

## Laryngeal contrast and tone in Tamang: A preliminary study

Hyun Kyung Hwang (RIKEN) Seunghun Lee (ICU)  
 Selin Grollmann, Pascal Gerber (Univ. of Bern)  
 hyunkyung.hwang@riken.jp, seunghun@icu.ac.jp,  
 {selin.grollmann, pascal.gerber}@isw.unibe.ch

### 1. Introduction

Tamang is a Sino-Tibetan language spoken mainly in Nepal. The Nepal census conducted in 2011 reports that there are 1.5 million speakers of Tamang language, which is the fifth largest among more than ninety languages spoken in Nepal.

Following the phonetic and phonological descriptions of Tamang by Mazaudon (1973), Tamang has been described as a language with contrastive tones. Interestingly, there are conflicting claims regarding laryngeal/tonal contrast of Tamang stops. Mazaudon describes that Tamang has four tones with distinctions in pitch and voice quality: Tone 1 (high falling pitch) and Tone 2 (high level) have modal whereas Tone 3 (low level) and Tone 4 (low-rising) have breathy phonation. Further, she claims that aspiration contrast is observed only in Tone 1 and 2 words, and Tone 3 and 4 can be voiced (Mazaudon 2014). Varenkamp (to appear, cited in Lee 2011), on the other hand, reports there are no phonemically voiced stops as none of Tamang stops are acoustically voiced. On the contrary, Yonjan (1993) argues that there is a 3-way contrast of aspirated/unaspirated/voiced stops without assuming phonological tones. Yet, the acoustic aspects of the Tamang sound system have not been fully investigated. In the current study, we focus on the phonetic realization of stops, and examine the laryngeal contrast of this relatively understudied language, with special attention to tones.

### 2. Method

#### 2.1. Recording

Seven stop-initial monosyllabic words were taken from Yonjan (1993), as shown in (1).

#### (1) Target words

be(ji)	‘done’
da(ji)	‘clean’
ge:	‘work’
ta(ji)	‘done’
to(ji)	‘beaten’
tha(ji)	‘to cut’
tho(ji)	‘to receive’

Four female speakers of Eastern Tamang took part in the recording session conducted in Kathmandu,

Nepal. All had no history of speech or hearing impairment. It should be noted that the participants speak Nepali as it is the official language in Nepal. However, all the speakers reported that the primary language they speak at home is Tamang. Table 1 summarizes the information of age and the home city that each speaker comes from.

Table 1: Demographic information of the speakers

	Age	Hometown
Speaker 1	18	Bhaktapur
Speaker 2	22	Dolkha
Speaker 3	23	Kavre
Speaker 4	25	Lalitpur

The target words were randomized and presented to the speakers using PowerPoint slides. Each slide was presented with a stimulus in the Nepali Devanagari script as well as in Roman alphabets. Participants were asked to identify words that are used in Tamang; only those words that were accepted by speakers were recorded. Five repetitions were made for each word in isolation and in a frame sentence. The current analysis focuses on the acoustic analysis of words in isolation and reports the results.

## 2.2. Analysis

Voice onset time (VOT) of the stops and  $f_0$  of the following vowels were measured using Praat. Time-normalized  $f_0$  contours of the vowels were created by averaging  $f_0$  values of 10 equally-timed windows within the vocalic intervals.

## 3. Results and discussion

Boxplots in Figure 1 presents VOT values of the stops depending on the stop categories: Aspirated (asp), voiced (vd), and voiceless (vl) categories. Overall, VOT values clearly separate the stops into three groups; long positive VOT for aspirated, short positive VOT for voiceless, and negative VOT for voiced stops. Except for some tokens of Speaker 2, the voiced stops always involve pre-voicing, which is contrary to the claim by Mazaudon (2014) or Varenkamp (to appear). Rather, our results seem to support the three categories of stop contrast argued by Yonjan (1993). Speaker 2 exhibits exceptionally large variations for aspirated and voiced categories. Perhaps the stop category of some words we recorded is actually different in the dialect that this particular participant speaks. However, this possibility is less likely since the five repetitions of the same word show a large variation in VOT. Alternatively, we can posit that some words beginning with an aspirated or a voiced stop undergoes a change-in-progress. More controlled data which are produced by speakers from the same area are required to confirm this alternative possibility.

A linear mixed-effects analysis was performed to compare the VOT values of stop contrast. We

entered into the model the speaker as random effects and the stop category as fixed effects. The confidence interval was set at 95%. The results revealed that the stop category has a significant effect on the VOT ( $F(2,123)=207.1019$ ,  $p<.001^*$ ).

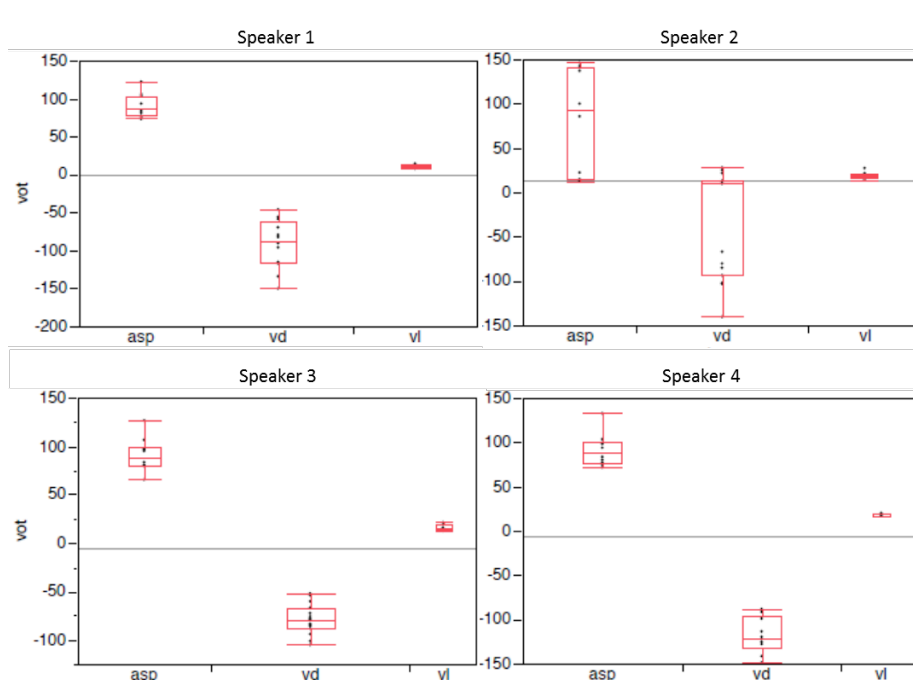


Figure 1: Stop VOT (ms) for each speaker. X-axis represents different stop categories

Turning to the  $f_0$  of the following vowels,  $f_0$  movements of the three categories do not exhibit appreciable differences. Figure 2 shows time-normalized  $f_0$  contours of the vowels for each speaker. Only the first 1 or 2 frames render noticeable differences where the voiced category is associated with lower  $f_0$  compared to the voiceless or aspirated category. It is not surprising that  $f_0$  following voiced obstruents typically produce lower  $f_0$  perturbation. Besides the vowel onsets, the contours seem to be extremely similar regardless of the stop category. Although Speaker 4 renders slightly higher  $f_0$  for voiceless stops, aspirated and voiced stops more or less pattern together, and the overall tonal shape of the voiceless category seems to be quite similar to the aspirated one. In order to test the statistical significance of this observation, a linear mixed-effects model was constructed with the speaker as random effects and the stop category as fixed effects. The results confirm that the stop category has no significant effect on the  $f_0$  ( $F(2,123)=1.6843$ ,  $p=0.19$ ).

The lack of difference in  $f_0$  among the stop categories suggests that the tonal distinction reported by Mazaudon (1973, 2014) may not be present, at least in the dialects we examined. It is worth reiterating that there is no phonologically contrastive voicing but it can be observed only in the two low tones (Tone 3 and Tone 4 in her terms). However, the results of the current acoustic study did not show supporting evidence for her analysis. The discrepancy between Mazaudon's description

and our results might be attributed to the recent loss of tones and the emergence of voicing perhaps due to sociolinguistic factors, such as the influence of Nepali or other non-tonal languages. The dominant language that the majority of Tamang speakers use in school or workplace is Nepali. Alternatively, dialectal differences might be responsible for the discrepancy. This issue needs to be investigated in more detail and with speakers living in the native Tamang area in future research.

