Lexical schwa and inserted schwa produced by Mandarin Chinese EAL learners

Xiaoqian Guo and Akitsugu Nogita
University of Victoria
guoxiaoqian.jan@gmail.com, akitsugu@uvic.ca

This study examines different types of vowel/schwa insertion in L1 illegal English consonant clusters by Mandarin Chinese (MC) English-as-an-additional-language (EAL) learners, as well as differences and similarities in phonetic qualities among inserted vowels by MC EAL learners, lexical schwas by MC EAL learners, and lexical schwas by native Canadian English (CE) speakers. In this study we have conducted a reading, a repetition and two syllabification tasks with 6 intermediate MC EAL speakers and 3 native CE speakers. There are three main findings: 1) Even with written cues, two MC EAL speakers likely have underlying vowels, which do not exist in native English speakers' underlying representation (UR); three MC speakers may have inserted excrescent schwas due to gestural mistiming rather than phonological schwas, and one MC speaker may have inserted excrescent schwas, because the individual likely has had extraprosodic consonants that are not linked to the syllable nodes. 2) English lexical schwas produced by CE speakers tend to be more variant in the second formant (F2) than those produced by MC learners, and lexical schwas by MC EAL learners have been occasionally rhotacized and deleted/devoiced. 3) MC EAL learners may not have explicitly understood the English syllable structures, even though some of them are aware of the presence or absence of vowels. Based on the findings, this paper proposes that it is important for instructors and learners to be aware that language learners may exhibit several different error types in the production of consonant sequences. Meanwhile, MC EAL learners may benefit from explicitly knowing the concept of English syllables.
Keywords: English consonant cluster; schwa insertion; excrescent schwa; extraprosodic consonant; Mandarin Chinese EAL learners

1 Introduction

In the field of second language (L2) phonology, a number of studies (e.g., Chan, 2006; Hansen, 2001; Miao, 2005) showed that Chinese EAL (English-as-an-additional-language) speakers may use vowel insertion as a common strategy to resolve English consonant clusters which are illegal in the first language (L1). Nogita and Fan (2012) found that Mandarin Chinese (MC) English-as-an-additional-language (EAL) speakers occasionally inserted a schwa-like sound in L1 illegal English consonant clusters. However, whether their vowel insertion was due to phonological vowel epenthesis or phonetic gestural mistiming was not fully figured out. Also, Nogita and Fan's model is unable to handle learners who are aware of the presence or absence of vowels but are not aware of syllabification as a higher prosodic unit than segments. This study examines the nature of MC EAL learners' vowel insertion, and adds another pattern of vowel insertion (i.e., extraprosodic consonants) to their model, as a follow-up study of Nogita and Fan (2012). This study replicates their study, in which the participants produced English nonsense words with L1 illegal consonant clusters and orally syllabified each stimulus word. The purpose of the design is to examine L2 learners' underlying representation (UR), and investigate acoustic properties of MC EAL learners' inserted schwa-like sounds, their lexical schwas, and native English speakers' lexical schwas.

2 Previous studies and research question

2.1 Four types of vowel insertion in consonant sequences in L2

Nogita and Fan (2012) proposed three different types in vowel insertion (see (a), (b), (c) in Figure 1): 1) L2 learners incorrectly memorize the underlying representation (UR) in their inter-language (IL) mental lexicon (e.g., /ɾʌgɯ̲bi/[1] in UR with an extra vowel instead of /ɾʌgbi/ *rugby*, /təɹant_/ with one vowel missing instead of /təɹanto/ *Toronto*); 2) L2 learners explicitly understand their L2 syllable structure and UR, but they still cannot automatize their proper production, so that they consciously or unconsciously insert a lexical vowel to repair L1 illegal syllable structure; in other words, a lexical vowel epenthesis; and 3) their UR is correct, but they fail to coordinate two consonants and result in a short schwa-like vocalic sound in the surface representation (SR), which is an excrescent vowel intrusion.

In the current study, we revised this model by adding the fourth type, an extraprosodic consonant followed by a non-lexical vowel. In the syllabification task in Nogita and Fan (2012), the participants were asked to orally divide each English word into syllables. One Japanese participant divided the word *webnet*, for example, into [wɛ-bɯ̆-nɛ-tɯ̥]. Apparently, the individual added extra syllables, but the inserted vowels (i.e. [bɯ̆] and [tɯ̥]) were notably short or devoiced. Presumably, the participant was aware of where to pronounce or not pronounce a vowel, but consonants without a following vowel were independent in their inter-language. Such consonants were linked to mora nodes but not to syllable nodes. Phonetic realizations of such unsyllabified consonants vary, but typically occur in conjunction with a short/voiceless vowel. A few similar occasions by MC participants were observed by Nogita and Fan (2012) as well, for example, /kokənʌt/ *coconut* →[ko-kə-nʌ-tə̥]. These EAL learners may have not been explicitly taught the rules of English syllables, specifically, how to assign consonants into syllables. The fourth type is shown in Figure 1(d) with another example, *subject*.

[1] /ɯ/ is the default epenthetic vowel in Japanese.

	(a)	(b)	(c)
Perceived L2 UR	/CVC/	/CC/	/CC/
	e.g. /ɹʌgɯbi/ *rugby*	/ɹʌgbi/	/ɹʌgbi/
	↓	↓	↓
Phonological adjustment based on L1	N/A	/CVC/	N/A
	/ɹʌgɯbi/	/ɹʌgɯbi/	/ɹʌgbi/
	↓	↓	↓
Phonetic SR	[CVC]	[CVC]	[C^{V}C]
	[ɹʌgɯbi]	[ɹʌgɯbi]	[ɹʌgᵊbi]

(d)

	σ			σ			
	μ		μ	μ		μ	μ
UR	/ s	ʌ	b	dʒ	ɪ	k	t /
SR	[s	ʌ	bɯ̆	dʒ	ɪ	kɯ̥	tɯ̥]

Figure 1. Differences among (a) misinterpreted L2 UR, (b) lexical vowel epenthesis, (c) excrescent vowel intrusion, and (d) extraprosodic consonants

Extraprosodic consonants are possible in actual languages. For example, in Nxaʔamxcín (Moses-Columbia Salish) the maximal syllable structure is CVC, but various consonant clusters occur with extraprosodic consonants, which are not incorporated into syllables, as in *scílksq’t* where only the underlined portion fits into the syllable template (Czaykowska-Higgins & Willett, 1997, p. 385). These unsyllabified consonant sequences are optionally along with a short transitional schwa or a voiceless schwa (Czaykowska-Higgins & Willett, 1997). This is much like some of the L1 illegal English consonant clusters produced by Japanese and MC EAL learners in Nogita and Fan’s study. Similarly, the initial /s/ in English as in *sky* is arguably “an appendix”, directly linked to the higher prosodic nodes (Sperbeck, 2010, p. 55). If this is the case, /s.CV/ might be more complex than /CV/ with a simplex onset, but less complex than /CCV/ with a true branching onset as in *cry* in terms of syllable structure (Sperbeck, 2010). As for L2 phonology, appearing as a production mistake, extraprosodic consonants were produced by Japanese and MC EAL learners, when they were dealing with true consonant clusters in English (Nogita & Fan, 2012). Therefore, it might be the case that L2 speakers are prone to less complex extraprosodic consonants rather than to more complex true tautosyllabic consonant clusters.

2.2 Excrescent schwa in L2 production

Some previous studies (e.g., Davidson, 2005 & 2006; Davidson & Stone, 2003) discovered that non-native speakers would insert a schwa-like sound in L1 illegal consonant clusters. These studies also demonstrated that there were acoustic distinctions between lexical schwas in the target languages and the inserted schwas. According to their explanations, the inserted schwas were mostly caused by the gestural mistiming, that is, an unsuccessful gestural coordination of consonant sequences. Gafos (2002) proposed three possible gestural coordination patterns (see Figure 2). In this figure, consonant segments are represented as plateaus with a horizontal line indicating the articulation target. In pattern (1), the articulation of two adjacent consonants has no gestural overlap at all; pattern (2) indicates a partial overlap of the two consonants with an intervening acoustic release between C1 and C2; pattern (3) shows a more substantial gestural overlap with no open vocal tract between the articulation of C1 and C2.

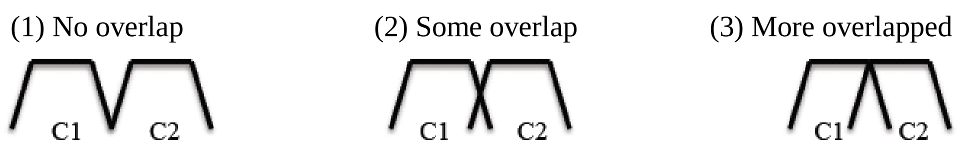

Figure 2. Patterns of gestural coordination in consonant clusters

These patterns are language specific (Gafos, 2002) and English consonant sequences typically follow pattern (3), more overlapped (Fan, 2011). Fan's (2011) research focused on MC EAL speakers' acquisition of English consonant clusters. By comparing the articulation data produced by 31 MC EAL speakers and 8 native English speakers from western Canada, Fan observed that MC speakers had on average less consonantal overlapping than English speakers, and that the performance of the advanced EAL speakers was more similar to that of native speakers, compared with low-intermediate speakers. Based on these findings, she proposed that the differences in the consonantal gestural overlap between native English speakers and MC EAL learners might contribute to Chinese speakers' foreign accent in English.

