiết tiếng Ba Na thuộc dạng cận âm tiết (sesqui-syllabic). Tiếng Ba Na thuộc nhóm Ba Na Trung nên có điểm chung về mặt cấu trúc âm tiết so với các ngôn ngữ như Ba Na, Mnông, Mạ, Cơ Ho. Xét ở góc độ tổng quát, cấu trúc âm tiết Ba Na đa phần giống với cấu trúc âm tiết của tiếng Ba Na, Mnông, Mạ, Cơ Ho. Từ ngữ âm (phonology word) của tiếng Ba Na bao gồm tiền âm tiết và âm tiết chính. Tiền âm tiết có thể xuất hiện hay vắng mặt nhưng âm tiết chính luôn luôn có mặt. Trong hai yếu tố của từ ngữ âm thì yếu tố thứ hai luôn mang trọng âm [11,12].

Mô hình cấu trúc tổng quát của tiếng Ba Na có dạng:

Cp1VCp2Cm(1,2)WVm Cf

* Cp1VCp2: là tiền âm tiết (presyllable), gồm có các phần như:

- Cp1 là phụ âm đầu tiền âm tiết gồm các âm: âm tắc /b/, âm xát /s/ hoặc /h/, âm vang /l/ hoặc âm tắc /g/.
- V là một nguyên âm trung hoà (schwa) /ə/, không đối lập về phẩm chất và trường độ.
- Cp2 là một âm vị âm tiết tính (syllabic phoneme), là một trong những âm sau: /m/, /n/, /p/, /ŋ/, /l/.

Cm(1,2)WVm Cf: là âm tiết chính (main syllable) bao gồm :

- Cm1 là một phụ âm bất kì trong hệ thống phụ âm đầu tiếng Ba Na.
- Cm2 là một trong các phụ âm như: /r/, /l/, /j/, ví dụ: [ʔlɔŋ] cây gỗ, [pran] khỏe.
- W là một âm đệm có đặc tính ngữ âm như một bán nguyên âm. Âm đệm là /u/ rất phổ biến trong tiếng Ba Na. Ví dụ: [swaŋ] mùa.
- Vm nguyên âm (vowel) là một nguyên âm thuộc hệ thống nguyên âm tiếng Ba Na.
- Cf là âm cuối (final). Âm cuối Ba Na có thể là một phụ âm, một bán nguyên âm, hoặc một tổ hợp gồm bán nguyên âm và một phụ âm /jh/.

Tiếng Ba Na Gia Lai có hai dạng từ ngữ âm: từ đơn tiết và từ đa tiết.

- **từ ngữ âm đơn tiết**: chỉ gồm một âm tiết (một tiếng), ví dụ:

Bảng 3.9: Ví dụ từ ngữ đơn âm tiết

tɛh	(đất)	ʔɔh	(em)
ʔdak	(nước)	Sɛm	(chim)
Hla	(lá)	Phɛ	(gạo)

- **từ ngữ âm đa tiết**: gồm một âm tiết đứng trước (gọi là "tiền âm tiết") và một âm tiết đứng sau được đọc nhấn mạnh hơn (gọi là "âm tiết chính"), ví dụ:

Bảng 3.10: Ví dụ từ ngữ đa âm tiết

Kɤmǎŋ	(đêm)	Lɤʔu	(dừa)
Pɤkaw	(hoa)	ʔakar	(da)
Kɤne	(chuột)	Bɤŋaj	(người)

Cấu trúc âm tiết tiếng Ba Na hầu như không có sự khác biệt khi so sánh các phương ngữ với nhau. Sự khác biệt chỉ thể hiện ở hệ quả của quá trình đơn tiết hoá. Đối với tiền âm tiết, âm cuối sẽ được lược bỏ để cấu trúc CVC chỉ còn lại CV.

2. Hệ thống nguyên âm

Hệ thống nguyên âm Ba Na Kriem gồm các nguyên âm "dài" và nguyên âm "ngắn". Đối lập "ngắn"/"dài" này chỉ có khi nguyên âm ở trong âm tiết chính.