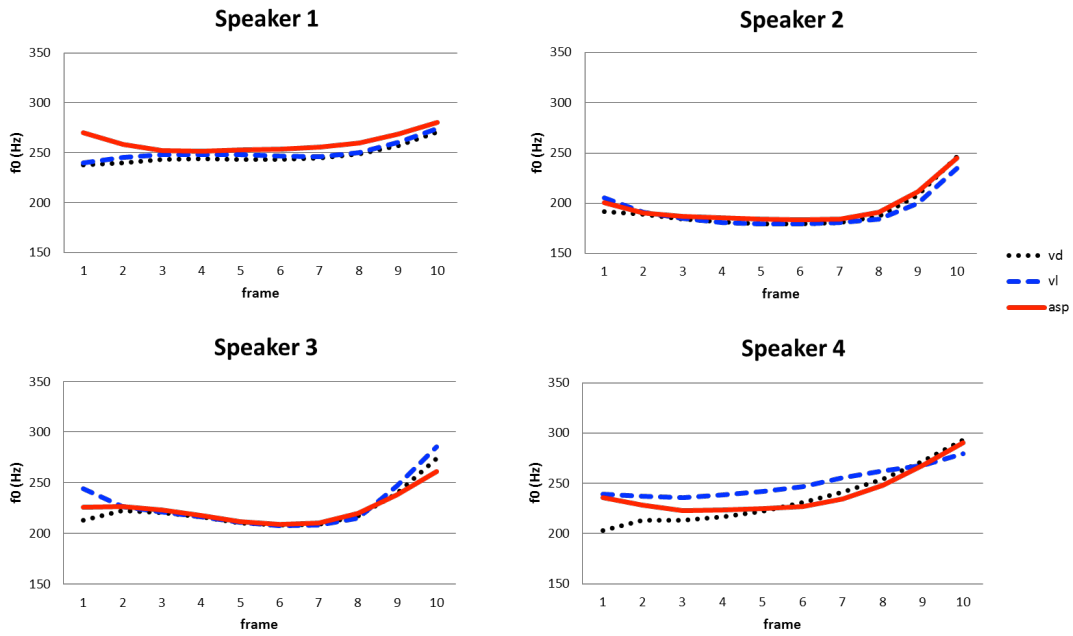


Figure 2: Time-normalized f0 contours of the following vowels for each speaker. Voiced (vd), voiceless (vl), and aspirated (asp) categories are represented by black dotted lines, blue dashed lines, and red solid lines, respectively.

#### 4. Conclusions

The present findings provide instrumental and empirical data to understand the sound system of (Eastern) Tamang. Our results reveal that the laryngeal contrast in Tamang is best characterized as a contrast among fully voiced, non-aspirated voiceless and heavily aspirated stops. Further, there was no close correlation between tonal patterns and VOT. Altogether, the acoustic data in the current study corroborate Yonjan (1993)'s phonological system. Still, it is not clear if this is a dialectal difference or a tone loss, which will be left for future studies.

#### References

- Lee, S.-W. (2011) Eastern Tamang Grammar Sketch. MA thesis, Graduate Institute of Applied Linguistics.
- Mazaudon, M. (1973) Phonologie Tamang: Étude phonologique du dialecte Tamang de Risiangku. Paris: Centre National de la Recherche Scientifique, Société d'Études Linguistiques et Anthropologiques de France.

- Mazaudon (2014) "Studying emergent tone-systems in Nepal: Pitch, phonation and word-tone in Tamang", *Language documentation & Conservation* 8.
- Varenkamp, B. (to appear) "Tamang grammar sketch", Unpublished manuscript.
- Yonjan, A. (1993) *Tamang bhasha bolchal* [तामाङ भाषा बोल्चाल]. Kathmandu: Nepal Tamang Damphu Samaj.

## Tonology of Daegu Korean and an OT solution

Jaehyun Son (Duksung women's University) & Yunjing Li (Tianjin Foreign Studies University)

In this paper we will present the pitch patterns of Daegu Korean words as well as the rule of how a compound gets its accent, and an OT solution of the data, among others.

### 1. Segments and syllable structure

The phonemes of this dialect are as follows: short vowels /i, u, e, o, ə, a/; long vowels /i:, u:, e:, o:, ə:, a:/; semivowels /w, j/; consonants /p, p<sup>h</sup>, p', t, t<sup>h</sup>, t', ʃ, ʃ<sup>h</sup>, ʃ', k, k<sup>h</sup>, k', m, n, ŋ, s, s', h, r/.

The syllable structure in Daegu Korean, and maybe also in other Korean languages, can be represented by a basic structure like CGVC, in which C stands for consonants, G for glides or semi-vowels, and V for vowels. V is believed indispensable, and all other elements are optional. Hence eight possible types of syllables allowed. They are V, VC, CV, CVC, GV, GVC, CGV and CGVC (Kim & Shibatani 1976).

### 2. Pitch pattern inventory

Several scholars have described the tone system of Daegu dialect (Kim 1994a, 1994b; Rah 1974; Son 2017; etc.). Our work will be mainly based the results of Son (2017), which is the latest one.

In Daegu Korean dialect, as in many other tonal Korean dialects, every word has a tonal pattern, which is contrastive. According to Son (2017), there are n+1 oppositions for n-syllable words. This accent type is the most widely distributed in the region of north Gyeongsang-do. These are shown in table (1) with examples.

(1)

syllables types	1	2	3	4	5
①	long vowel pe: 'twofold' R	u:.pi 'umbrella' H:F	ma:.nu.ra 'wife' H:HL	i:.sɔn.sɛŋ.nim 'Mr. Lee' H:HLL	pa:n.to.ʃ <sup>h</sup> e.san.ɔp 'semiconductor industry' H:HLLL
	short vowel pe 'boat' F	ka.ʃi 'branch' HF	mu.ʃi.ke 'rainbow' HHL	jo.ʃa.ʃ <sup>h</sup> in.ku 'girlfriend' HLLL	kuk.p'i.ju.hak.s'ɛŋ 'student with government expence' HLLLL
②	pe 'pear' F	u.ri 'cage' HL	mjo.nu.ri 'daughter-in-law' HLL	nam.ʃa.ʃ <sup>h</sup> in.ku 'boyfriend' HLLL	son.ʃa.mjo.nu.ri 'granddaughter-in-law' HLLLL
③		u.ri 'we' LF	mi.na.ri 'watercelery' LHL	sin.hon.jo.heŋ 'honeymoon' LHLL	ʃaŋ.ko.ri.jo.heŋ 'long trip' LHLLL
④			sa.ta.ri 'ladder' LLF	kim.sɔn.sɛŋ.nim 'Mr. Kim' LLHL	a.k <sup>h</sup> a.si.a.k <sup>h</sup> ot 'acacia blossoms' LLHLL
⑤				ka.ur.p'a.ram 'autumn breeze' LLLF	ɔm.ʃi.son.k'a.rak 'thumb' LLLHL
⑥					pi.heŋ.ki.jɔr.rjo 'plane fuels' LLLLF

In table (1), H, L, R and F stand for high tone, low tone, rising tone and falling tone respectively. Each syllable should and can only be linked with one of these four tones.

The pitch patterns for words are dubbed ①, ②, ③, ④ and so forth. Words of Type ① consisting of one syllable can be realized into two surface forms as respect to the length of the vowel in it. For words with a long vowel, the tone will be R, while for those with a short vowel, the tone will be F. Words of Type ① consisting of two and more syllables are not affected by vowel length other than the duration itself, and the pitch patterns are HF, HHL, HHLL and so forth by adding L to the end of it to make it longer.

Words of the types other than Type ①, i.e. Type ②, ③ and so forth, can have only one H or F in each of them. The number for their naming indicates the exact position of the syllable with an H or F in the word counted from the left.

The H/F element in a tone of non-① type an accent kernel. Hence Type ① is a kernel-less tone, and all other types are kernel-having tones.

The syllable of a monosyllabic words with a short vowel of both Type ① and Type ② should be surfaced as an F tone. They are differentiated based on what they will behave in compound words, which we will talk about in later sections.

### 3. The compound rule

When two words form a compound, mostly the word on the left keeps its accent, and the word on the right loses its accent, except when the first word is final-accented and the second word is not a ① type word, in which case, the word on the right keeps its accent, and the word on the left loses its accent. If there are three or more componential parts in a compound word, they will be combined each one to the neighboring one cyclically.

### 4. Phonological analysis

#### 4.1 Making the system simpler

It will be clearer to observe the system if we eliminate the examples as shown in (2).

The circled numerals at the heads of the rows indicate the position of the accent kernel. The accent kernel takes place between the accented syllable and the next syllable.

We have interpreted words of Type ① as kernelless in Daegu dialect. While kernel may remain in the second elements of compound nouns with other types, ① does not keep its pitch patterns.

A word with a long tone at the beginning of the word (which is referred to as a long vowel syllable) appears in a slightly ascending tone in a single syllable word alone, but appears in a high tone in more than one syllable. It is similar to the tone of HF, HHL, HHLL, ... at the point of tonal type. The differences can be attributed to the difference of the vowel length, and is combined to make ① types. The long vowel is regarded as a phoneme, and not the reflection of accent difference.

(2)

syllables types		1	2	3	4	5
①	long vowel	R				
	short vowel	F	HF	HHL	HHLL	HHLLL
②		F	HL	HLL	HLLL	HLLLL
③			LF	LHL	LHLL	LHLLL
④				LLF	LLHL	LLHLL
⑤					LLLF	LLLHL
						LLLLF

By observation, we can easily see that a) in this system, R, F and H are not contrastive; b) if we assume there is always a boundary tone L% attached to the end of a word, then F at the end may be

actually H. (H+L% linked to the same syllable.) and c) if we assume there is a boundary tone %L attached to the beginning of a word, and can only be realized on a long vowel, in isolation, then R may be also H actually. (%L+H linked to the same syllable.) So we can combine R, F and H together as H, and the system will be simplified as having only H and L, as shown in (3)

(3)

syllables types	1	2	3	4	5
①	H(H)	HH	HHL	HHLL	HHLLL
②	H	HL	HLL	HLLL	HLLLL
③		LH	LHL	LHLL	LHLLL
④			LLH	LLHL	LLHLL
⑤				LLLH	LLLHL
⑥					LLLLH

After doing this, we can find some traits of the system easily. a) Every word must have one and only one **accent**. b) Tone bearing unit (TBU) should be syllable. c) In words of Type ①, the accent is realized as HH on the first two syllables, or an H if monosyllabic. d) In words of other types, the accent is realized as an H on a specific syllable. e) Unaccented syllables are always linked with an L tone.

