Article

Language *and* Speech

Effects of Vowel Duration and Vowel Quality on Vowel-to-Vowel Coarticulation

Language *and* Speech
54(4) 527–545

DOI: 10.1177/0023830911404961
las.sagepub.com
SAGE

Peggy P. K. Mok
Chinese University of Hong Kong

Abstract
This work investigates how vowel duration and vowel quality affect degrees of vowel-to-vowel coarticulation. The effects of these two factors on vowel-to-vowel coarticulation have previously received little study. Phonological durational differences due to vowel length distinction were examined in Thai. It was hypothesized that shorter vowel duration could result in more vowel-to-vowel coarticulation, and further that the vowel /a/ would allow more vowel-to-vowel coarticulation than /i/ or /u/ cross-linguistically. Thus, the susceptibility of different vowel qualities to vowel-to-vowel coarticulation was examined using Thai data. Results show that shorter vowel duration did not affect vowel-to-vowel coarticulation; and the lower the vowel, the more susceptible it is to coarticulation. Possible factors contributing to such patterns are discussed.

Keywords
duration, Thai, vowel length, vowel quality, vowel-to-vowel coarticulation

Introduction

Vowel-to-vowel (V-to-V) coarticulation, referring to the coarticulatory effects of one vowel on another across one or more intervening consonants, was first described by Laclotte (quoted in Scripture, 1902, p. 372), and then discussed by Joos (1948). It was after the seminal paper by Öhman (1966) on Swedish, Russian, and American English that more research interest was drawn to this phenomenon. Since then, V-to-V coarticulation has been studied in many languages (e.g., Butcher & Weiher, 1976; Fowler, 1977; Hawkins & Slater, 1994; Magen, 1984; Manuel, 1990; Purcell, 1979; Recasens, 1984). The effects of various factors on V-to-V coarticulation were also investigated, for example, stress (Fowler, 1981), coarticulatory direction (Recasens, 2002), coarticulatory resistance (Fowler & Brancazio, 2000; Recasens, Pallarès, & Fontdevila, 1997), extent of coarticulation (Magen, 1997; Recasens, 1989), prosody (Cho, 2004), vowel harmony (Beddor & Yavuz, 1995; Boyce, 1990; Przezdziecki, 2005), syllable structure (Modarresi, Sussman,

Corresponding author:
Peggy P. K. Mok, Department of Linguistics and Modern Languages, Leung Kau Kui Building, Chinese University of Hong Kong, Shatin, Hong Kong
Email: peggymok@cuhk.edu.hk

Lindblom, & Burlingame, 2004; Mok, 2007a, 2010), and language differences based on vowel inventory density (Manuel, 1990, 1999; Mok, 2007b).

This paper investigates how two additional factors which were little explored before, vowel duration and vowel quality, affect V-to-V coarticulation. There are only very few studies investigating the effects of durational differences on V-to-V coarticulation, although there are many studies on duration and vowel reduction and coarticulation in general. Additionally, consistent patterns of different susceptibility of /i a u/ to V-to-V coarticulation have been observed in various languages, but the susceptibility of other vowels to V-to-V coarticulation is less clear. This study explores both these issues using data from an Asian language (Thai), hoping to widen the perspective on V-to-V coarticulation.

1.1 Effects of duration on vowel production and coarticulation

There are at least two causes of systematic vowel duration variation: stylistic (due to speaking rate variation) and phonological (due to vowel length distinctions). Many studies have investigated the effects of duration on vowel production and coarticulation (e.g., Engstrand, 1988; Fourakis, 1991; Gay, 1968, 1978, 1981; Lindblom, 1963; Moon & Lindblom, 1994; Tsao, Weismer, & Iqbal, 2006), but only a few studies concern duration and V-to-V coarticulation specifically. A literature search suggests that, so far, only one study has dealt directly with this topic. Hertrich and Ackermann (1995) looked at how a slow speaking rate would affect V-to-V coarticulation in German. While previous studies on duration and coarticulation usually investigated the effects of fast speaking rates, Hertrich and Ackermann compared normal, slow, and slower speaking rates, and found that anticipatory V-to-V coarticulation did not significantly depend on speech rate, but carryover coarticulation did reduce at a slower speech rate. So far, there is no study investigating how vowel length distinction can affect V-to-V coarticulation.