2.3 Mandarin Chinese phonotactics vs. English phonotactics

The possible syllable shapes in Mandarin Chinese (MC) include (C)(G)V(X)[2] (San, 1990). An example of a maximal syllabic structure is *nian* [njæn] "Year"[3]. Some researchers (e.g., Fan, 2011; Lin, 2001) proposed that the pre-nucleus glide

[2] C=Consonant; G=Glide; V=Vowel; X=Nasal, Glide or [r]; bracketed segments are optional

[3] The examples of disyllabic words are provided in Pinyin, IPA, and their meanings in English.

(G) should be included in the rhyme, while other researchers (e.g., San, 1990) argued that no phonetic evidence to date could support that the pre-nuclear CG were actually two segments in MC phonotactics and that the initial part is a single onset C^G. Both these assumptions hold that there are no unambiguous consonant clusters in either onset or coda position in MC phonotactics. On the other hand, heterosyllabic consonant clusters are possible: e.g., *be<u>nb</u>u* [pɤ<u>np</u>u] 'headquarters', *benlai* [pɤnlai] 'original' (Lin, 2001). As for English, it allows complex, or branching, onsets and codas (e.g., *<u>str</u>ing*, *si<u>xth</u>*), which is distinct from MC. Thus, the error types of MC EAL speakers listed in Figure 1 may be due to the phonological difference between MC and English.

2.4 Acoustic properties of lexical schwas in North American English

The lexical or phonological English schwa, a mid central vowel, is characterized as a short and reduced vowel, which is restricted to unstressed syllables. A schwa in English easily assimilates to its segmental contexts and its second formant (F2) frequencies especially tend to vary (Flemming & Johnson, 2007; Kondo, 1994). Flemming and Johnson (2007) analyzed the formant structures of word-final and non-final schwas in a carrier sentence 'Say ___ to me' by nine native American English speakers. The finding is that word-final schwas, as in *Ros<u>a</u>, sof<u>a</u>, comm<u>a</u>,* and *umbrell<u>a</u>,* consistently maintained the quality as a mid-central vowel (average F1 = 665Hz; average F2 =1772Hz). However, word-medial schwas, as in *s<u>u</u>ppose, t<u>o</u>day,* and *prob<u>a</u>ble,* showed various qualities determined by its surrounding context; their average F1 (428Hz) indicated a relatively high vowel and F2 covered a wide range depending on the adjacent consonants. In addition, Klatt (1976) observed that in English connected discourse the average duration of stressed vowels was about 130 milliseconds (ms), while that of schwas was about 70ms, which was very similar to the average duration for consonants (71ms). Similarly, the average duration for non-final schwa in Flemming and Johnson's (2007) study was 64ms. The tendency for schwas to assimilate to their neighboring contexts may be caused by the fact that schwas have short duration, in which there is insufficient time for the tongue to arrive at the target position (Flemming, 2004; Lindblom, 1963). The failure of achieving the target position is regarded as the phenomenon of "undershoot" (Lindblom, 1963). This explains Flemming and Johnson's (2007) findings that word-medial schwas assimilate more easily to its segmental context compared with word-final schwa, which had longer average duration.

Schwas in the word-medial position may have the characteristics of a relatively high vowel (Flemming & Johnson, 2007), but they are not necessarily targetless (Browman & Goldstein, 1992). If they are, the tongue should smoothly move from the preceding sound to the following one (Van Bergem, 1995). In Browman and Goldstein's study (1992), they used X-ray micro-beam technology to examine the articulatory data of the stimuli [pV_1pəpV_2pə] produced by an American English speaker, and V_1 and V_2 were selected from a vowel set [i, ɛ, ɑ, ʌ, u]. The results indicated that the tongue body positions from V_1 to V_2

sometimes did not show a linear movement and did try to reach a target position of the first schwa. Echoing what Flemming and Johnson (2007) discovered about schwa's assimilation to its surrounding sounds, the stimulus [pipəpipə] in Browman and Goldstein's study (1992) was produced with the first schwa as a relatively high vowel, instead of a mid-vowel. The possible reasons included the influence of a neighbouring vowel sound [i], as well as the inherent target of schwa. It is also worth noting that the speakers' tongue moved slightly downwards between two high vowels [i], instead of maintaining a high position during for the first schwa. For this reason, it is not wise to view schwas as completely targetless sounds.

2.5 Research questions

The present study aims to answer the following three research questions:
1. Are there any differences or similarities between lexical schwas produced by native English speakers and lexical schwas produced by MC EAL speakers? If yes, what are they?
2. Are there any differences or similarities between lexical schwas and inserted vowels produced by MC EAL speakers? If yes, what are they?
3. Ultimately, what is the nature of MC EAL speakers' vowel insertion? Is it (a) misinterpreted L2 UR, (b) lexical vowel epenthesis, (c) excrescent vowel intrusion, or (d) extraprosodic consonants?

3 Methodology

3.1 Participants

The present study involved nine voluntary participants: six MC EAL learners, and three native Canadian English (CE) speakers. As shown in Table 1, all the MC speakers were exchange students at the University of Victoria with a length of residence (LOR) in Canada of two to three months. According to their self-reported language test scores (see Table 1) and the rubrics posed on the official websites of *TOEFL*[4] and *IELTS*[5], they were intermediate level English learners. The CE speakers, who were graduate students in the Department of Linguistics at the University of Victoria, all came from western Canada. None of the MC and CE participants reported any history of speech or hearing impairments.

[4] *TOEFL* score interpretation: http://www.ets.org/toefl/institutions/scores/interpret/
[5] *IELTS* score interpretation:
http://www.ielts.org/researchers/score_processing_and_reporting.aspx

Table 1. Participant information

		Chinese speakers *N* = 6	English speakers *N* = 3
Gender		3M3F	1M2F
Age in Years	Range	18-26	26-52
	Mean	23.3	36.7
Length of Residence	Range	2-3 months	N/A
	Mean	2.5 months	
Reported *TOEFL* Scores	*N*	4	N/A
	Range	87-95	
	Mean	90.5	
Reported *IELTS* Scores	*N*	2	N/A
	Range	5.0-6.5	
	Mean	5.75	

3.2 Speech stimuli

The speech materials included two groups of sequences (Target Stimuli: CVCC and CCVC; Control Stimuli: CVCəC, CVCCə, CəCVC, and əCCVC) both written with IPA symbols. There were 18 target stimuli, in which the consonant clusters in onset or coda positions were allowed in English phonotactics, but illegal in MC. As the comparing stimuli, Control Stimuli included lexical schwa where MC EAL learners were expected to insert a vowel in CVCC and CCVC (Hansen, 2001; Miao, 2005). The purpose of using IPA representations was to minimize the influence of English orthographic representations, and to make sure that all the participants would know that there were lexical schwa sounds in Control Stimuli. Table 2 shows the speech stimuli. A cross-linguistically common vowel sound [a][6] was used as the V in all the stimuli. C_1C_2 in the coda position included 12 types of consonant sequences: stop + stop (/pt/, /kt/), stop + fricative (/ks/, /ts/, /dz/, /gz/), fricative + stop (/sp/, /ʃt/), and /l/ + stop (/lk/, /lt/, /lb/, /ld/); C_1C_2 in the onset position included six consonant sequences: /s/ + voiceless stops (/sp/, /sk/) and stops + /l/ (/pl/, /kl/, /bl/, /gl/). A common word-initial as well as word-final consonant [k] was used as C_0 and C_3 in the stimuli. There were also 16 fillers with no consonant clusters.

[6] In Canadian English it is typically realized as a low back vowel [ɑ]. Because Canadian English does not phonologically contrast the central [a] and the back [ɑ], in this study we used the broad transcription [a], which is more familiar to the linguistically naïve participants.

Table 2. Stimuli words

Target Stimuli (18)		**Control Stimuli (36)**	
$C_0VC_1C_2$	/kápt/ /kákt/ /káks/ /káts/ /kádz/ /kágz/ /kásp/ /káʃt/ /kálk/ /kált/ /kálb/ /káld/	$C_0VC_1əC_2$	/kápət/ /kákət/ /kákəs/ /kátəs/ /kádəz/ /kágəz/ /kásəp/ /káʃət/ /kálək/ /kálət/ /káləb/ /káləd/
		$C_0VC_1C_2ə$	/káptə/ /káktə/ /káksə/ /kátsə/ /kádzə/ /kágzə/ /káspə/ /káʃtə/ /kálkə/ /káltə/ /kálbə/ /káldə/
$C_1C_2VC_3$	/spák/ /skák/ /plák/ /klák/ /blák/ /glák/	$C_1əC_2VC_3$	/səpák/ /səkák/ /pəlák/ /kəlák/ /bəlák/ /gəlák/
		$əC_1C_2VC_3$	/əspák/ /əskák/ /əplák/ /əklák/ /əblák/ /əglák/
Fillers (16)			
CVC	/kát/ /kád/ /kák/ /kág/ /káf/ /káv/ /káz/ /káʃ/		
CVCə	/kápə/ /kábə/ /kátə/ /kádə/ /kákə/ /kágə/ /kávə/ /kásə/		

3.3 Data collection

Employing the same procedure from Nogita and Fan's (2012) study, the current study required participants to perform four experiment tasks (see Table 3): a reading task, a syllabification task with written stimuli, a repetition task, and a syllabification task with sound stimuli.

