DURATION FEATURE LEARNING FOR AN INTERVOCALIC VOICELESS VELAR STOP CONSONANT IN L2 JAPANESE

Akiko Kato[a] *& Felicity Cox*[b]

[a]Takushoku University, Japan; [b]Centre for Language Sciences, Macquarie University, Australia
katoakiko20@hotmail.com; felicity.cox@mq.edu.au

ABSTRACT

This study investigates how beginning learners of Japanese language (L2 Japanese) develop the durational features of an intervocalic consonant. Acoustic analyses were conducted on speech data collected from Australian learners of Japanese over a sixteen month period at the early stages of learning. The durations of the intervocalic consonant, and the component durations, voice onset time (VOT) and closure, were analysed with respect to preceding vowel length.

The results suggest that beginning learners have the ability to learn gross segmental L2 durational characteristics but have difficulty acquiring the phonetic detail appropriate for native-like production. The learning of durational features varied according to preceding vowel length suggesting the influence of L1 syllable structure on L2 learning.

Keywords: L2 Japanese, durational feature learning, voiceless stop consonant

1. INTRODUCTION

Recent studies of second language (L2) learning have extended their investigations from simple segmental features to sub- and supra- segmental levels (e.g. [2], [3], [8], [10]). Wang and Behne [10] observed successful attainment of L2 English stop consonant duration by Chinese learners but found that learners could not correctly adjust subsegmental components such as VOT and closure durations. The study indicated that the learning of segmental duration and the adjustment of its temporal components are not simultaneous. Trofimovich and Baker [8] investigated the attainment of stress-timing in the L2 English of Korean learners. They found a strong influence of L1 syllable structure on temporal control in L2 speech even after a long period of L2 experience.

The current study examines whether beginning learners adjust the durational features of an intervocalic (IV) consonant from L1, Australian English (AusE), to L2, Japanese (L2J). The two languages differ in the durational characteristics of voiceless stops. A tendency for longer VOT for voiceless stop in English (American English) than Japanese at the syllable initial position has been documented [7]. AusE and Japanese also differ in the minimum meaningful speech unit within the prosodic hierarchy, that is, the syllable in AusE and the mora in Japanese.

In this paper we will examine the IV voiceless stop /k/ in L2J to explore the development of the language-specific durational characteristics of VOT and closure. We will also consider the influence of L1 syllable structure on the development of these features.

2. METHODOLOGY

2.1. Speakers

Two groups of speakers provided speech data. The first group consisted of five female and three male monolingual AusE speaking learners of Japanese language. All began learning Japanese at university without any previous experience with this language. Speech data was obtained from the Australian participants in their native language (AusE) and in their L2 language (L2J). These participants attended four recording sessions; a single AusE session and three L2J recording sessions. The L2J recordings were conducted 4, 9, and 16 months after the onset of Japanese learning.

The second group consisted of four female and one male Japanese native speaker. All were speakers of standard Japanese from Tokyo or the surrounding area. These participants produced native Japanese (NJ) speech data.

2.2. Speech materials and recording procedures

This study focuses on words with the structure /k/ V_1 /k/ V_2, where V_2 is [iː] in AusE, and [i] in Japanese (see Table 1). The first vowel, V_1, is one of four vowels: [ɔ, oː] and [ʊ, ʉː] in AusE; and [o, oo] and [ɯ, ɯɯ] in Japanese. Lexical stress for AusE and pitch-accent for Japanese are each

assigned to the first vowel. In English, stressed syllables require a coda consonant when the syllable nucleus is a short vowel because only heavy syllables can carry lexical stress [4]. Therefore, the putative status of the IV consonant in AusE is a coda of the first syllable in the short vowel condition. On the other hand, it is ambisyllabic or onset in the long vowel condition. In Japanese, an IV stop consonant is always the onset of the following mora.