Previous studies on vowel duration and vowel production were often inspired by the work of Lindblom (1963) on target undershoot, which suggested that a shorter duration would result in more vowel reduction. Although target undershoot is not directly related to V-to-V coarticulation, its principle can provide some useful insights for our purpose. The target undershoot model proposed by Lindblom (1963) suggested that duration is the main determinant of vowel reduction. He found that formant frequencies in CVC syllables varied as a function of vowel duration and consonantal context. As rate increased and vowels became shorter, the formants tended to undershoot their acoustic targets, and tended to be displaced towards the frequency values of the adjacent consonants. He suggested that target undershoot occurs because the articulators have less time to complete the invariant motor commands of successive segments so they overlap more at a fast rate.

There are many studies evaluating the notion of target undershoot, but results are not consistent. Some studies found vowel undershoot at fast speaking rate (e.g., Gay, 1968; Gay, Ushijima, Hirose, & Cooper, 1974) or with a shorter vowel duration (Moon & Lindblom, 1994), while others found no difference in formant frequency as a function of speaking rate (e.g., Engstrand, 1988; Fourakis, 1991; Gay, 1978; Nord, 1987; Tsao et al., 2006; Van Son & Pols, 1992). Some methodological issues may explain why there is discrepancy in the above studies, for example, whether the speakers were phonetically trained; whether they were asked to speak clearly; and the use of different vowel contexts (e.g., wVl versus hVd).

Besides methodological differences, vowel reduction can be affected by many other factors in addition to duration. According to Lindblom (1963), target undershoot occurs as a result of more overlap of invariant motor commands. He assumed that velocity of the tongue movement would not change. Gay (1981), however, stated that although target undershoot was common at a fast speaking rate, it was by no means ubiquitous because speakers could control speaking rate by more than one mechanism. Speakers can use different strategies to cope with the rate change, for example, by increasing movement velocity of the articulators (e.g., Cho, 2008; Flege, 1988; Kuehn & Moll, 1976; Matthies, Perrier, Perkell, & Zandipour, 2001).

Lindblom (1983, 1990) later also suggested that at a faster speaking rate, speakers can increase articulatory effort to cope with the rate change. A more recent revised version of the target undershoot model (formant undershoot by Moon & Lindblom, 1994) still proposed that vowel duration and context are the main factors contributing to vowel reduction, but the rate of formant frequency change (related to articulatory effort) can modify the amount of formant undershoot: a faster transition can entail less undershoot.

The studies reviewed above suggest that duration (mainly by manipulating speaking rates) can affect vowel production, although it is not the sole determinant of vowel reduction. A logical question to ask, then, is how duration would affect coarticulation, and for our purpose, V-to-V coarticulation specifically. Gay (1978) suggested that the coordination of articulatory movements at fast rates involves a nonlinear compression along the time dimension. This could lead to decreased duration within segments and increased coarticulation between adjacent segments. His EMG data also showed a greater degree of CV coarticulation during fast speech (see Gay, 1981). Matthies et al. (2001) also hypothesized that a faster speaking rate would result in a general reduction of segmental duration, which could then induce more overlap of the movements for adjacent segments and an increase in coarticulation. They found that coarticulation was somewhat stronger at a fast rate compared to a normal speaking rate, but the effect was small. In addition to coarticulation between consonant and vowel, Hardcastle (1985), Zsiga (1994), and Byrd and Tan (1996) all found an increase in gestural overlap for consonant sequences in both tautosyllabic and heterosyllabic structures (VCCV) spoken at faster rates. Their results show that a shorter duration can induce more overlap between different segments. Therefore, it seems reasonable to assume that a shorter duration can increase coarticulation in general.

However, there are very few studies investigating the effects of durational variation on V-to-V coarticulation specifically. So far, only one study dealt directly with this topic (Hertrich & Ackermann, 1995). Various studies reviewed above showed that a shortened segmental duration could lead to more coarticulation. Hertrich and Ackermann (1995) also showed that an increased duration caused by slower speaking rates could reduce V-to-V coarticulation. Based on these results, one may hypothesize that a shorter duration of the target vowel can result in more V-to-V coarticulation.