#### 4.2 Formalizing the compound rule

We have described the compound rule in words. To make it phonologically more formalized, we still can do something.

In the compound rule, words are actually classified into three groups. They are words of Type ①, words of other types with an accent or H at the last syllable, and words of other types other than Type ① with an accent at syllables of another position than the last one. If we name these three groups of words respectively as A, B and C, as shown in (4), the compound rule can be rewritten into phonological rules as in (5).

(4)

A	words of ① type	H of Type ①, HH, HHL etc.
B	final-accented words	H of type ②, LH, LLH, etc
C	other words	HL, HLL, LHL, LLHL, LHLL

(5) (O means the accent of the word is lost in the compound)

A+A → AO e.g. HH+HH → HHLL	B+A → BO e.g. LH+HH → LHLL	C+A → CO e.g. HL+HH → HLLL
A+B → AO e.g. HH+LH → HHLL	B+B → OB e.g. LH+LH → LLLH	C+B → CO e.g. HL+LH → HLLL
A+C → AO e.g. HH+HL → HHLL	B+C → OC e.g. LH+HL → LLHL	C+C → CO e.g. HL+HL → HLLL



#### 4.3. An Optimality Theory (OT) analysis of the compounding rule

OT works with a set of ordered constraints, an input and the candidates generated from the input, in a tableau. The constraints we are supposing for the present analysis are as follows.

##### (6) Constraints

**ALIGN-A-LEFT:** If an Type A word keeps its accent, but not on the left side of a compound, assign a \* to the candidate.

**ALIGN-B-RIGHT:** If an Type B word keeps its accent, but not on the right side of a compound, assign a \* to the candidate.

**MAX-ACCENT-LEFT:** The left part of a compound should keep its accent in a compound.

**\*LAPSE:** if neither of the two parts of a compound has an accent, assign a \* to the candidate.

**\*CLASH:** if both of the two parts of a compound has an accent, assign a \* to the candidate

Among the five constraints we have proposed, there are three faithfulness constraints, ALIGN-A-LEFT, ALIGN-B-RIGHT and MAX-ACCENT-LEFT, and two markedness constraints, \*LAPSE and \*CLASH.

ALIGN-A-LEFT and ALIGN-B-RIGHT were proposed based the observation that in a word of Type A, the accent is always on the left most syllables, while in a word of Type B, the accent is always on the right most syllable. MAX-ACCENT-LEFT was proposed because among all the nine possible combinations of the three types, A, B and C, seven produce compound words with the left part keeping its accent, and the right part losing its accent. \*LAPSE and \*CLASH were proposed based on the observation that in the process of compounding two words together into one, only one of the two componential words can keep its accent. The five constraints are then arranged in a hierarchy like the following.

##### (7)

\*LAPSE, \*CLASH >> ALIGN-A-LEFT, ALIGN-B-RIGHT >> MAX-ACCENT-LEFT

All the nine possible combinations go through this grammar and get their respective correct output. Here are some of them as example.

##### (8) A + C → A-O

input: A+C	*LAPSE	*CLASH	ALIGN-A-LEFT	ALIGN-B-RIGHT	MAX-ACCENT-LEFT
→ A-O					
O-C					*!
O-O	*!				*
A-C		*!			

##### (9) B + A → B-O

input: B+A	*LAPSE	*CLASH	ALIGN-A-LEFT	ALIGN-B-RIGHT	MAX-ACCENT-LEFT
→ B-O				*	
O-A			*		*!
O-O	*!				*
B-A		*!	*	*	

## (10) B+B → O-B

input: B+B	*LAPSE	*CLASH	ALIGN-A-LEFT	ALIGN-B-RIGHT	MAX-ACCENT-LEFT
B-O				*!	
→ O-B					*
O-O	*!				*
B-B		*!		*	

## (11) B+C → O-C

input: B+C	*LAPSE	*CLASH	ALIGN-A-LEFT	ALIGN-B-RIGHT	MAX-ACCENT-LEFT
B-O				*!	
→ O-C					*
O-O	*!				*
B-C		*!		*	

## (12) C+B → C-O

input: C+B	*LAPSE	*CLASH	ALIGN-A-LEFT	ALIGN-B-RIGHT	MAX-ACCENT-LEFT
→ C-O					
O-B					*!
O-O	*!				*
C-B		*!			

## 5. Discussion

## 5.1 Basic vs. Non-Basic

From the above analysis, it appears that Type A and B are more basic than Type C, which is always derivable from the former two.

Evidence 1:

Monosyllabic words are either Type A or Type B, which can be identified by adding a suffix or making a compound.

Evidence 2:

C is derivable from B and A's combinations according to the compound rule as in , while A is not derivable from other B and C, and B is not derivable from A and C.

## (13)

B+A → BO e.g. LH+HH → LHLL (Type C)

Evidence 3:

In OT grammar, A and B are actively trying to keep their accent (ALIGN-A-LEFT & ALIGN-B-RIGHT), while C is not specified in any constraints.

So, if we are forced to differentiate basic ones from non-basic ones, Type A and B seem more basic than Type C.

## 5.2 Left accented, Right accented, and Non-directionally accented

Enlightened by the OT analysis, we might be safe to claim that the three Types of words, A, B and C are having accents of different traits. The accent associated with words of Type A is left oriented,

and always tends to stick to the left end. The accent associated with words of Type B is right oriented, and always tends to stick to the right end. And the accent associated with words of Type C is non-oriented. The left accent surface itself as HH on the left end of a word, and the right accent surface itself as an H on the right end of a word. The non-oriented accent is surfaced as an H on any one syllable except the last one.

### 5.3 Pitch pattern inventory revisit

If we start deriving the pitch patterns of this language from a monosyllabic word of Type A and another monosyllabic word of Type B through the compound rule, we **can and only can** derive the whole inventory of the pitch patterns shown in (14). No more, no less.

In table 14, A stands for a monosyllabic Type A word, B for a monosyllabic Type B word, X for a monosyllabic word of Type A or B, XX for a disyllabic word of any kind, and etc.

(14)

types \ syllables	1	2	3	4
① Combinations which can producing the pattern	H(H) A	HH A+A A+B	HHL A+XX HH+X	HLLL A+HLL HH+XX HHL+X
① Combinations which can producing the pattern	H B	HL B+A	HLL B+HH HL+X	HLLL B+HHL HL+XX HLL+X
② Combinations which can producing the pattern		LH B+B	LHL LH+A B+HL	LHLL B+HLL LH+HH LHL+X
③ Combinations which can producing the pattern			LLH B+LH LH+B	LLHL B+LHL LH+HL LLH+A
④ Combinations which can producing the pattern				LLLH B+LLH LH+LH LLH+B

So, it might be safe to claim that the whole inventory of pitch patterns in this language is built on three things: a) A left accent, b) A right accent, and c) A compound rule.

Multisyllabic non-compound simple words **can and only can** pick a pitch pattern from this inventory. Pitch patterns not found in this inventory like \*LHHLL, \*LLHH, \*HHHL etc are not grammatical.

### References

- Kim, Chakyun. 1994a. 60tay Taykwu pangengkwa 20tay Taykwu pangeng sengcouy tayco punsek. [Comparative analysis of the tone of 60's and 20's Daegu dialect]. Enehak 16: 23-80. The Linguistics Society of Korea.
- Kim, Chakyun. 1994b. Taykwu pangen sengco cheykyeyuy pyenchen. [A comparative analysis of tones between the older generation's Daegu dialect and the younger one]. Ene 15: 101-366. The Linguistics Society of Korea.
- Kim, K-O & Shibatani, M. 1976. Syllabification phenomena in Korean. 『言語研究』 12: 91-98.
- Rah, Sungsook (羅聖淑) 1974. 「韓国語大邱方言の音韻 — アクセントを中心に」 『言語研究研究』 66, pp.1-44.
- Son, Jaehyun (孫在賢). 2017. 『韓国語諸方言のアクセント体系と分布』, Chaek-Sarang, Seoul, Korea.

