

## 配布資料はこちら

- PDFファイル(2.1 M)



PPTXファイル(11.7 M)



# 母語話者が有するアクセント感覚の統計的モデリングとそれを用いた学習者のアクセント感覚の定着に関する定量的分析

李俏（大連理工大学・東京大学）    liqiao@gavo.t.u-tokyo.ac.jp

李墨彤（大阪大学）    ribokuto@gmail.com

峯松信明（東京大学）    mine@gavo.t.u-tokyo.ac.jp

## 本発表の流れ

- 1. 先行研究
- 2. 本研究の目的
- 3. 母語話者のアクセント感覚の統計的モデル化
- 4. 中国人学習者を対象としたデータ収集とアクセント感覚のモデル化
- 5. 母語話者と学習者のアクセント感覚の乖離に関する分析
- 6. 今後の課題

**先行研究**

首都圏方言話者による架空外来語のアクセント型の生成  
および自然度評価(李, 2018)

## 先行研究：外来語アクセントの予測可能性

-3の規則 (McCawley, 1968) : 語末から3つ目のモーラを含む音節にアクセント核を置く

(川越, 2007): 語末音節を切り離し、その次から数えて2つ目のモーラを含む音節に核を置く

調査の目的 (a) ゆか (わ) ひで き え ざ き れ お な ロ サン ゼ ル ス  
リハ プール オー ス ト リ ア オー ス ト ラ リ ア

上記のような規則を踏まえ、母語話者が未知なる外来語にどのようなアクセントを振るのか、どのようなバリエーションがあるのかをアンケートにより、調査した。この時、特定の形態素（「**さ**ひ<sup>ハ</sup>た<sup>タ</sup>ま<sup>マ</sup>」**ばん**ご<sup>ゴ</sup>」など<sup>坂東</sup>）および音節解釈の曖昧性（促音など）による影響を抑えて架空外来語を作成し、利用されるアクセントを調査した。

サッカー サイダー

生成(読み上げ)と知覚(アクセント型の確然度評価)の両方から考察法則』, p20)

\* (b)は-3の位置に特殊モーラを持つ語である

「.」は音節境界を表す

モーラは、基本的にかな文字一文字=1モーラである

# 架空外来語リスト

モーラ数	音節構造 (17種)	1	2	3	4	5	6	7	8	9
2	LL	ペザ	コデ	ギパ	ジテ	ピゾ	ガゾ	パキ	ゴヨ	ザミ
3	LLL	ピゴピ	パヘミ	ナザエ	メネピ	ゾゲマ	ロデセ	オデハ	ミポメ	ペパニ
	HL	ジーハ	ヒーポ	ローベ	デーベ	ポーマ	ボンメ	ドンサ	ノンシ	ワンマ
	LH	マベン	レポン	ゲムン	ジダー	ミダン	ゾボー	コガー	ジルー	ロモー
4	LLLL	ワネポシ	ピネテニ	ラポネソ	ピギチネ	メヨゼソ	マガデオ	ピニギノ	セネヘゾ	テタゾモ
	HLL	デーヘモ	ナービマ	ベンタポ	コンレピ	ローゲワ	ベーパノ	アンロピ	ムンベハ	パーネザ
	LHL	ヨダーピ	メボンヨ	モゾーハ	ゾモーボ	ミゾンナ	ミアンゾ	ジラーミ	キナーメ	ソルンヒ
	LLH	コポレー	ボミナー	テポマー	ミハナン	タポソン	ゾテルー	リペレン	ネテシー	セネノン
	HH	ゾンエー	サーペー	ミードン	グーレン	ザーセン	ズンロー	ギーレン	ランボン	トンジー
5	LLLLL	ペハヨピモ	セソポメヨ	エゼガヘオ	ヤデピギノ	ウゼテベマ	ミジナペシ	ビヒペミボ	ヤテソヘピ	ウパヨゼホ
	HLLL	カーシポヤ	ドンラゼコ	ヘンタゾネ	ボンセポニ	スンゾナポ	バーマバヨ	ウージモザ	ラーゼワメ	スンパザヨ
	LHLL	ニスンパガ	ザポーモベ	テルンポハ	タズーラカ	チマーペヒ	カザーパセ	ベラーゴラ	ガパーピワ	パドンザテ
	LLHL	マピガンエ	オデヌンザ	レゾガーメ	ピメモーナ	ビハポーレ	ベハワーモ	ペサギンモ	ゼポネーベ	レデソンノ
	LLLH	ガテピゼン	セネハヨン	ポダザユー	ロレボシー	オメナズー	ゴラソガー	ポナパヒー	テザポユー	ネナモチン
	HHL	メージーモ	ダンツープ	ハーバンポ	ヘンヌーシ	ターシーベ	ビンサーデ	ヒンパーゼ	タンチーマ	チンミンネ
	HLH	ウーポエー	ローパルン	ソンビツー	ドーペズン	ビンヘゾー	アンメクー	ムーモネン	テーレズー	ブーラケン
	LHH	ラソープン	テモハー	ヨモンゴー	ゲワーリン	ピノータン	オグーノン	サゴントー	ヨリンビー	エターモン

既存の形態素による影響をできるだけ避けるため、N-gram（音節単位で2音節から）でかく調査語の音節連続を走査し、いずれも辞書（『日本語の語彙特性』（天野・近藤, 1999））に登録されていないことを確認した

- **調査対象**

全員首都圏方言話者

- **人数**

読み上げ調査と自然度評価調査両方参加：22人

読み上げ調査のみ参加：1人

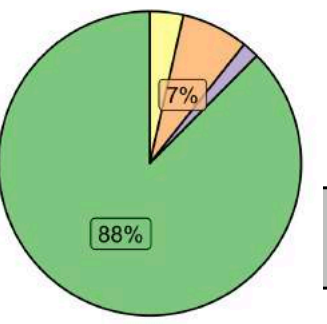
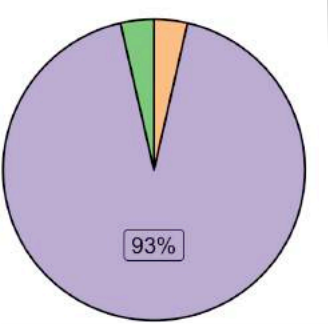
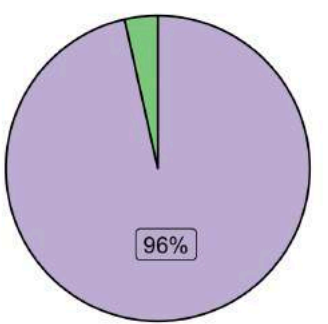
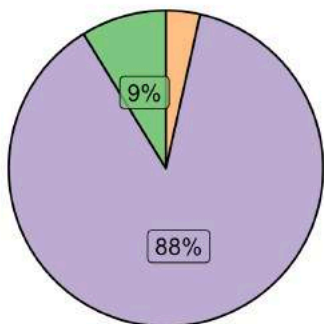
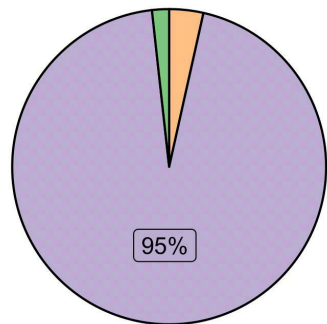
自然度評価調査のみ参加：3人

- **読み上げ調査**：各単語につき、一回発声させ、その音声から付与したアクセントを同定。
- **聞き取り調査**：可能な全てのアクセント型で、音声提示。各アクセントについて自然性を評価。
- 読み上げ時に特定のアクセント型が極めて優勢であっても、聞き取り時には、それ以外のアクセントに対しても自然性=0とならず、アクセントのバリエーションの分析が可能

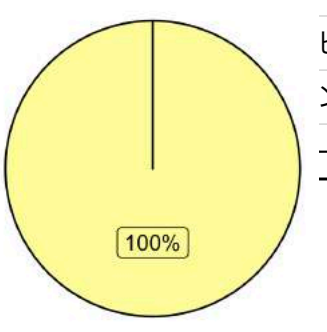
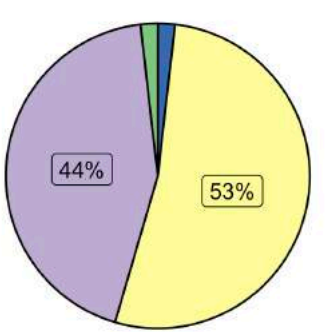
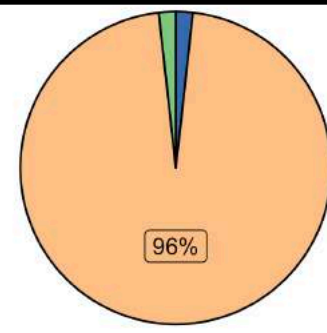
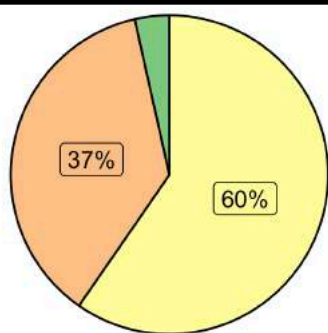
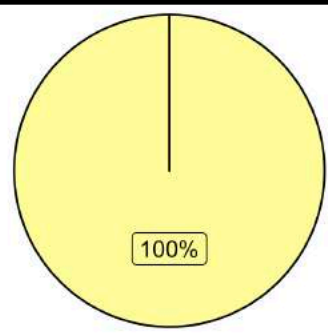
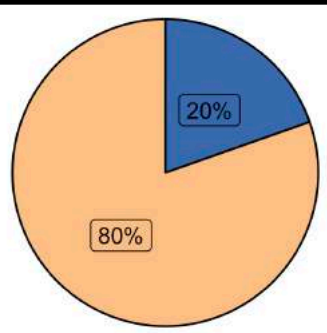
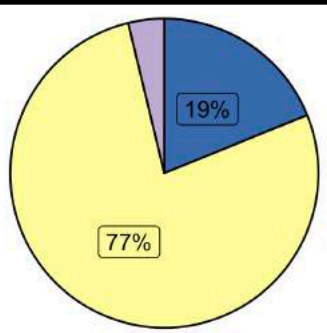
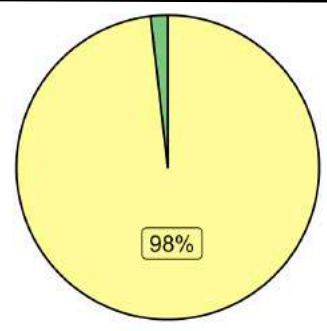
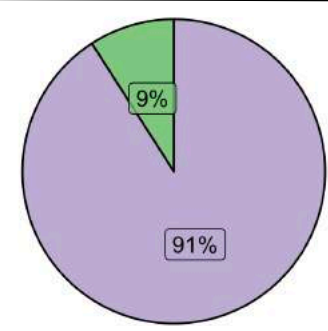
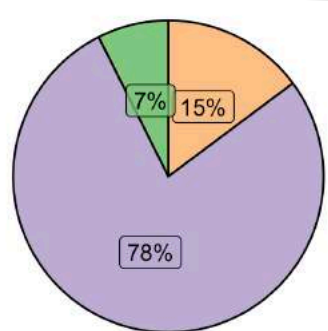
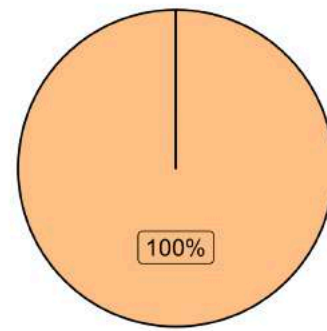
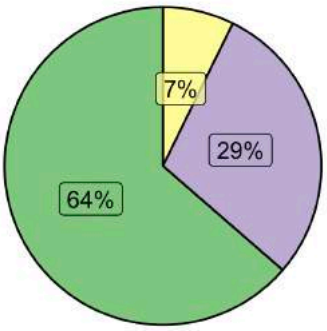
例：これは、ペハヨピモです。

- 0 型      これは、ペハヨピモ<sup>◦</sup>です。
- 1 型      これは、<sup>◦</sup>ペハヨピモです。
- 2 型      これは、ペ<sup>◦</sup>ハヨピモです。
- 3 型      これは、ペハ<sup>◦</sup>ヨピモです。
- 4 型      これは、ペハヨ<sup>◦</sup>ピモです。
- 5 型      これは、ペハヨピ<sup>◦</sup>モです。





9



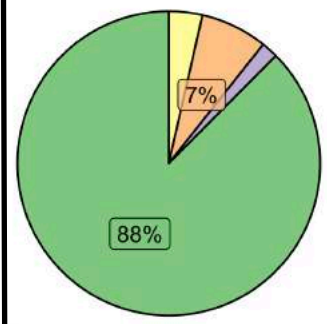
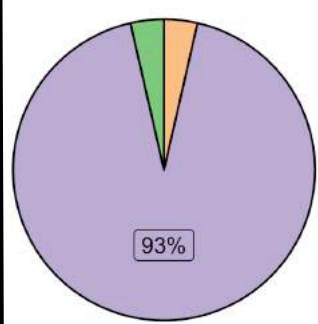
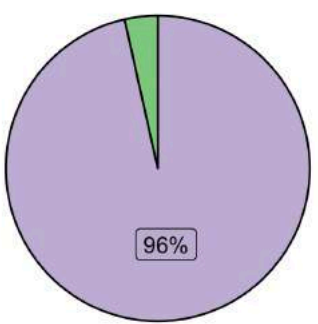
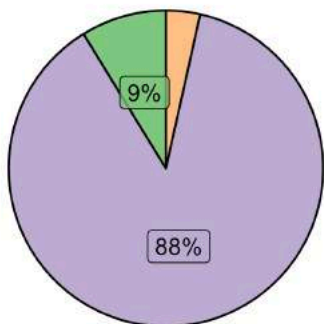
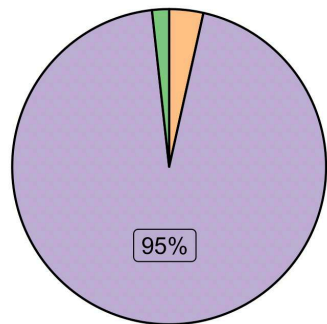
3の規則に基づくアサモシト型が優勢だが、別の型が利用されることもある

Category	Percentage
ハーバンボ	~1%
ヘンヌーシ	~1%
ターシーベ	~1%
ビンサーデ	~1%
ヒンパーゼ	~1%
タンチーマ	~1%
チンミンネ	~1%
ソソビツ	~1%
ドーベズン	~1%
アサモシト	~1%
メクム	~1%
ムネモネ	~1%
テトズ	~1%
ドーラケン	~1%
ヨモンゴ	~1%
ゲワール	~1%
ペンタン	~1%
オグ	~1%
サゴ	~1%
ヨリンビ	~1%
エターモン	~1%

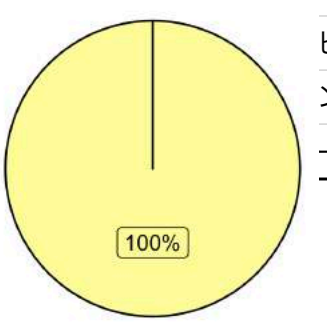
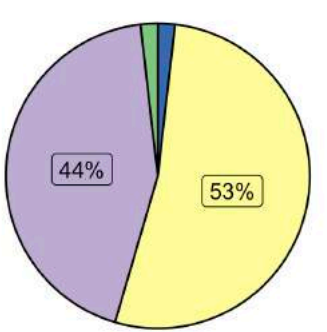
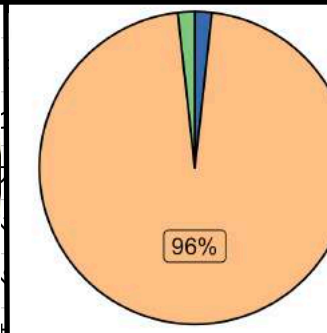
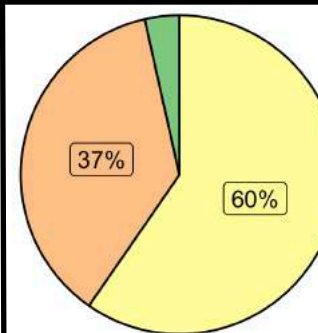
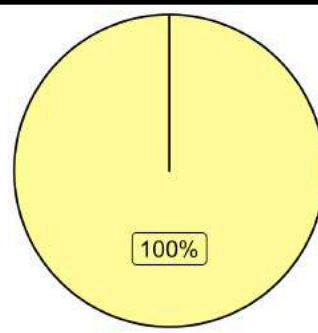
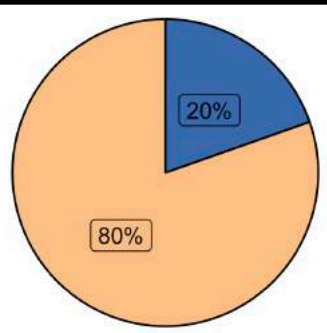
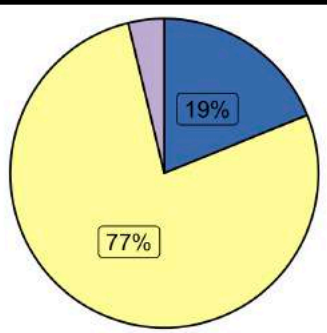
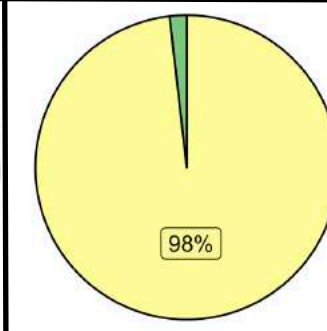
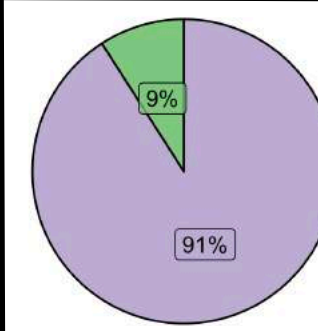
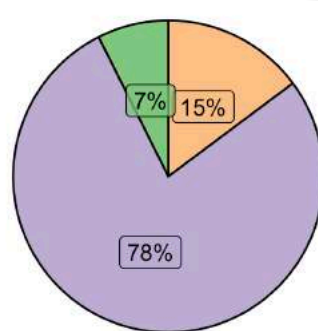
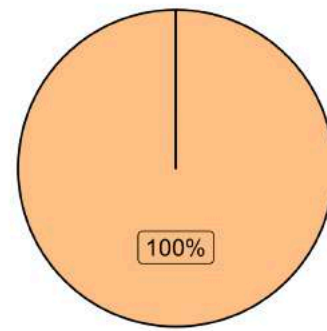
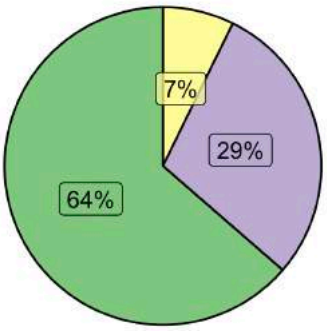
## 本研究の目的

- 外国人学習者の単語アクセント感覚の定着を調査したい
- 外国人の日本語学習者を対象として、同じような読み上げ調査を行う
- 単語グループごとに、母語話者のアクセント感覚を統計的にモデル化
- 単語グループごとに、学習者のアクセント感覚を統計的にモデル化
- 母語話者と学習者の間で、アクセント感覚の乖離が大きいグループを特定
- 母語話者と特定の学習者の間で、アクセント感覚の乖離が大きいグループを特定

# 母語話者のアクセント感覚の統計的モデル化

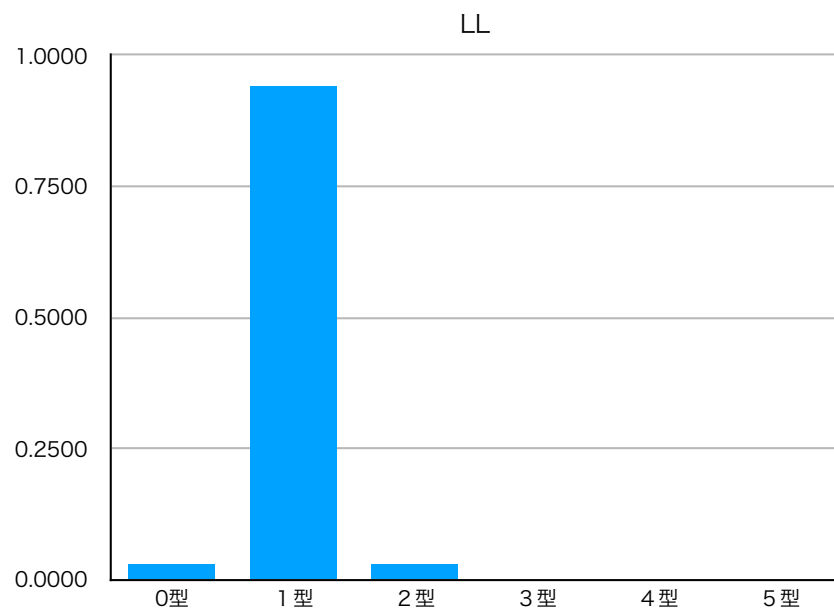
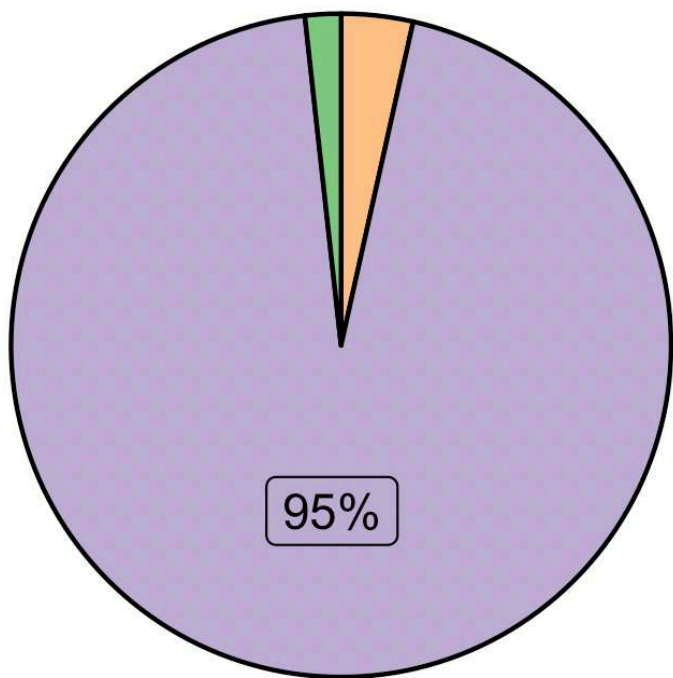