Most of the studies discussed above manipulated vowel duration by using speaking rate differences. Vowel duration difference can be phonological too, for example, long versus short vowels. Vowel length distinctions are quite common in the world's languages (Maddieson, 1984). The magnitude of durational difference between long and short vowels can differ among languages (Lehiste, 1970). It is not uncommon to find that the ratio of long-to-short vowel duration can be larger than 2.5 when words containing these vowels are spoken in isolation, which means that short vowels can be 60% shorter than long vowels. This magnitude of durational differences is larger than the stylistic speaking rate manipulation reported in many studies, for example, Gay (1978) (22%), Van Son and Pols (1992) (13%) and Fourakis (1991) (29%), which suggests the possibility that short vowels due to vowel length distinction may be more susceptible to coarticulation than short vowels resulting from speaking rate changes. There are many studies on how speaking rate changes affect vowel length distinctions, for example, Japanese (Hirata, 2004), Korean (Magen & Blumstein, 1993), and Thai (Svastikula, 1986). These studies suggested that rate changes can affect the vowel length distinctions differently in different languages. However, no one has examined the susceptibility of long and short vowels to V-to-V coarticulation. Since short vowels can be up to around 60% shorter than long vowels, the target undershoot model discussed earlier would predict that short vowels are more susceptible to reduction and coarticulation than long vowels. This prediction is partially supported by the literature because short vowels are often more centralized than their long counterparts in many languages (e.g., Abramson & Ren, 1990; Johnson & Martin, 2001; Keating & Huffman, 1984; Lee, 1983; Lehiste, 1970). Based on these findings, short vowels may also exhibit more V-to-V coarticulation than long vowels, but the effects are still unknown, since the majority of the studies on durational effects on vowel production manipulated vowel duration using speaking rates. Phonological durational effects (long versus short vowels) on V-to-V coarticulation are investigated using Thai data in this study.

Nevertheless, there are a number of differences between a shorter vowel duration caused by speaking rate change and one caused by vowel length distinctions. First of all, rate change potentially affects all speech sounds in an utterance, although the magnitude can be different for consonants and vowels. Vowels are proportionally shortened more than consonants (Gay, 1981). Vowel length distinctions affect only the target vowels. Therefore, rate change has a global effect while vowel length distinction has a local effect. Second, rate change is under the subjective control of the speakers, while vowel length distinction is determined by the language. Third, as suggested by one reviewer, rate change is gradient but vowel length distinctions are potentially categorical. These differences make the investigation using phonologically short vowels even more interesting. Moon and Lindblom (1994) manipulated vowel duration not by speaking rates but by using the so-called "word-length effect" in English. For example, the vowel /i/ gets increasingly shortened in *speed*, *speedy*, *speedily*. Their use of the "word-length effect" ensured that vowel shortening is not caused by speakers subjectively controlling their speaking rates. They found strong undershoot at varying vowel duration at a constant speaking rate. Vowel shortening induced by the "word-length effect" is not unlike vowel shortening caused by phonological vowel length distinctions. Their results can be taken to support the hypothesis that phonologically short vowels may allow more V-to-V coarticulation than long vowels.

1.2 Effects of vowel quality on V-to-V coarticulation

Many studies in the literature show that high vowels like /i/ and /u/ are more resistant to V-to-V coarticulation than /a/ cross-linguistically (e.g., Beddor, Harnsberger, & Lindemann, 2002; Beddor & Yavuz, 1995; Cho, 2004; Kondo & Arai, 1998; Recasens, 1987). Since the same phenomenon is reported in many languages, it is possible that there is a physiological explanation. However, it is unclear what determines the susceptibility of different vowels to V-to-V coarticulation since most of the studies on V-to-V coarticulation only used a limited number of vowels, most notably /i a u ə/.