## 日本語母語話者の L2 英語発音評価を構成する音声特性

小西隆之(早稲田大学大学院)・近藤真理子(早稲田大学)  
tkonishi@aoni.waseda.jp

### 1. はじめに

#### 1.1. 概要

本研究は、大規模 L2 英語音声コーパスを用い、日本語を母語とする英語学習者(以下「日本語母語話者」)が日本語訛りの英語(以下「日本語英語」)の発音評価において、分節音と韻律を相対的にどの程度重視しているかを検証する。

#### 1.2. 習熟度による日本語英語の発音の変化

日本語話者にとって、英語の分節音は、知覚・産出の両面において、韻律よりも習得が困難であることが先行研究により示唆されている。特に緊張・弛緩母音(/i/-/ɪ/)対立の聞き分け(Morrison 2002 等)や曖昧母音(schwa, /ə/)産出時の音質(Lee et al. 2006, Konishi & Kondo 2015 等)は母語の影響を強く受け、上級話者になってもその克服が困難であることが多い。

Saito et al.(2015)が英語母語話者による日本語英語の音声の評定値を分析した結果、「母語訛りの強さ(accentedness)」については全習熟度の学習者の発話の評定において分節音と韻律両方の正確さが影響を与えていたが、「理解容易度(comprehensibility)」については、韻律の正確さは全習熟度の学習者の発話の評定に影響を与えていたものの、分節音の正確さが影響を与えていたのは中・上級学習者の発話のみであった。

以上から、日本語母語話者の英語に関して、韻律は分節音よりも習得が容易で、かつ分節音よりも重要な音声特性であるという傾向が見られる。

#### 1.3. 本研究の目的

本研究は、先行研究により示された習熟度別の発音変化と同様の変化が、日本語母語話者の知覚においても起こっているという仮説を検証する。すなわち、韻律の正確さは初級段階からある程度判断される一方、分節音の発音の正確な判断は上級学習者のみになされるという仮説を検証する。

英語母語話者が学習者英語(以下「L2 英語」)を評定し、母語訛りや理解容易度に影響を与える音声特性を分析した先行研究は存在するが、英語学習者が自身と同じ母語訛りを持つ L2 英語の発音の良し悪しを判断した研究は管見の及ぶ限り存在しない。

本研究の結果は、日本語母語話者がそれぞれの習得段階において、L2 英語のどのような音声特性に注目し、それを習得目標としているか、また、その習得目標が英語コミュニケーションの見地から正しいかどうかという問題に対する示唆を与え、英語教育等の諸分野に応用されることが期待される。

## 2. 分析

### 2.1. 音声

評定対象の音声には、大規模 L2 英語音声コーパスである J-AESOP コーパスの中の *The North Wind and the Sun* の読み上げ文(男声 68 名分、女声 115 名分)を用いた。J-AESOP コーパスは日本、台湾、韓国、タイなどの研究機関が共同で行っている L2 英語音声コーパスプロジェクト AESOP(*Asian English Speech cOrpus Project*)によって構築された日本語英語のコーパスである。AESOP は各国で共通のプラットフォームの下にアジア英語音声コーパスを構築している。被験者に呈示した音声は、*The North Wind and the Sun* を三分割して(表 1)ランダム化したものである。三分割したのは、3 回の評定結果を平均することでデータの精度を上げるためと、局所的な読み間違いやつかえなどの非流暢性の印象が発話全体に及ぶのを避けるためである。また、音声の発話者全員が非母語話者であるというバイアスを与えないために、ダミーとして 25 名の英語母語話者の音声を同様に三分割したものを加えた。従って評定対象とした音声は 624 トークン(日・英語母語話者計 208 名 x 3 トークン)である。

表 1 三分割された *The North Wind and the Sun*

<b>Section 1</b>	The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other.
<b>Section 2</b>	Then the North Wind blew as hard as he could, but the more he blew, the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt.
<b>Section 3</b>	Then the Sun shone out warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

### 2.2. 評定者

本研究では、日常的に英語を用い、音声学または関連分野の修士号以上の学位を持つ評定者(n = 16; 以下「専門家群」)の評定値を基準とし、音声学の知識のない様々な習熟度の日本語母語話者(n = 27; 以下「被験者群」)の評定値が専門家群のものとの程度一致しているかを以て被験者群の評定の正確さとした。

専門家群は、1)アメリカ英語母語話者 4 名、2)日本語母語話者 4 名および 3)その他の言語の母語話者 8 名(ドイツ語、フランス語、スペイン語、ポーランド語、中国語(北京語)、広東語、韓国語、パンジャブ語)により構成される。その他の言語の母語話者を評定者に加えたのは、近年の L2 英語話者の増大に伴う世界英語(*English for International Communications*,

World Englishes, English as a Lingua Franca 等)の概念を反映してのものである<sup>1</sup>。

被験者群は、3 ヶ月以上の海外滞在経験を持たず、インターナショナルスクール等における日本語以外の言語での教育を一切受けていない国内大学の学部1,2年生で構成され、その全員が評定作業の前後5週間以内に Educational Testing Service 社の TOEFL-ITP<sup>®2</sup>を受験している。習熟度別の分析を行うため、ヨーロッパ言語共通参照枠(Common European Framework of Reference for Languages; 以下CEFR)に基づいた被験者群の分類を行った。CEFRは、言語能力を、C2を最高レベルとしてC2, C1, B2, B1, A2, A1の6段階に分けている<sup>3</sup>。ETS Global<sup>4</sup>のCEFR換算表(表2)を用いると、被験者群はListeningセクションのスコア38以上47未満(A2レベル)の初級群(n=7)、47以上54未満(B1レベル)の中級群(n=12)、54以上64未満(B2レベル)の上級群(n=8)の3群に分かれた。なお、Listeningセクションのスコアが38未満もしくは64以上の被験者はいなかった。また、比較対象として、音声学の知識のない英語母語話者(n=2; 以下「母語話者群」)も被験者群に加え、同様の評定作業を依頼した。

表 2 TOEFL-ITP スコアの CEFR 換算表 (ETS Global のホームページを参考に作成)

CEFR level	Total score	Listening	Structure	Reading
C1	627	64	64	63
B2	543	54	53	56
B1	460	47	43	48
A2	337	38	32	31

### 2.3. 評定基準

専門家群の評定値は J-AESOP コーパスの正式な評定値となっており、評定項目は a)分節音の正確さ、b)韻律の正確さ、c)流暢さ、d)英語母語話者度(nativeness<sup>5</sup>)の4項目で、いずれも1から10の10尺度(10が最良)である。今回は4つの評定項目のうち、分節音と韻律に該当する a)と b)を用いた。なお、専門家群の各々の評定者の評定値について多重相関分析を行ったところ、概ね相関係数0.6から0.8の高い一致率があることが示された。

被験者群の評定値は「発音が良いか悪いか」という1項目のみで、評定項目名によりバイ

<sup>1</sup> 詳細は Crystal (2003)等参照。

<sup>2</sup> [https://www.ets.org/toefl\\_itp/](https://www.ets.org/toefl_itp/)

<sup>3</sup> C2, C1 は「熟達した言語使用者」、B2, B1 は「自立した言語使用者」、A2, A1 は「基礎段階の言語使用者」とされている。詳細は Council of Europe のホームページ(<https://www.coe.int/en/web/common-european-framework-reference-languages>)参照。

<sup>4</sup> ETS Global (<https://www.etsglobal.org/Tests-Preparation/The-TOEFL-Family-of-Assessments/TOEFL-ITP-Assessment-Series/Scores-Overview>)

<sup>5</sup> 一般的には「母語訛りの強さ(accentedness)」という用語が用いられるが、今回の評定でこの項目名を用いると、高い評定値が訛りの度合いが「低い」ことを表すのか「高い」ことを表すのかが曖昧で、評定者が混乱する恐れがあったため、対立概念の「英語母語話者度(nativeness)」を用いた。

アスが生じるのを避けるため、「10(Good)-1(Not good)の 10 尺度で評定するように」という指示を与え、その他に「発話全体の印象で評定する」、「読み間違いに起因する繰り返しによって評価を下げない」等の詳細な説明を加えた(表 3)。

表 3 被験者群の評定基準

10 (Good) - 1 (Not good)	局所的な発音ではなく発話全体の印象で評定する (必ず発話全体を聴き終えてから評定する) Rate with global, rather than local, impressions. (Listen to the entire utterance before you rate.)
	個人の声質(たとえばアナウンサーのような声)に由来する聞きやすさは考慮しない Disregard voice quality (e.g. news caster-like pronunciation) in evaluating the pronunciation.
	読み間違いに起因する繰り返しによって評価を下げない ただし、流暢でないと感じられる場合はこの限りではない Ignore repetitions as far as the whole utterance sounds fluent.
	なるべく全ての数字を使って評定する (ネイティブでなくても発音がよければ10にする) Use the whole scale: 1 to 10. 10 can be used for nonnative speakers, too.

## 2.4. 評定作業

評定作業は専用のプログラムを作成して被験者に配布し、自宅、大学等の任意の静穏環境下で行ってもらった。専門家群の評定所要時間は平均 7 時間程度、被験者群の評定所要時間は平均 3.5 時間程度だった。評定基準が揺らぐのを防ぐため、被験者群には作業開始から 3 日以内に評定を完了してもらった。専門家群には大学の専任教員も含まれ、作業の所要時間も長かったため同様の制約は設けなかった。

## 2.5. 評定値に影響する要因の分析

各被験者群の評定平均値を従属変数とし、専門家群の「分節音の正確さ」と「韻律の正確さ」各々の評定平均値およびそれらの交互作用を独立変数とする重回帰分析を行った。分析には R 3.5.0 の *lm()* 用いた。

## 3. 結果

### 3.1. 初級群

「韻律の正確さ」が有意な予測変数となっていた( $t(545) = 15.340, p < .001$ )。「分節の正確さ」および交互作用は有意ではなかった( $ps > .1$ )。

### 3.2. 中級群

初級群と同様に、「韻律の正確さ」が有意だった( $t(545) = 11.750, p < .001$ )。「分節音の正確さ」および交互作用は有意ではなかった( $ps > .6$ )。

### 3.3. 上級群

「分節音の正確さ」、「韻律の正確さ」の両方が有意な予測変数となっていた(それぞれ  $t(545) = 2.960, p < .005$ ;  $t(545) = 14.991, p < .001$ )。また、交互作用も有意だった( $t(545) = -3.545$ ,

$p < .001$ ).