Table 3. Procedure

	MC EAL Speakers *N* = 6	Native CE Speakers *N* = 3
Task 1	Reading	Reading
Task 2	Syllabification with written stimuli	N/A
Task 3	Repetition	Repetition
Task 4	Syllabification with sound stimuli	N/A

First, MC and CE participants individually performed Task 1, the reading task. They looked at 70 written stimuli on PowerPoint Slides and read them aloud. After the reading task, only MC participants did Task 2, in which they verbally syllabified each stimulus by pausing between syllables. If they thought there was only one syllable in the sequence, they did not make a pause. Task 3 was a repetition task, in which participants listened to each stimulus once from the pre-recorded clip without any written cues. The pre-recorded clip was provided by a female graduate student in the Linguistics Department at University of Victoria. Immediately after listening to each stimulus, participants repeated it. After that, only MC participants performed Task 4, in which they listened to each stimulus again and then verbally syllabified it by pausing between two syllables. All the verbal data were recorded with the software Audacity set at 44100Hz and 32-bit float. The recordings were saved as .wav files. The reading task was done before the repetition task in order to extract the MC participants' own productions without being influenced by the native English speakers' productions. The reading task was also conducted before the syllabification task in order not to get the participants to focus on the syllable structures of the stimuli.

3.4 Data analysis

3.4.1 Coding and measurement of inserted vocalic elements

We examined whether the six MC participants had inserted schwa-like vowels in consonant clusters. We analyzed the waveform and spectrogram for the 108 (6×18) tokens from the reading task and the 108 (6×18) tokens from the repetition task using the software PRAAT. Vowel insertion was judged under the following criteria in accordance with Davidson (2006): 1) in Stop/fricative + stop/fricative, any periods of voicing with formant structure that appeared between two obstruents, and 2) in Liquid /l/ + stop, the voice bar of vocalic elements tended to be darker than that of liquid /l/. With an inserted vowel after /l/, there is typically an abrupt raising intensity, a clear change in the wave form patterns. The duration and F1 and F2 at midpoint of inserted vocalic elements were measured based on Davidson's (2006) and Fan's (2011) criteria. The duration was manually measured from the first zero-crossing point of the first glottal pulse of the vowel to the last zero-crossing point of the last glottal pulse. Figure 3 illustrates a coded display of the token /kats/, which had two inserted vowels (i.e., [katə̠sə̠]) produced by a female MC participant labelled C3. The first and the second authors independently coded and measured all inserted vowels. The inter-rater reliability was 82.73%. The two coders discussed all the disagreement codes until 100% agreement was achieved.

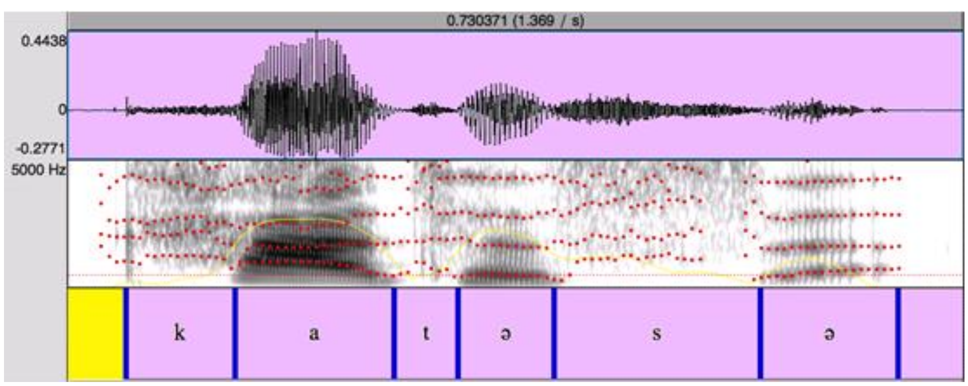

Figure 3. /kats/ produced by a female MC speaker

We also counted the number of deleted lexical schwas (e.g. /kákə̠s/ realized as [káks]), and rhotacized lexical schwas (e.g. /kákə̠s/ realized as [kákə̠ɹs] or [kákɚ̲s]) in all 70 tokens including target stimuli, control stimuli, and fillers, in both tasks. Figure 4 shows a male participant C4's production of the stimulus /kalət/ realized as [kaləɹt] with a rhotacized schwa whose F3 was lowering.

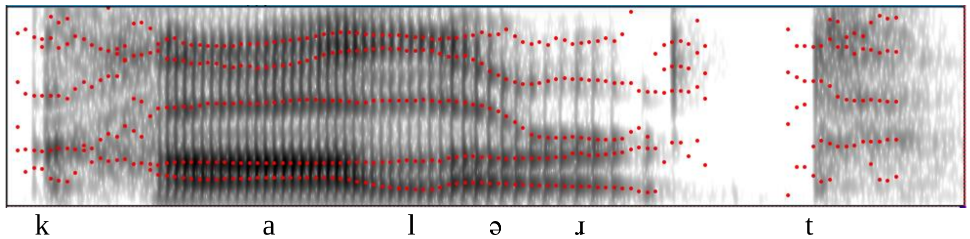

Figure 4. /kalət/ produced by a male MC speaker (showing up to 5000Hz)

3.4.2 Measurement of lexical schwas

Since MC participants did not show insertion in all the sequences in Target Stimuli, this study only measured the lexical schwa in the sequences corresponding to the target sequences, which had been coded with a vowel insertion (Davidson, 2006). For example, as mentioned above, /kats/ was pronounced as [katə̲sə̲], so that the lexical schwas in /katə̲s/ and /katsə̲/ produced by both the MC and the English speakers were measured. Descriptive statistics for the lexical schwas and the inserted schwa-like vowels will be presented in section 4.

3.4.3 Performance in syllabification tasks

The two syllabification tasks were designed to examine whether MC speakers had extra lexical vowels in their UR of their inter-language mental lexicon. We also examined the number of extra vowels and missing lexical schwas in the participants' syllabification.

4 Results

4.1 Vowel insertion, lexical schwa deletion/rhoticization patterns

In the reading task, all six MC EAL learners had vowel insertion errors. 23 tokens out of 108 (21.29%) were found as vowel insertion cases. In contrast, in the repetition task, only 4 insertion errors out of 108 (3.7%) by three participants were found. Table 4 presents the frequencies of vowel insertion in the reading and repetition tasks. Insertion errors decreased dramatically from the reading task to the repetition task, which agrees with Nogita and Fan's (2012) findings, but the general pattern remained consistent across the two tasks. The participant labeled C3 had the most insertion errors and C5 had the least insertion errors in both tasks. Table 5 shows vowel insertion patterns in specific sequences. This revealed that most insertions happened in the onset sequences of stop + /l/ and the coda sequences of stop + fricative. MC participants did not insert a vowel in onset /s/ + stop sequences in both production tasks. As for the position of insertion, among the 23 tokens with insertion, 16 of them were CVCəC cases, nine of them were CəCVC cases, one was CVCCə case, and one was CVCəCə case with two inserted vowels. Such differences might be an effect of the different number of onset vs. coda conditions. Table 6 shows lexical vowel deletion and lexical schwa rhotacization in all the 70 stimuli × 6 participants in the reading and repetition tasks. Schwa deletion and rhotacization also generally decreased in the repetition task except for C3. Table 7 shows lexical schwa deletion and lexical schwa rhotacization based on the phonetic contexts. Deletion occurred almost universally. Again, the apparent distribution may be an effect of the different number of conditions. For example, 6 stimuli have the /Sə/ ending and 2 have the /Fə/ ending. Rhotacization occurred only once in the first syllable and the rest occurred in the second syllable. Neither deletion nor rhotacization occurred in word-initial schwas, as in /əklák/.

Table 4. Frequencies of vowel insertion by MC speakers in 108 tokens

C1		C2		C3		C4		C5		C6	
Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp
2	0	3	1	13	2	2	0	1	0	2	1

Note. Rd=the reading task; Rp=the repetition task

Table 5. Vowel insertion frequencies in specific sequences in 108 tokens

	Onset (CCVC)						Coda (CVCC)					
	/s/ + S		S + /l/		S + S		S + F		F + S		/l/ + S	
	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp
Insertion	0	0	7	2	2	0	8	2	1	0	5	0

Table 6. Lexical schwa deletion/rhotacization by MC speakers in 420 tokens

	C1		C2		C3		C4		C5		C6	
	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp
De	13	10	3	2	1	4	0	1	3	1	1	1
Rh	0	0	3	1	0	4	31	3	2	1	3	0

Note. Rd=the reading task; Rp=the repetition task; De schwa deletion; Rh schwa rhotacization.