9

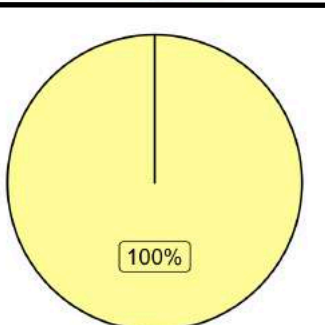
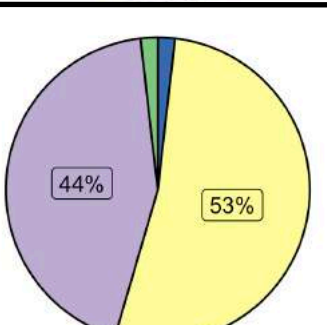
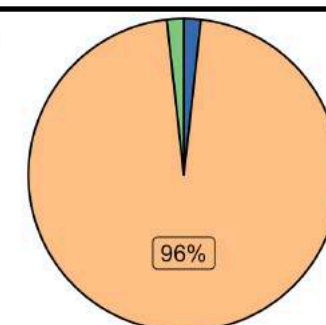
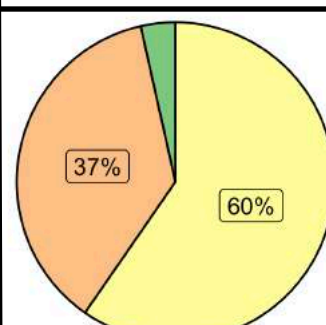
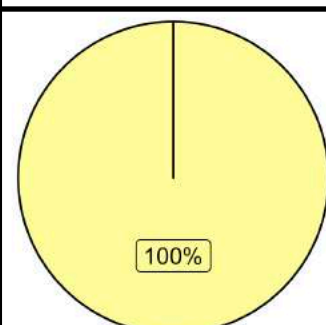
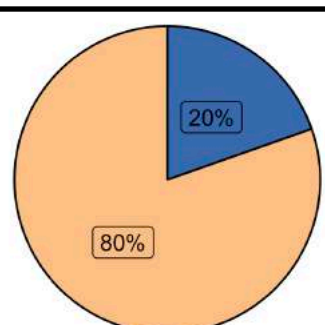
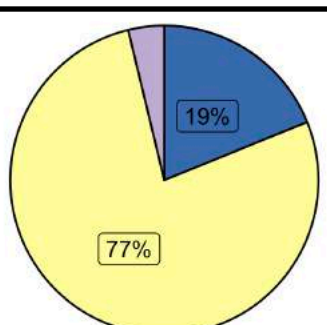
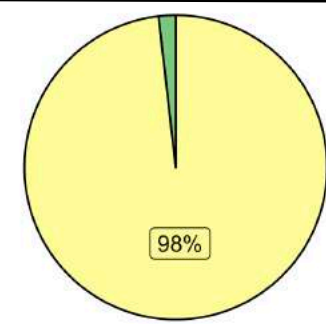
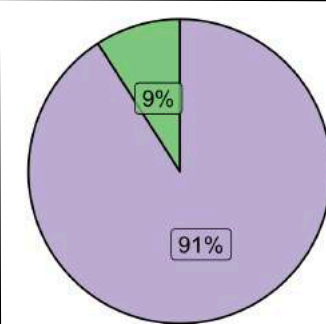
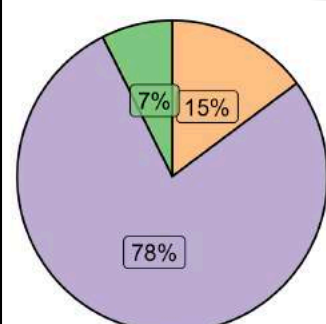
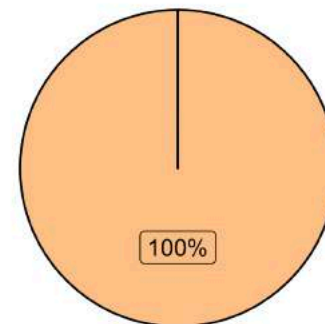
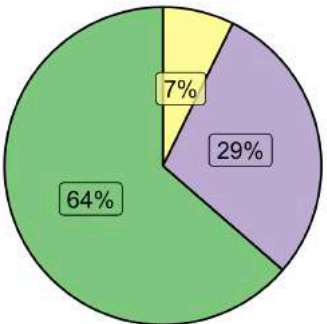
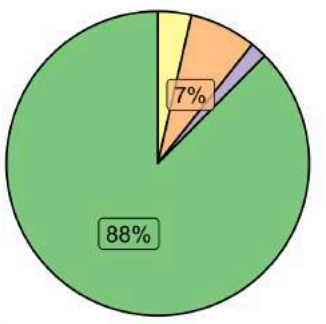
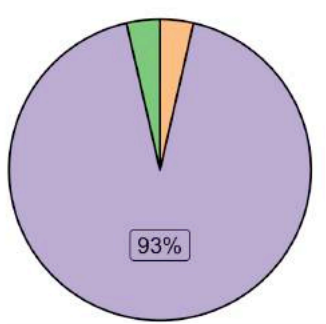
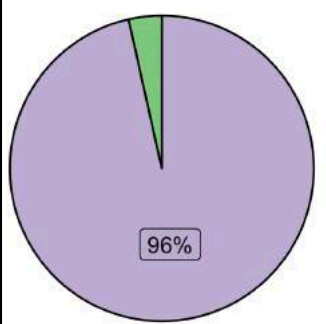
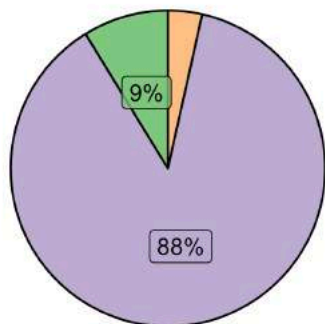
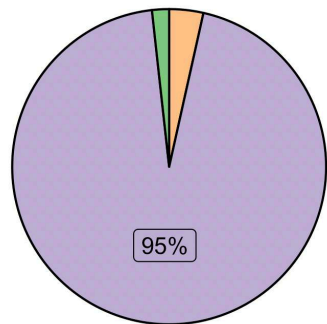


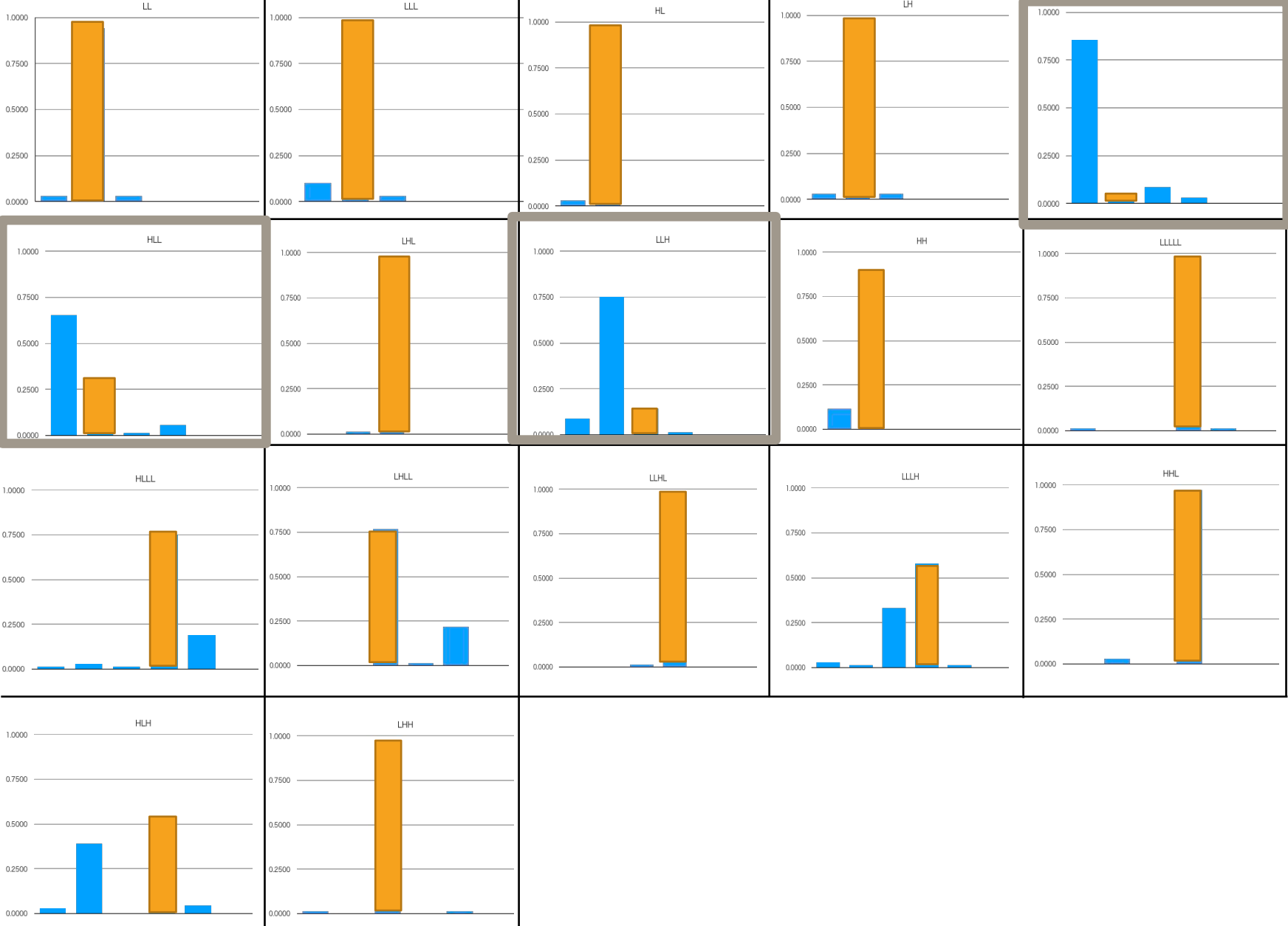
ビ	ハーバンボ	ヘンヌーシ	ターシーベ	ビンサーデ	ヒンパーゼ	タンチーマ	チンミンネ
ン	ソソビツ	ドーペズン	ビンヘゾー	アンメクー	ムーモネン	テーレズ	ブーラケン
ー	ヨモンゴ	ゲワーリン	ピノータン	オグーノン	サゴントー	ヨリンビー	エターモン

ザミ  
パニ  
ンマ  
1モー  
クゾモ  
ーネザ  
ルンヒ  
クノン  
ンジー  
ヨゼホ  
パザヨ  
ンザテ  
ソソノ  
モチン

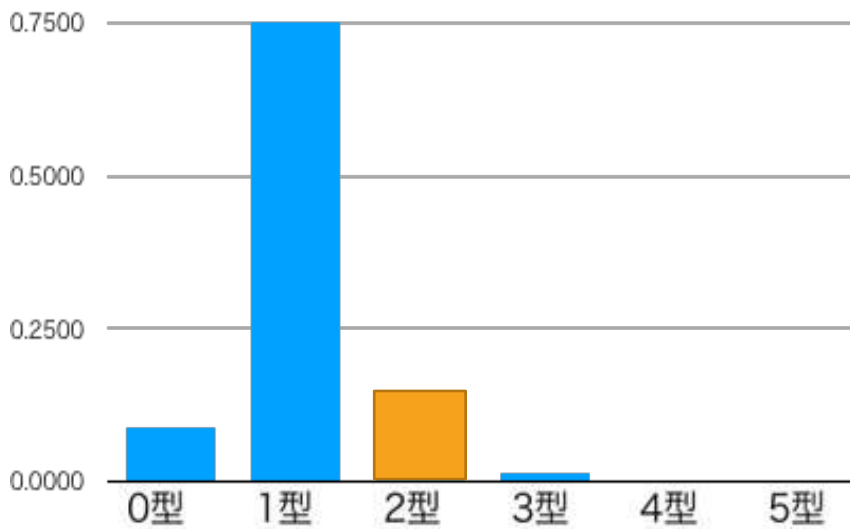
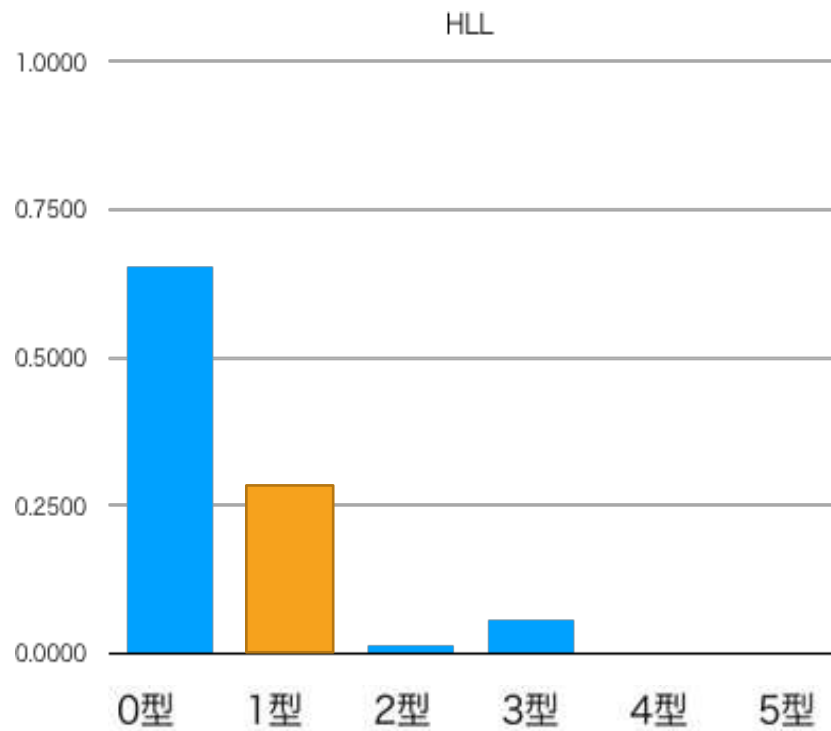
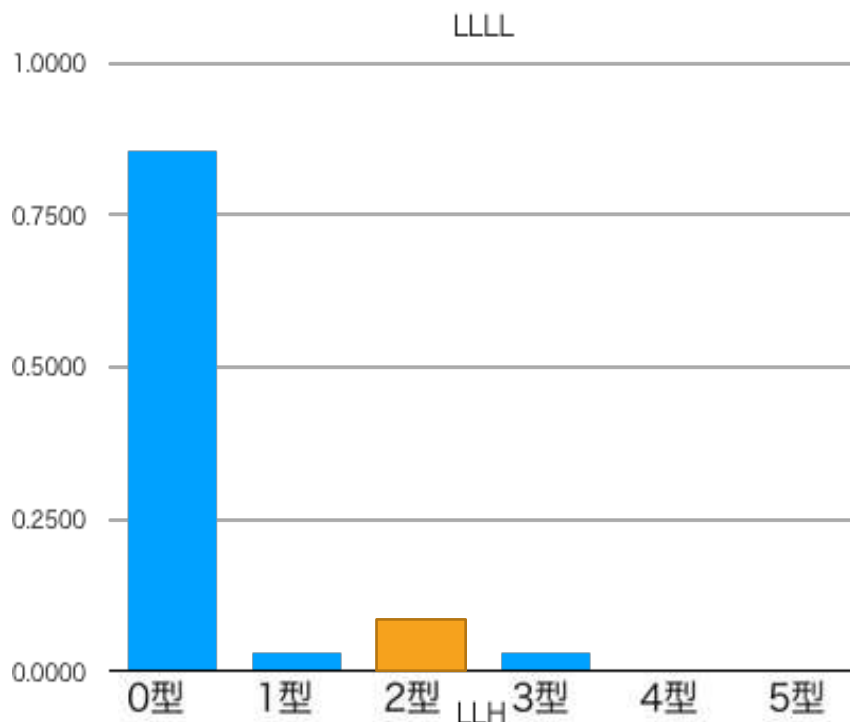


定量化





縦軸：確率



-3の規則に基づくアクセント型が優勢だが、別の型が利用されることもある



# 中国人学習者を対象としたアクセント付与実験

- **調査対象**

全員N1能力試験に合格

- **人数**

15人

- **教科書**

『総合日本語』（大連理工大学出版社）

# 架空外来語リスト

モーラ数	音節構造 (17種)	1	2	3	4	5	6	7	8	9
2	LL	ペザ	コデ	ギパ	ジテ	ピゾ	ガゾ	パキ	ゴヨ	ザミ
3	LLL	ピゴピ	パヘミ	ナザエ	メネピ	ゾゲマ	ロデセ	オデハ	ミポメ	ペパニ
	HL	ジーハ	ヒーポ	ローベ	デーベ	ポーマ	ボンメ	ドンサ	ノンシ	ワンマ
	LH	マベン	レポン	ゲムン	ジダー	ミダン	ゾボー	コガー	ジルー	ロモー
4	LLLL	ワネポシ	ピネテニ	ラポネソ	ピギチネ	メヨゼソ	マガデオ	ピニギノ	セネヘゾ	テタゾモ
	HLL	デーヘモ	ナービマ	ベンタポ	コンレピ	ローゲワ	ベーパノ	アンロピ	ムンベハ	パーネザ
	LHL	ヨダーピ	メボンヨ	モゾーハ	ゾモーボ	ミゾンナ	ミアンゾ	ジラーミ	キナーメ	ソルンヒ
	LLH	コポレー	ボミナー	テポマー	ミハナン	タポソン	ゾテルー	リペレン	ネテシー	セネノン
	HH	ゾンエー	サーペー	ミードン	グーレン	ザーセン	ズンロー	ギーレン	ランボン	トンジー
5	LLLLL	ペハヨピモ	セソポメヨ	エゼガヘオ	ヤデピギノ	ウゼテベマ	ミジナペシ	ビヒペミボ	ヤテソヘピ	ウパヨゼホ
	HLLL	カーシポヤ	ドンラゼコ	ヘンタゾネ	ボンセポニ	スンゾナポ	バーマバヨ	ウージモザ	ラーゼワメ	スンパザヨ
	LHLL	ニスンパガ	ザポーモベ	テルンポハ	タズーラカ	チマーペヒ	カザーパセ	ベラーゴラ	ガパーピワ	パドンザテ
	LLHL	マピガンエ	オデヌンザ	レゾガーメ	ピメモーナ	ビハポーレ	ベハワーモ	ペサギンモ	ゼポネーベ	レデソンノ
	LLH	ガテピゼン	セネハヨン	ポダザユー	ロレボシー	オメナズー	ゴラソガー	ポナパヒー	テザポユー	ネナモチン
	HHL	メージーモ	ダンツープ	ハーバンポ	ヘンヌーシ	ターシーベ	ビンサーデ	ヒンパーゼ	タンチーマ	チンミンネ
	HLH	ウーポエー	ローパルン	ソンビツー	ドーペズン	ビンヘゾー	アンメクー	ムーモネン	テーレズー	ブーラケン
	LHH	ラソープン	テモハー	ヨモンゴ	ゲワーリン	ピノータン	オグーノン	サゴントー	ヨリンビー	エターモン

既存の形態素による影響をできるだけ避けるため、N-gram（音節単位で2音節から）でかく調査語の音節連続を走査し、いずれも辞書（『日本語の語彙特性』（天野・近藤, 1999））に登録されていないことを確認した

- 学習者が意図したアクセント型と実際に音声として生成された（聴取者が知覚する）アクセント型が異なる場合がある
- 本研究は学習者が持つアクセントの暗黙知を対象としているので、読み上げではなく、アクセント型を直接答えさせる形とした

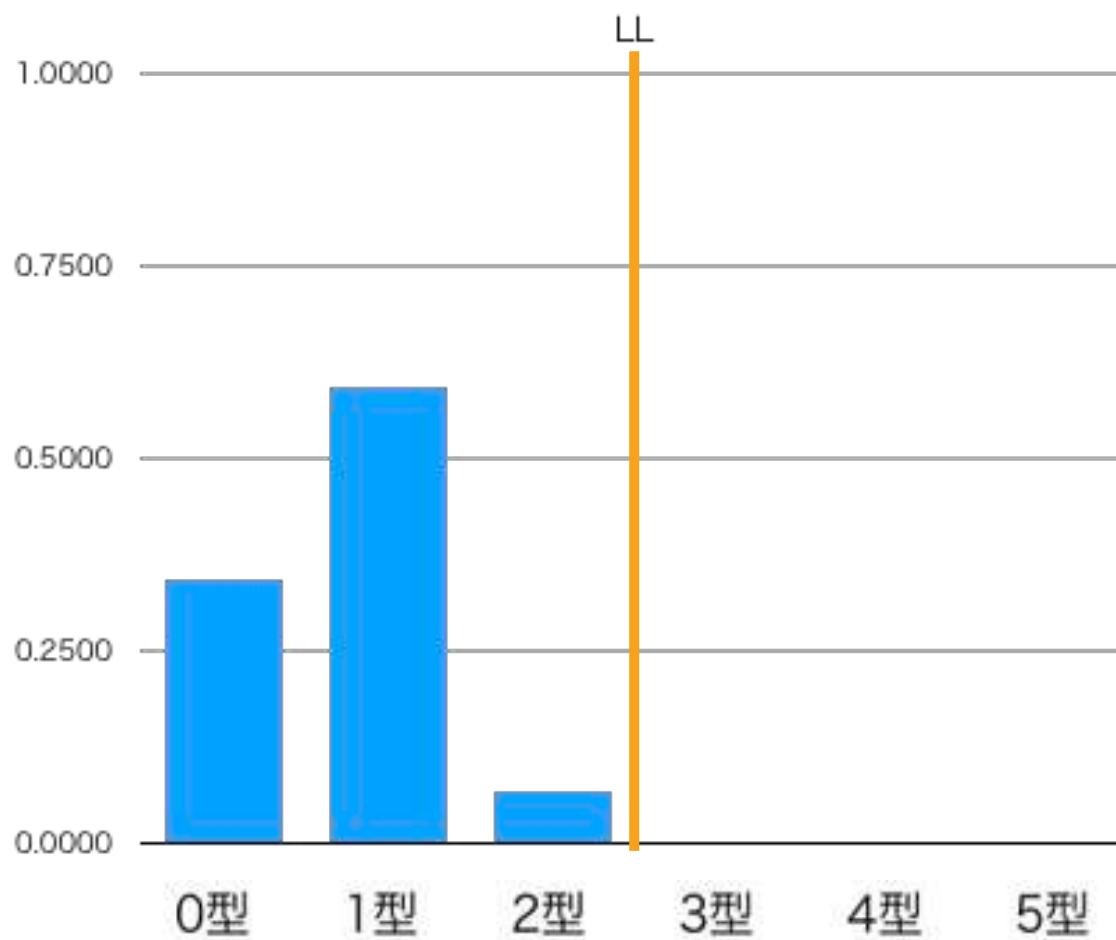
音節構造	架空外来語	被験者数	アクセント型_比率					
			0	1	2	3	4	5
LL	ペザ	15	0.13	0.8	0.07	0	0	0
	コデ	15	0.74	0.13	0.13	0	0	0
	ギバ	15	0.4	0.53	0.07	0	0	0
	ジテ	15	0.33	0.47	0.2	0	0	0
	ピゾ	15	0.27	0.73	0	0	0	0
	ガゾ	15	0.47	0.53	0	0	0	0
	パキ	15	0	1	0	0	0	0

⋮      ⋮      ⋮      ⋮      ⋮      ⋮      ⋮      ⋮

LHH	オグーノン	15	0.2	0	0.53	0.13	0.13	0
	サゴントー	15	0.47	0	0.47	0	0	0.06
	ヨリンビー	15	0.33	0	0.33	0.13	0.13	0.08
	エターモン	15	0.27	0	0.47	0.2	0.06	0

# 学習者のアクセント感覚

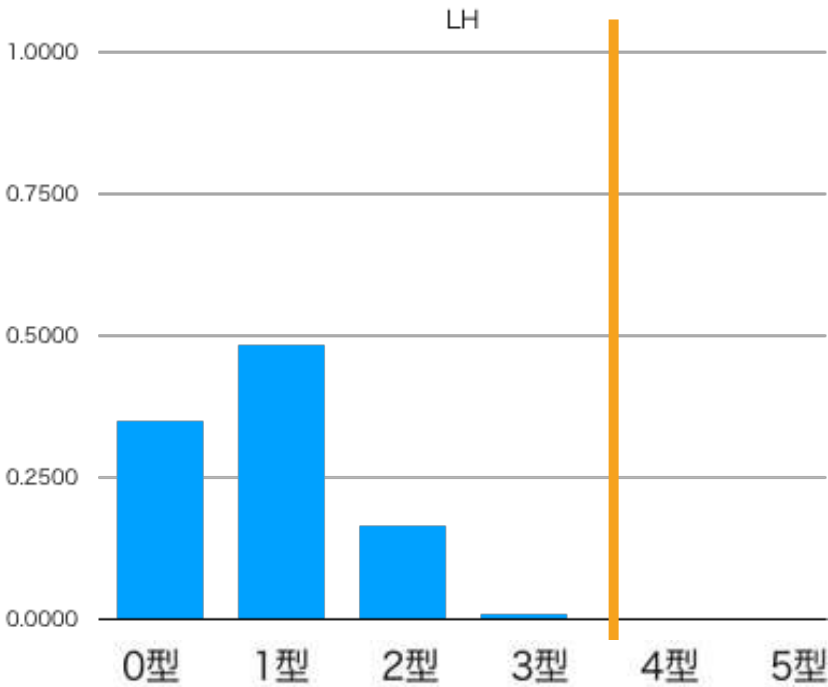
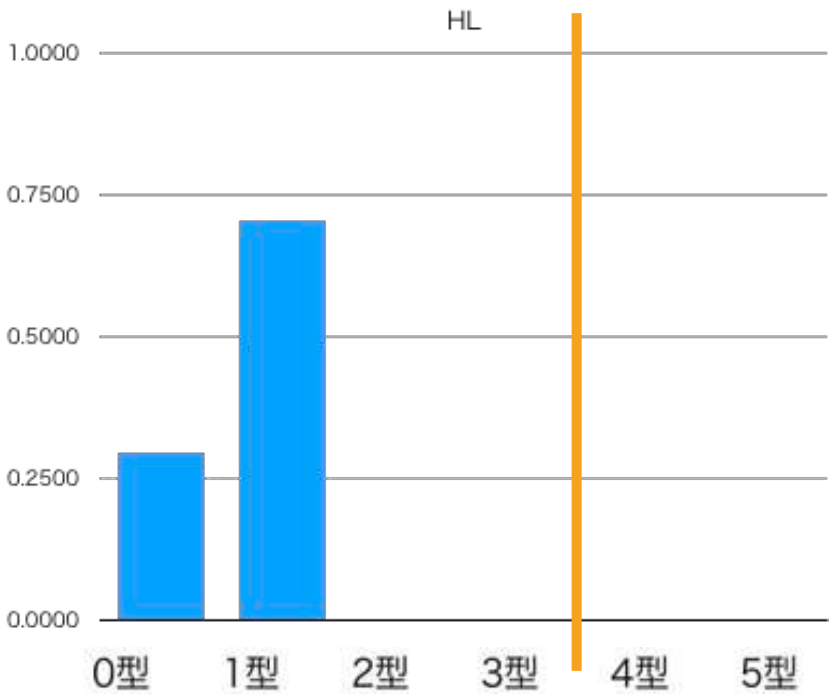
## 2モーラ LL