Recordings were carried out in a recording studio in the Centre for Language Sciences in the Department of Linguistics at Macquarie University. The participant produced the isolated words following a randomly presented orthographic cue on the computer screen. The utterances were directly recorded via an AKG C414 B microphone onto hard disk. The sound file of each session was digitised at 44.1 kHz with 16 bit resolution. Low-pass filter was set as 5000Hz for male and 5500Hz for female speakers

Table 1: The words used to obtain IV stop voiceless consonant.

Vowel type	Japanese words			AusE words	
	Cues	Alphabetic transcription	IPA	Cues†	IPA
o	こき	*koki*	koki	cocky	kɔkiː
	こーき	*kooki*	kooki	corky	koːkiː
u	くき	*kuki*	kɯki	cookie	kʊkiː
	くーき	*kuuki*	kɯɯki	kooky	kʉːkiː

2.3. Speech data, measurements and variables

A total of 140 tokens (2 vowel type × 2 proceeding vowel length × 5 repetitions × 7 speakers = 140) were obtained from the AusE learners for each recording (AusE, L2J 4, 9, 16 months after starting L2J learning) and a total of 80 tokens (2 vowel type × 2 proceeding vowel length × 5 repetitions × 4 speakers = 80) were obtained from NJ speakers for their one-off recording session.

Each token was manually labelled based on the acoustic waveform and spectrogram in *Praat* Version 4.3.16 [1].

The closure of the IV consonant was defined from the offset of the preceding vowel, determined by the end of the second formant, to the release burst. For the VOT measurement, an acoustic waveform was used to determine a release-burst. The offset of VOT was determined by the beginning of voicing for the following vowel [5]. The duration of the IV consonant was measured from the offset of the preceding vowel to the voicing of the following vowel. In Japanese language, high front vowels after /k/ are sometimes voiceless [9]. Any cases of voiceless vowels were excluded from the dataset.

The values used in the analyses are as follows:

- Duration of the IV consonant (period from end of V_1 to beginning of V_2)
- Closure duration of the IV consonant (from the end of V_1 to the consonant burst)
- VOT duration of the IV consonant (from burst to beginning of V_2)

2.4. Statistical analyses

Mixed model analyses were conducted on the data obtained across multiple recordings. Factors employed in the analyses were as follows.

- Recording conditions (**Rec**: AusE, L2J at 4, 9, and 16 months, and NJ)
- Length of vowel nuclei in the preceding syllable/mora (**Vowel**: short and long vowel)
- Repetitions (**Run**: Run1, Run 2, Run 3, Run 4 and Run 5)
- Type of vowel nuclei (**Type**: o [ɔ, oː, o, oo] and u [ʊ, ʉː, ɯ, ɯɯ])
- Individual speakers (**Speakers**)

The **Rec**, **Vowel**, **Type** and **Run** main effects and their interactions were tested for each of the dependent variables. **Speaker** served as a random factor in the model. Univariate tests and pairwise comparisons with Bonferroni correction were conducted as post hoc tests when required.

3. RESULTS

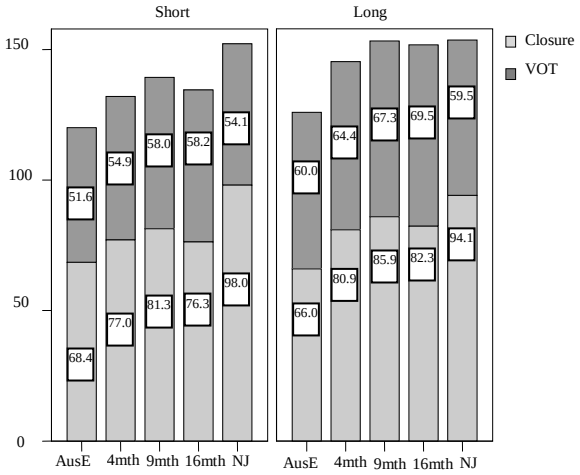
3.1. Data overview

The duration of the IV consonant, the VOT and closure were analysed to determine how these features differed between AusE and NJ and whether learners were able to change the durational characteristics of the IV stop from AusE to L2J.