Table 7. Lexical schwa deletion/rhotacization in specific contexts in 420 tokens

	Cə̲CVC				CVCə̲C							
	/səS/		/Səl/		/SəS/		/SəF/		/FəS/		/ləS/	
	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp
Deletion	0	2	1	7	2	0	4	1	1	0	1	0
Rhoticize	1	0	0	0	2	0	4	1	3	0	6	0
	CVCə̲				CVCCə̲							
	/Sə/		/Fə/		/SSə/		/SFə/		/FSə/		/lSə/	
	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp
Deletion	9	2	0	1	2	2	0	1	0	2	1	1
Rhoticize	6	2	2	0	2	0	3	0	5	4	5	2

4.2 Acoustic properties of inserted vowels

The mean duration and midpoint F1 and F2 of the inserted vowels by the six MC speakers are presented in Table 8. These acoustic data indicate that most of the inserted sounds were mid central vowels, i.e. schwas. On average, schwas in the repetition task were higher in place, that is more like [ɨ], and shorter in duration than those in the reading context. The acoustic data of the inserted schwas by individual participants are presented in Table 9. The formant frequencies and duration values of the inserted schwas have variations within and across speakers. For the MC speakers (C2, C3, C6) who had insertion errors in both tasks, the mean durations and F1 frequencies decreased in the repetition context. Note that C4 inserted vowels only twice but both were fairly long, and one of them was rhotacized. This rhotacized schwa may have been caused by the participant's misreading of the written stimulus /kádzə/ as /kádəz/, which was realized as [kádɚːzə].

Table 8. Averaged acoustic properties of inserted vowels

		Reading Task		Repetition Task	
		Female	Male	Female	Male
F1(Hz)	Mean	608	454	613	395
	SD[a]	93	62	27	68
F2(Hz)	Mean	1623	1310	1737	1465
	SD	121	134	69	33
Duration (ms)	Mean	72	89	42	29
	SD	41	87	11	6

[a]Note. SD = standard deviation

Table 9. Individual acoustic properties of inserted vowels by MC speakers

			C1 F	C2 M	C3 F	C4 M	C5 F	C6 M
Reading	F1(Hz)	Mean	532	500	622	468	579	370
		SD	15	21	99	4		33
	F2 (Hz)	Mean	1613	1331	1625	1216	1624	1372
		SD	183	177	126	29		134
	Dur. (ms)	Mean	82	34	74	214	20	46
		SD	54	10	40	27		1
Repetition	F1(Hz)	Mean		443	613			347
		SD			27			
	F2 (Hz)	Mean		1441	1737			1489
		SD			69			
	Dur. (ms)	Mean		25	42			33
		SD			11			

Note. See Table 4 for the total number of vowel insertion.

The acoustic data of the inserted schwa elicited from the reading task were used for the data analysis in this paper. The first reason was to minimize the effect of task variables (written stimuli vs. sound stimuli) on MC speakers' production. Second, the MC participants had much more insertion instances in the reading task than that in the repetition task (23 vs. 4), and not all six MC participants made insertion errors in the repetition task. Third, the MC participants' productions – without being influenced by sound cues – were most likely their "genuine" inter-language productions.

A scatter plot of the midpoint F1 and F2 frequencies of the inserted 23 sounds in the reading task is displayed in Figure 5 (females) and Figure 6 (males). As a reference, seven American English vowels (i.e., /i/, /ɪ/, /æ/, /ɑ/, /ʌ/, /ʊ/, /u/ from Hillenbrand et al.'s (1995) data from 45 English speaking males, 48 females, and 46 children) are also indicated. As seen in these figures, most of the inserted vowels fell within the area of mid central vowels. Two inserted vowels were produced with a higher tongue position in the contexts /ʃt/ and /gz/ involving coronal consonants, which would raise F1. One inserted vowel in /gl/ was relatively back, where /l/ might lower F2.

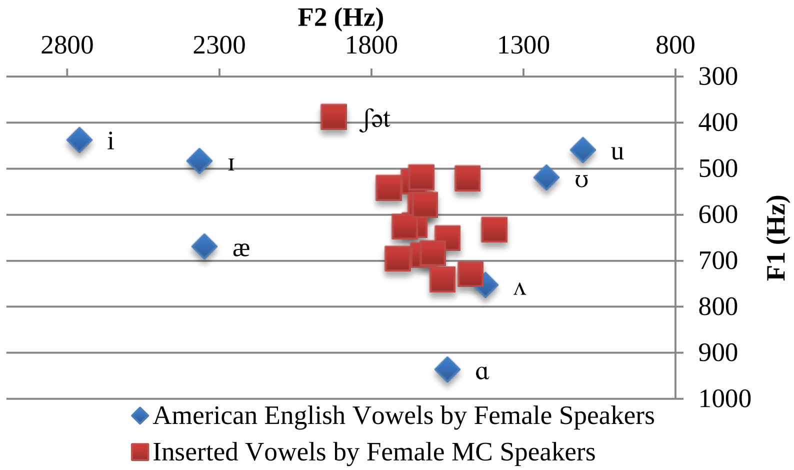

Figure 5. Inserted vowels by **female** MC speakers

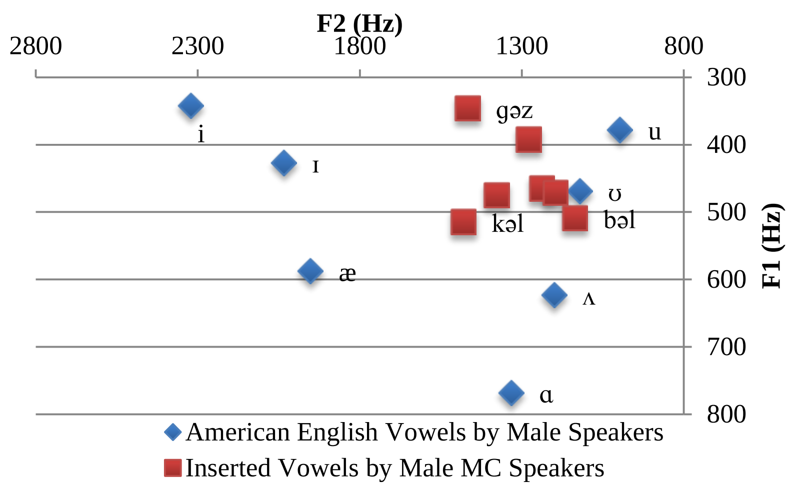

Figure 6. Inserted vowel by **male** MC speakers

4.3 Lexical schwas by MC and English speakers

We analyzed lexical schwas in MC EAL and native CE speakers' productions of 14 consonant sequences (see Table 10), whose corresponding sequences in the Target Stimuli had been coded as vowel insertion in the reading context. For example, the acoustic properties of the lexical schwa in /kə̲lák/ were examined because one or some MC participant(s) inserted a vowel between /k/ and /l/ in /klák/.

Table 10. The 14 sequences with lexical schwa

		Sequences with Lexical Schwa
Onset	Stop + /l/	/kəlak/ /pəlak/ /bəlak/ /gəlak/
Coda	Stop + Stop	/kákət/ /káptə/
	Stop + Fricative	/kátəs/ /kágəz/ /kádəz/
	Fricative + Stop	/káʃət/
	/l/ + Stop	/kálət/ /káləd/ /káləb/ /kálək/

The data from the reading task showed that the lexical schwas produced by the MC speakers and the English speakers were similar in terms of F1, F2, and duration (see Table 11). Generally, these were mid-central vowels.

Table 11. Acoustic properties of lexical schwa in the reading task

		MC Speakers *N* = 6	CE Speakers *N* = 3
F1 (Hz)	Mean	556	571
	SD	85	120
F2 (Hz)	Mean	1479	1599
	SD	164	193
Duration (ms)	Mean	101	93
	SD	53	44

In Table 12, the acoustic properties of the lexical schwas by individual participants are presented. In terms of F1 average, female MC speakers were similar to female CE speakers, while the male MC speakers' lexical schwas tended to be lower in position (or higher in F1) (C2: 539Hz, C4: 505Hz, C6: 496Hz) than the male CE speaker's (E3: 467Hz). As for F2, two female MC speakers performed (C1: 1581Hz, C2: 1588Hz) quite similar to one female CE speaker (E1: 1583Hz), but different from the other female CE speaker (E2: 1778Hz). The duration values (Minimum: 73ms; Maximum: 142ms) were quite different among the nine speakers. The values were not similar either within the same language group or within the same gender group. In more details, two MC EAL speakers' standard deviation values were quite bigger than those of CE speakers. Within the groups, the CE speakers' mean values ranged between 77 and 112ms, whereas that of the MC speakers was between 73 and 142ms. This suggests that MC EAL speakers' schwas were even more varied in duration than those by CE speakers. Still, the measured data (i.e., duration, midpoint F1 and F2) were quite different even within the CE group. For example, E3 produced a relatively long lexical schwa compared to the production of E1 and E2.