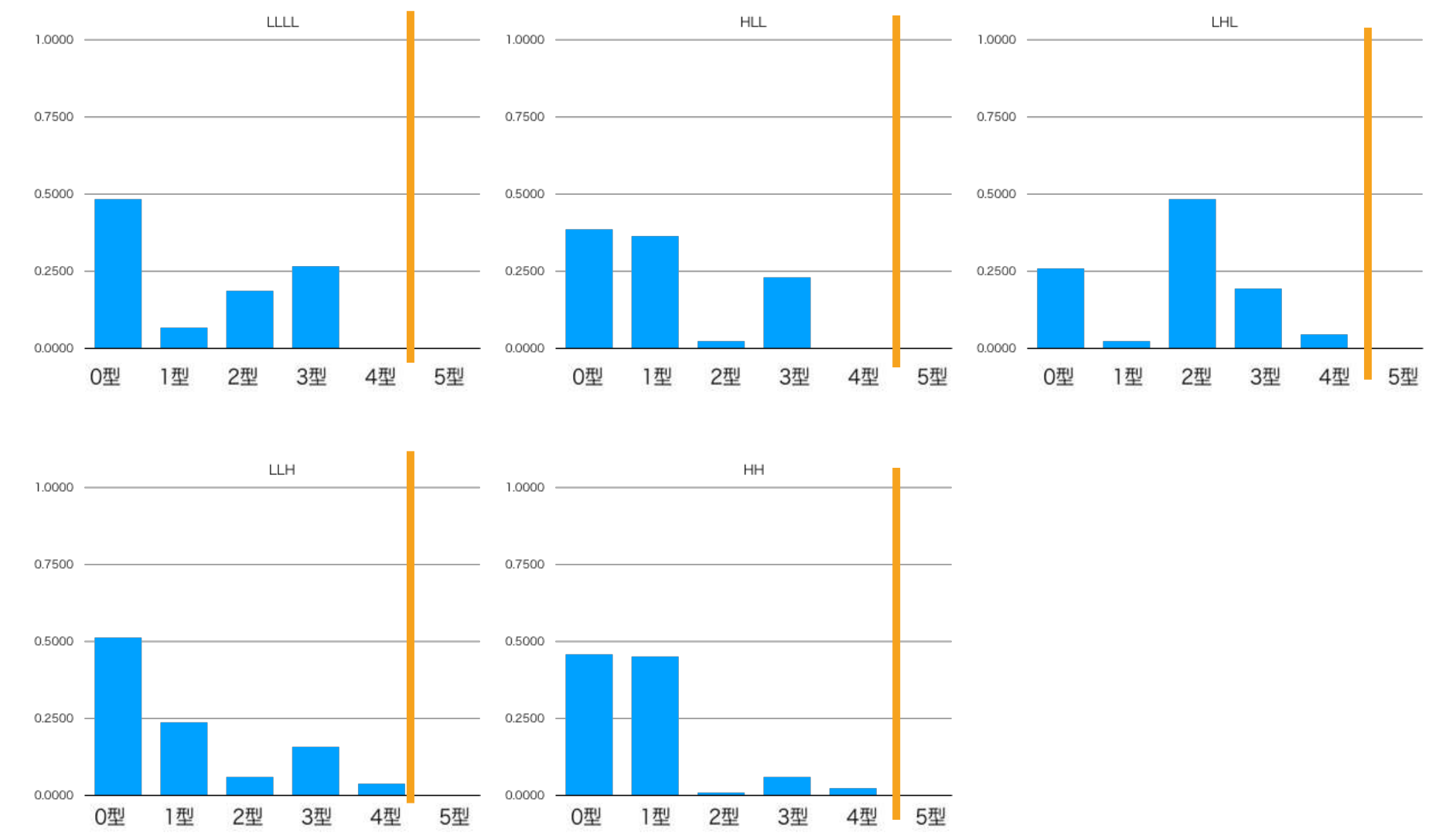
縦軸：確率

3 モーラ LLL HL LH

縦軸：確率

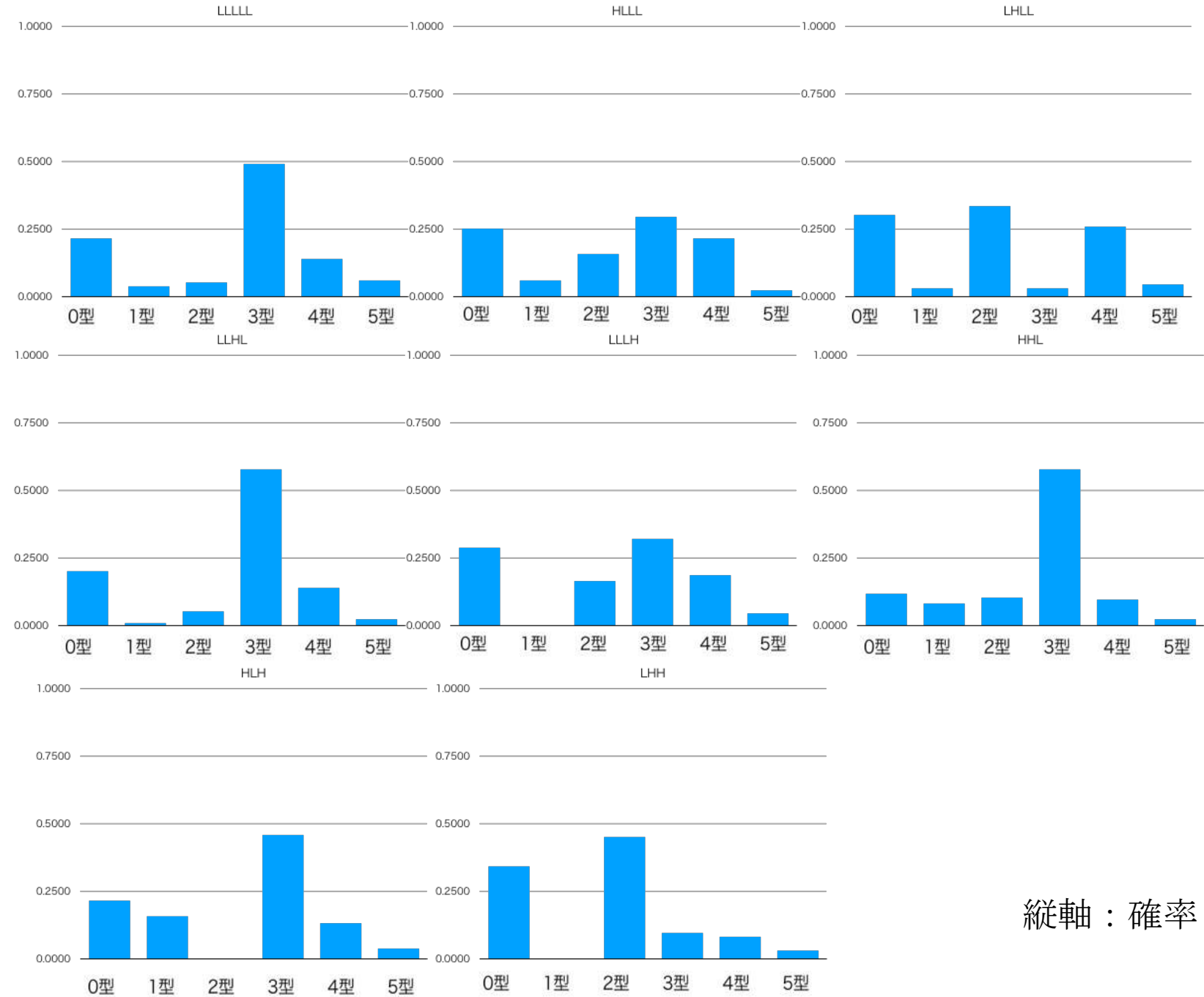


# 4 モーラ LLLL HLL LHL LLH HH



縦軸：確率

# 5モーラ LLLL HLLL LHLL LLHL LLLH LLLH HHL HLH LHH

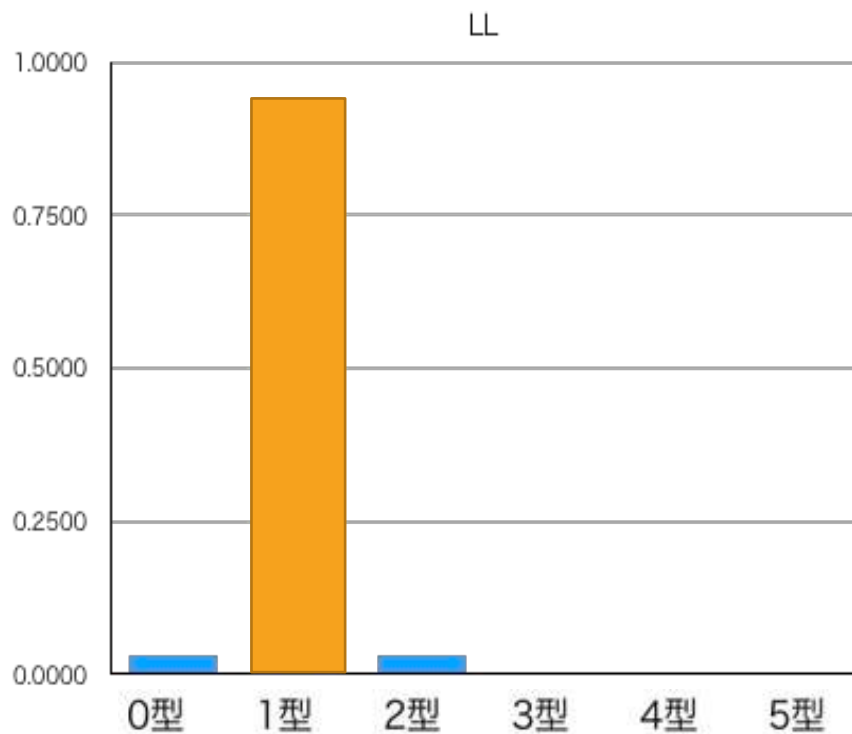


縦軸：確率

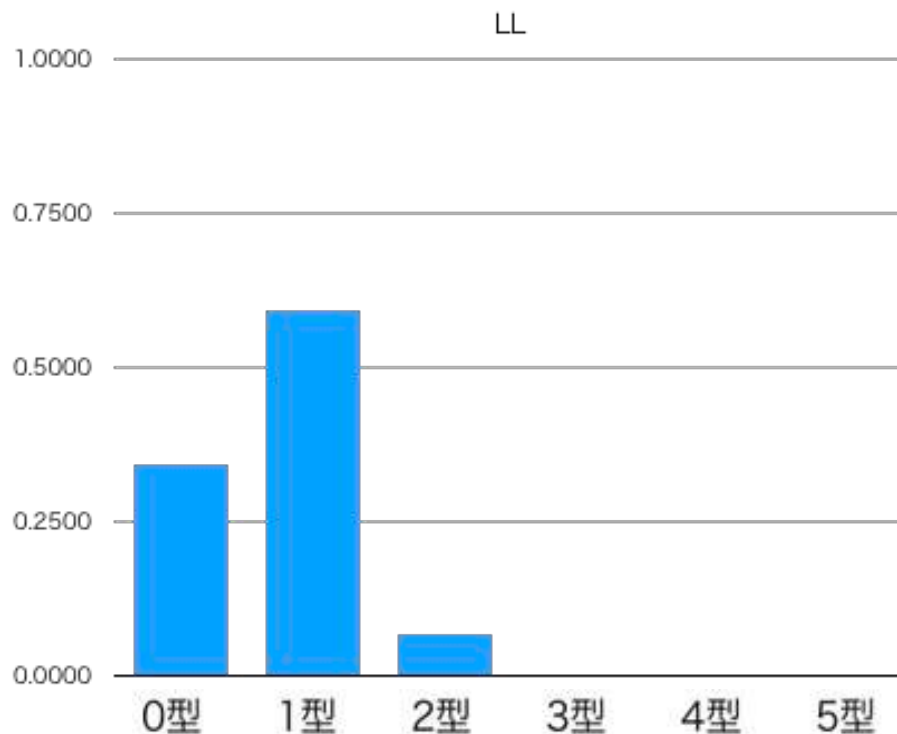


母語話者VS學習者

## 母語話者



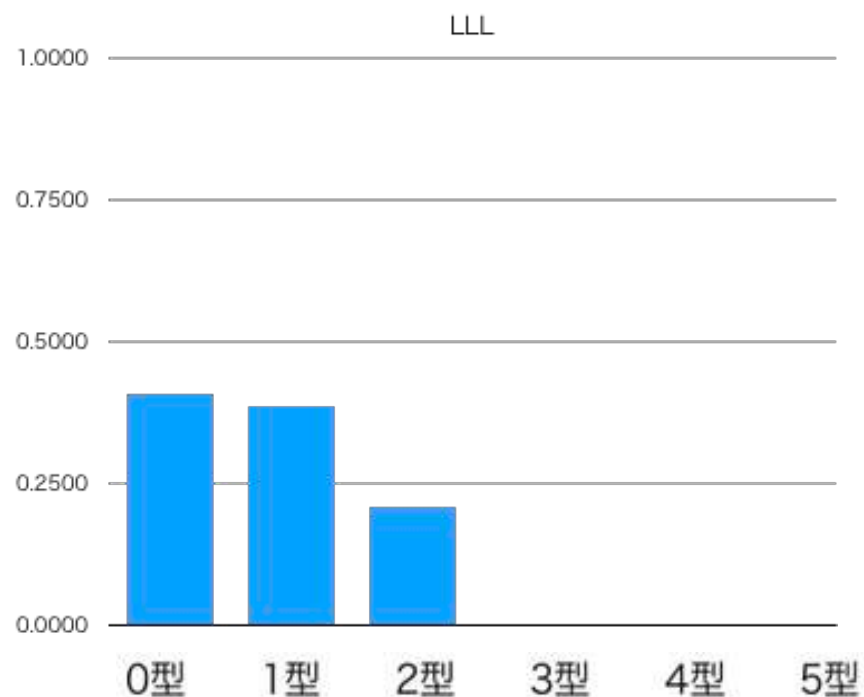
## 学習者



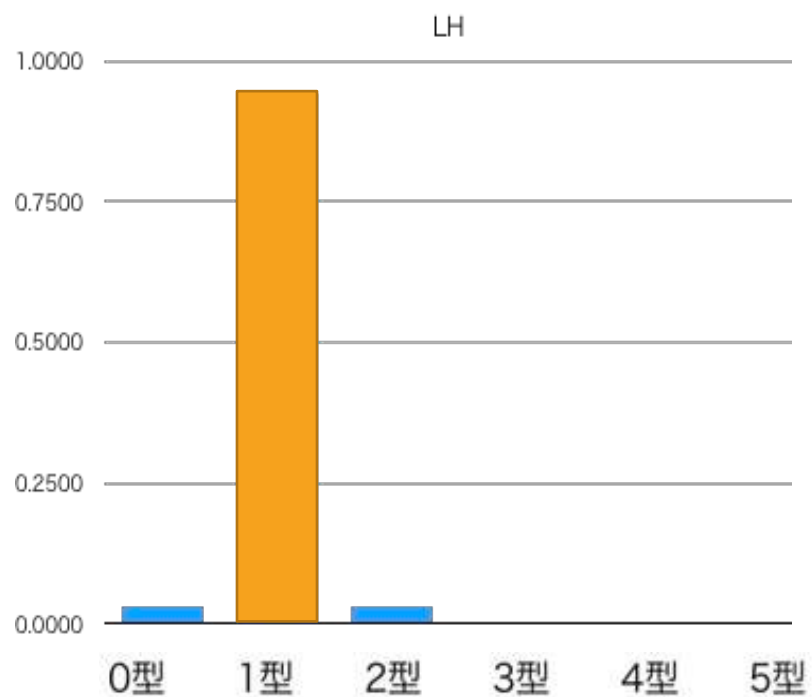
## 母語話者



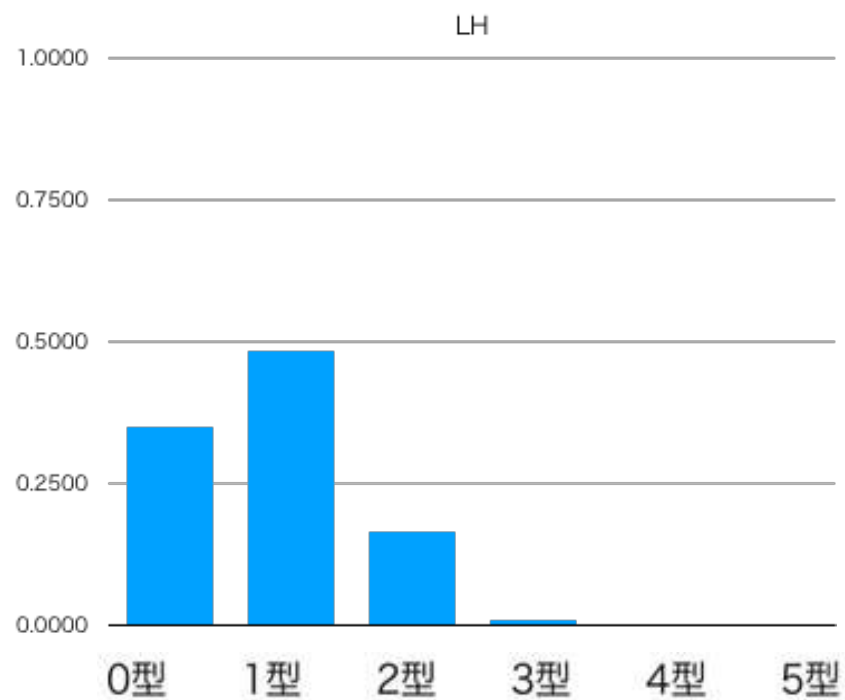
## 学習者



## 母語話者



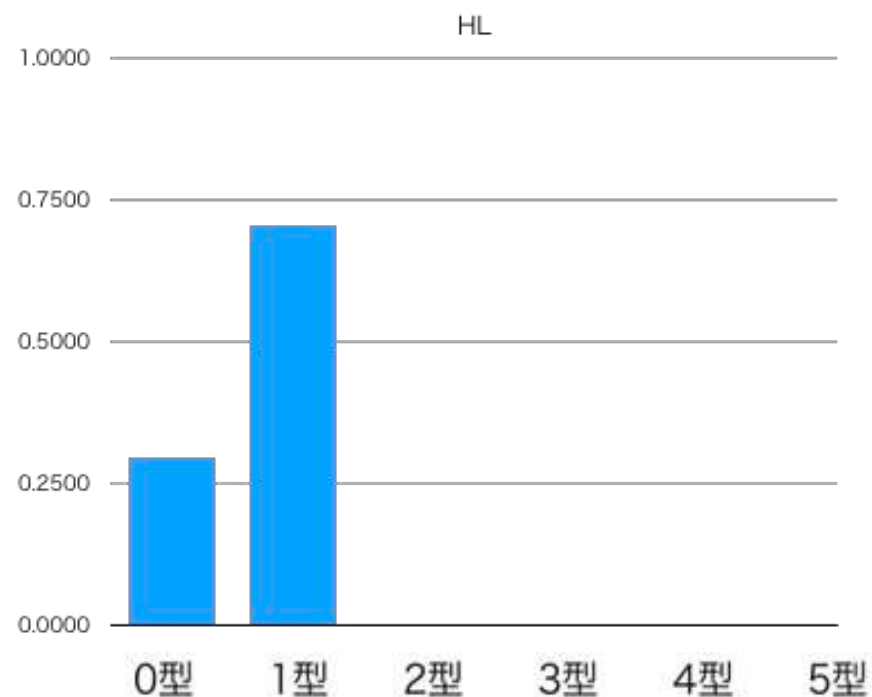
## 学習者



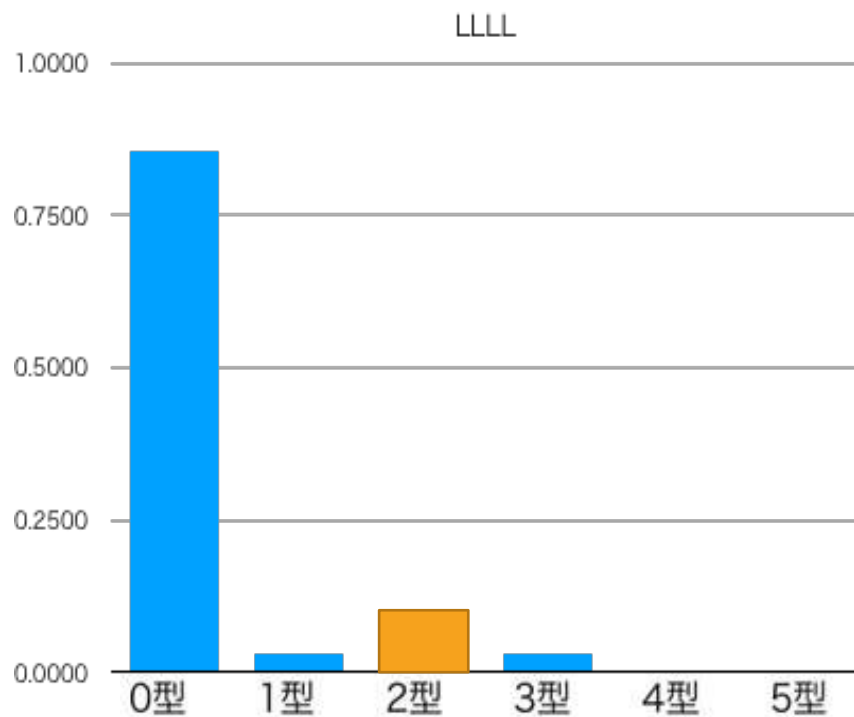
## 母語話者



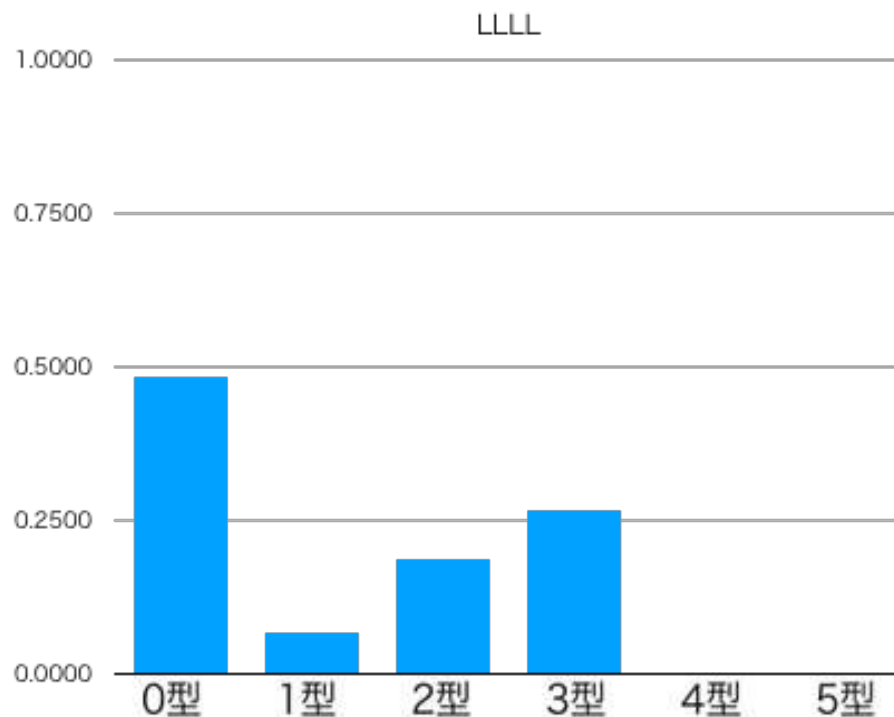
## 学習者



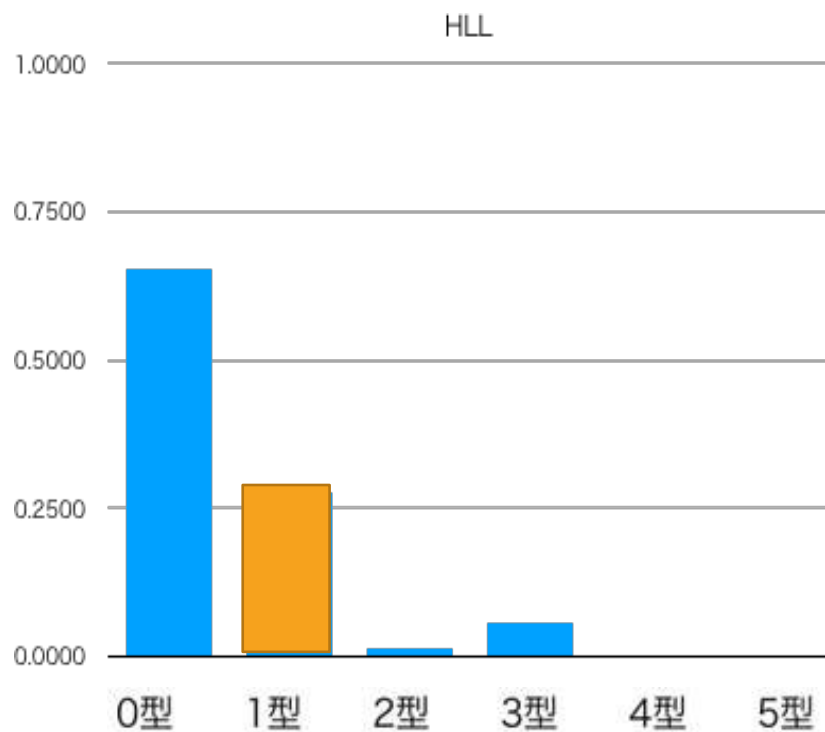
## 母語話者



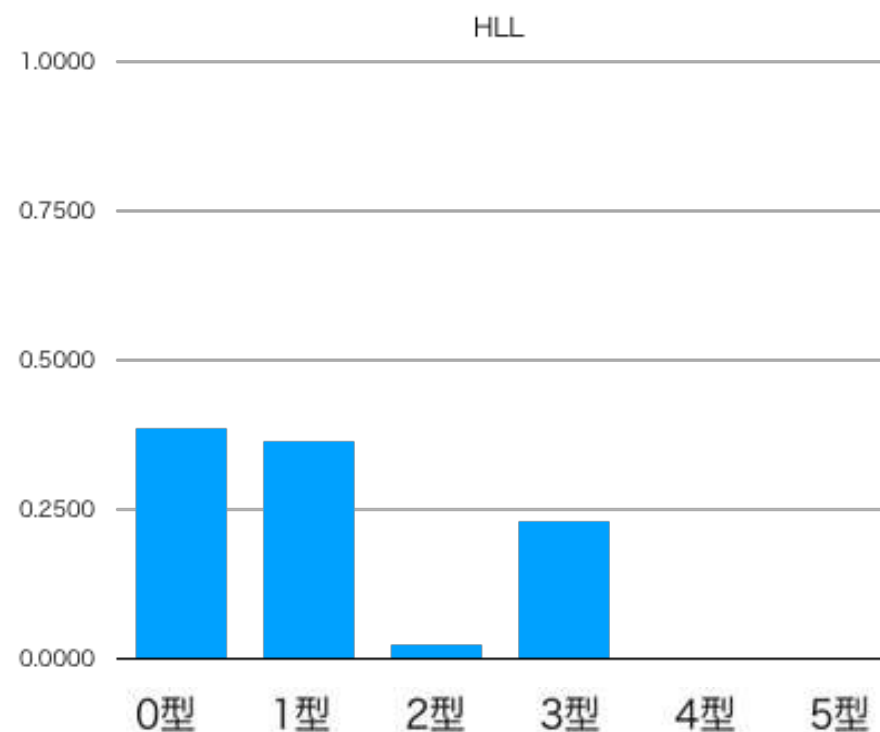
## 学習者



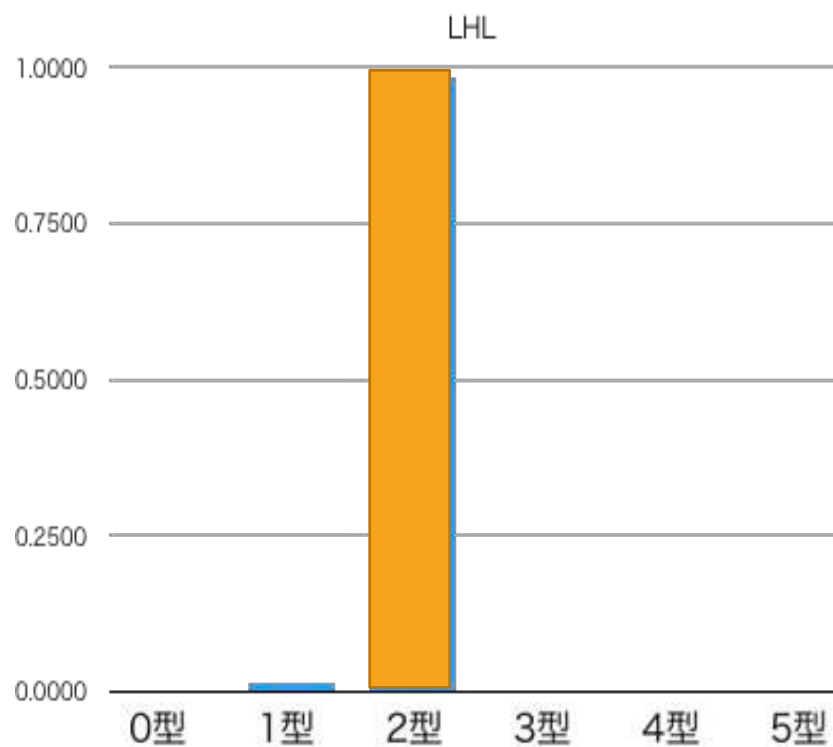
## 母語話者



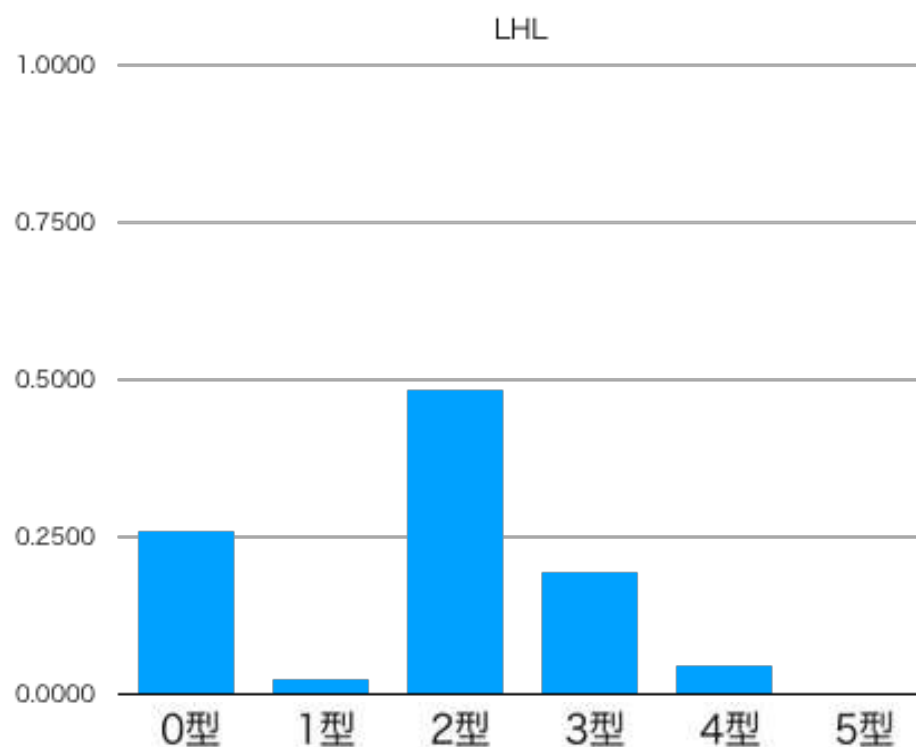
## 学習者



## 母語話者



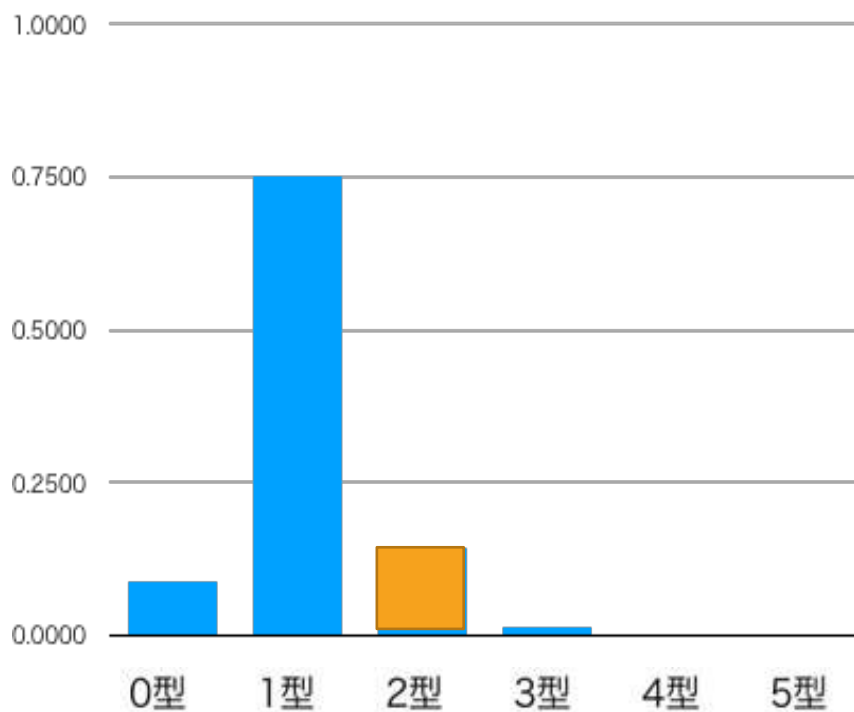
## 学習者





## 母語話者

LLH

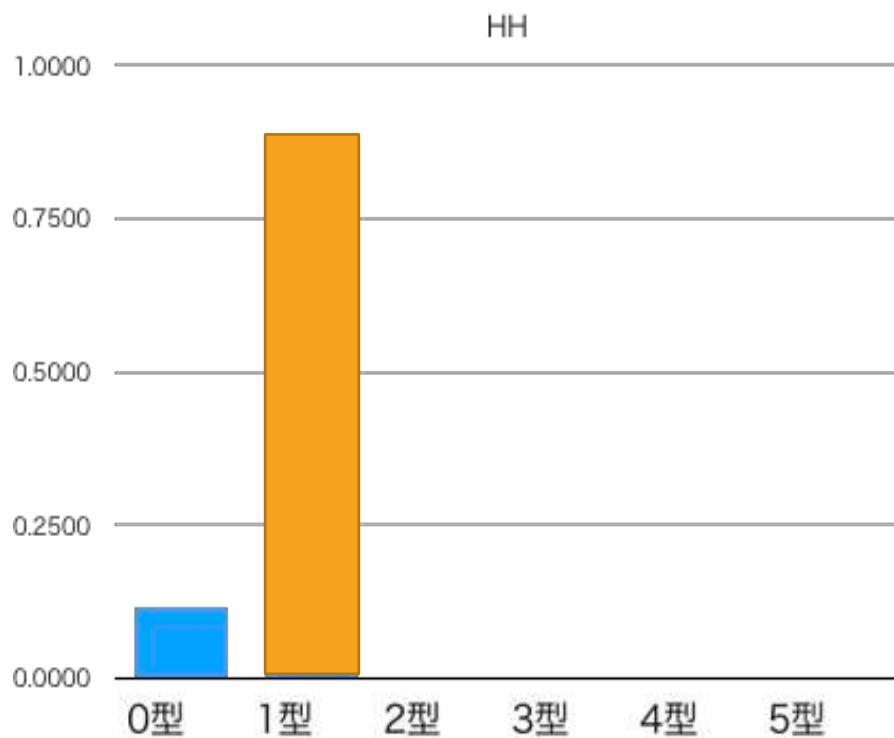


## 学習者

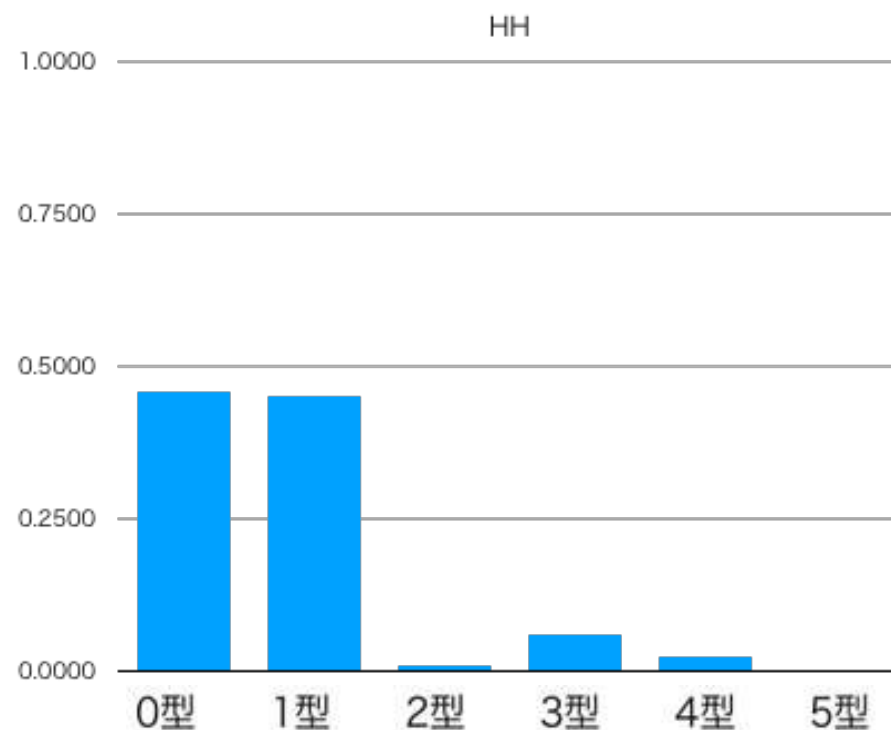
LLH



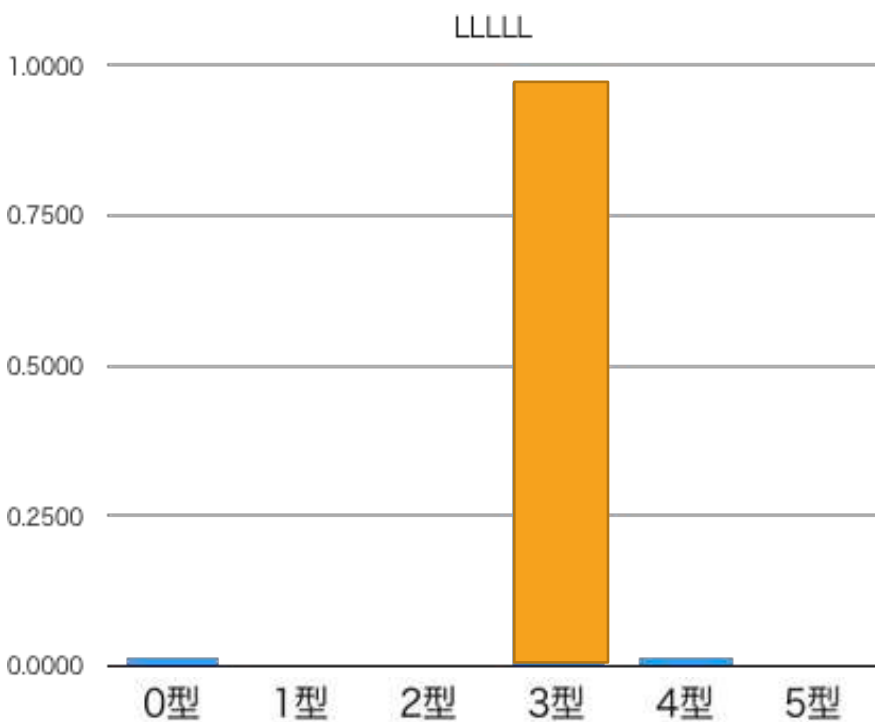