### 3.4. 母語話者群

上級群と同様に、「分節音の正確さ」、「韻律の正確さ」の両方が有意な予測変数となっており(それぞれ  $t(545) = 5.354, p < .001$ ;  $t(545) = 2.560, p < .05$ )、交互作用も有意だった( $t(545) = -3.081, p < .005$ ).

### 3.5. 上級群と母語話者群の比較

上級群と母語話者群は両方の独立変数および交互作用が有意であるという点で共通しているが、独立変数の相対的な寄与度に差が見られた。両群の結果を示した表 4 における各々の独立変数の予測値を比較すると、上級群においては「韻律の正確さ」の寄与度が大きかったのに対し、母語話者群においては「分節音の正確さ」の寄与度が大きかったことがわかる。

表 4 上級群と母語話者群の回帰分析結果

		予測値	標準誤差	t 値
上級群	分節音の正確さ	0.247210	0.083514	2.960 **
	韻律の正確さ	0.948843	0.063293	14.991 ***
母語話者群	分節音の正確さ	1.13931	0.21280	5.354 ***
	韻律の正確さ	0.41286	0.16128	2.560 *

\*  $p < .05$     \*\*  $p < .01$     \*\*\*  $p < .001$

## 4. 考察

まず、「韻律の正確さは初級段階で既にある程度正確に判断される一方、分節音の正確さは上級学習者のみに正確に判断される」という仮説が正しかったことが分析結果により証明された。「韻律の正確さ」は全ての被験者群の評定値に有意に寄与していた一方で、「分節音の正確さ」が評定値に影響を与えたのは上級群と母語話者群のみであった。このことから、「韻律の正確さは初級レベルから上達していく一方、分節音の正確さは上級レベルにならないと習得されない」という Saito et al. (2015)の分析結果が、本研究により知覚において再現されたと言える。

また、上級学習者は L2 英語における分節音と韻律両方の正確さを判断できていたが、相対的に韻律の正確さを重視する傾向にある、もしくは、分節音よりも韻律の正確さの判断に長けているという傾向が示された。一方で、母語話者は韻律よりも分節音を重視して評定を行っていた。

今回の評定には読み上げ音声を用いられたため、理解容易度(comprehensibility)だけでなく英語母語話者度(nativeness/accentedness)に関する印象が評定に影響を与える可能性を排除できていない。今後、自発音声を用いた後続研究が行われることが期待される。

## 5. 結論

本研究の結果から、日本語母語話者は、初級レベル段階で既に正確な韻律を判断する能力を有していることがわかる。一方で分節音の正確さの判断は上級レベルにならないと習得



されないことが示唆された。この結果は産出に関する先行研究の結果とも一致する。本研究の結果は特に日本語母語話者への英語教育に応用されることが期待される。先行研究と本研究の結果から、日本語話者への英語発音指導において、韻律の指導は知覚・産出ともに初級段階から開始するのが効果的であるのに対し、分節音の指導は中上級段階において重点的に行うのが効果的であることが示唆される。

## 6. 謝辞

本研究は科研費若手研究(B) (17K13513)、基盤研究(B) (15H02729)および早稲田大学特定課題研究助成費(2018K-390)の助成を受けている。

## 参考文献

- Crystal, D. (2003). *English as a Global Language*, Cambridge: CUP.
- Konishi, T. & Kondo, M. (2015). Developmental Change in English Stress Manifestation by Japanese Speakers. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS XVIII)*.
- Lee, B., Guion, S. G., & Harada, T. (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition*, 28(3), 487-513.
- Morrison, G. S. (2002). Perception of English /i/ and /ɪ/ by Japanese and Spanish listeners: Longitudinal results. In G. S. Morrison & L. Zsoldos (Eds.), *Proceedings of the Northwest Linguistic Conference 2002* (pp. 29–48). Burnaby, BC: Simon Fraser University Linguistics Graduate Student Association.
- Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37(2), 217-240.
- Visceglia, T., Tseng, C. Y., Kondo, M., Meng, H., Sagisaka, Y. (2009). Phonetic Aspects of Content Design in AESOP (Asian English Speech cOrpus Project). In 2009 Oriental COCOSDA International Conference on Speech Database and Assessments. 60- 65. IEEE.

## 日本語母語話者による英語高前舌母音/i:, ɪ/の発話習得に関する研究

矢澤 翔・小西 隆之（早稲田大学大学院）・近藤 眞理子（早稲田大学）  
 k.yazawa3@kurenai.waseda.jp, tkonoshi@aoni.waseda.jp,  
 mkondo@waseda.jp

### 1. はじめに

#### 1.1. 概要

本研究は第二言語（L2）英語音声コーパス J-AESOP を用いて、日本語母語話者による英語高前舌母音/i:, ɪ/の発話が英語習熟度に応じてどのように変化していくかを検証し、同母音の知覚パターンとの関連性を考察する。

#### 1.2. 日本語母語話者による/i:, ɪ/の知覚と習得

英語の高前舌母音/i:, ɪ/の音響特性は質的・量的ともに異なり、緊張母音/i:/は周辺的で長く、弛緩母音/ɪ/は中心的で短い。しかし、英語母語話者は時間長を知覚の手がかりとせず、主に質で両母音を聞き分けることが知られている（Hillenbrand, Clark, & Houde, 2000）<sup>1</sup>。すなわち/i:, ɪ/の対立において時間長は余剰な音声的特徴であるが、日本語を母語とする英語学習者は時間長を主な手がかりとして両母音を区別する傾向にある。これは、日本語には長短母音/i:, ɪ/の区別が存在するため、緊張母音/i:/が長母音/i:/に（例：“leave” → 「リーブ」）、弛緩母音/ɪ/が短母音/i:/に（例：“live” → 「リブ」）それぞれ同化されるためと考えられる。日本語母語話者の/i:, ɪ/の知覚は第一言語（L1）が L2 に影響を及ぼす好例としてしばしば取り上げられるが、その習得に関しては明らかになっていない点も多い。Morrison（2002）によれば、日本語母語話者の/i:, ɪ/の知覚判別パターンは5ヶ月の英語圏滞在経験を経てもほとんど変わらず、時間長に依存していたことから、母語の影響は根強く残ることが予測される。その一方で、短期集中のトレーニングにより質に基づく知覚が習得可能であるとする研究（Fox & Maeda, 1999）や、十分なインプットを与えれば知覚は徐々に質を用いるように変化していくという研究（Yazawa, Kondo, & Escudero, 2017）も存在する。また、現在進めている別の研究は質的な知覚が習得可能であることを示唆しながらも、知覚パターンには個人差が大きく、英語習熟度が高くとも必ずしも質的な区別を行うとは限らないという結果となった。これらをまとめると、/i:, ɪ/の知覚に関して母語の影響は根強く、ネイティブらしい質的な知覚の習得は可能であるものの容易ではないようである。

#### 1.3. 研究目的

日本語母語話者による/i:, ɪ/の習得に関する発話研究は知覚研究と比べて乏しく、同母音の発話が発達的にどのように変化していくかはよく分かっていない。本研究では英語習熟度の異なる多数の日本語母語話者の発話を収録した L2 英語音声コーパス J-AESOP を用いて、英語母音/i:, ɪ/の質的・量的な実現方法と英語習熟度の関連性を調べる。

<sup>1</sup>質と量の両方を知覚の手がかりとする英語方言も存在するが（例：南イギリス英語）、本稿では日本における英語学習者が最も馴染みのあるであろう北米英語を前提とする。

## 2. 手法

### 2.1. データ

J-AESOP コーパス内の日本語母語話者 183 名（男性 68 名・女性 115 名）の音声データを用いて分析を行った。J-AESOP コーパスは、アジア諸国の研究機関と共同で進行中の L2 英語音声コーパス構築プロジェクト AESOP (Asian English Speech cOrpus Project; Meng et al., 2009) の一部で、日本語母語話者の英語発話音声データが収録されている。分析対象としたのはイソップ寓話の「北風と太陽」(”The North Wind and the Sun”; International Phonetic Association, 1999) の読み上げ音声で、音素バランス文となっている。

対象の音声データには、Praat (Boersma & Weenink, 2018) の TextGrid 形式で分節音及び語のアノテーションが付与されている。これは HTK (<http://htk.eng.cam.ac.uk/>) と TIMIT コーパス (<https://catalog.ldc.upenn.edu/ldc93s1>) を用いた自動アラインメントの出力結果に、J-AESOP チーム内の音声学に通ずる作業者が手修正を加えたものである。

### 2.2. 評定値

J-AESOP コーパスの各話者には英語レベルの評定値が付与されている。これは、16 名の評定者が各話者の「北風と太陽」の発話を聞き、(ア) 分節音的正確さ、(イ) 韻律、(ウ) 流暢さ、(エ) ネイティブらしさの各基準に関して、他の 25 名の英語母語話者の発話を参照しながら 10 段階で評価した値である。評定者は音声学及び関連分野を専門とする修士号または博士号の取得者で、うち 4 名は日本語母語話者、4 名は英語母語話者、8 名は他言語話者である。