A scatter plot (see Figure 7 and 8) presents the midpoint F1 and F2 frequencies of the lexical schwa produced by MC and CE speakers. Overall, most of the production fell into the mid-central vowel area. The three CE speakers produced lexical schwas with more F2 variations than the six MC EAL speakers did, while they were performing the same reading task.

Table 12. Acoustic properties of lexical schwa by MC and English Speakers in Reading Task

		C1 F	C2 M	C3 F	C4 M	C5 F	C6 M	E1 F	E2 F	E3 M
F1	Mean	571	539	630	505	606	496	677	577	467
	SD	100	23	85	63	88	59	63	123	53
F2	Mean	1581	1371	1588	1315	1630	1486	1583	1778	1473
	SD	155	126	147	62	76	100	121	212	116
Dur	Mean	73	94	75	132	142	81	86	77	112
	SD	36	46	27	62	70	21	42	35	46

Note. *N* = 14, F1 and F2 measured in hertz, and Dur (duration) measured in milliseconds

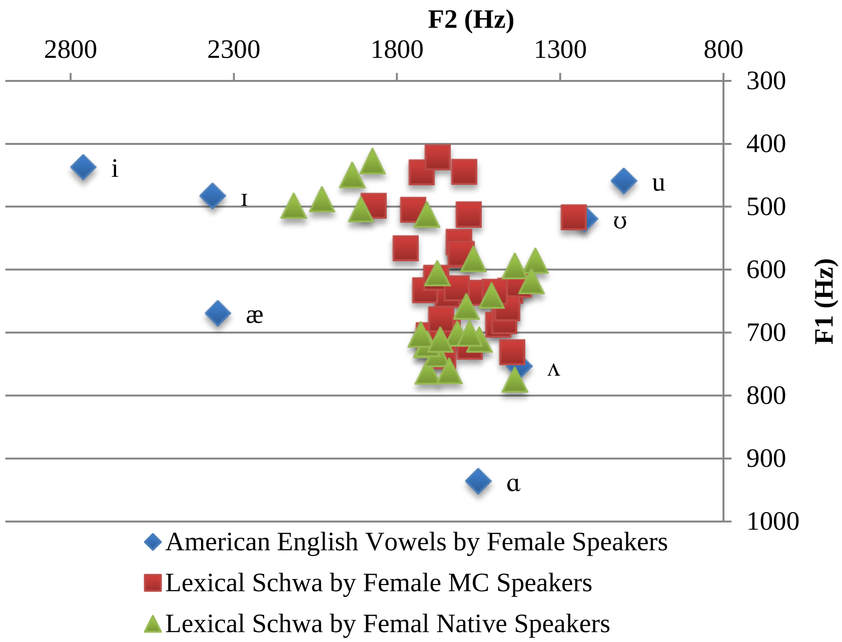

Figure 7. Lexical schwas produced by **female** MC and CE speakers

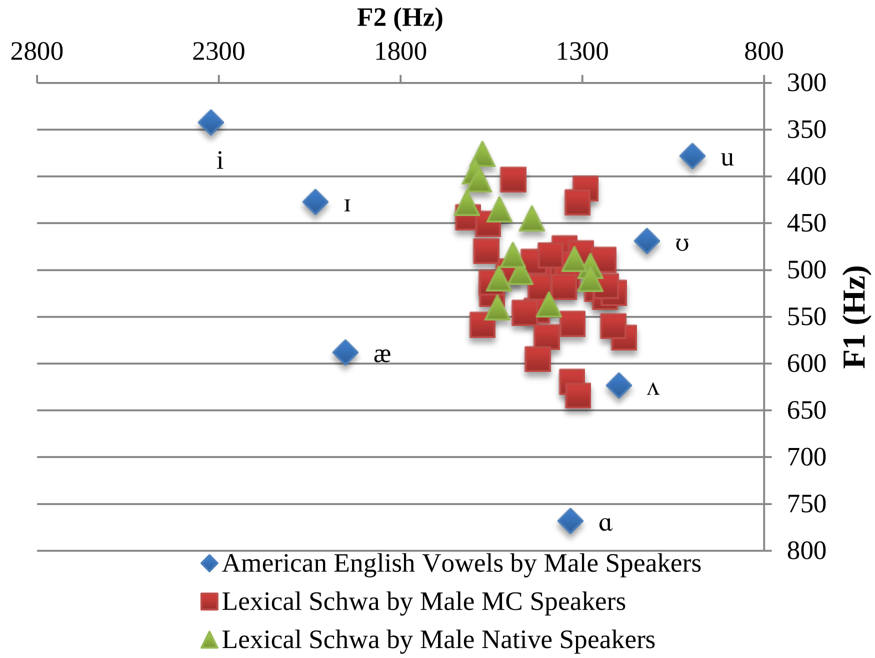

Figure 8. Lexical schwas produced by **male** MC and CE speakers

4.4 Comparison between lexical and inserted schwas

In terms of the mean F1 frequencies, all the MC EAL speakers articulated the inserted vowels with a higher tongue position than they did with the lexical schwa. As for the duration values, three MC EAL speakers (C2, C5, C6) inserted clearly shorter vowels than their lexical schwas. Table 13 shows the overall results.

Table 13. Inserted vowels and lexical schwas in the reading task

		C1 *N* = 2	C2 *N* = 3	C3 *N* = 13	C4 *N* = 2	C5 *N* = 1	C6 *N* = 2
Lexical	F1 (Hz)	571	539	630	505	606	496
	F2 (Hz)	1581	1371	1588	1315	1630	1486
	Dur (ms)	73	94	75	132	142	82
Inserted	F1 (Hz)	532	500	622	468	579	370
	F2 (Hz)	1613	1331	1625	1216	1624	1372
	Dur (ms)	**82**	34	74	**214**	20	46

Note. Bold numbers are unexpected results.

C5 and C6 had only three schwa insertion errors in the reading task. Table 14 shows the details. Compared with their lexical schwa counterparts produced in the same consonant sequences, the F1 and F2 frequencies and the duration values (bold numbers in Table 14) of the inserted vowel sounds were all lower.

Table 14. Inserted vowels and lexical schwas by participants C5 and C6

	C5		C6			
	[kálᵊb]	[káləb]	[gᵊlák]	[gəlák]	[kágᵊz]	[kágəz]
F1 (Hz)	**579**	606	**393**	541	**346**	451
F2 (Hz)	**1624**	1630	**1278**	1413	**1467**	1559
Duration (ms)	**20**	142	**46**	75	**47**	89

To examine the differences and similarities between the lexical schwas and the inserted vowels by MC speakers in terms of the place of articulation, we classified the acoustic data of the schwas into two categories: involving /l/ vs. only obstruents. Figure 9 (females) and Figure 10 (males) display the lexical schwas and the inserted schwas produced in /l/ contexts.

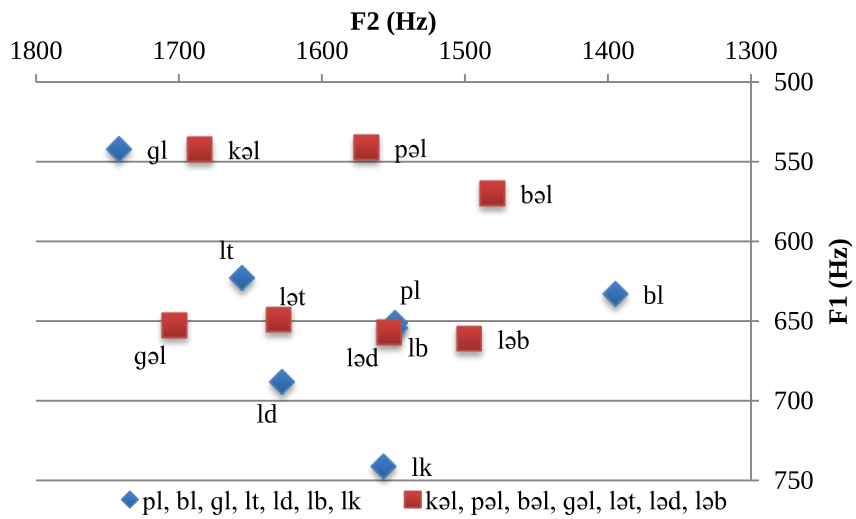

Figure 9. /l/ contexts: Inserted vs. lexical schwas by **female** MC speakers

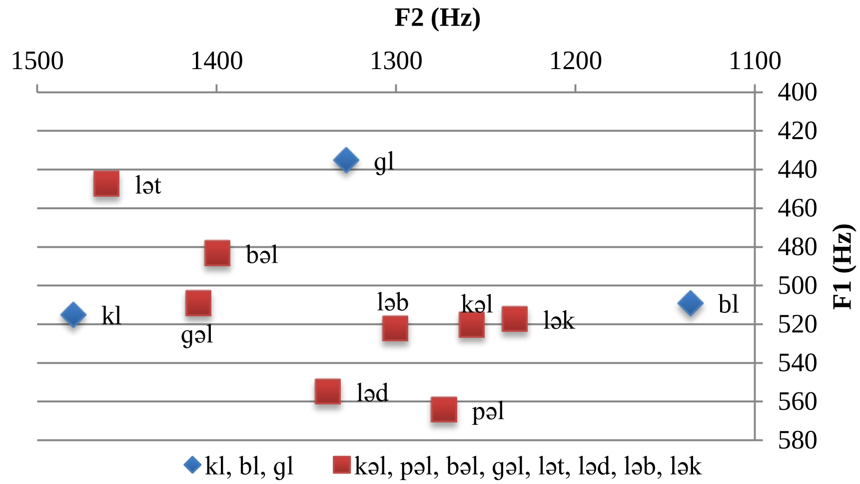

Figure 10. /l/ contexts: Inserted vs. lexical schwas by **male** MC speakers

Figure 11 (females) and Figure 12 (males) display the lexical schwas and the inserted vowels in obstruent contexts. In the females' productions, except for the inserted vowel in [ʃt], which sounded like a high front vowel, the average midpoint F1 and F2 frequencies of the lexical schwas and the inserted vowels were similar. In the males' productions, both F1 and F2 values of the inserted vowels were lower than those of the lexical schwas.