## 母語話者



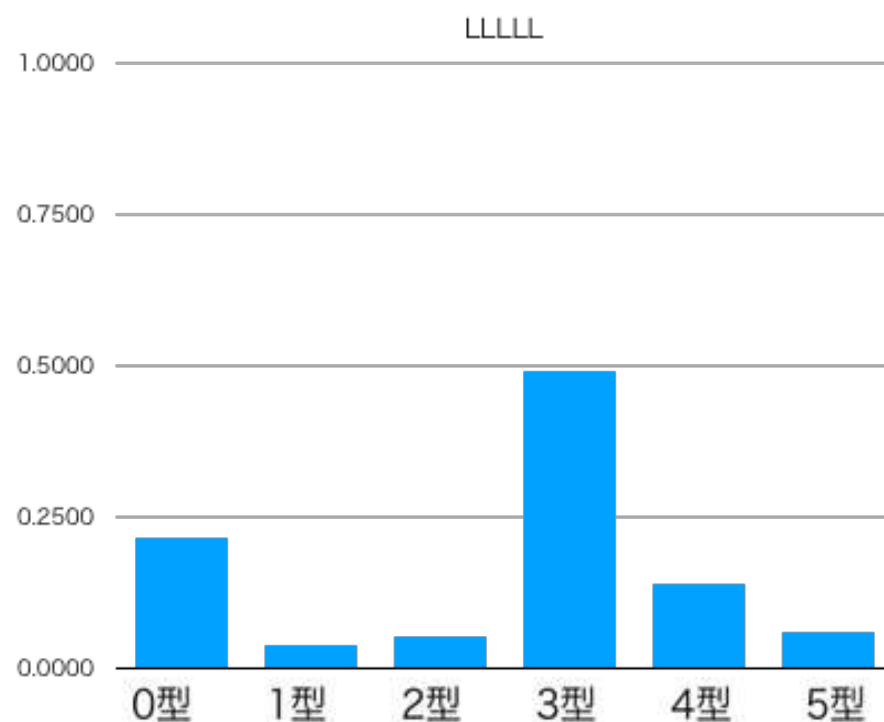
## 学習者



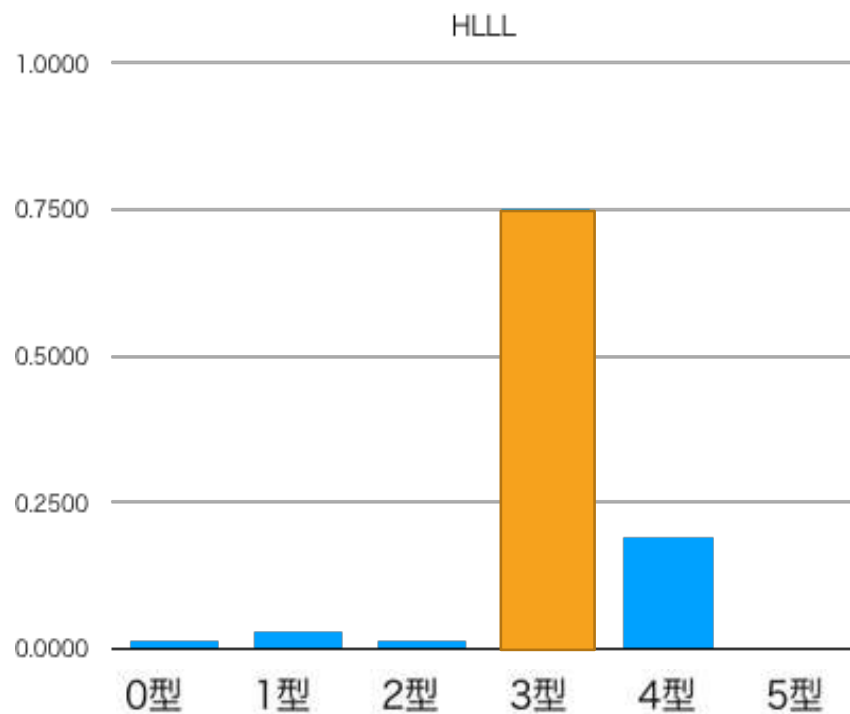
## 母語話者



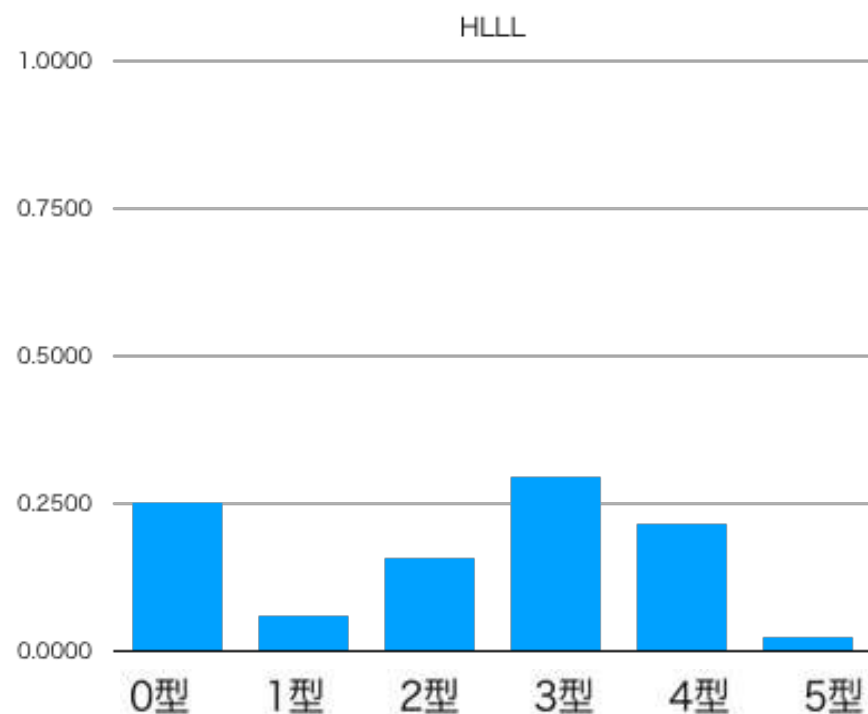
## 学習者



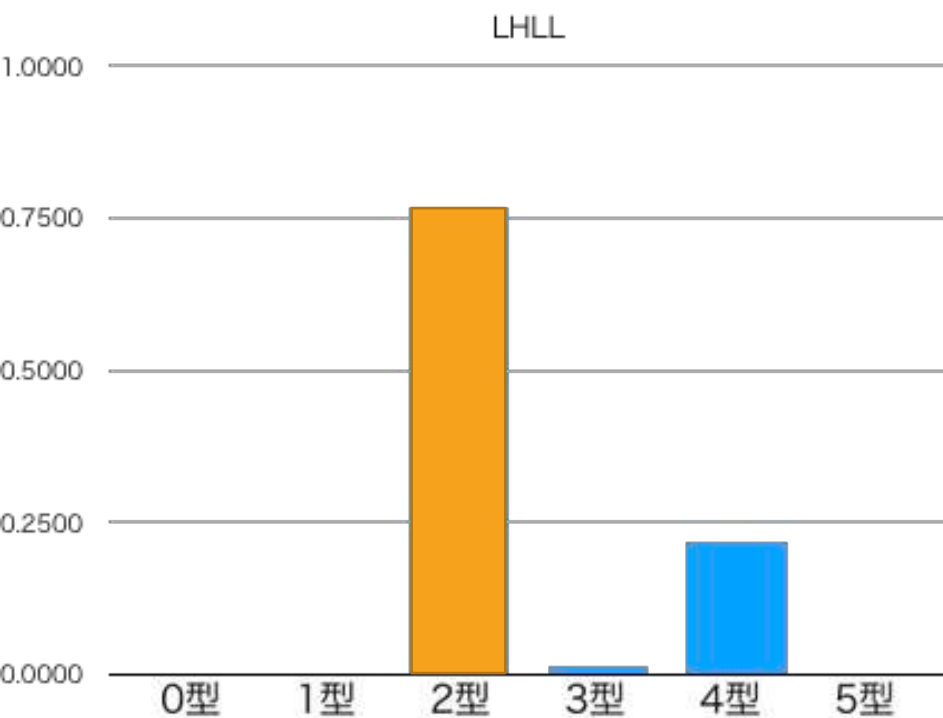
## 母語話者



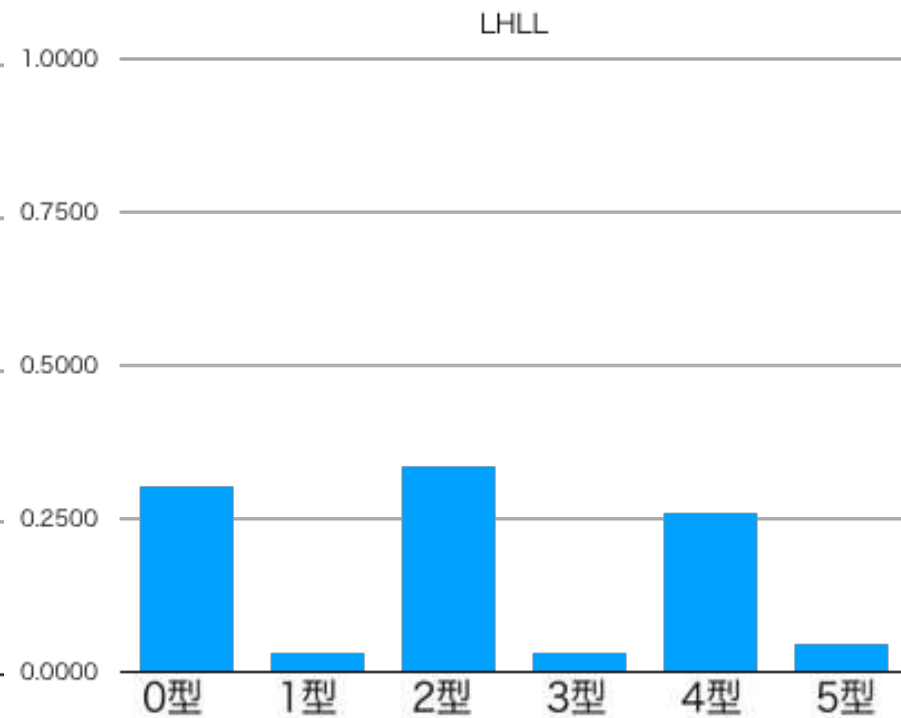
## 学習者



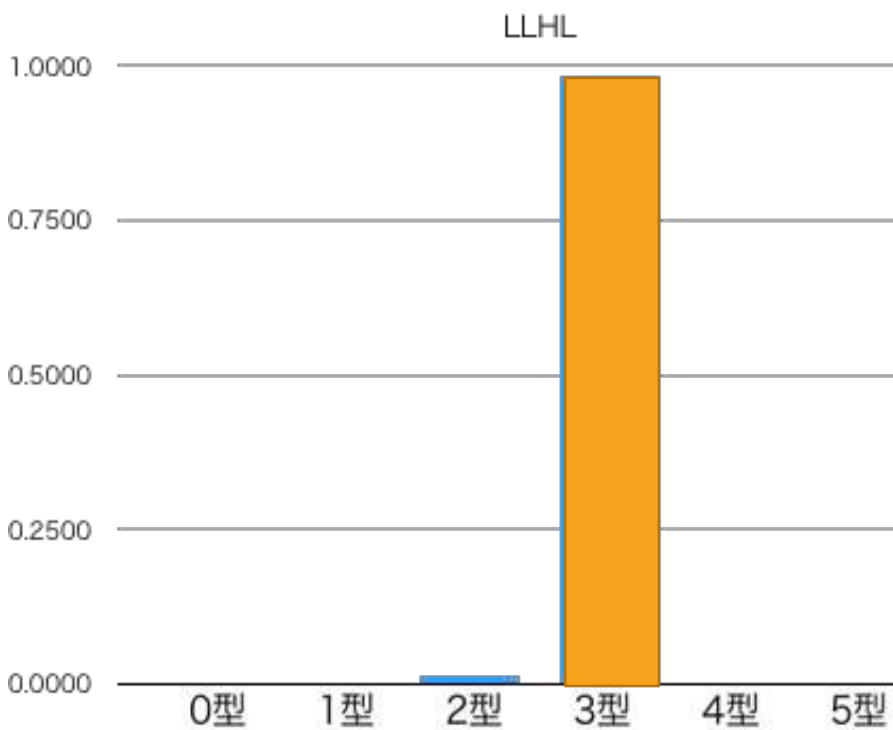
## 母語話者



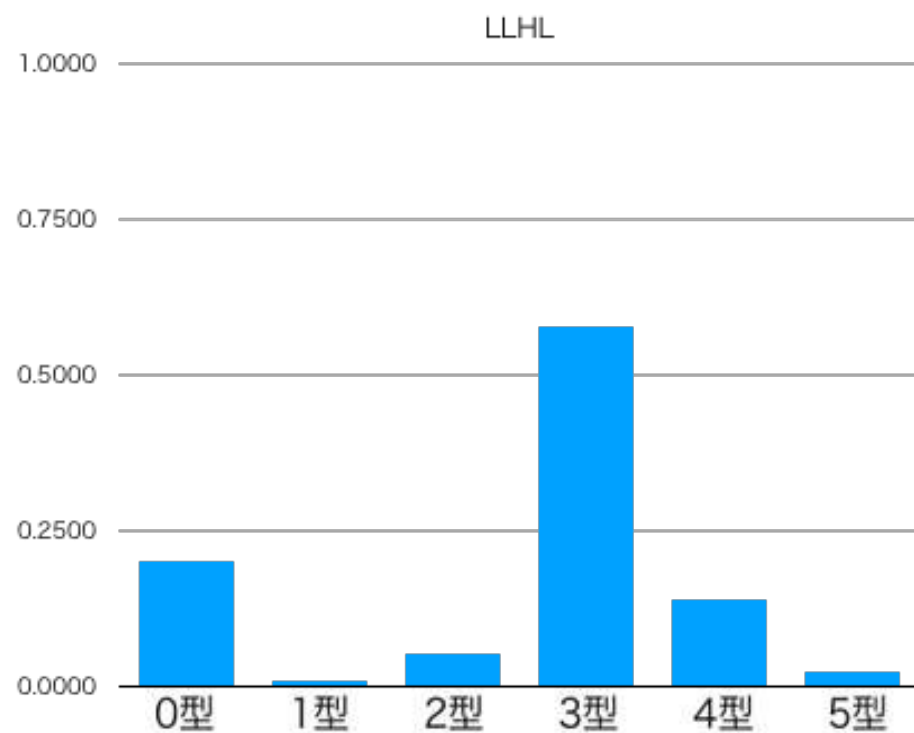
## 学習者



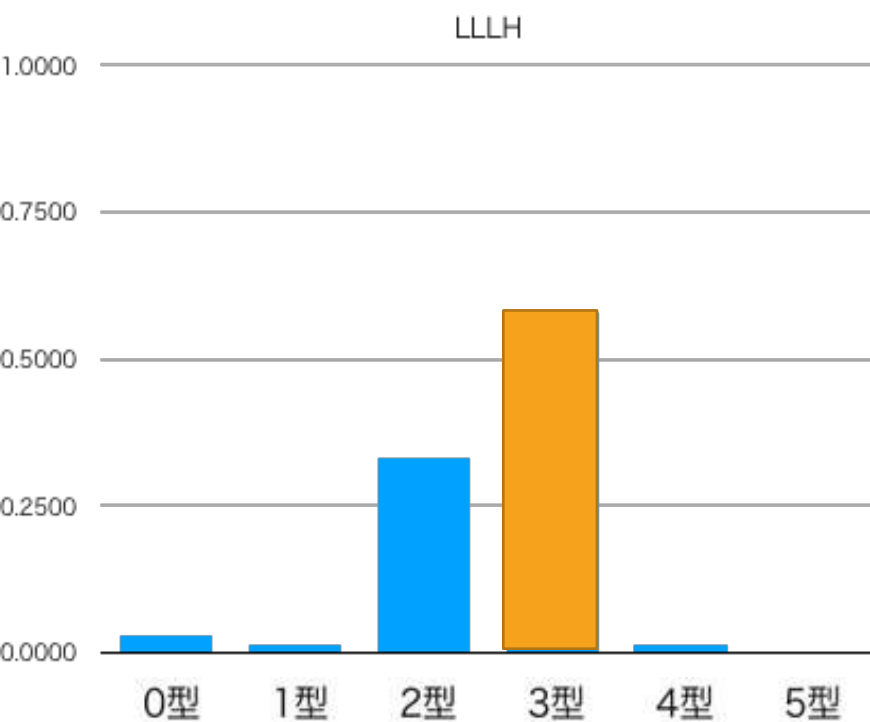
## 母語話者



## 学習者



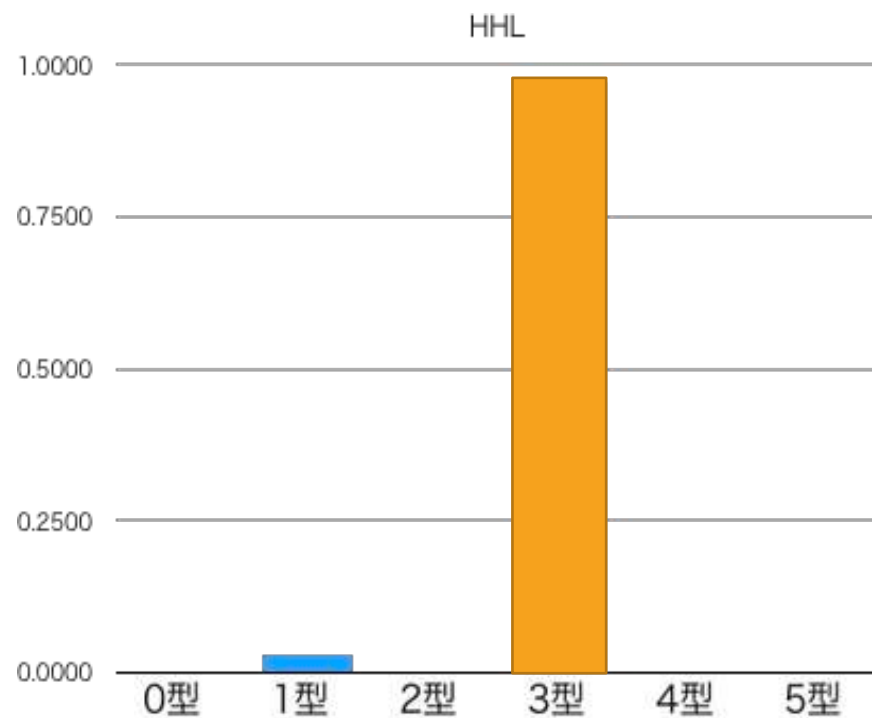
## 母語話者



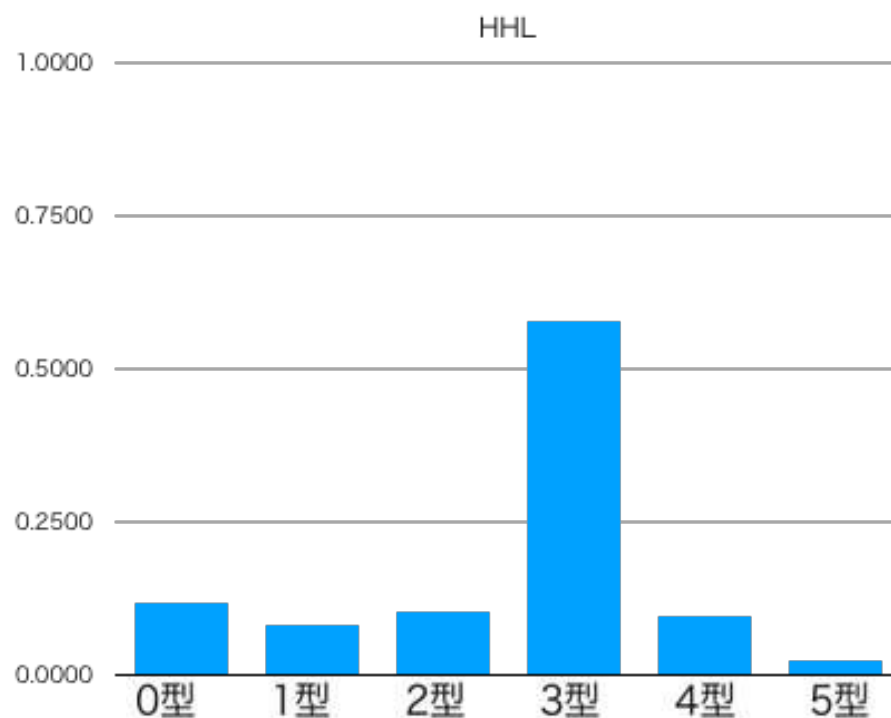
## 学習者



## 母語話者

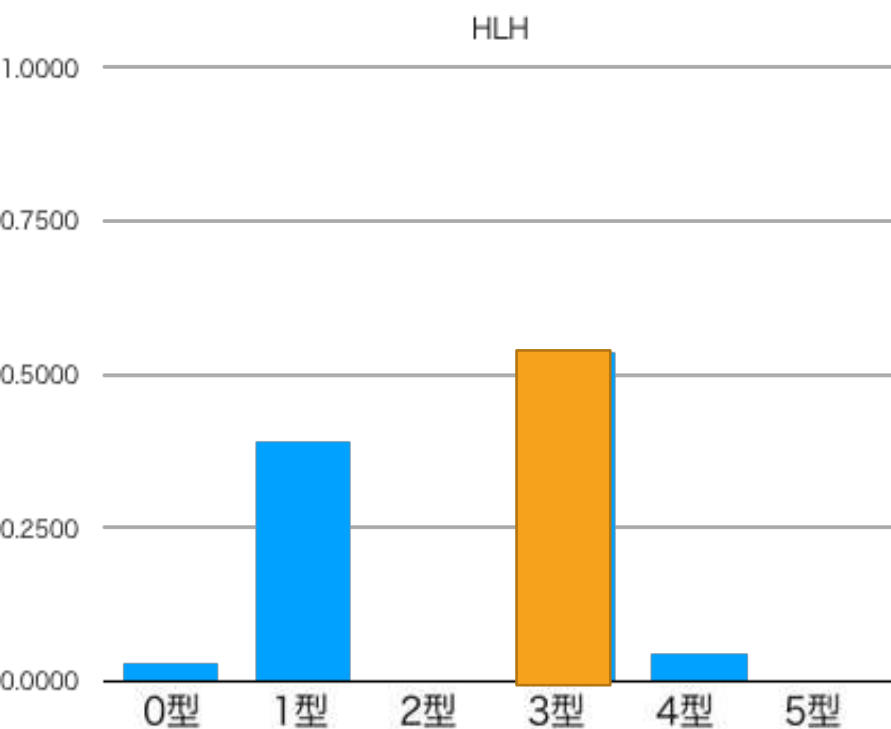


## 学習者

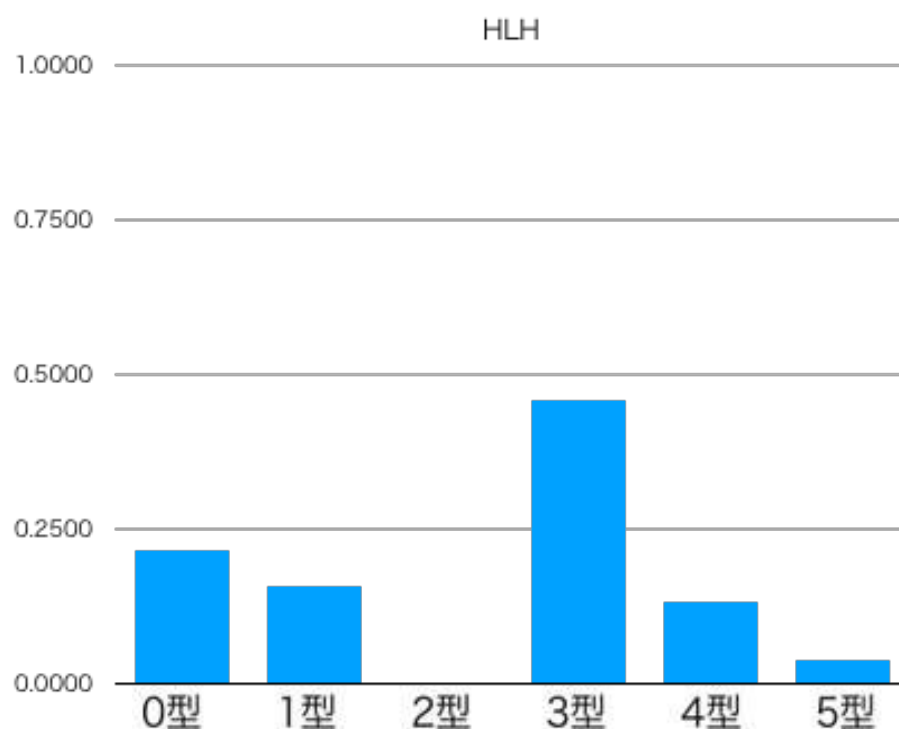




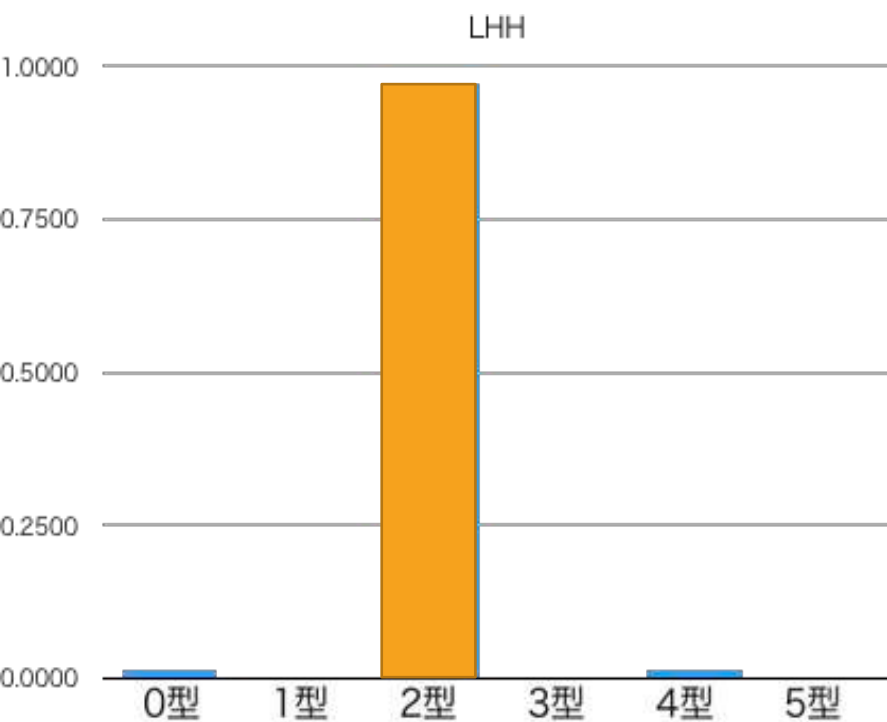
## 母語話者



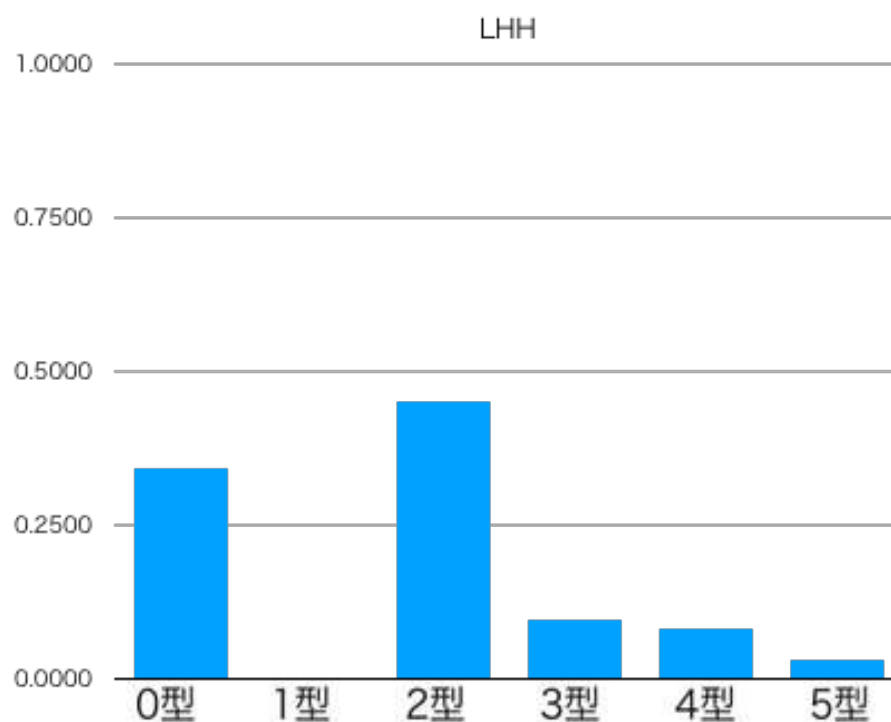
## 学習者



## 母語話者



## 学習者



縦軸：確率

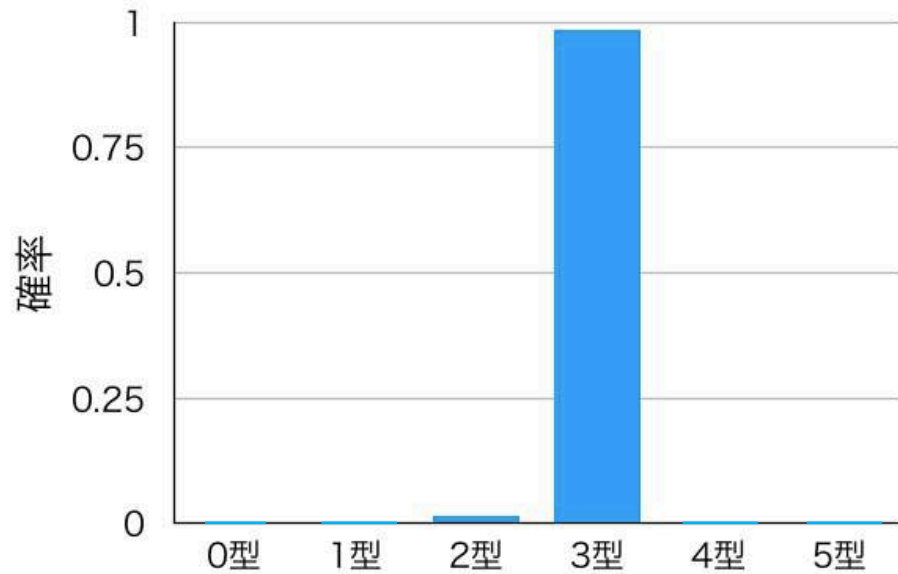


図 1

5 モーラLLHLに対して母語話者が  
付与したアクセント型の分布

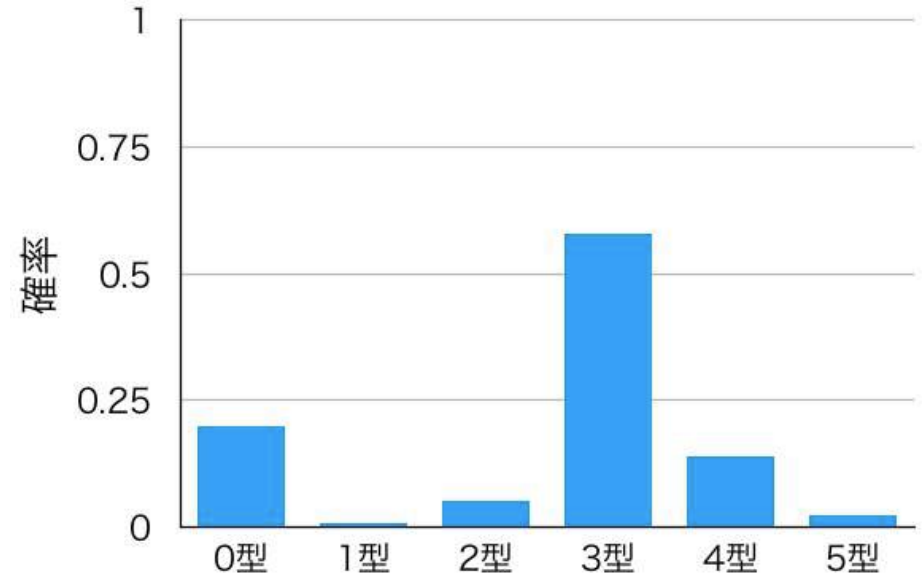


図 2

5 モーラLLHLに対して中国人学習  
者が付与したアクセント型の分布

LLHLの例：ビハポーレ

縦軸：確率



図3

4モーラLLHに対して母語話者が付与したアクセント型の分布

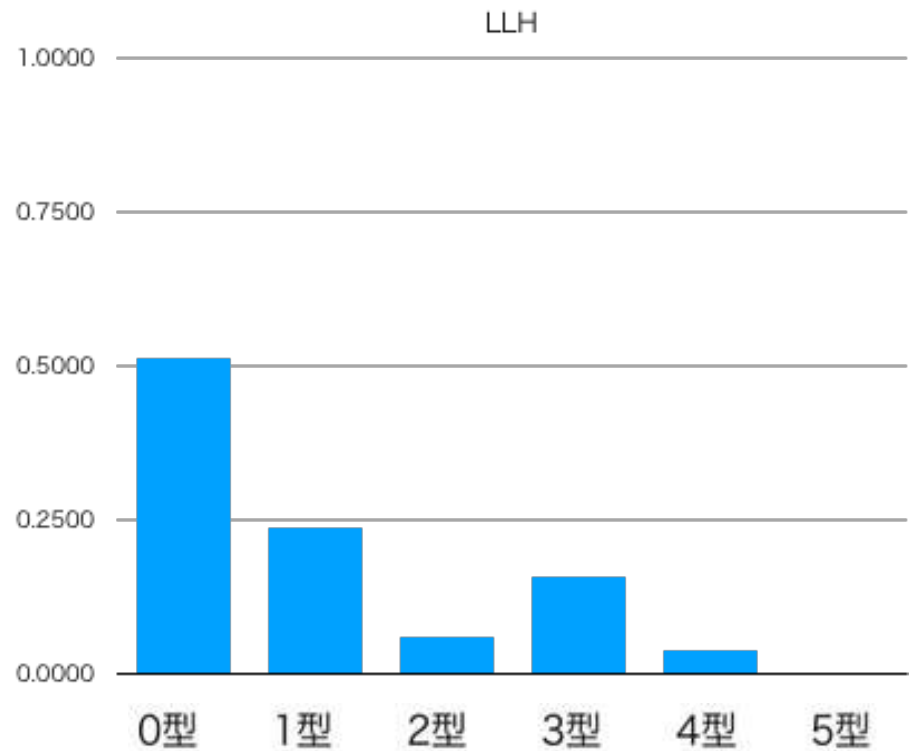


図4