There are two accounts in the literature which may provide some ideas for the different susceptibility of vowel qualities to V-to-V coarticulation. The three vowels /i a u/ are considered to be more stable than other vowels by the Quantal Theory (Stevens, 1989) because they are being produced in acoustically stable regions of the articulatory space. They are the so-called "quantal vowels". The theory suggests that non-quantal vowels are more variable than quantal vowels. Therefore, it is possible that non-quantal vowels could be more susceptible to coarticulation than quantal vowels. However, Diehl (1989) questioned the stability of quantal vowels because stability according to one articulatory configuration can mean instability in other vocal tract parameters. Furthermore, the three quantal vowels themselves also differ in coarticulatory susceptibility. It seems that the Quantal Theory alone cannot fully account for the susceptibility of different vowel qualities to V-to-V coarticulation.

Another model that addresses why vowels may differ in their "coarticulability" is Recasens' degree of articulatory constraint (DAC) model. The DAC model suggests that how "coarticulable" a vowel is can be predicted by the degree of tongue dorsum involvement during its production (Recasens, 2002; Recasens & Pallarès, 2001; Recasens et al., 1997). That is, segments with the tongue dorsum raised would constrain coarticulatory effects, and are themselves resistant to coarticulation. Each phonetic segment is assigned a DAC value ranging from 1 (minimum) to 3 (maximum). Coarticulatory sensitivity is inversely related to DAC. The DAC model is more concerned with consonants than vowels, but as stated in Recasens et al. (1997, p. 545), the assignment of DAC values to vowels does not differ from that of consonants. They proposed such a ranking from maximally to minimally constrained segments: /ʃ ɲ i k ɫ s/ (DAC = 3) > /n l a u/ (DAC = 2) > /p ə/ (DAC = 1) (Recasens, 2002; Recasens et al., 1997). They suggested that V-to-V coarticulation involves the DAC specification for both the intervocalic consonants and the target vowels.

The DAC model provides a simple descriptive account of coarticulatory sensitivity of vowels, but must be tested on a wider range of languages and stimuli. The DAC model has been developed primarily from Catalan data, although it claims to account for universal patterns. It is more concerned with consonants than vowels. For example, Recasens and Pallarès (2001) tested many consonants, such as /p t n ɫ s r ʃ ʎ ɲ j k/, but only two vowels /i/ and /a/, although more vowel data were reported by the same author before the model was proposed (Recasens, 1985). The model also does not specify the DAC value assignment of vowels other than /i a u ə/. In addition, the relationship between DAC values and coarticulatory sensitivity is more complex for V-to-V coarticulation since it involves both consonants and vowels, and the patterns are also less consistent (Recasens, 2002; Recasens et al., 1997). The discussion about V-to-V coarticulation and DAC values by these authors appears to be more concerned with the direction of coarticulation rather than the sensitivity of different vowels to coarticulation since mainly two or three vowels were used in their studies. Therefore, although the DAC model provides a general framework to account for the susceptibility of vowel qualities to V-to-V coarticulation, much work must still be carried out before clear predictions can be made.

In this paper, the susceptibility to coarticulation of six selected vowels is examined in Thai. There are altogether nine vowel qualities in Thai, each with long and short counterparts: /i e ɨ æ a ɤ ɔ o u/. Some short vowels are more open than the long counterparts, for example, /i e o u/ (Abramson & Ren, 1990). So in order to keep the analyses comparable, /e o/ were excluded but /i u/ were retained because the two vowels /i u/ are often used in V-to-V coarticulation research. The vowel /ɨ/ was not used because short /ɨ/ is not combined with /p/ as a monosyllabic real word in Thai (for testing the durational effects of vowel length).

In summary, two little-explored factors on V-to-V coarticulation are tested using Thai data in this study: vowel duration and vowel quality. It is hypothesized that phonologically short vowels would allow more V-to-V coarticulation than long vowels. The "coarticulability" of six vowel qualities /i æ a ɤ ɔ u/ is also compared.

2 Method

2.1 Speakers

Six native Bangkok Thai speakers (three male and three female) were used. They were students at the University of Cambridge for less than three years. They were all born in Bangkok and spent most of their lives there. All speakers were in their 20s or 30s and reported no language or hearing impairment. They were paid for participating in the experiment.