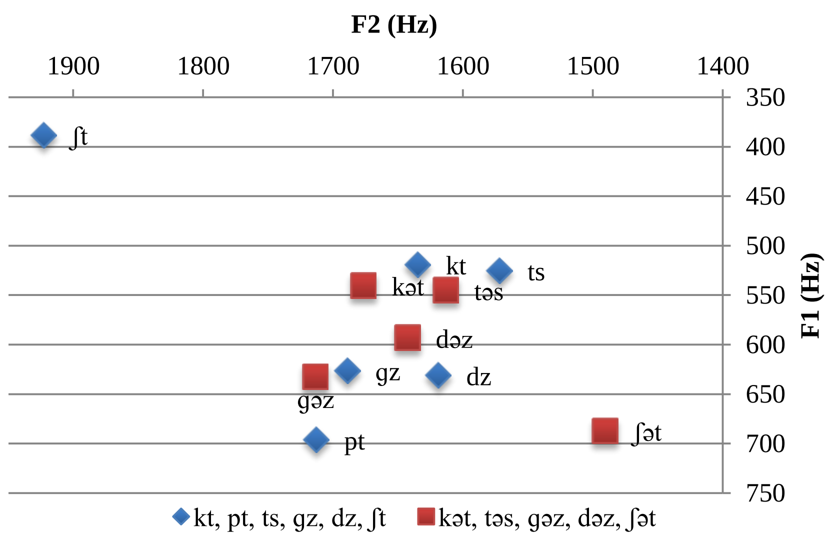

Figure 11. Coronal contexts: Inserted vs. lexical schwas by **female** MC speakers

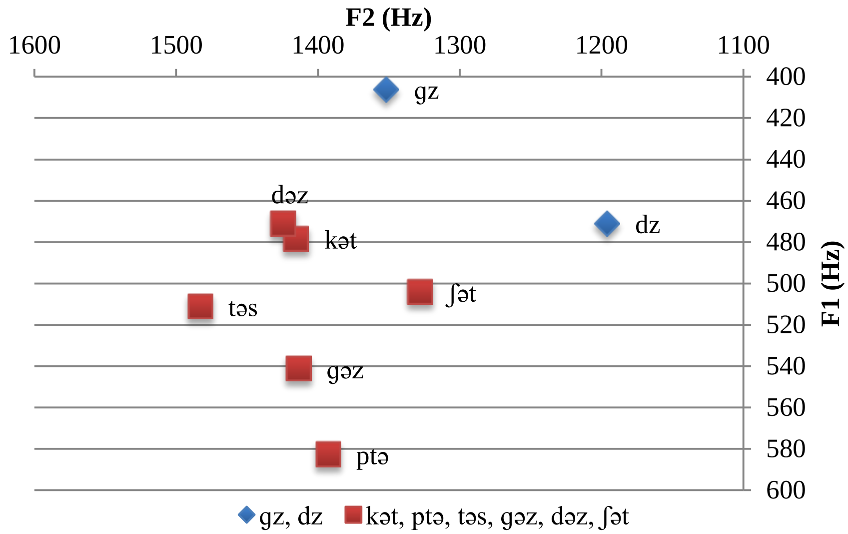

Figure 12. Coronal contexts: Inserted vs. lexical schwas by **male** MC speakers

Figure 13 shows the mean duration values of the inserted vowels and the lexical schwas by all participating speakers. In the examined 14 consonantal contexts, the mean duration of the inserted vowels were shorter than that of the lexical schwas.

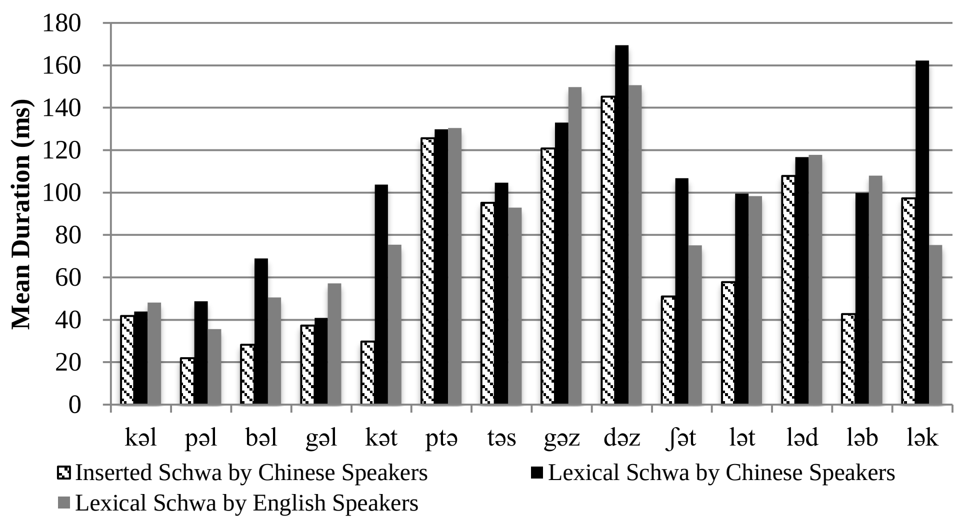

Figure 13. Mean duration values of inserted vowels and lexical schwas

4.5 Performance in syllabification tasks

Table 15 shows the frequencies and the percentages of vowel insertion errors the MC participants made in Target Stimuli in the two syllabification tasks. There were individual differences. C1, C3, and C4 often inserted vocalic sounds into CVCC and CCVC stimuli items to syllabify them as two- or three-syllable sequences. Participants performed consistently across Task 2 (only written stimuli) and Task 4 (only sound stimuli). For example, C2, C5, and C6, who performed with fewer vowel insertions in Task 2 tended to have fewer vowel insertions in Task 4.

Table 15. Vowel insertion errors in four tasks

	C1		C2		C3		C4		C5		C6	
	N[a]	%[b]	N	%	N	%	N	%	N	%	N	%
Task1[c]	2	11%	3	17%	13	72%	2	11%	1	6%	2	11%
Task2	11	61%	0	0%	18	100%	15	83%	0	0%	2	11%
Task3	0	0%	1	6%	2	11%	0	0%	0	0%	1	6%
Task4	12	67%	2	11%	18	100%	18	100%	3	17%	1	6%

Note. [a]N refers to the number of vowel insertion errors in a task. [b]% refers to the percentage of the insertion errors in relation to the target 18 CVCC and CCVC tokens. [c]Task1=reading task, Task2=syllabification task with written stimuli, Task3=repetition task, Task4= syllabification task with sound stimuli.

In the two syllabification tasks, some interesting patterns were observed. Sometimes, vowels were inserted and these were syllabified; at other times, vowels were inserted but not syllabified. It is also interesting to find that, in some cases, only consonants were syllabified without an inserted vowel. Some lexical vowels were not syllabified, a few were not pronounced, and a few were devoiced; some lexical vowels were syllabified, but not pronounced. Table 16 shows the details of MC speakers' performance in the syllabification tasks with all 70 stimuli × 6 participants, including Target Stimuli, Control Stimuli, and Fillers. Table 16 also shows each participant's error type judging from all the data of this study. The error patterns will be discussed in §5.1.4.

Table 16. Performance of MC speakers in the syllabification tasks

	C1		**C2**		**C3**		**C4**		**C5**		**C6**	
Task	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp	Rd	Rp
σ error	20	27	4	15	60	54	51	54	8	18	25	34
ins.V̩ σ	11	8		1	73	71	17	16	1	2	4	1
C̩ σ	11	15	1	6	4	1	50	57	1	2		1
ins.V̯	3	4	7	3	7				2	1	8	7
lex.V̯ (V̥)		6	1	11 (3)	1				7 (5)	15 (3)	21	33 (3)
lex.V̥ σ	14	17	2	1			1	3				
Type	Incorrect UR?		Excres-cent		Incorrect UR		Extra Metric		Excres-cent		Excres-cent?	