4モーラLLHに対して中国人学習者が付与したアクセント型の分布

LLHの例：コポレー

縦軸：確率

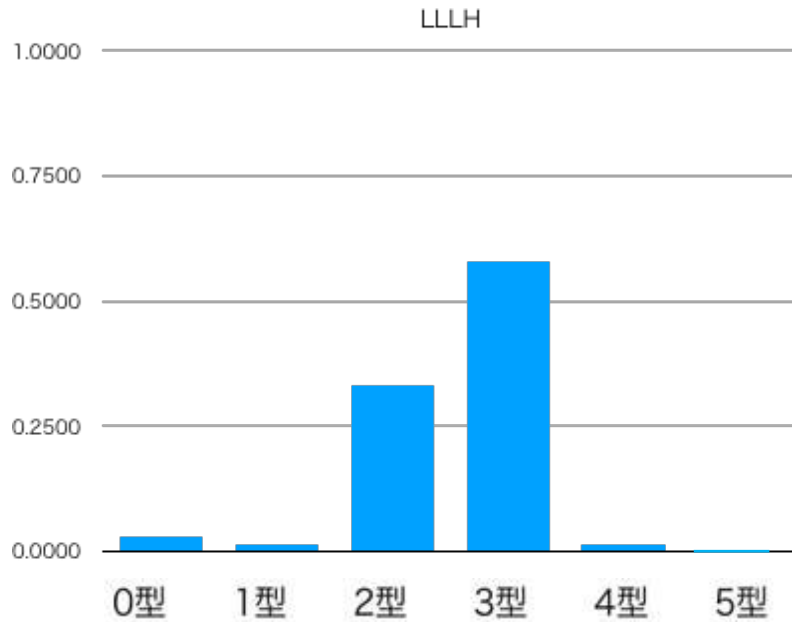


図5

5 モーラLLLHに対して母語話者が  
付与したアクセント型の分布

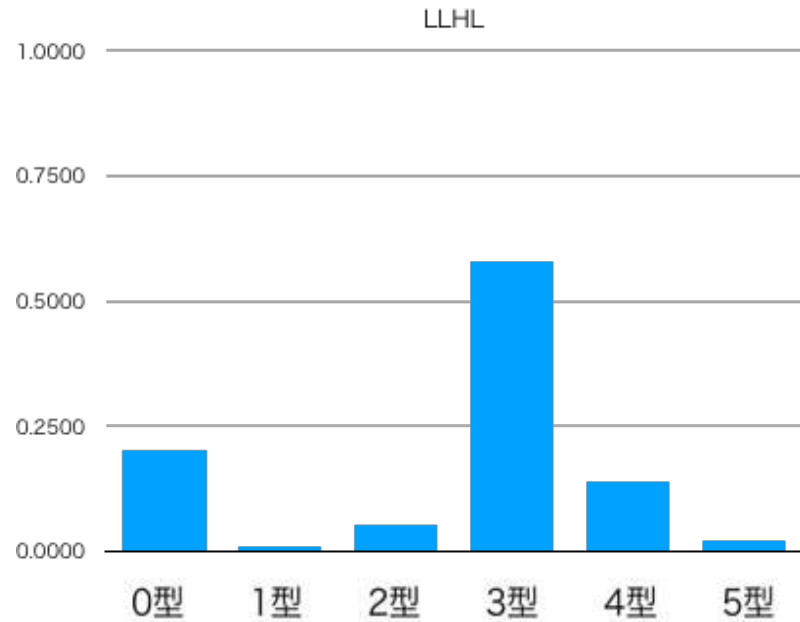


図6

5 モーラLLLHに対して中国人学習  
者が付与したアクセント型の分布

LLLHの例：ロレボシー

乖離を数値化？

母語話者も中国人学習者も3型を付与したの  
はほとんどだが、偏っているところもある  
(乖離が大きい)

# バタチャリヤ距離

- $BD(p, q) = -\log_e \sum_{i=1}^N \sqrt{p_i} \sqrt{q_i}$  但し,  $p_i \geq 0, q_i \geq 0, \sum_{i=1}^N p_i = 1.0, \sum_{i=1}^N q_i = 1.0$

二つのヒストグラム間の類似度を表示する

架空外来語「コデ」を例として

音節 構造	架空外来語	被験 者数	アクセント型_比率(日本人)					
			0	1	2	3	4	5
LL	コデ	9	0	0.89	0.11	0	0	0

音節 構造	架空外来語	被験 者数	アクセント型_比率(中国人)					
			0	1	2	3	4	5
LL	コデ	15	0.74	0.13	0.13	0	0	0

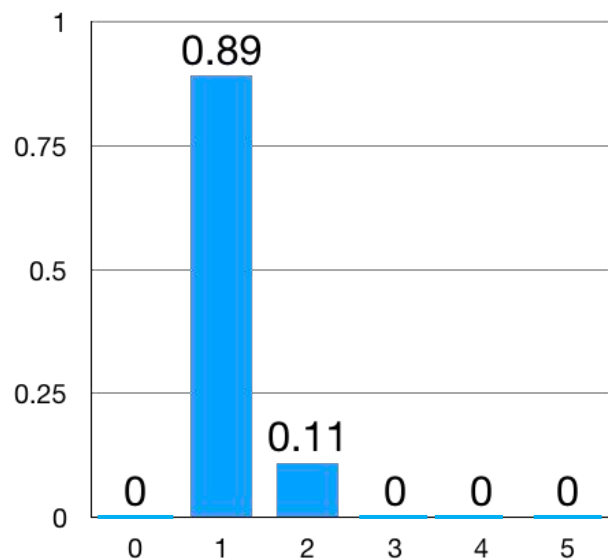
- $BD(p, q) = -\log_e \sum_{i=1}^N \sqrt{p_i} \sqrt{q_i}$  但し,  $p_i \geq 0, q_i \geq 0, \sum_{i=1}^N p_i = 1.0, \sum_{i=1}^N q_i = 1.0$

音節構造	架空外来語	被験者数	アクセント型_比率(日本人)					
			0	1	2	3	4	5
LL	コデ	9	p1	p2	p3	p4	p5	p6

音節構造	架空外来語	被験者数	アクセント型_比率(中国人)					
			0	1	2	3	4	5
LL	コデ	15	q1	q2	q3	q4	q5	q6

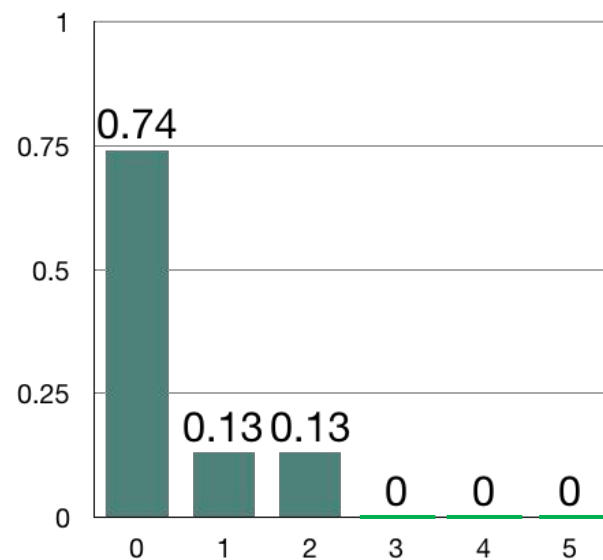
音節構造	架空外来語	被験者数	アクセント型_比率(日本人)					