Figure 1 presents the mean durations of the IV consonant, the closure and VOT.

The absolute duration of the IV consonant was longest for NJ compared to AusE and the three L2J recordings in both the short and long vowel conditions. The results from the LJ2 recordings show that the learners lengthened the duration of the IV consonant compared to their AusE production. The learners achieved this lengthening by increasing the duration of both the closure and the VOT.

Figure 1: Average durations of VOT and closure: AusE, L2J recordings and NJ by preceding vowel length. Left panel - short vowel condition. Right panel - long vowel condition. Y-axis is in milliseconds.



3.2. AusE and NJ comparison

A total of 440 tokens (280 for AusE and 160 for NJ) were analysed in the AusE-NJ comparison. Three durational variables — duration of the IV consonant, closure and VOT — were separately submitted to a mixed model to investigate the effects of recording condition (**Rec**: AusE and NJ), preceding vowel length (**Vowel**: short and long vowel conditions), and their interactions.

Table 2: Main effects and interactions for IV consonant duration, VOT and closure: NJ and AusE. Significances are in bold ($p < 0.05$).

	df	F	p
IV consonant duration			
Rec (AusE and NJ)	**(1, 9)**	**11.08**	**0.009**
Vowel (short and long)	**(1, 411)**	**12.52**	**0.000**
Rec × Vowel	(1, 411)	0.30	0.583
VOT			
Rec (AusE and NJ)	(1, 9)	0.05	0.833
Vowel (short and long)	**(1, 411)**	**36.71**	**0.000**
Rec × Vowel	(1, 411)	1.65	0.199
Closure			
Rec (AusE and NJ)	**(1, 9)**	**14.73**	**0.004**
Vowel (short and long)	**(1, 411)**	**7.32**	**0.007**
Rec ×Vowel	(1, 411)	0.40	0.526

Both the total duration of the IV consonant and its closure were longer for NJ than AusE but the VOT did not yield any significant difference between the two languages. Therefore the cross-language difference in the IV consonant duration was due to the difference in the closure duration. The preceding vowel length significantly affected these durational features in both languages with longer closures in the short vowel context.

3.3. Development of durational features for L2J voiceless stop consonant

A total of 1000 tokens (840 for L2J and 160 for NJ) were analysed by a mixed model to investigate **Rec** and **Vowel** main effects and their interaction. Table 3 presents the results of the analyses. For the IV consonant, closure duration and VOT, significant differences were obtained for **Rec** and **Vowel** main effects. Interactions between **Rec×Vowel** were also found for consonant and closure but not for VOT.

Table 3: Main effects and interactions for an IV consonant duration, VOT and closure: L2J 4, 9, 16 months and NJ. Significances are in bold ($p < 0.05$).

	df	F	p
IV consonant duration			
Rec (NJ & L2J 4, 9, 16 mth)	**(3, 23)**	**4.93**	**0.009**
Vowel (short and long)	**(1, 948)**	**85.55**	**0.000**
Rec × Vowel	**(3, 948)**	**3.40**	**0.017**
VOT			
Rec (NJ & L2J 4, 9, 16 mth)	**(3, 23)**	**5.50**	**0.005**
Vowel (short and long)	**(1, 948)**	**110.72**	**0.000**
Rec × Vowel	(3, 948)	1.77	0.151
Closure			
Rec (NJ & L2J 4, 9, 16 mth)	**(3, 23)**	**6.67**	**0.002**
Vowel (short and long)	**(1, 948)**	**8.55**	**0.004**
Rec × Vowel	**(3, 948)**	**4.59**	**0.003**

Table 4: Results of univariate tests on consonant duration and closure by Vowel: L2J recordings and NJ. Significances are in bold ($p < 0.013$).