Note. σ error = The number of errors in syllabification out of 70 stimuli words; ins.V̩ σ = syllabified inserted vowel: e.g., /kalb/ → [kal-bə]; C̩ σ = syllabified consonant without vowel insertion: e.g., /kapt/ → [kap-t̩]; ins.V̯ = unsyllabified inserted vowel: e.g., /kalb/ → [kalbə̯]; lex.V̯ (V̥) = unsyllabified lexical vowel, which is supposed to be syllabified including devoiced/deleted ones (the bracketed numbers are the only devoiced/deleted ones among all the unsyllabified lexical vowels): e.g., /ka-pət/ → [kapə̯t] ([kapt]); lex. V̥ σ = syllabified but devoiced/deleted lexical vowels: e.g., /ka-pət/ → [ka-pt] or [ka-kə̥t]; Type = presumed error types shown in Figure 1.

5 Discussion

5.1 Key findings

5.1.1 Vowel insertion patterns in CVCC and CCVC

All six MC EAL learners inserted a vowel in L1 illegal consonant clusters in CVCC and CCVC at least once, but the frequency of insertion was low, which was consistent with Nogita and Fan's (2012) results. The vowel insertion errors decreased greatly from the reading task to the repetition task in the current study, similar to the conclusions by Funatsu et al. (2008) and Nogita and Fan (2012). Contrary to vowel insertion, lexical schwas were occasionally deleted (or devoiced) and rhotacized. This may be because [ə] is difficult to produce for MC speakers, so they either deleted or rhotacized it.[7] Interestingly, lexical vowel deletion/rhotacization patterns also decreased from the reading task to the repetition task, except for C3, who deleted and rhotacized more lexical vowels in the repetition task. All the participants may have attempted to make their productions closer to the native CE speakers' productions when they listened to the sound stimuli as the model. C3's vowel insertion dramatically decreased in the repetition task, and the participant might have over-deleted vowels, including lexical vowels. These results suggest that the EAL speakers could successfully

[7] Note that [ɹ] is a common sound in MC, or if they had an impression that consonant clusters and rhotic vowels were common in North American English, they might have hyper-corrected their productions.

perceive the consonant clusters produced by a native speaker and could successfully imitate the target overlap patterns at the phonetic level to some degree.

More specifically, MC speakers in this study seemed to experience difficulties in pronouncing the sequences of stop + /l/ and /l/ + stop. They inserted vowels in all /kl/, /pl/, /bl/, /gl/, /lt/, /ld/, /lb/, and /lk/ combinations. The frequency of insertion errors in these eight sequences was 14, out of the total 27 insertion cases. Similarly, Nogita and Fan (2012) observed that vowel insertion by MC speakers tended to occur before /l/.

5.1.2 Acoustic properties of inserted vowels

Most of the inserted vowels were mid-central vowels and met the description of English schwa according to previous studies (e.g., Flemming, & Johnson, 2007; Kondo, 1994). The average F1 of 27 inserted schwas was 536 Hz; their average F2 was 1565 Hz, and the mean duration was 56 ms.

5.1.3 Native vs. non-native lexical schwas in English

There were no notable differences between MC speakers' and CE speakers' schwas in terms of F1 and F2. However, this does not necessarily mean that MC speakers' schwas are almost native-like. In fact, as Table 6 and 7 above showed, schwas were occasionally rhotacized. Specifically, C4 frequently rhotacized schwas in the reading task. From an acoustic point of view, even though F1 and F2 values of schwas are similar across MC and CE situations, F3 can be quite different if schwa rhotacization occurs because [ɹ] lowers F3. Since /ə/ and /ɚ/ are phonemic in the majority of North American English dialects, schwa rhotacization by MC EAL speakers may need a further investigation. Also, in very careful speech, as in the syllabification tasks in this study, some schwas (especially by C3) sounded like [ɨ] after /s/, and were diphthongized after /g/ and /k/. Certain speech styles may affect the degree to which MC speakers' schwa productions are influenced by their L1 phonology. As well, although schwa deletion/devoicing was possible by native CE speakers in a fast speech, such as /pə̲teᶦɾoᶷ/ *potato* → [pʰteᶦɾoᶷ], in MC learners' productions quite a few schwas were deleted/devoiced in a relatively formal speech style, i.e. word reading tasks. This difference is noteworthy. In terms of the F2 values, native speakers' schwas were more varied than that in MC speakers' productions. We propose two possible reasons. First, the sample size of native speakers was only three. Second, L2 learners might not have native intuition or experience of allophonic variations in English schwas. Future studies can work on this assumption.

5.1.4 Lexical schwa vs. inserted schwa

Recall the four types of L2 vowel insertion shown in Figure 1: a) lexical vowels in misinterpreted L2 UR, b) lexical epenthetic vowels to repair L1 illegal syllable structures, c) non-lexical excrescent vowels caused by gestural mistiming, and d) non-lexical short (or excrescent) vowels along with unsyllabified extraprosodic consonants in misinterpreted L2 UR. When the participants made mistakes in the syllabification tasks, they likely had problems in their UR, suggesting either a) lexical vowels in UR, or d) non-lexical vowels with extraprosodic consonants in UR.

Because C3 almost consistently inserted a vowel in the syllabification task, C3's type of error was assumed to be a misinterpretation of L2 UR, and the inserted vowels existed in their mental lexicon. An interesting case was in the reading task: C3 pronounced /káks/ as [káːkəsɨ], but pronounced /kákəs/ as [káːks]. In C3's IL, both /káks/ and /kákəs/ may be stored as /kákəsə/ (or /kákəsɨ/ or better described as /kákəsi/) in UR, and in SR, she allophonically deleted the weak vowel(s), like /ə/ (and /ɨ/), just as Japanese EAL learners did in Nogita and Fan's (2012) study. C1 also made quite a few errors in the syllabification tasks, suggesting that her problem was also likely a UR issue. C1 deleted lexical schwas more often than inserting extra vowels in the reading and repetition tasks. Also, C1 often syllabified consonants alone in the syllabification tasks. These suggest that in this participant's IL [ə] and [ɨ] may be allophonic variations, but whether [ə] is underlyingly present or absent is not certain. Both C3 and C1's inserted vowels and lexical vowels did not show a difference in duration. This may support the analysis that their inserted vowels tend to be lexical vowels in nature.

One might argue that it is unusual that the written stimuli clearly show the presence or absence of vowels, but the L2 learners still cannot correctly interpret their phonological representations. However, such misinterpretation would be completely possible if learners assume that English orthography or IPA is an abugida or alphasyllabary system, in which each consonant letter is along with a default vowel sound. In fact, Matsumoto (2011) reported that many Japanese learners of Spanish in a Spanish class claimed that *frío* /frío/ and *julio* /xúljo/ were complete homophones, even when the learners looked at the spelling as well as listened to the sounds. Aside from the consonant confusions in /ɾ/ vs. /l/ and /f/ vs. /x/, *xul* and the consonant cluster *fr* confused the learners. Some Japanese EAL learners in Nogita and Fan's (2012) research showed a similar tendency. Previous reports and the current study suggest that spelling cannot necessarily help L2 learners be aware of the difference between CC and CVC.

C2 made the fewest errors in the syllabification tasks. The F1 and F2 values of C2's inserted vowels were lower than those of their lexical vowels, and the duration was much lower in the inserted vowels than that in the lexical schwas as shown in Table 13. According to Davidson (2006), if the inserted schwa is an excrescent sound, its F1 and F2 values can be predicted to be lower than those of lexical schwas. Thus, C2's inserted vowels may be considered as excrescent vowels caused by the gestural mistiming. C2 made much more errors in the syllabification task with the sound stimuli as shown in Table 16. His error patterns were mostly unsyllabified lexical schwas. This might be because he could not detect schwas in the sound stimuli as syllabic unit. In fact, the Japanese participants in Sperbeck's study (2010) often could not detect schwas, so that in the repetition task, the sound stimulus [Cəl], for example, was often produced as [Cl], even though the [Cl] is illegal in Japanese. As shown in Table 14, we conclude that C5's and C6's inserted vowels may be excrescent vowels. Although these two, especially C6, made many errors in the syllabification tasks as shown in Table 16. They mostly did not syllabified lexical schwas, which was very similar to C2's performance. Thus, C5 and C6 may be aware of the presence or absence of vowels in English, but they (especially C6) may not have the knowledge of a higher prosodic structure.

C4's production errors fit in the extraprosodic consonant type. Interestingly, he inserted a vowel only twice in the reading task, but these vowels were fairly long as shown in Table 13. These vowels may be because he simply misread the sound stimuli. In the syllabification tasks, he mostly syllabified consonants alone without inserting a vowel: e.g. /əklák/ → [ə-k^{h}-la-k^{h}]. Inserted vowels in the syllabification tasks mostly followed the release of voiced stops or /l/s: e.g. /blák/ → [b^{ə}-la-k^{h}]. This is similar to the case of aforementioned Japanese participants in Nogita and Fan's (2012) study. Thus, participant C4 may be well aware of the presence or absence of vowels; in his IL, each consonant not followed by a vowel was independent and was not attached to the syllable node.