			0	1	2	3	4	5
LL	コデ	9	0	0.89	0.11	0	0	0

各アクセント型の比率(日本人)



音節構造	架空外来語	被験者数	アクセント型_比率(中国人)					
			0	1	2	3	4	5
LL	コデ	15	0.74	0.13	0.13	0	0	0

各アクセント型の比率(中国人)



$$BC = \sqrt{0 \times 0.74} + \sqrt{0.89 \times 0.13} + \sqrt{0.11 \times 0.13}$$

$$BD = -\ln(BC)$$

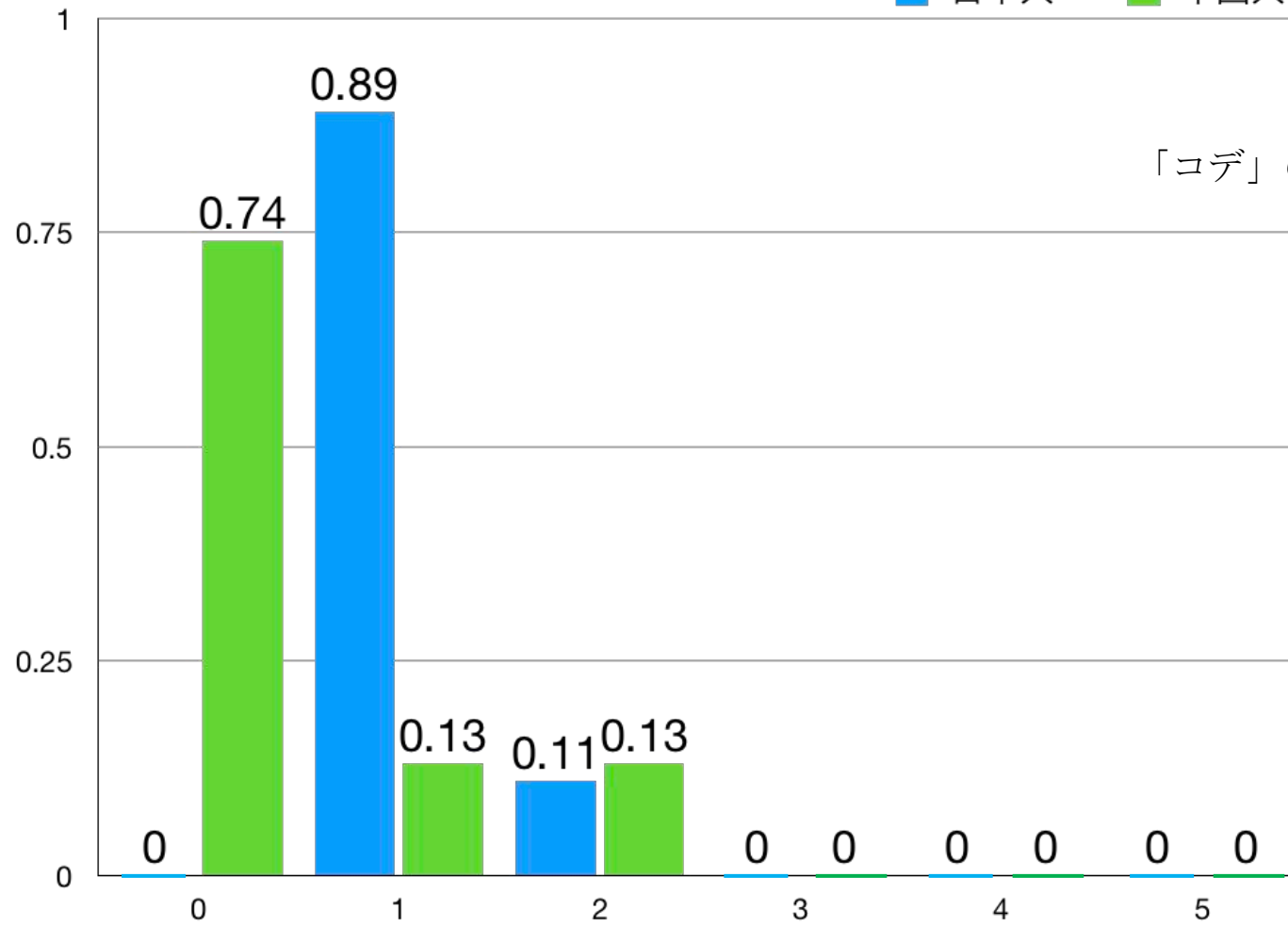
$$\Rightarrow BD = 0.7771$$



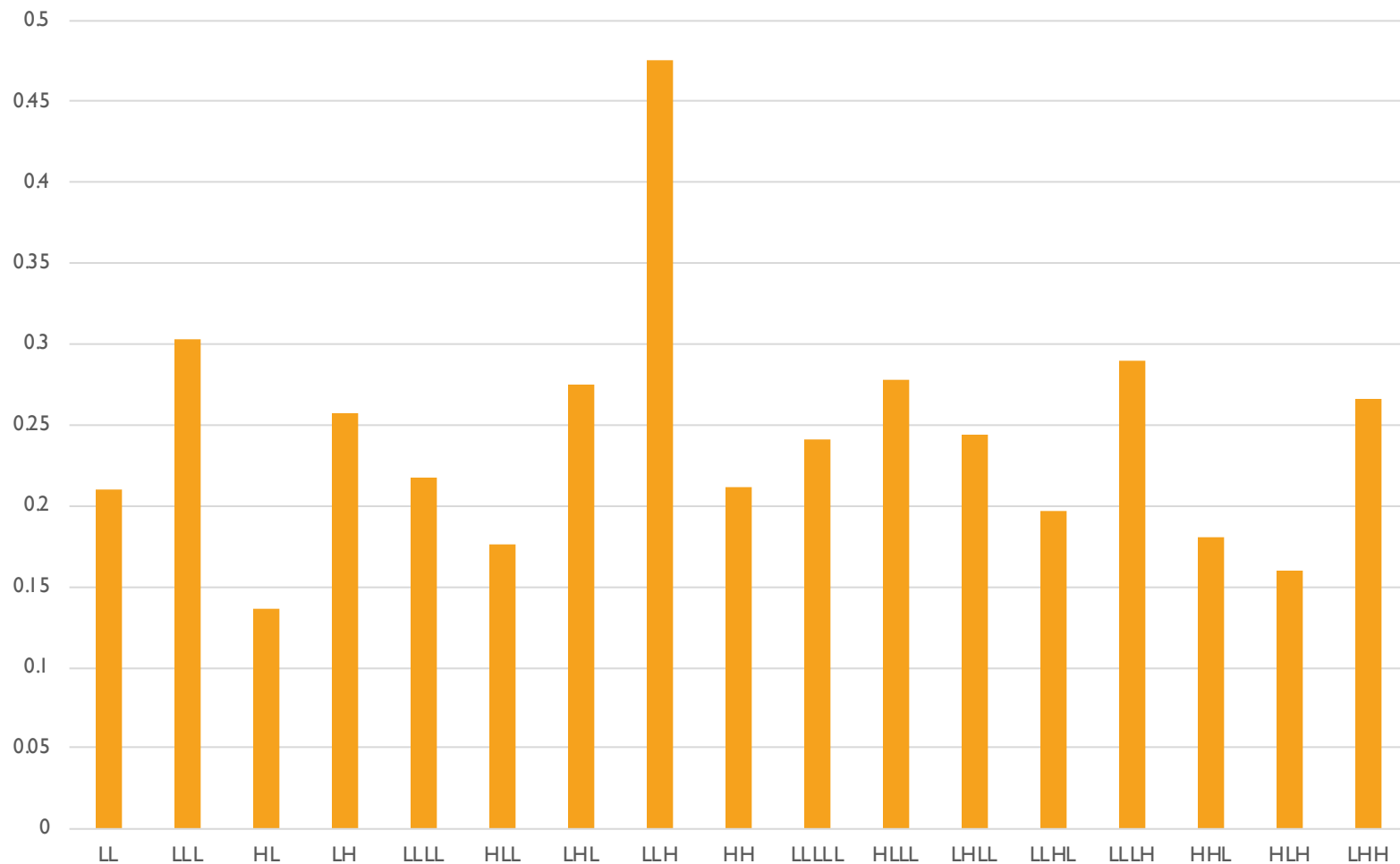
日本人 中国人

**BD=0.7771**

「コデ」の読み方に大きな違いがある



縦軸：カタチャリヤ距離



音節構造別に計算した母語話者と中国人学習者の  
アクセント暗黙知の乖離

→学習者別にも応用できる

## 一番見たくない状況

音節 構造	架空外来語	被験 者数	アクセント型_比率(日本人)					
			0	1	2	3	4	5
XXX	YYY	P	0	1	0	0	0	0

音節 構造	架空外来語	被験 者数	アクセント型_比率(個人)					
			0	1	2	3	4	5
XXX	YYY	Q	0	0	1	0	0	0

一人は一回しか答えないから、各アクセント型の比率は0と1しかない

$$BC = \sqrt{0 \times 0} + \sqrt{1 \times 0} + \sqrt{0 \times 1} = 0$$

$$BD = -\ln(BC) \Rightarrow \underline{BD \rightarrow \infty}$$



εを使う

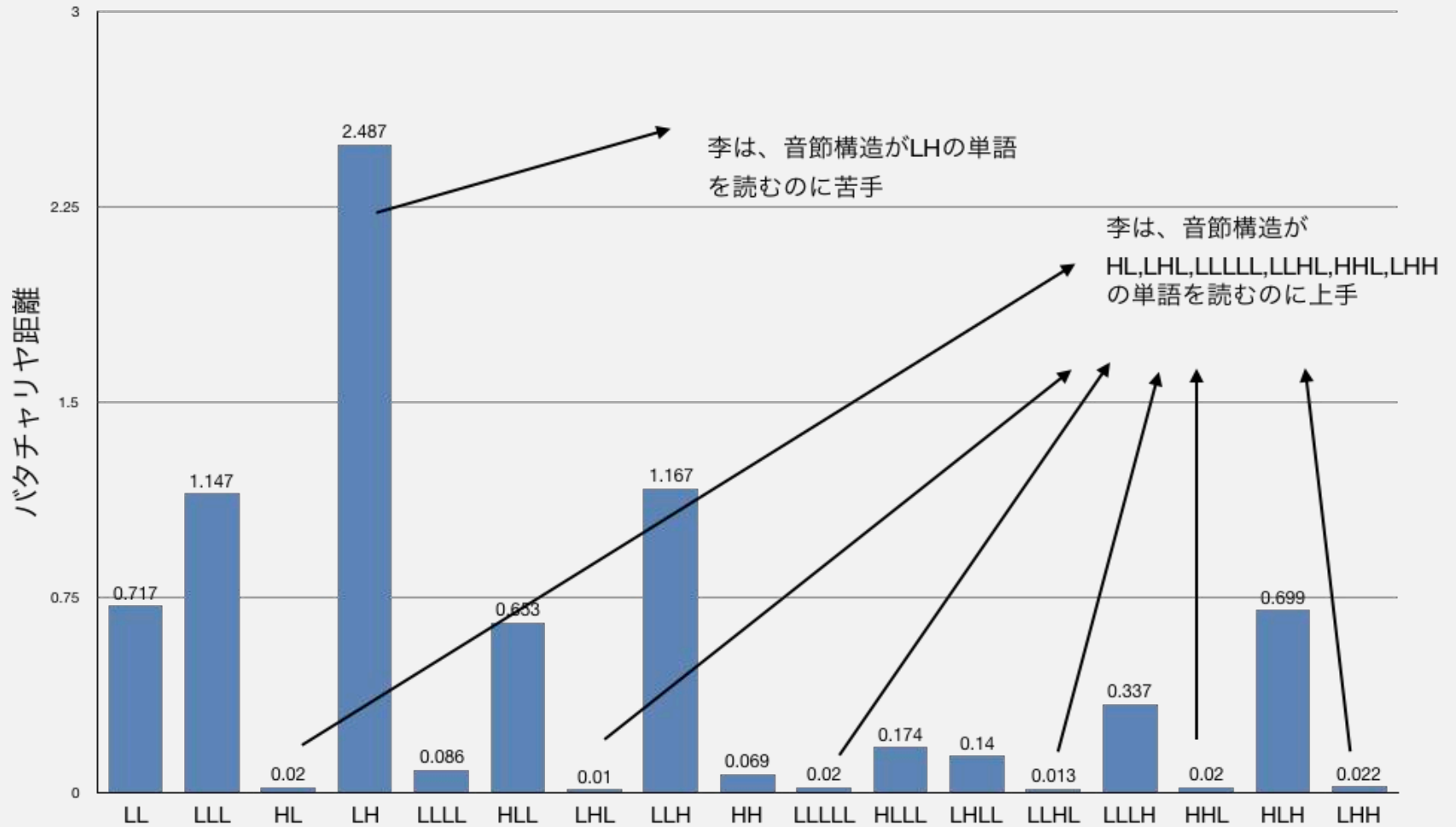
BDの平均値は計算できなくなる！

$$\varepsilon = 1/200 = 0.005$$

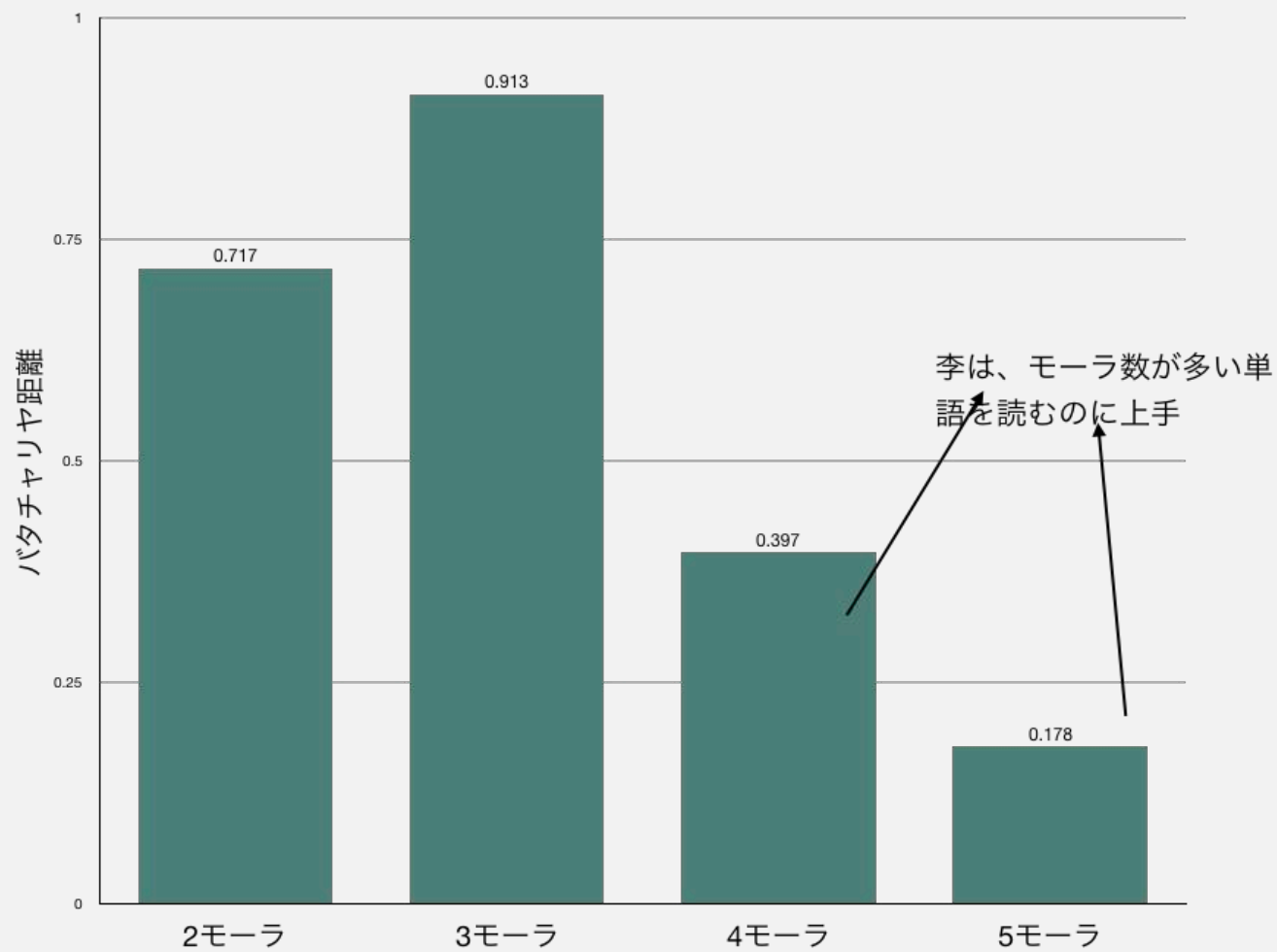
音節構造	架空外来語	被験者数	アクセント型_比率(日本人)					
			0	1	2	3	4	5
XXX	YYY	P	0.005	0.975	0.005	0.005	0.005	0.005