	df	F	p
IV consonant duration			
Short vowel	(3, 26)	3.208	0.040
Long vowel	(3, 26)	2.321	0.099
Closure			
Short vowel	**(3, 25)**	**4.554**	**0.011**
Long vowel	(3, 25)	3.492	0.031

Univariate (post hoc) tests were conducted separately by preceding vowel length on the duration of the IV consonant and closure across the recordings (L2J 4, 9, 16 months and NJ). These results are summarised in Table 4. The total consonant duration did not differ between NJ and the L2J sessions according to preceding vowel condition, while the closure differed only in the short vowel condition. When taken together with

the multivariate results, this means that the learners were able to obtain the increased consonant duration only in the long vowel condition but not in the short vowel condition.

3.4. Summary

The AusE and NJ comparison showed that the two languages have different consonant durations and that this difference was a result of Japanese displaying greater closure duration. Therefore, the learners were required to lengthen only closure to achieve the desired duration of the IV consonant. However, the learners increased both the VOT and closure durations. In addition, preceding vowel length affected the degree of enhancement of closure duration. Only in the long vowel contexts were learners able to increase closure to achieve the NJ values.

4. DISCUSSION

The results show that beginning learners had the ability to change absolute consonant duration. However they could not adjust its subsegmental durational features, VOT and closure in the correct proportions. This result is similar to that of [10] for Chinese learners of English who succeeded in producing L2 stop consonant duration but did not achieve native-like durational coordination for the sub-segmental durational features.

We observed the lengthening of the IV consonant as a result of lengthening its sub-segmental components. The closure lengthening towards the NJ norm was observed while the degree of VOT lengthening was excessive.

These lengthening may be due to the slower speech rate in L2 compared to L1 (AusE). It has been reported that L2 speech tends to be slower than L1, because L1 phonetic encoding is automatic [6]. Much attention is required in the L2 speech production, especially, for learners at the beginning stage. However the degree of lengthening observed here was dependent on preceding vowel length suggesting that higher order phonological factors are at play. The learners' L2 closure was only different from that of NJ in the short vowel condition but not in the long vowel condition suggesting that it is more difficult for the learners to change the durational features of the IV consonant in the short vowel condition. In this study the IV consonant in the AusE data is the coda in the short vowel contexts but ambisyllabic or onset in the long vowel context. In NJ the IV consonant is always the onset. The fact that learners could change the durational features of the IV consonant in the long vowel context may relate to the shift from ambiyllabic to onset. We propose that the structure of the L1 syllable containing the short vowel seemed to impede L2 learning. The results in the current study support previous findings that adult learners' L2 production is strongly influenced by their L1 prosodic structure [2, 3, 8].

5. REFERENCES

[1] Boersma, P., Weenink, D. 2005. Praat: doing phonetics by computer (Version 4.3.16) [Computer program]. *http://www.praat.org/*

[2] Guion, S. 2005. Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition* 27, 503-533.

[3] Guion, S., Harada, T., Clark, J. 2004. Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition* 7, 207-226.

[4] Hayes, B. 2008. *Introductory Phonology*. Oxford: Wiley-Blackwell.

[5] Klatt, D. 1975. Voice onset time, frication and aspiration in word-initial clusters, *Journal of Speech, Language and Hearing Research*. 18, 686-706.

[6] Kormos, J. 2006. *Speech Production and Second Language Acquisition.* Mahwah, NJ: Lawrence Erlbaum Associates.

[7] Riney, T., Takagi, N., Ota, K., Uchida, Y. 2007. The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics* 35, 439-443.

[8] Trofimovich, P., Baker, W. 2006. Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech and fluency characteristics of second language speech. *Studies in Second Language Acquisition* 28, 1-30.

[9] Tsujimura, N. 2007. *An Introduction to Japanese Linguistic* (2nd ed). Wiley-Blackwell.

[10] Wang, Y. Behne, D. 2007. Temporal remnants from Mandarin in non-native English speech. In Bohn, O.-S., Munro, M. (eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege.* Amsterdam: John Benjamins, 167-184.