Tiếng Ba Na cũng có 16 nguyên âm: 8 nguyên âm dài /i, e, ɛ, ə, a, u, o, ɔ/, 6 nguyên âm ngắn /ĩ, ɛ̃, ə̃, ă, ʊ̃, ɔ̃/ và 3 nguyên âm đôi /iə/, /uə/, /uə/.

Nguyên âm âm tiết chính tiếng Ba Na gồm có 18 âm vị. Trong đó có 9 nguyên âm dài, 7 nguyên âm ngắn và 2 nguyên âm đôi. Nguyên âm đơn tiếng Ba Na gồm có 16 âm vị: /i, ĩ, e, ɛ, ɛ̃, u, ũ, ə, ə̃, a, ă, u, ʊ̃, o, ɔ, ɔ̃/.

Bảng hệ thống nguyên âm đôi tiếng Ba Na được mô tả theo độ nâng của lưỡi: cao, vừa, thấp. Vị trí của lưỡi có 3 vị trí: trước, giữa và sau. Các cặp nguyên âm theo thể đối lập về trường độ bao gồm: [i, ĩ], [u, ũ], [ə, ə̃], [ɛ, ɛ̃], [a, ă], [ɔ, ɔ̃]. Trong các ngôn ngữ cùng nhóm Ba Na Nam đều có sự đối lập tương tự.

2.1. Nguyên âm đơn

Bảng nguyên âm âm tiết chính tiếng Ba Na

Bảng *Error! No text of specified style in document.*1: Danh sách nguyên âm chính trong tiếng Ba Na

	Trước		Giữa		Sau	
Cao	ĩ	i	ụ	u	ũ	u
Vừa		e	ə̣	ə		o
Thấp	ɛ̣	ɛ	ă	a	ɔ̣	ɔ
Nguyên âm đôi		ie		uə		uo

- Các nguyên âm dòng trước gồm có: /i, ĩ, e, ɛ, ɛ̣/

/i/: nguyên âm dòng trước, cao, trường độ dài, ví dụ: [gi] “lắc lư”, [hil] “trống rỗng”.

/ĩ/: nguyên âm dòng trước, cao, trường độ ngắn, VD: [rĩh] “sống”, [lĩh] “chậm”.

/e/: nguyên âm dòng trước, vừa. VD: [chet] “nhún nhảy”, [ncem] “dự trữ”

/ɛ/: nguyên âm dòng trước, thấp, VD: [chɛ] “dây”, [dapɛ] “men”.

/ɛ̣/: nguyên âm dòng trước, thấp, ngắn, VD: [atɛ̣h] “phía dưới”.

- Các nguyên âm dòng giữa /u, ə, ə̣, a, ă/

/u/: nguyên âm dòng giữa, cao, dài, VD: . [chup] “con gián”, [vur] “con bò”

/ə/: nguyên âm dòng giữa, vừa, dài, VD: [cavəm] “dơ bẩn”, [ganəl] “gót chân”

/ə̣/: nguyên âm dòng giữa, vừa, ngắn, VD: [glə̣ŋ] “chim sơn ca”, [dagə̣t] “lạnh run”

/a/: nguyên âm dòng giữa, thấp, dài, VD: [dagat] “lừa dối”, [dapa] “rùa”

/ă/: nguyên âm dòng giữa, thấp, ngắn, VD: [dă?] “ho”, [drăng] “cạnh cứng”

- Các nguyên âm dòng sau: /u, ũ, o, ɔ, ɔ̣/.

/u/: nguyên âm dòng sau, cao, dài, VD: [chut] “khịt”

/ũ/: nguyên âm dòng sau, cao, ngắn, VD: [chũ?] “dăm”

/o/: nguyên âm dòng sau vừa, VD: [choŋ] “nhánh (sông)”

/ɔ/: nguyên âm dòng sau, thấp, dài, VD: [choɭ] “mài”.