BDは永遠に $\infty$ にならない  
 日本人のアクセント型に完全にずれても  
 BDの平均値は計算できるようになる

■ 各音節構造によるバタチャリヤ距離

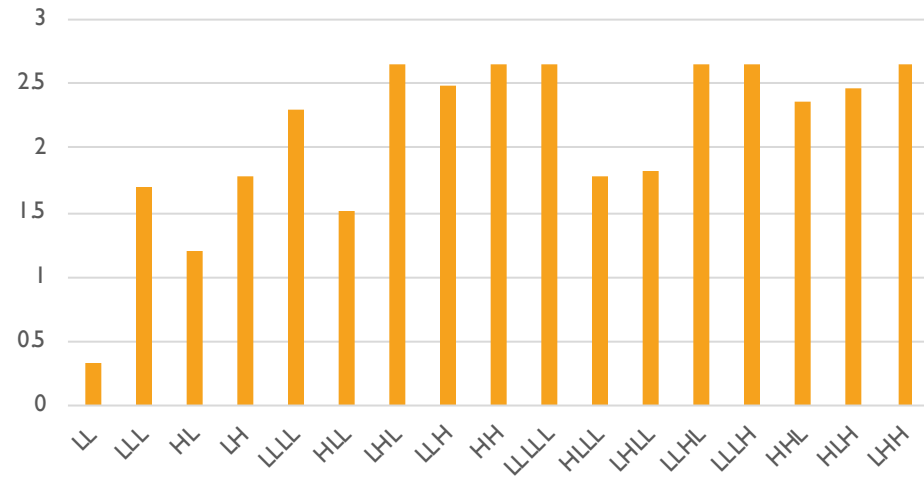


■ 各モーラ数によるバタチャリヤ距離

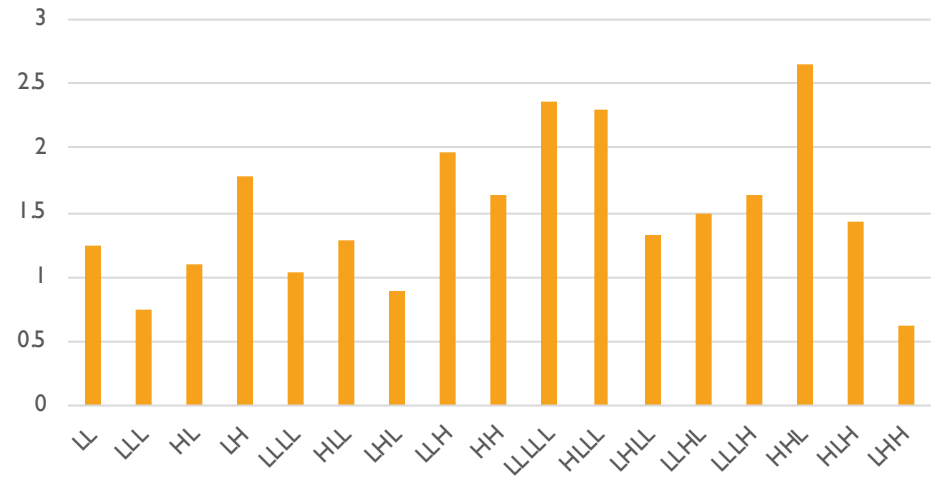


学習者別のBD

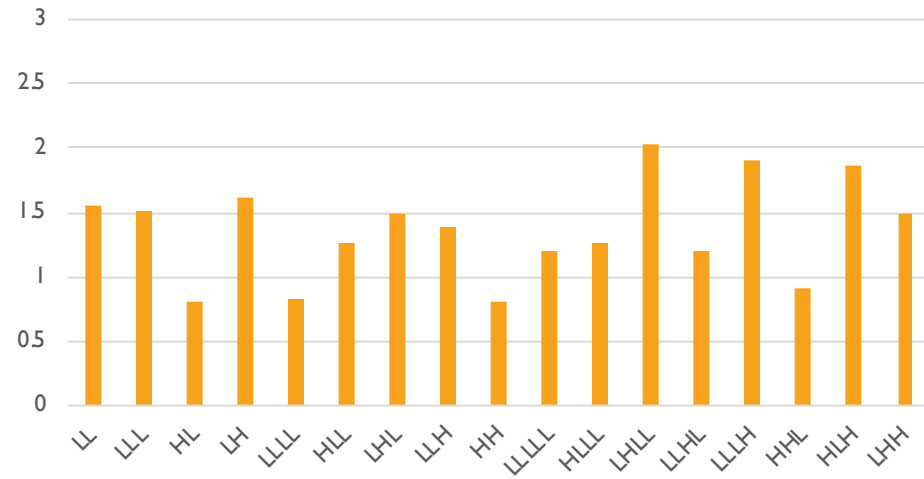
1 (女性、内モンゴル、3年)



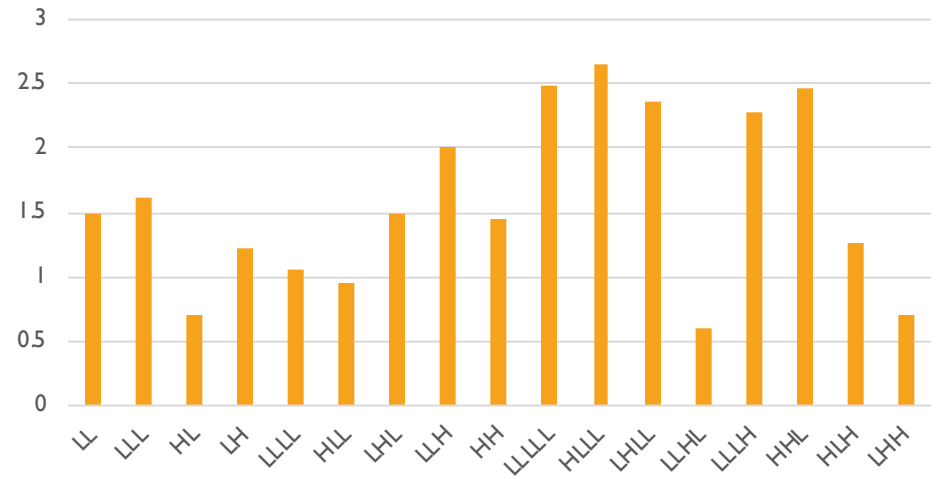
2 (男性、安徽省、3年)



3 (女性、遼寧省、3年)



4 (女性、遼寧省、3年)

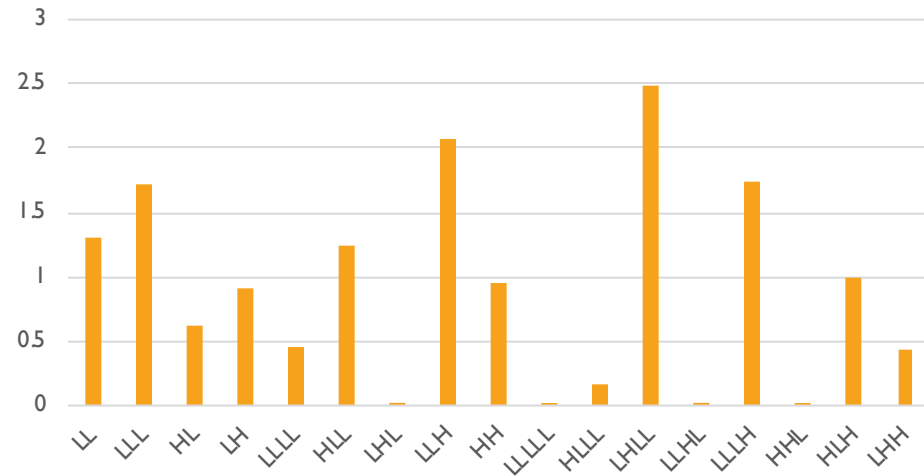


縦軸：バタチャリヤ距離

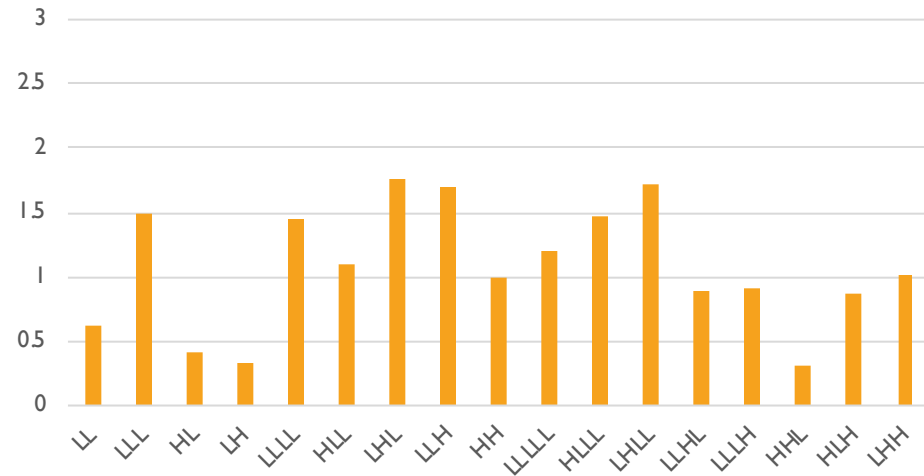
(性別、出身地、学習歴)



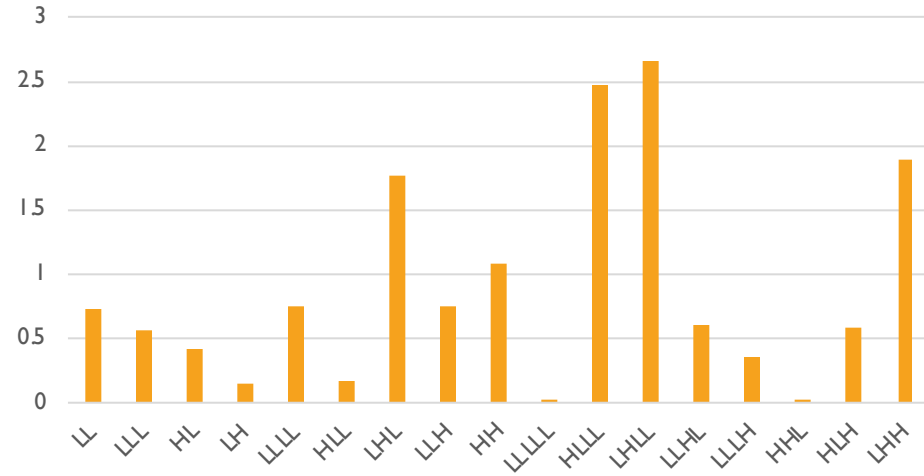
5 (男性、遼寧省、3年)



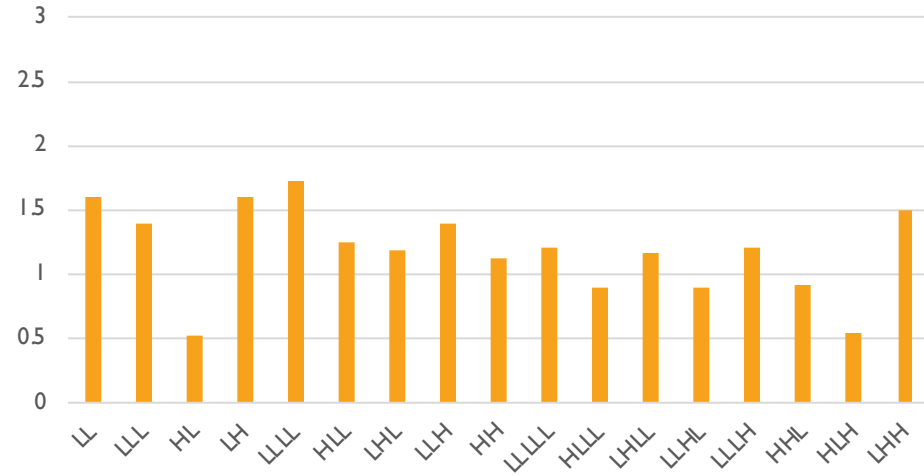
6 (女性、遼寧省、3年)



7 (男性、黒竜江省、2年)



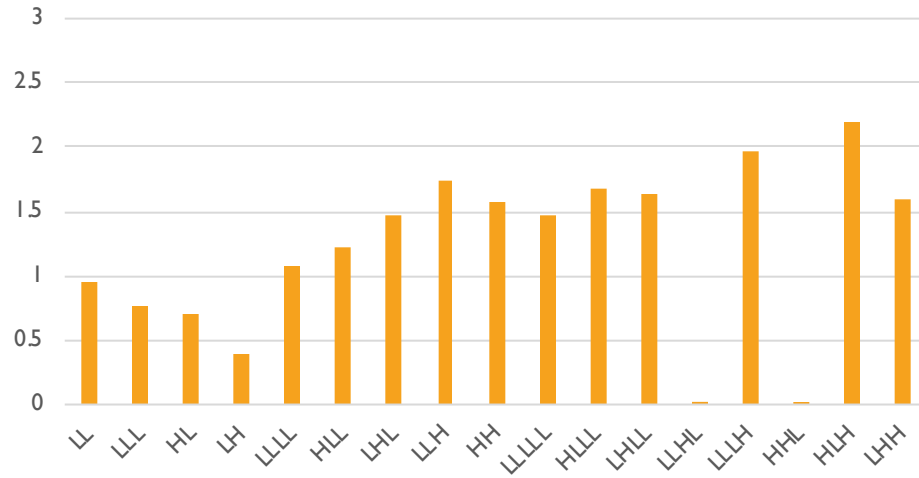
8 (女性、遼寧省、3年)



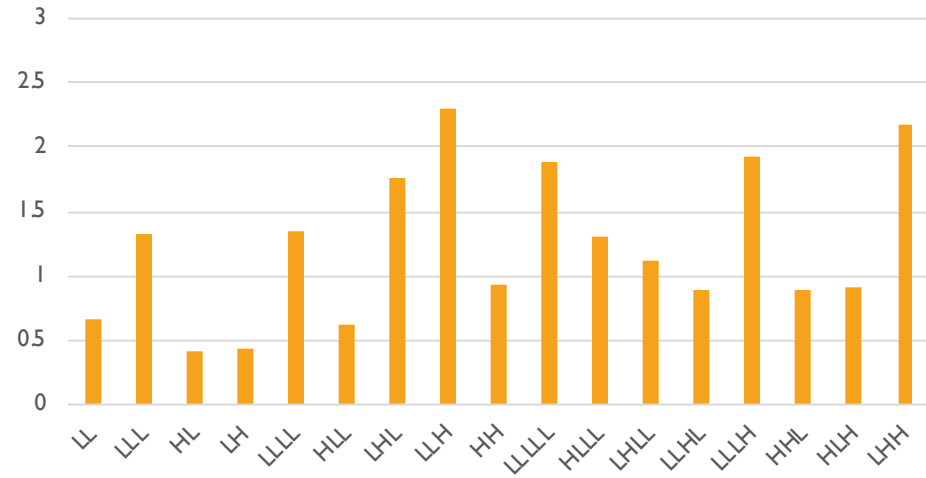
縦軸：バタチャリヤ距離

(性別、出身地、学習歴)

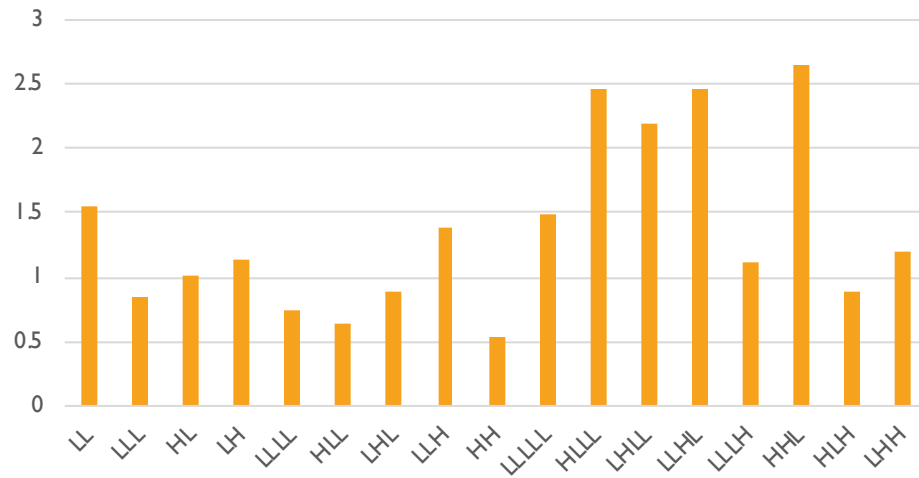
9 (男性、遼寧省、2年)



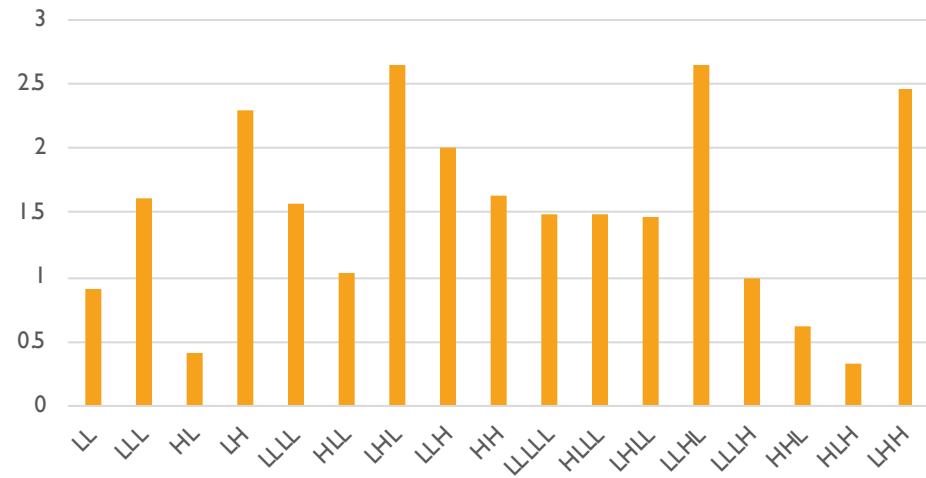
10 (男性、遼寧省、3年)



11 (女性、山東省、3年)



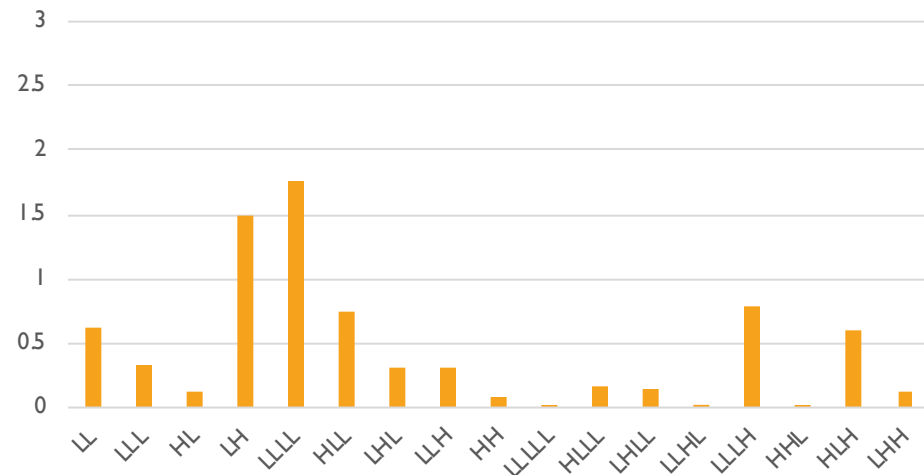
12 (男性、遼寧省、3年)



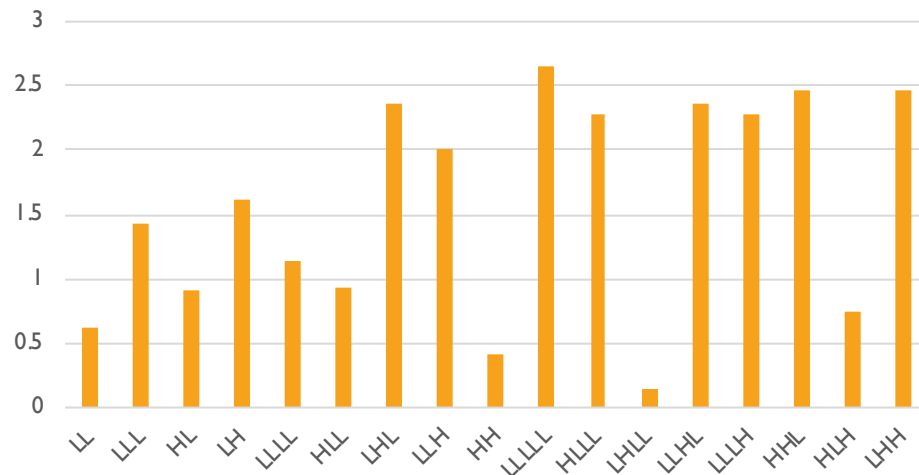
縦軸：バタチャリヤ距離

(性別、出身地、学習歴)

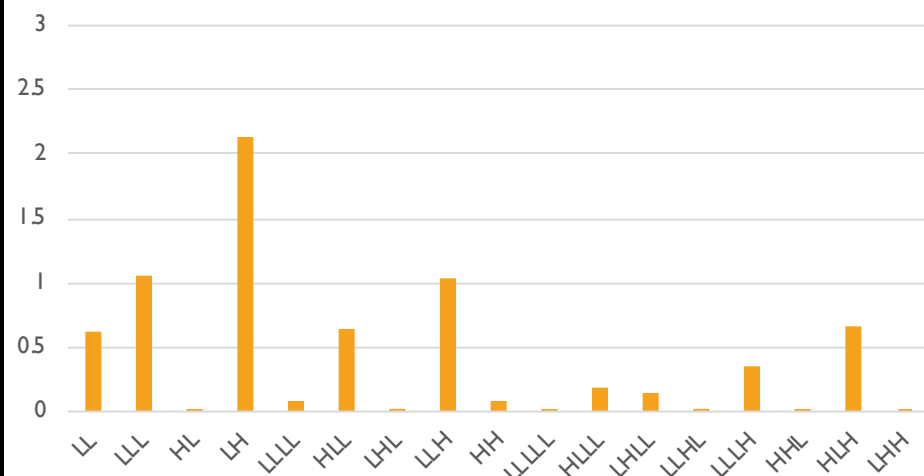
13 (女性、遼寧省、3年)



14 (男性、遼寧省、3年)



15 (女性、遼寧省、3年)