今回は分節音的正確さの評定値（以下「分節音スコア」）のみを分析に用いた。図 1 は各話者の平均分節音スコア（評定者ごとに平均）の分布で、平均値は 5.23、中央値は 4.65 となっている。

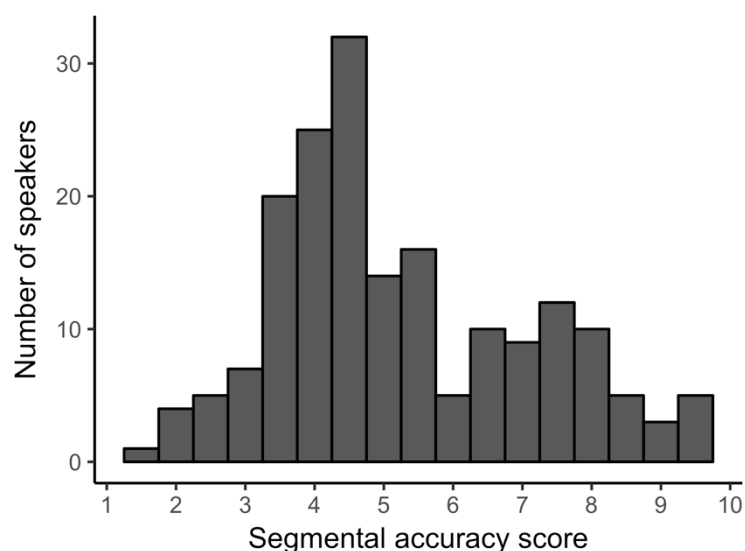


図 1: 分節音スコアの分布 ( $N = 183$ )

### 2.3. 音響値の測定と正規化

音声データとアノテーションに基づき、分析対象の/i:/ (n=2086) と/ɪ/ (n=3196) 及び他の単母音 (/ɛ, æ:, ʌ, ɔ:, u:, ʊ/) の第一・第二フォルマント (F1・F2) と時間長を Praat で測定した。フォルマント値は性別による個人差が大きく、時間長も習熟度による話速の変化に影響を受けやすいことから、両者とも正規化を行うことが望ましい。そこで、F1・F2・時間長それぞれに関して、全ての単母音の平均と標準偏差を求め、各母音の音響値を Z スコアで表す Lobanov 法 (Lobanov, 1971) を適用した。Lobanov 法は古典的なフォルマント正規化方法であるが、他の手法と比べても遜色なく効果的であり (Adank, Smits, & van Hout, 2004)、かつ時間長にもそのまま用いることができる利点がある。

### 2.4. 統計解析

各母音カテゴリ (/i:, ɪ/) に対して、分節音スコアを従属変数、正規化した F1・F2・時間長を固定効果、語と評定者をランダム効果とする以下のような線形混合モデルを作成した。モデルの作成には R (R Core Team, 2017) の lme4 パッケージ (Bates, Mächler, Bolker, & Walker, 2015) を、有意差の算出には lmerTest パッケージ (Kuznetsova, Brockhoff, & Christensen, 2017) を用いた。

```
lmer(score ~ F1.norm + F2.norm + dr.norm + (1|word) + (1|rater), data)
```

## 3. 結果

線形混合モデルによる解析結果を表 1 に示す。まず母音の質について、/i:/は F1 が低く F2 が高いほど分節音スコアが高くなる傾向にあった一方、/ɪ/に関しては F1 が高く F2 が低いほどスコアが高いという真逆の傾向が見られた。次に時間長に関して、両母音とも短いほど分節音スコアは高かったが、推定係数を比べるとスコアに対する時間長の貢献度は/i:/よりも/ɪ/の方が大きいことが分かる。図 2 と図 3 は正規化済みの F1・F2 値と時間長の分布で、習熟度別の変化を可視化するために話者を分節音スコアの高い順に「high」「mid」「low」の 3 郡 (各 61 名) に分割してある。これらの結果をまとめると、話者の分節音的正確さに関する習熟度が高いほど、①両母音は母音空間上で相対する動きをし、次第に離れていくこと、そして②時間長は両母音とも短くなるが、/i:/と比べて/ɪ/の方が減少の程度が大きいことが明らかになった。

表 1: 線形混合モデルの結果 (\*\*\*) =  $p < .001$

	F1			F2			Duration		
	estimate	S.E.	t	estimate	S.E.	t	estimate	S.E.	t
/i:/	-0.48	0.03	-17.69***	0.91	0.02	39.60***	-0.01	0.01	-6.62***
/ɪ/	0.27	0.02	15.79***	-0.86	0.02	-51.39***	-0.21	0.02	-13.80***

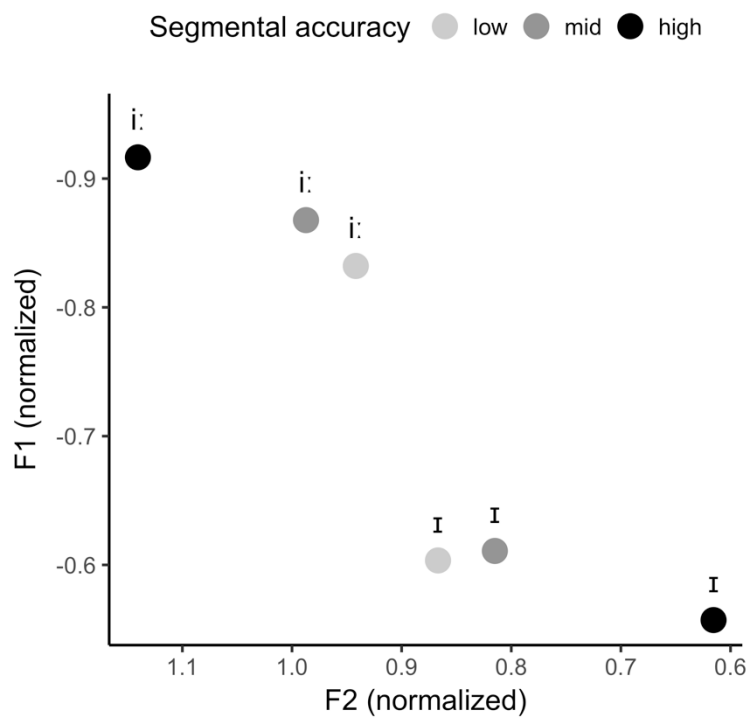


図 2: 習熟度別の正規化 F1・F2 値の平均

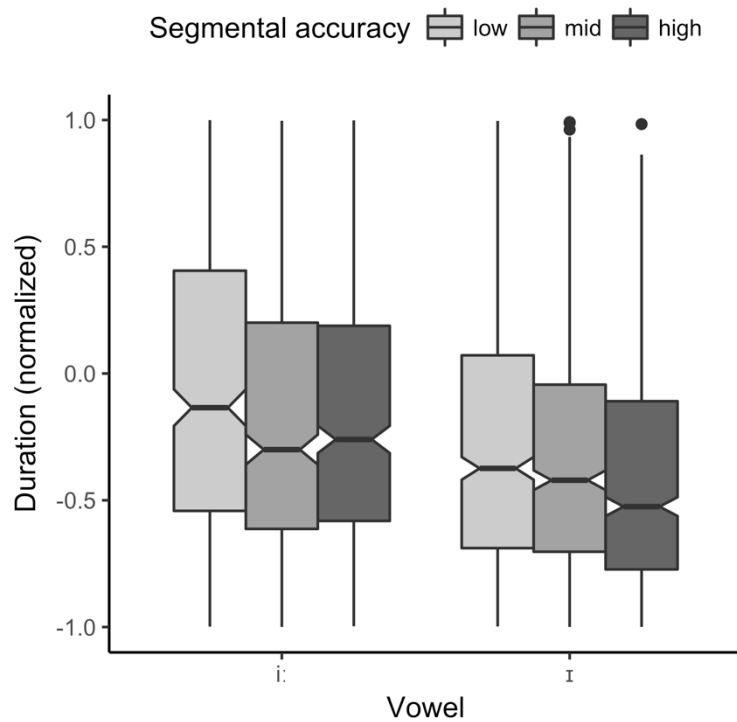


図 3: 習熟度別の正規化時間長の分布

#### 4. 考察とまとめ

本稿では J-AESOP コーパスを用いて、日本語母語話者による英語高前舌母音/i:, ɪ/の発話における発達的变化を調べた。その結果、英語習熟度が高くなるにしたがって、①両母音の質は次第に離れていき、②弛緩母音/ɪ/は特に短くなることが明らかになった。①の結果は、学習初期段階では/i:, ɪ/の音質は似ているが、習熟度が上がるに応じて英語母語話者に近い質的な対立が確立されていくことを示唆する。これは、知覚と産出の少なくとも部分的な重複 (Buchsbaum, Hickok, & Humphries, 2001)を踏まえると、同母音の知覚的手がかりが時間長から質に変化していくとする先行研究 (Yazawa, Kondo, & Escudero, 2017) の結果を支持するものである。