/ɔ̣/: nguyên âm dòng sau, thấp, ngắn, VD: [chọh] “đá”

Các nguyên âm dài làm thành hệ thống sau:

Bảng Error! No text of specified style in document..2: Hệ thống các nguyên âm dài trong tiếng Ba Na

i	ʔbiɰ	(con rắn)
	Kɣtiŋ	(xương)
e	Peŋ	(ba)
	Theŋ	(đuôi)
ɛ	Sɛm	(chim)
	Kɣpɛn	(cái khố)
u	ʔɯmɯl	(mây)
	Gruum	(sấm)
ɤ	Tɤm	(gốc)
	Klɤm	(gan)
a	Akar	(da)
	ʔdak	(nước)
u	ʔbum	(củ)
	ʔɲuɟ	(khói)
o	ʔapom	(tròn)
	Goh	(đỏ)
ɔ	ʔboɰ	(muối)
	Krɔŋ	(sông)

Các ví dụ nguyên âm ngắn bao gồm:

Bảng Error! No text of specified style in document..3: Hệ thống các nguyên âm ngắn tiếng Ba Na

ĩ	tĩh	(to, lớn)
---	-----	-----------

	sĩt	(khâu, may)
ε̣	mε̣ t	(mặt)
	ʔabε̣ n	(váy)
ɾ̣	ʔbɾ̣ l	(mồm)
	ʔdɾ̣ k	(cám)
ặ	ʔnặ r	(ngày)
	tɿlặ m	(trong, bên trong)
ụ̃	cụ̃ r	(vôi)
	ɲụ̃ ɲ	(con lợn)
ɔ̣	ʔdọ k	(con khi)
	ʔalọ ɲ	(cổ)

2.2. Nguyên âm đôi

Nguyên âm đôi: ngoài các nguyên âm đơn, trong tiếng Ba Na Kriem còn có các nguyên âm đôi: /iə/, /uə/ và /ʉə/. Đây là các nguyên âm đôi có yếu tố thứ nhất là nguyên âm khép /i, u, ʉ/ yếu tố thứ hai là nguyên âm trung hoà /ə/

Trong tiếng Ba Na có 3 nguyên âm đôi /ie, uə, uo/. Nguyên âm đôi /ie/ có vị trí cấu âm dòng trước, hẹp vừa, bổng. Âm sắc của /ie/ không cố định, là nguyên âm đôi đơn âm vị tính. Nguyên âm đôi /ie/ có hai biến thể:

- Đối với âm tiết mở và âm tiết có âm cuối /ɲ/ thì yếu tố thứ hai của /ie/ được thể hiện rộng hơn, như một nguyên âm hàng giữa có âm sắc trung hoà [ə]. Ví dụ: ??] “lợn rừng”, [nriəɲ] “cái rây”, [ratiəɲ] “củ năng”.
- Trong các loại âm tiết còn lại thì /ie/ được bắt đầu bằng yếu tố [i] có độ mở hẹp, sau đó chuyển sang yếu tố [e] có độ mở lớn hơn một bậc và có cùng âm sắc bổng. Ví dụ: [katiet] “vẹt”, [dier] “quay vòng”, [siem] “bón cơm”.

Nguyên âm đôi /uo/: có vị trí dòng sau tròn môi, hẹp vừa, âm sắc trầm, không cố định, âm lượng trầm, VD: [lapuol] “bí đao”. Nguyên âm đôi [uo] chỉ xuất hiện trước các phụ âm cuối như sau: [t, n, l, j] ví dụ, [duon] “mũ”, [lapuon] “cổ họng”. Biến thể tự do của nguyên âm đôi /uo/ là âm tố [uə], ví dụ [tuə] “bớt”.

Ví dụ nguyên âm đôi:

/iə/ ví dụ: /təgliəp/ bong gân
/kədiəh/ búng tay.

/uə/ ví dụ: /puən/ bốn
/kruəʔ/ cam
/luən/ giết

/ʊə/ ví dụ: /gruəʔ/ bấy,
/jʊəŋ/ giường.