As for the consonantal contexts, the inserted vowel in [ʃt] produced by C3 was odd, as shown in Figure 11 – it was a high front vowel and its mean duration was short (51 ms). Its highness and frontedness are likely due to the high and front tongue positions of [ʃ] ([Dorsal, Coronal]). Another interesting phenomenon is that the inserted vowel in [ʃt] has clear formant structures and a clear voice bar in the spectrogram, even though its surrounding consonants are both voiceless. In the same fashion, a voicing bar can occur between voiceless consonants (i.e., [k^{ə}t, p^{ə}t, t^{ə}s]). Vowels in such conditions were produced with a target, or at least the speaker actively vibrated his/her vocal fold. In such cases, there are two possibilities. First, based on Lindblom's (1963) Undershoot Hypothesis, such lexical schwas are undershot. For example, C3's short inserted vowels were considered to exemplify this phenomenon, and even C3's devoiced lexical schwas may be undershot. Additionally, one of the reasons that the number of vowel insertions in the repetition task decreased can be explained by lexical vowels being undershot. In fact, Nogita and Fan (2012) concluded that many of their Japanese participants undershot underlying vowels to produce apparent consonant clusters in SR. Second, if the speaker is aware of the absence of vowels but not the syllable structure, he/she might pronounce a consonant as an independent unit. Then, in order to make each consonant more perceptible, he/she might end up with a short schwa-like vowel. Similar phenomena can be observed in other languages, such as in Salish languages, in which non-lexical schwas occur in extraprosodic voiceless consonant sequences (Czaykowska-Higgins & Willett, 1997).

5.2 Implications

Pedagogically, the findings of the present study provide insight into the performance of intermediate level MC EAL learners while pronouncing English consonant clusters. It is important for instructors and learners to be aware that language learners may have specific error types (see Figure 1) in the production of consonant sequences. Moreover, none of the MC participants correctly syllabified all the stimuli. Rather, there were quite a few errors. They might not have known the generlaization that one vowel is assigned to one syllable. It would be helpful in English learning if EAL speakers could explicitly learn the concept of English syllabification.

5.3 Limitations and Future Studies

There are several limitations of this study. First, this preliminary study has only involved six intermediate MC EAL learners. The findings may not be generalized to learners at other English proficiency levels. Meanwhile, only three native CE speakers have participated in this study and their production of English lexical schwas may not be representative of the schwa production by a larger native CE speaker population.

Second, the data of schwas generated from this study are not sufficient in number to run factorial ANOVA tests, so it is hard to know the effects of different factors on the acoustic properties of schwa. For example, the lexical/inserted conditions (e.g., reading vs. listening) may exert a bigger influence on the nature of schwa than the segmental contexts do (Davidson, 2006).

In the theoretical aspect, this study observes that English lexical schwas produced by non-native speakers are more consistent in F2 than the lexical schwa produced by native English speakers. Studies (Flemming & Johnson, 2007; Kondo, 1994) have revealed that the English lexical schwas produced by native speakers fail to be targeted compared with other English vowels. More studies are needed to explore the nature of English lexical schwas produced by non-native speakers from different linguistic backgrounds to see whether non-native English lexical schwas are more targeted than the native English lexical schwa.

6 Conclusion

This study provides insight into the acoustic characteristics of schwas in three different conditions: native English lexical schwa, non-native English lexical schwa, and non-native inserted schwa.

In terms of the nature of inserted schwa, at least two MC speakers (C1 and C3) may have intentionally inserted a target schwa into L1 illegal English consonant sequences CVCC and CCVC because schwas likely exist in their incorrectly memorized inter-language UR. The other three MC speakers (C2, C5, C6) may have the tendency to produce an excrescent insertion. The excrescent schwas are by-products, when the speakers are unable to sufficiently overlap the consonant gestures. Another possibility is that when the speaker is not aware of syllable nodes, some consonants would be left unsyllabified. Since these consonants are treated as one unit and it may be perceptually salient for the speaker, an excrescent vowel may occur. If the speaker becomes aware of syllable nodes, such excrescent vowels may be less likely to occur. One MC participant (C4) shows extraprosodic consonant error patterns: he would potentially pronounce excrescent vowels, although he has not pronounced unambiguous excrescent vowels in the reading and repetition tasks. In any case, it is unnecessary and unwise to define or claim that all the insertion behaviours by non-native speakers are a phonological insertion or a transitional sound. Even within the same L1 background, learners use different production strategies.

As for the lexical schwas, CE speakers and MC speakers have had similar production, in terms of midpoint F1 and duration. Nevertheless, this does not necessarily mean that MC speakers' schwas are native-like. In fact, lexical schwa rhoticization and deletion have been observed. As for the F2 values, lexical schwas produced by MC speakers tend to be more stable in F2 than those produced by native CE speakers. One of the possible explanations is that MC speakers are not as sensitive to allophonic variations as native CE speakers do. As stated before, few empirical studies to date have compared English lexical schwas produced by native and non-native speakers. It is meaningful to find that there are both similarities and differences in the English lexical schwa produced by these two groups. To explain the similarities between native and non-native English lexical schwa found in this study, future L2 studies could conduct the phonetic comparison between the MC lexical schwa by native MC speakers and the English lexical schwa by native English speakers.

Acknowledgements

Many thanks to Dr. Sonya Bird for her instruction of this study and her helpful comments on all the drafts of this paper. Thanks are also due to each of the nine participants, Dr. Ewa Czaykowska-Higgins, Yanan Fan, and Alesia Malec. This study could not have happened without them.

References

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6*, 201-252.

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 341-376). Cambridge: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). “Targetless” schwa: An articulatory analysis. In D. Docherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II* (pp. 26-56). London: Cambridge University Press.

Chan, A. Y. W. (2006). Cantonese ESL learners' pronunciation of English final consonants. *Language, Culture and Curriculum*, *19*(3), 296-313.

Czaykowska-Higgins, E. and M.L. Willett (1997) Simple Syllables in Nxa’amxcin. *International Journal of American Linguistics 63*(3), pp. 385-411.

Davidson, L. (2005). Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics, 19*(6-7), 619-633.

Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics, 34,* 104-137.

Davidson, L., & Kevin, R. (2008). Durational correlates for differentiating consonant sequences in Russian. *Journal of the International Phonetic Association*, *38*(2), 137-165.

Davidson, L., & Stone, M. (2003). Epenthetsis versus gestural mistimimg in consonant cluster production: An ultrasound study. In G. Garding & M. Tsujimura (Eds.), *WCCFL 22 Proceedings* (pp. 165-178). Cambridge, MA: Cascadilla Press.

Fan, Y. (2011). *Articulatory timing of English consonant clusters in the coda positions: A study of Chinese-English interlanguage* (Master’s Thesis). University of Victoria, Victoria, Canada.

Flemming, E. (2004). Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner & D. Steriade (Eds.), *Phonetically-Based Phonology* (pp. 232-276). Cambridge: Cambridge University Press.

Flemming, E., & Johnson, S. (2007). Rosa’s roses: Reduced vowels in American English. *Journal of the International Phonetic Association, 37*, 83-96.

Funatsu, S., Imaisumi, S., Fujimoto, A., Hashizume, M., & Kurisu, K. (2008). Do Japanese speakers perceive nonexistent vowels in non-native consonant clusters? *The Journal of the Acoustical Society of America*, *123*(5), 3072.

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory*, *20*(2), 269–337.

Hansen, J. G. (2001). Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics, 22,* 338-365.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am. 97*(5), 3099-3111.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *J. Acoust. Soc. Am., 59*(5), 1208-1221.

Kondo, Y. (1994). Targetless schwa: Is that how we get the impression of stress timing in English? *Proceedings of the Edinburgh Linguistics Department Conference, 1994*, 63-76.

Lin, H. (2001). *A grammar of Mandarin Chinese.* Munic, Germany: Lincom Europa.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *J. Acoust. Soc. Am.*, *35*, 143-162.

Matsumoto, J. (2011). Percepción de los fonemas secuenciales iniciales /CCV/ y /CVCV/ por los estudiantes japoneses de E/LE con enfoque a las comparaciones entre los grupos y el analisis de los datos individuales, *HISPANICA/HISPÁNICA*, *2011*(55), pp.115-133.

Miao, R. (2005). *Loanword adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factors* (Doctoral Dissertation). Stony Brook University, New York, USA.

Nogita, A., & Fan, Y. (2012). Not vowel epenthesis: Mandarin and Japanese ESL learners' production of English consonant clusters. *Working Papers of the Linguistics Circle (WPLC)*, *22*(1), 1-26.

San, D. (1990). *A formal study of syllable, tone, stress and domain in Chinese languages* (Doctoral Dissertation). MIT, Massachusetts, USA.

Sperbeck, M. (2010). *The acquisition of consonant clusters by Japanese learners of English: interactions of speech perception and production*. Ph.D. dissertation. The City University of New York.

Storkel, H. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language and Hearing Research*, *44*, 1321–1337.

Van Bergem, D. R. (1994). A model of coarticulatory effects on the schwa. *Speech Communication*, *14*(2), 143-162.