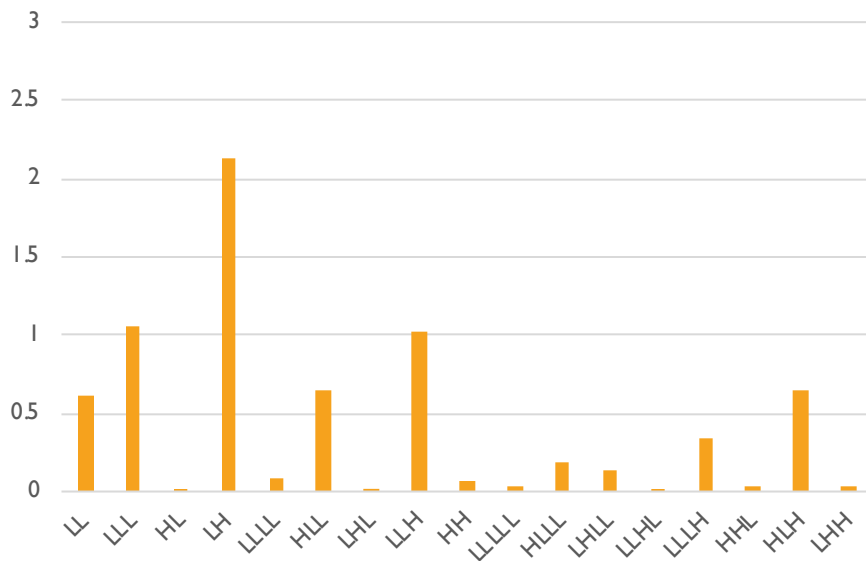
アクセント感覚は学習者によって大きく異なる

縦軸：バタチャリヤ距離

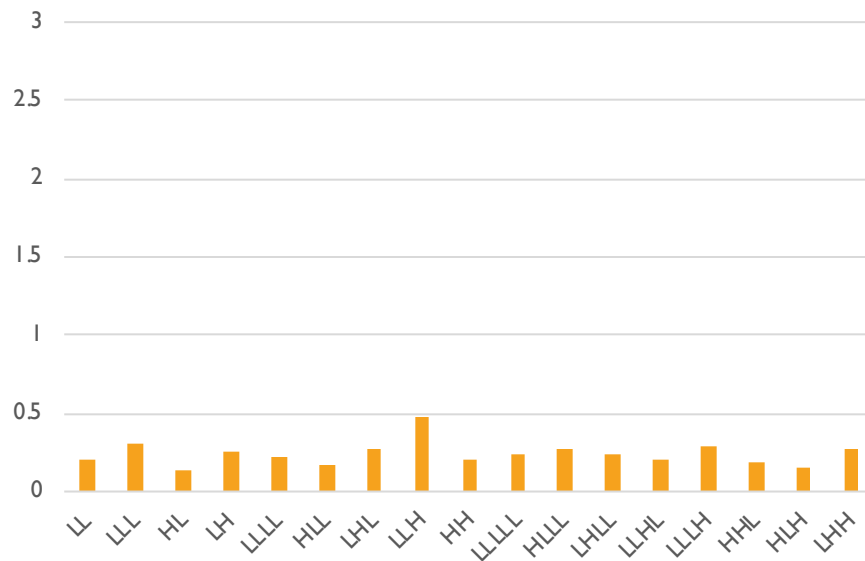
全体的に見れば一番上手な（BDの値が一番小さい）学習者は誰？

（性別、出身地、学習歴）

縦軸：バタチャリヤ距離



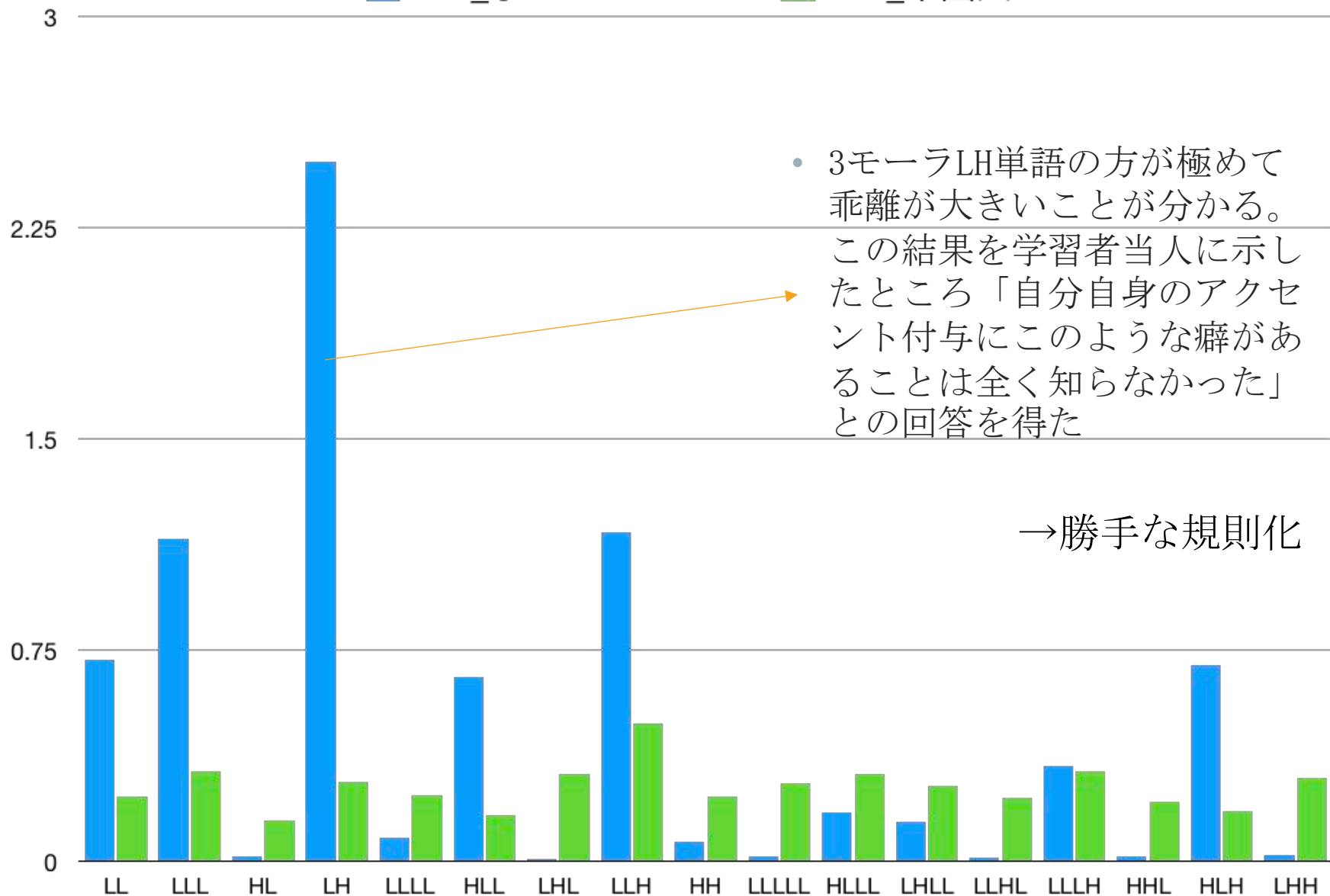
音節構造別に計算した母語話者とある  
学習者のアクセント暗黙知の乖離



音節構造別に計算した母語話者と中国人学習者のアクセント暗黙知の乖離

■ BD\_李

■ BD\_中国人



- 3モーラLH単語の方が極めて乖離が大きいことが分かる。この結果を学習者当人に示したところ「自分自身のアクセント付与にこのような癖があることは全く知らなかった」との回答を得た

→勝手な規則化

## 今後の課題

より多くの学習者からデータを集めたい

学習者が勝手に規則化してしまう要因（出身、母国語、教科書…）を知りたい

学習者ごとの癖を直す方法を考えたい

# ウェブサイト計画

日本語アクセント調査にご協力頂き、  
誠に有難うございます

**\* 説明 \***

(ここに説明を書く)これは、XXです。

あなたの母国語は？

☐日本語 ☐中国語 ☐韓国語 ☐その他

LL

- |       |                         |                         |                         |
|-------|-------------------------|-------------------------|-------------------------|
| 1. ペザ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |
| 2. コデ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |
| 3. ギバ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |
| 4. ジテ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |
| 5. ビゾ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |
| 6. ガゾ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |
| 7. パキ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |
| 8. ゴヨ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |
| 9. ザミ | <input type="radio"/> 0 | <input type="radio"/> 1 | <input type="radio"/> 2 |

- HTML+PHP+SQL
- データ収集&分析
- 6月末に公開予定

学習者の音節構造によるBD  
はグラフの形で表示される

# 癖が判明した学習者への対応

李は、3モーラLHに対して勝手な規則化をしている  
→3モーラLHの既知語を使ってアクセントを再度練習する

検索・表示条件

検索の条件

教科書   から  まで

品詞

難易度  から  まで

旧能力試験  から  まで

表示の条件

表示順序

カスタマイズ 1.  2.  3.

オプション

アクセントの揺れ

教科書と難易度

単語の検索



名詞	辞書形
間	あいだ
五日	いつか
田舎	いなか
後ろ	うしろ
売り場	うりば
上着	うわぎ
英語	えいご
おかず	おかず
お金	おかね
置き場	おきば
お客	おきゃく



# まとめ

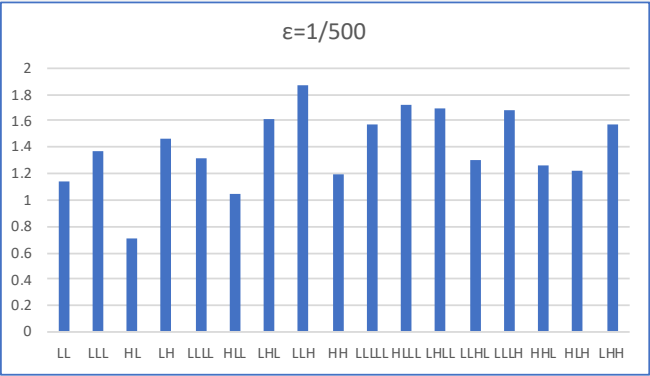
- 母語話者のアクセント感覚：規則に従ったアクセント付与が優勢なことが多いが、そうでないケースも散見される
- 学習者のアクセント感覚：優勢な度外が大きく低下する  
– 3の規則は定着していない
- 母語話者と中国人学習者のアクセント感覚の比較：3モーラLLL、4モーラLLH、5モーラLLLHで乖離が大きい
- 母語話者と特定中国人学習者のアクセント感覚の比較：学習者によって、アクセント感覚は大きく異なる。本分析によってどのような単語グループで自分勝手な規則化をしていることが明確にできる

- 学習者の音声学習は、教科書の音声CDを真似ることが基本的であるが、どのような韻律規則を（無自覚的に）獲得してしまっているのか、学習者自身も把握することは困難である
- 本研究の試みによって、学習者が自分勝手な規則化（思い込み）をしているのかを指摘し、正しい韻律知識を効率的に獲得する手助けとなれば、と考えている

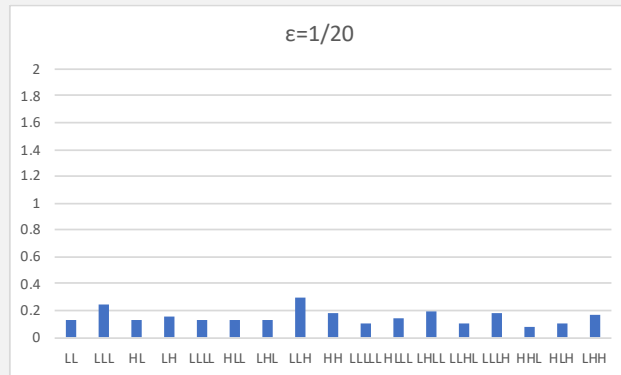
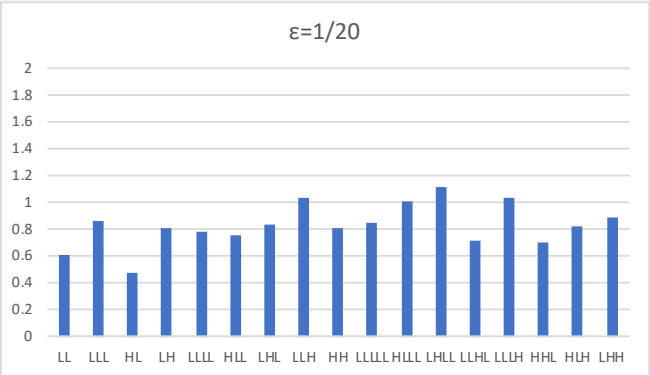
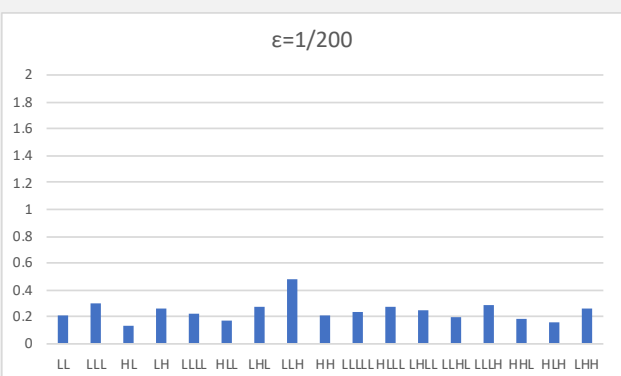
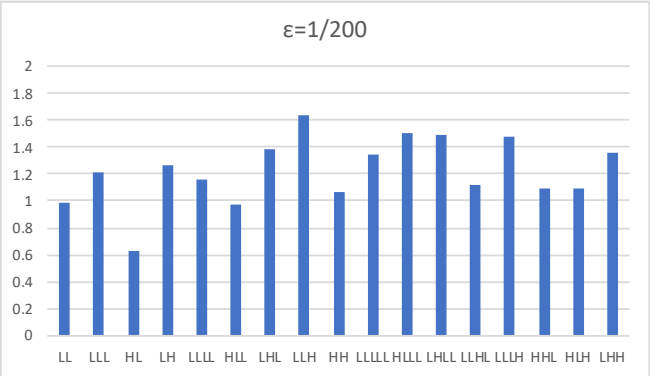
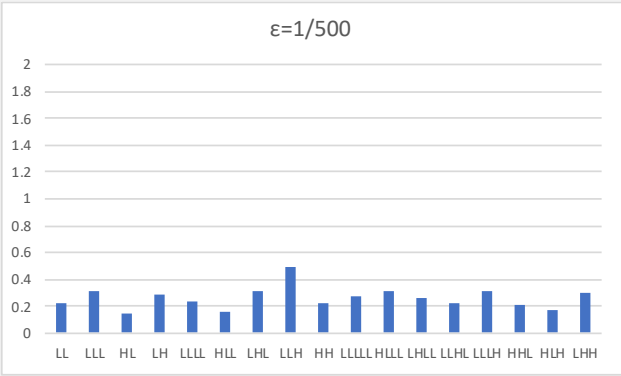
ご静聴ありがとうございました

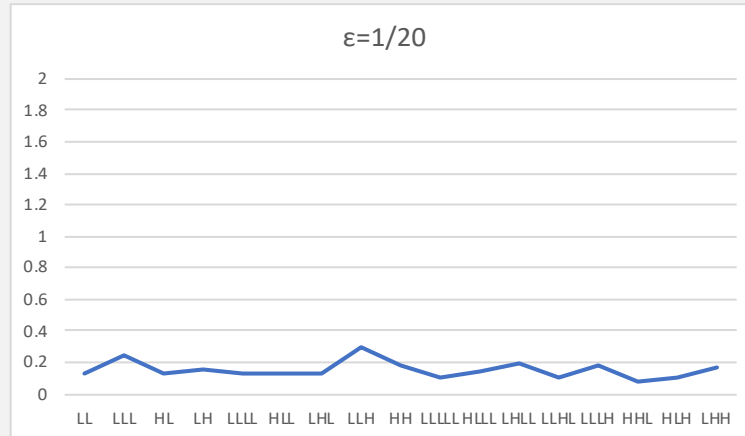
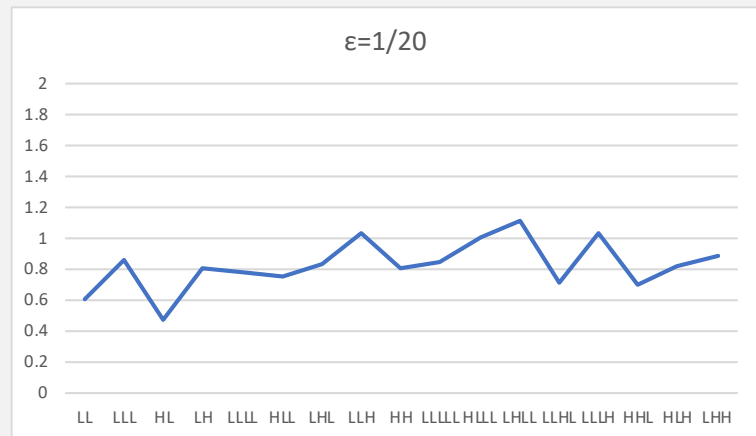
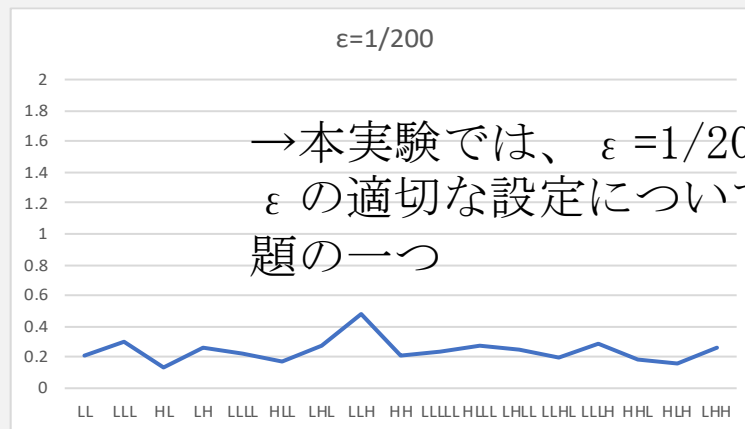
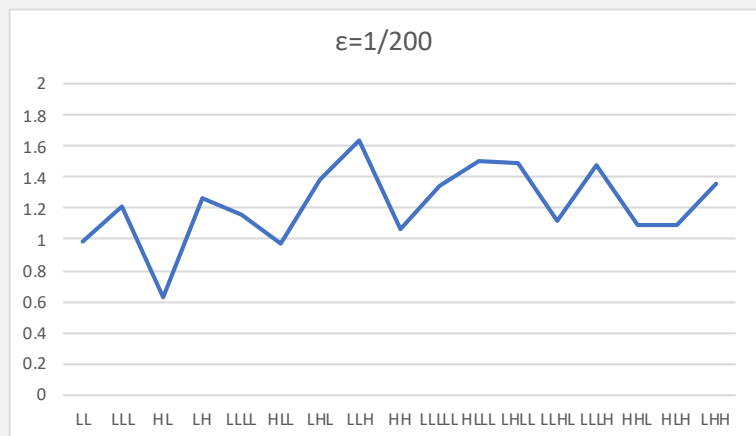
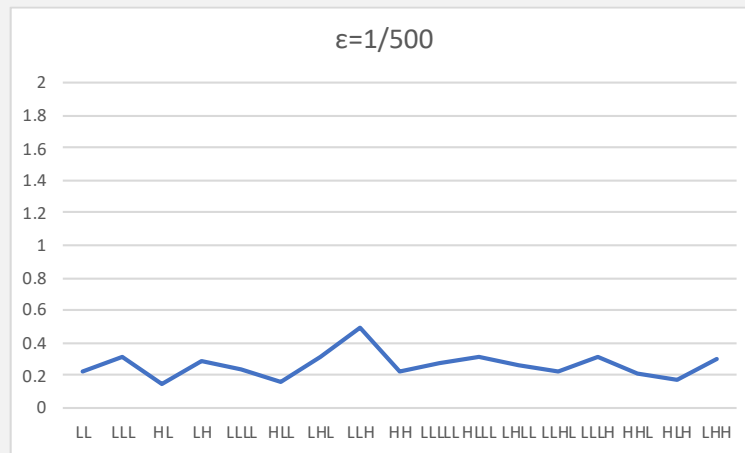
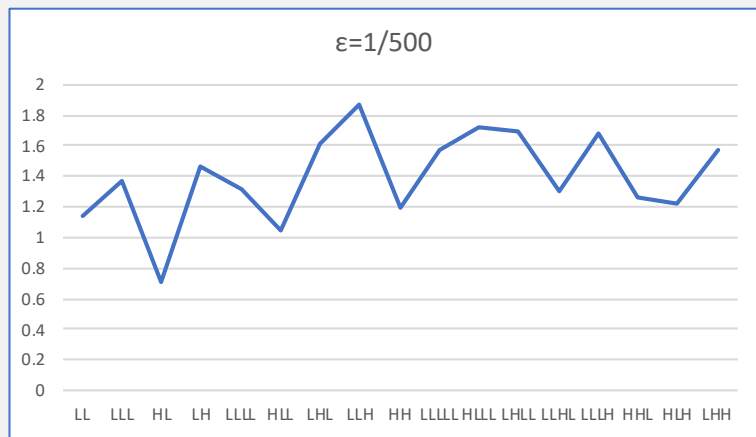
$\varepsilon$  の値

Average(BD\_of\_chinese)



BD(average\_of\_chinese)





バッテリー距離

# バタチャリヤ距離

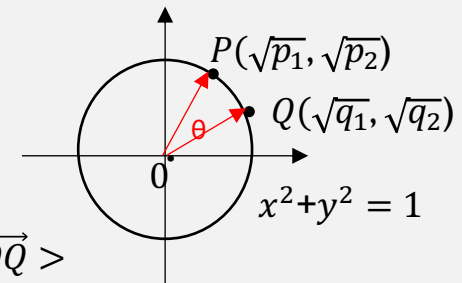
- $BD(p, q) = -\log_e \sum_{i=1}^N \sqrt{p_i} \sqrt{q_i}$  但し,  $p_i \geq 0, q_i \geq 0, \sum_{i=1}^N p_i = 1.0, \sum_{i=1}^N q_i = 1.0$

$$\begin{aligned} (p_1, p_2, \dots, p_n) \quad & \sum_{i=1}^N p_i = 1 \\ (q_1, q_2, \dots, q_n) \quad & \sum_{i=1}^N q_i = 1 \end{aligned}$$

N=2の場合

$$\begin{aligned} p_1 + p_2 = 1 \quad \text{即ち} \quad (\sqrt{p_1})^2 + (\sqrt{p_2})^2 = 1 & \Rightarrow x_1^2 + y_1^2 = 1 \\ q_1 + q_2 = 1 \quad \text{即ち} \quad (\sqrt{q_1})^2 + (\sqrt{q_2})^2 = 1 & \Rightarrow x_2^2 + y_2^2 = 1 \end{aligned}$$

$$\sqrt{p_1}\sqrt{q_1} + \sqrt{p_2}\sqrt{q_2} = \overrightarrow{OP} \cdot \overrightarrow{OQ} = |OP||OQ| \cos \angle \overrightarrow{OP}, \overrightarrow{OQ} = \cos \angle \overrightarrow{OP}, \overrightarrow{OQ}$$



$$BD(p, q) = -\log_e \cos \angle \overrightarrow{OP}, \overrightarrow{OQ} = -\log_e \cos \theta$$

$\cos \theta < 1$ のため、 $\log_e \cos \theta < 0$

$\theta$ が大きければ、 $\cos \theta$  と  $\log_e \cos \theta$  が小さくなり、 $-\log_e \cos \theta$  が大きくなる



# バタチャリヤ距離

- $N=3$  の場合
- $(p_1, p_2, \dots, p_n) \quad \sum_{i=1}^N p_i = 1$
- $(q_1, q_2, \dots, q_n) \quad \sum_{i=1}^N q_i = 1$
- $p_1 + p_2 + p_3 = 1$  即ち  $(\sqrt{p_1})^2 + (\sqrt{p_2})^2 + (\sqrt{p_3})^2 = 1 \Rightarrow x_1^2 + y_1^2 + z_1^2 = 1$
- $q_1 + q_2 + q_3 = 1$  即ち  $(\sqrt{q_1})^2 + (\sqrt{q_2})^2 + (\sqrt{q_3})^2 = 1 \Rightarrow x_2^2 + y_2^2 + z_2^2 = 1$
- $\sqrt{p_1}\sqrt{q_1} + \sqrt{p_2}\sqrt{q_2} + \sqrt{p_3}\sqrt{q_3} = \overrightarrow{OP} \cdot \overrightarrow{OQ} = |OP||OQ| \cos \angle \overrightarrow{OP}, \overrightarrow{OQ} = \cos \angle \overrightarrow{OP}, \overrightarrow{OQ}$
- $BD(p, q) = -\log_e \cos \angle \overrightarrow{OP}, \overrightarrow{OQ} = -\log_e \cos \theta$
- $\cos \theta < 1$  のため、 $\log_e \cos \theta < 0$
- $\theta$  が大きければ、 $\cos \theta$  と  $\log_e \cos \theta$  が小さくなり、 $-\log_e \cos \theta$  が大きくなる