②に関して、弛緩母音/ɪ/により大きな変化が見られた理由としては英語の強勢リズムの習得が挙げられる。分析に用いた「北風と太陽」の読み上げ文には、/ɪ/に強勢が置かれず弱化する語が含まれる (例: “making”)。分節音スコアの高い話者は韻律的な習熟度も相対的に高く、非強勢母音を短く発話する強勢リズムを習得していると考えられることから、/ɪ/に見られた時間長の変化は、分節音的習得のみならず韻律的習得にも起因する可能性がある。今後の課題としては、分節音スコアに加えて韻律の評定値も分析の対象とし、超分節音的な要因も考慮した包括的な習得の軌跡を明らかにしたい。

#### 謝辞

本研究は、日本学術振興会特別研究員制度 (DC2) 及び科学研究費補助金 (15H02729、17K13513、18J11517) の助成を受けたものである。

#### 参考文献

- Adank, P., Smits, R., & van Hout, R. (2004) "A comparison of vowel normalization procedures for language variation research." *The Journal of the Acoustical Society of America* 116(5), 3099-3107.
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015) "Fitting linear mixed-effects models using lme4." *Journal of Statistical Software* 67(1), 1-48.
- Boersma, P., & Weenink, D. (2018) *Praat: doing phonetics by computer* (Version 6.0.40). <http://www.fon.hum.uva.nl/praat/>
- Buchsbaum, B. R., Hickok, G., & Humphries, C. (2001) "Role of left posterior superior temporal gyrus in phonological processing for speech perception and production." *Cognitive Science* 25(5), 663–678.
- Fox, M. M., & Maeda, K. (1999) "Categorization of American English vowels by Japanese speakers." *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1437-1440).
- Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000) "Some effects of duration on vowel recognition." *The Journal of the Acoustical Society of America* 108(6), 3013–3022.
- International Phonetic Association. (1999) *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017) "lmerTest package: Tests in linear mixed effects models." *Journal of Statistical Software* 82(13), 1-26.
- Lobanov, M. Boris (1971) "Classification of Russian vowels spoken by different listeners." *The Journal of the Acoustical Society of America* 49(2B), 606-608.
- Meng, H., Tseng, C., Kondo, M., Harrison, A., & Visceglia, T. (2009) "Studying L2 suprasegmental features in Asian Englishes: A position paper." *Proceedings of the 10th Annual Conference of the International Speech Communication Association* (pp. 1715-1718).
- Morrison, G. S. (2002) "Perception of English /i/ and /ɪ/ by Japanese and Spanish listeners: Longitudinal results." *Proceedings of the Northwest Linguistic Conference 2002* (pp. 29-48).
- R Core Team (2017) *R: A language and environment for statistical computing*. Vienna, Australia: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Yazawa, K., Kondo, M., & Escudero, P. (2017) "Modelling Japanese speakers' perceptual learning of English /i:/ and /ɪ/ within the L2LP framework." *Proceedings of the Phonetic Teaching and Learning Conference 2017* (pp. 115-119).

## 米ペンシルベニア州における英語の後舌狭母音/u/の前舌化： 地理的伝播と道路交通の関係を探る

木村 公彦（東京外国語大学大学院）  
kimura.kimihiko.q0@tufs.ac.jp

### 1 研究背景と目的

アメリカ英語の方言発展を考えるうえで母音の変化は特に重要な意味を持っており、例えば Labov, Ash, & Boberg (2006) は母音変化の地理的分布に注目することで方言区分を定めている。大きく分けると母音変化には 1. 方言区分を定める地域に特有なもの、2. 方言区分を超えて広域に分布するものの 2 種類がある。後舌狭母音/u/の前舌化は元々は前者に当たる母音変化であったが、時代を経るにつれてアメリカ全土に拡大し後者へと変化した。後舌狭母音の前舌化は他母音の連鎖シフトを引き起こすことがあり、以上のような地理的拡大は母音変化のトリガーとなり得る変化がアメリカ全土に拡大していることを示唆している。後舌狭母音の前舌化はこのように重要な変化であるが、ペンシルベニア州内部において 1940 年代～1990 年代の伝播の詳細は不明であった。本研究では特にペンシルベニア州南部のフィラデルフィアとピッツバーグの 2 都市に挟まれた地域（以下、ペンシルベニア中南部と呼ぶ）に注目して、研究史上の空白の一部を埋めることを試みる。

#### 1.1 後舌狭母音の前舌化

後舌狭母音の前舌化は/u/の調音位置が前寄りに変化する現象であり、後舌母音/oo/の前舌化をはじめとする英語の母音体系内での連鎖的な音変化を引き起こすトリガーとなり得る。前舌化の度合い（音響的には第 2 フォルマント（F2）の大小に対応）は前後の音環境に依存することが知られており、先行する子音が舌頂音の場合は前舌化の度合いは大きくなり、舌頂音以外が先行する場合にはその度合いは比較的小さくなる。また、/l/が後続する場合、前舌化は抑制されて起こらない。



図 1: アメリカの後舌狭母音の前舌化の歴史的発展（赤：1940 年代、紺：1990 年代）(Labov et al., 2006)



歴史的には後舌狭母音の前舌化はアメリカの一部地域でしか見られなかったものであり、Kurath & McDavid (1961) によると、1940 年には図 1 の赤色で示した領域 (Mid-Atlantic と South) でのみ報告されていた。しかし時代が下るにつれて変化が見られる地域が拡大していき、1990 年代には図 1 の紺色で示した領域 (West, North, Midland) にまで拡大した (Fridland & Macrae, 2008)。

## 1.2 ペンシルベニア州における後舌狭母音の前舌化

上記のように、今日では後舌狭母音の前舌化はアメリカ全土に拡大したが、歴史的な詳細を見ると各方言地域内での前舌化の伝播には未だ不明な点が存在する。本研究ではペンシルベニア州の南部に注目した調査を行った。これはこの地域に次のような研究史上のギャップがあるためである。

ペンシルベニア州内部では、既に 1940 年の時点には東部のフィラデルフィアと西部のピッツバーグで前舌化が確認されていた (Kurath & McDavid, 1961)。また当時の観察によると、フィラデルフィアの周辺に関していえば、フィラデルフィアから遠いほど前舌化の度合いは小さくなる傾向があった。このような記述はペンシルベニア内部には 2 つの前舌化拡大のノードが存在し、そこから周囲の地域へと変化が伝播していったことを示唆するものである。

1990 年代になると、それら 2 つのノードに挟まれたペンシルベニア中南部でも前舌化が確認されるようになった (Kopp, 1999)。しかし、そこに至るまでの 1940 年代～1990 年代の期間に渡って、具体的な伝播の様相に関しては観察されておらず、研究史上の空白となっている。そこで本研究ではこの約 50 年に渡る音韻史上の空白の一部を解明することを目指し、当該期間の半ばに当たる時期の後舌狭母音の前舌化に関する調査を行った。

## 2 調査手法

本研究では 2017 年に新たに拡張・公開されたインタビュー録音を使用し、この録音データの F1, F2 をパラメータとする音響分析を行った。以下、分析手法の詳細について述べる。

### 2.1 使用した録音データ

本研究で用いたデータは *Dictionary of American Regional English* 編纂で使用されたインタビュー録音であり、ウィスコンシン大学のデジタルアーカイブ (<https://uwdc.library.wisc.edu/collections/amerlangs/>) 上で、2017 年に新たに大幅拡充・公開されたものである。使用したデータは、すべて 1967 年～1969 年に録音されたものであり、文章読み上げと自由発話の 2 通りの形式がある。本発表では読み上げ録音と自由発話の両方に関する分析を行い、考察を行った。

本研究で使用した録音音声のインフォーマントに関する詳細を表 1 に示す。都市名に付けられた番号は図 2 上の番号と対応しており、インフォーマントは全員、録音当時で 50～70 代の女性である。表 1 の一番右側には自由発話録音の有無を示した。○が示されたインフォーマントに関しては文章読み上げと自由発話の両方の音声分析が可能だったことを意味する。×が付いているインフォーマントに関しては文章読み上げにのみが分析可能であった。

図 2 に示されているように、本研究で分析を行った地点はペンシルベニア中南部を東西および南北に分割する直線上に沿って分布する。このようなデータ点の取り方をしたことで、ペンシルベニア中南部を縦横に走る直線に沿った比較が可能となり、対象地域の地理的背景との関連性を複数の側面から考察できるようになった。

表 1: 本研究のインフォーマント

都市名	性別	年齢	録音年	自由発話録音
① エヴェレット	女性	54	1967	○
② カーライル	女性	68	1967	○
③ ニューブルームフィールド	女性	73	1969	×
④ ゲティスバーグ	女性	64	1969	○
⑤ サンベリー	女性	62	1968	○
⑥ マンハイム	女性	56	1967	×
⑦ ランカスター	女性	72	1967	○
⑧ テレヒル	女性	60	1967	×
⑨ フィラデルフィア	女性	79	1968	○

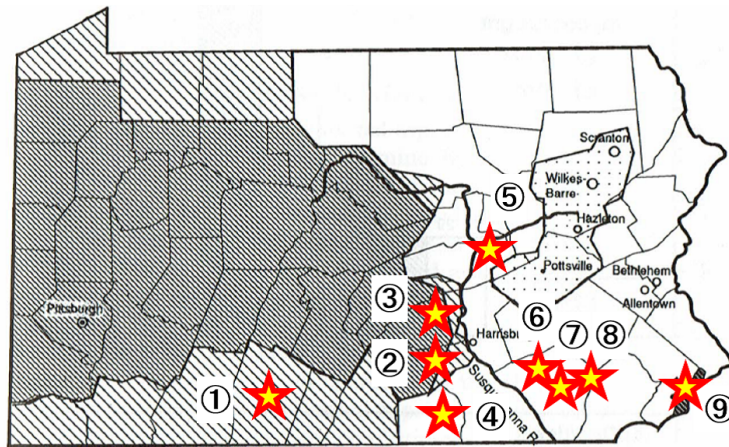


図 2: インフォーマントの地理的分布 (?)