(nguyên âm đôi ʊə xuất hiện trong một số ít từ)

/jer/ gà

Các nguyên âm (trừ /o và e/) tham gia vào sự đối lập về trường độ (nguyên âm ngắn - nguyên âm dài)

Trước phụ âm cuối tắc họng, nguyên âm mất đối lập về trường độ. Tuy nhiên, sự khác biệt về trường độ ngắn/dài của nguyên âm có thể có giá trị chỉ “sắc thái” nhấn mạnh trong cặp [pruʔ] “rất xấu” và [pruːʔ] hơi xấu.

Trong tiếng Ba Na vần [ăj] theo chữ viết truyền thống ghi bằng **ei**, thực chất đây là nguyên âm dòng giữa /ă/ kết hợp âm cuối /-j/, tương tự vần /ăj/ tiếng Việt (chữ quốc ngữ ghi **ay**).

Tham khảo thêm tại:

https://docs.google.com/document/d/1eUJJaRHxxZshxS6DIziHUln5nifhArk-/edit?usp=share_link&ouid=106195316964367283586&rtpof=true&sd=true

Cấu trúc âm tiết của tiếng Ba Na như vậy: Cp1VCp2Cm(1,2)WVm Cf

mình có nguyên âm của tiền âm tiết: V và nguyên âm của âm tiết chính Vm. Anh gửi bảng liệt kê con chữ của hai nguyên âm này rồi em đánh dấu.

Ở dạng đầy đủ nhất thì âm tiết sẽ được viết đủ hết theo mẫu trên. Mình căn cứ vào mẫu này để lọc.

các con chữ dùng cho V: i, a, ơ

Đợi chút anh liệt kê các con chữ dùng cho Vm:



Vm: i, ỉ, iê, iễ, ê, ể, e, ẻ, ư, ử, ươ, ướ, ơ, ỡ, a, ă, u, ũ, uô, uỗ, ô, ồ, o, ồ

17:12



CHAPTER 2. TASK

1. Required Tech Knowledge:

- Python
- Jupyter Notebook

2. Tasks:

Google Drive của project:

https://drive.google.com/drive/folders/1MLL7PIRjssxPfb6ZE3cOoxz_wTkES59?usp=share_link

Step 1: Vào drive đó và tải các files trong drive về máy local

Step 2: Bắt đầu với file `Kriem_common_checked.xlsx`, cần kiểm tra âm tiết dài nhất và âm tiết ngắn nhất:

ÂM TIẾT = Nguyên âm + phụ âm

- **Âm tiết dài nhất:** Âm tiết có nhiều kí tự nhất. Để tìm âm tiết dài nhất:

* **Idea:** Thiết kế giải thuật **tìm ra nguyên âm trước**, sau đó span ra các characters bên cạnh (tất nhiên sẽ phải tuân thủ theo cấu trúc âm tiết bahnaric ở Chapter 1).

Prototype

```
Found_vowels = dict()
```

```
Found_consonants = []
```

```
Max = 0
```

NOTE: mỗi tuple trong dict sẽ có 2 thuộc tính là key và value. Cần key là vowel tìm được, value là index của nó (for easier debugging later on)

while(not end of word)

1.1 Tìm được nguyên âm, lưu vào Found_vowels, khởi tạo
consonant = <nguyên_âm_vừa_tìm_được>

1.2 Cứ span rộng ra (nếu các character kề bên thỏa được cấu trúc âm tiết thì cứ span, nhưng lưu ý phải span cả bên trái lẫn bên phải), span thì cộng vào consonant đang xét

1.3 Khi không thể span được nữa thì push vào Found_consonants, và so sánh max với len của consonant vừa tìm thấy. Nếu $\text{max} > \text{len}(\text{consonant})$ thì $\text{max} = \text{len}(\text{consonant})$

- **Âm tiết ngắn nhất:** Âm tiết có nhiều ít tự nhất. Giải thuật tương tự, chỉ khác chỗ set min = 100000 rồi so sánh xuống.

* Các kết quả thu thập được cần trình bày ở jupyter notebook để mọi người có thể thẩm định tính chính xác của giải thuật.

Step 3: File Tu_dien_Bahnar_Viet_corrected_by_human.xlsx tương tự.

Step 4: Record các thông tin thu thập được vào 2 file _stat.xlsx trong Drive