フィラデルフィアは話者間の比較をする際の参照点として用いた。これはフィラデルフィアでは 1940 年代以降、後舌狭母音が前舌化することが知られているためである (e.g. Kurath & McDavid (1961))。

## 2.2 音響分析と母音空間の正規化

音響分析用ソフトウェア Praat を用いて、第 1、第 2 フォルマントをパラメータとする音響分析を行い、各インフォーマントの母音空間を可視化した。以下、分析に用いた手法についての詳細を述べる。分析の手順は大きく、(1) トークンの選別、(2) 第 1・第 2 フォルマントの測定、(3) 各話者の母音空間の正規化、の 3 段階に分かれる。

### 2.2.1 トークンの選別

最初に以下の 3 つの基準にしたがってトークンの選別を行った。

- 後舌低母音 /u/、/i/、/æ/ を含む
- 鼻子音に隣接しない
- 接近音に隣接しない

1 番目の基準は分析対象の後舌狭母音/u/に加えて、後述の母音空間の正規化に必要な/i/、/æ/が含まれるトークンを分析に使用したことを意味する。2 番目、3 番目の基準は前後の子音の影響で調査対象の母音の質が変化してしまうこと、すなわち鼻母音化および入渡り/出渡りの測定点への影響を防ぐこと目的に設定した。尚、本研究ではトークン数の制限から舌頂音が先行する/u/に関する場合に限って分析を行った。

### 2.2.2 フォルマンの測定

次に、分析対象の母音の第 1、第 2 フォルマンを以下の手順に従って測定した。

1. 波形の特徴的ピークが全て見えるように母音の始点と終点を決定
2. 隣接子音からの影響を避けるため、母音の持続時間の中央で測定
3. 聴覚的な Bark 尺度への変換 (Traunmüer, 1990)

1 番目と 2 番目の基準は、前後の子音の影響が出やすい遷移領域での測定を避けるために設定した。最後に、[Hz] の次元を持つ測定値を、Traunmüer (1990) の方法にしたがって、より聴覚に近い音響尺度である Bark へと変換した。

### 2.2.3 母音空間の正規化

2.2.2 で測定し、Bark に変換した各母音のフォルマンの平均値を出し、話者毎に母音空間図を作成した。しかし、こうして作成された母音空間図には話者の声道の長さをはじめとする個人差の影響があり、そのままでは話者横断的な母音空間の比較はできない。そこで本研究では、以下に示すような Watt & Fabricius (2002) の手法を用いて、各話者の母音空間を正規化した。

1. 仮想的な音素 /u' / (F1、F2 共に最小値を取ると仮定) を再構成
2. 3 点 /i/、/æ/、/u' / (母音空間上の隅) の重心 S を求める
3. フォルマンの測定値 (Bark) を重心 S で割り、相対フォルマンとして使用

本研究では、この正規化の手順を踏むことにより、母音空間の話者横断的な比較が可能となった。

## 3 結果と考察

音響分析の結果、文章読み上げ、自由発話それぞれに関して以下の図 3 と図 4 に示すような母音空間図を得た。これらの図中の黒塗りの星印は参照点であるフィラデルフィアを示している。i 以下、フィラデルフィアを基準として、前舌化の度合いを 3 つのクラス (i) 「小」: フィラデルフィアよりも F2 が低い、(ii) 「中間」: フィラデルフィアと同程度の F2 を持つ、(iii) 「大」: フィラデルフィアよりも F2 が高い、に分けて考える。

読み上げ録音の分析ではエヴェレットのみが「大」の前舌化を示し、マンハイム、サンベリー、ゲティズバーグは「小」に分類された。他の地点での前舌化の度合いは「中間」であった。一方、自由発話の分析ではエヴェレットとサンベリーが「大」、ゲティズバーグが

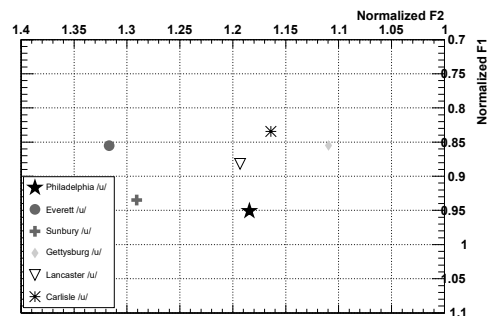
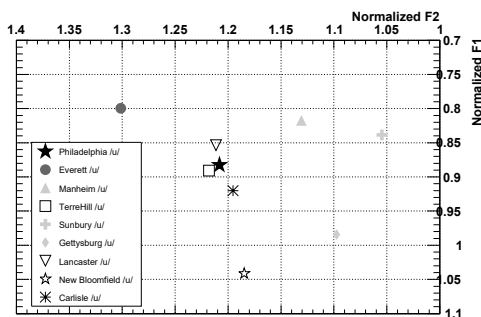


図 3: 読み上げ録音における/u/のフォルマント  
図 4: 自由発話録音における/u/のフォルマント

表 2: 文章読み上げと自由発話の間での前舌化度合いの比較

都市名		後舌狭母音の前舌化度合い	
		文章読み上げ	自由発話
①	エヴェレット	大	大
②	カーライル	中間	中間
③	ニューブルームフィールド	中間	-
④	ゲティスバーグ	小	小
⑤	サンベリー	小	大
⑥	マンハイム	小	-
⑦	ランカスター	中間	中間
⑧	テレヒル	中間	-
⑨	フィラデルフィア	中間	中間

「小」、それ以外は「中間」という結果になった。読み上げと自由発話との比較のために、これらの結果を表 2 にまとめる。

文章読み上げと自由発話の比較をすると、両者の前舌度合いは話者毎に概ね一致していることが分かる。より自然な発音を引き出しやすい自由発話の録音において、文章読み上げと同様の結果が得られたことから、1960年代では既に普通の会話でも前舌化を示している地域があったことが示唆される。

一方、サンベリーのインフォーマントについてのみ、読み上げでは「小」、自由発話では「大」という不一致が生じた。この話者は普通の発音では自由発話で見られたような「大」の前舌化を示すが、読み上げの際には注意深くなり、保守的な「小」の前舌化を示したと解釈できる。

今回、フィラデルフィア以上の前舌化（「大」と「中間」）が見られた地域は、概ねフィラデルフィアとピッツバーグを結ぶ直線状に並ぶことが分かった。この直線は幹線道路の道路交通量が多い地域と一致する傾向があり、このことより前舌化が交通に伴う人口の流れによって伝播したという、これまで不明であった前舌化の伝播要因の一部が示唆された。

#### 4 結論と今後の課題

本発表ではペンシルベニア中南部に注目をし、そこでの後舌狭母音 /u/ の前舌化の研究史における空白を埋めることを目指した。本研究では 1960 年代の録音音声を用いて、F1・F2 をパラメータとする音響分析を行うことで、この目的の一部を達成した。

本研究の結果から言えることは以下の 3 点である。

- 自由発話でも後舌狭母音の前舌化が観察されたことから、1960 年代の時点でペンシルベニア中南部に前舌化が既に伝播していたことが示唆された
- サンベリーの話者は読み上げで保守的な傾向を示したが、自由発話の結果から前舌化は既にサンベリーにも拡大していると考えられる
- 前舌化が見られた地域は道路交通量が多い地域に一致する傾向があり、このことから道路交通に伴う人口の流れが前舌化拡大の一因となった可能性が示唆された

本研究によって、道路交通の前舌化拡大への寄与の可能性が示されたが、これは両者が直接的に関与していることを示したのではない。このため、今後考察を深めるためには通勤・通学など、道路交通による具体的な人口の流れを考慮に入れる必要がある。また、サンベリーに関しては道路交通網だけでは説明できない傾向を示しており、地理的要因など、道路交通網以外の影響を考慮に入れる必要があると考えられる。

#### 参考文献

- Fridland, V., & Macrae, T. (2008) "Patterns of /uw/, /u/, and /ow/ fronting in Reno, Nevada." *American Speech*, 83, 432-454.
- Kopp, A. (1999) *The Phonology of Pennsylvania German English as Evidence of Language Maintenance and Shift*. Selinsgrove, PA: Susquehanna University Press.
- Kurath, H., & McDavid, R. I. (1961) *The pronunciation of English in the Atlantic States: based upon the collections of the linguistic atlas of the Eastern United States*. Ann Arbor: The University of Michigan Press.
- Labov, W., Ash, S., & Boberg, C. (2006) *The Atlas of North American English: Phonetics, Phonology, and Sound Change*. Berlin: Walter de Gruyter.
- Trautmüer, H. (1990) "Analytical expressions for the tonotopic sensory scale." *Journal of the Acoustical Society of America*, 88, 97-100.
- Watt, D., & Fabricius, A. (2002) "Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1 F2 plane." *Leeds Working Papers in Linguistics and Phonetics*, 9, 159-173.