2.2 Materials

The experimental materials were trisyllables in the form of /pV1pV2paː/. Both long and short counterparts of the six chosen vowel qualities were used, resulting in twelve target vowels: /iː i, æː æ, aː a, ɤː ɤ, ɔː ɔ, uː u/. Context vowels were of the same qualities but always long, that is, /iː æː aː ɤː ɔː uː/. Target vowels can be V1 (for investigating anticipatory coarticulation) or V2 (for investigating carryover coarticulation). The twelve target vowels were put into all possible combinations with the context vowels. Table 1 shows all the combinations for target /a/ as an example. The last syllable in the trisyllables was

Table 1. All combinations of experimental trisyllables for target /a/ (in bold) in Thai

Context vowel	Long target /aː/		Short target /a/	
	Anticipatory	Carryover	Anticipatory	Carryover
/iː/	/**paː**piːˈpaː/	/piː**paː**ˈpaː/	/**pa**piːˈpaː/	/piː**pa**ˈpaː/
/æː/	/**paː**pæːˈpaː/	/pæː**paː**ˈpaː/	/**pa**pæːˈpaː/	/pæː**pa**ˈpaː/
/aː/	/**paː**paːˈpaː/	/paː**paː**ˈpaː/	/**pa**paːˈpaː/	/paː**pa**ˈpaː/
/ɤː/	/**paː**pɤːˈpaː/	/pɤː**paː**ˈpaː/	/**pa**pɤːˈpaː/	/pɤː**pa**ˈpaː/
/ɔː/	/**paː**pɔːˈpaː/	/pɔː**paː**ˈpaː/	/**pa**pɔːˈpaː/	/pɔː**pa**ˈpaː/
/uː/	/**paː**puːˈpaː/	/puː**paː**ˈpaː/	/**pa**puːˈpaː/	/puː**pa**ˈpaː/

'Anticipatory' and 'carryover' refer to the direction of contextual variation.

always /paː/, but it was not analyzed. It was only there to take the stress, since primary stress falls on the final syllable of a word or a phrase in Thai. The individual target syllables (/pV/) carry a mid tone (unmarked in Thai script), and many of them appear in Thai either as a monosyllabic word or as part of a polysyllabic word. The resultant trisyllables are nonsense sequences. There were altogether 108 trisyllables (6 vowel qualities × 2 lengths × 2 coarticulatory directions × 6 contexts minus 36 repetitive ones for long target vowels). Five repetitions of each trisyllable were collected, which were embedded in carrier phrases with similar syntactic structures and stress patterns as shown below. They were of medium length in order to elicit more natural speech. The trisyllables were presented as foreign names in the carrier phrases. All five repetitions were used for analysis.

The five carrier phrases in broad phonemic transcription are as follows. The numbers show the tones of the syllables: 1 = mid tone (32), 2 = low tone (21), 3 = high tone (51), 4 = rising tone (45), 5 = falling tone (213):

1. /saw5 t^hiː3 læw4 aː1 tɕaːn1 ___________ paj1 p^huː1 ket2/
   (Last Saturday Professor ______ went to Phuket.)
2. /p^hut4 t^hiː3 læw4 aː1 tɕaːn1 ___________ paj1 duː1 naŋ1/
   (Last Wednesday Professor ______ went to a movie.)
3. /suk2 t^hiː3 læw4 aː1 tɕaːn1 ___________ paj1 sɨː4 k^hɔŋ5/
   (Last Friday Professor ______ went shopping.)
4. /tɕan1 t^hiː3 læw4 aː1 tɕaːn1 ___________ paj1 wai3 nam4/
   (Last Monday Professor ______ went swimming.)
5. /a1 t^hit4 t^hiː3 læw4 aː1 tɕaːn1 ___________ paj1 ŋan1 sam5 ma4 naː1/
   (Last Sunday Professor ______ went to a seminar.)

2.3 Procedures

All speakers were recorded in a sound-treated room at the phonetics laboratory at the University of Cambridge. Before the actual recording, they practiced by reading a randomized list of some of the materials as many times as they liked. They read the materials in Thai script with a normal speaking rate. Their speech was recorded using a DAT tape with a Sennheiser MKH 40 P48 microphone and a Symetrix SX 202 amplifier into a Sony DTC-60ES recorder and later down-sampled using *Xwaves* with a sampling frequency of 16 kHz. If the speakers made a mistake or hesitated when reading the sentences, they were asked to re-read the whole sentence again immediately. Short breaks were also given during the recording.

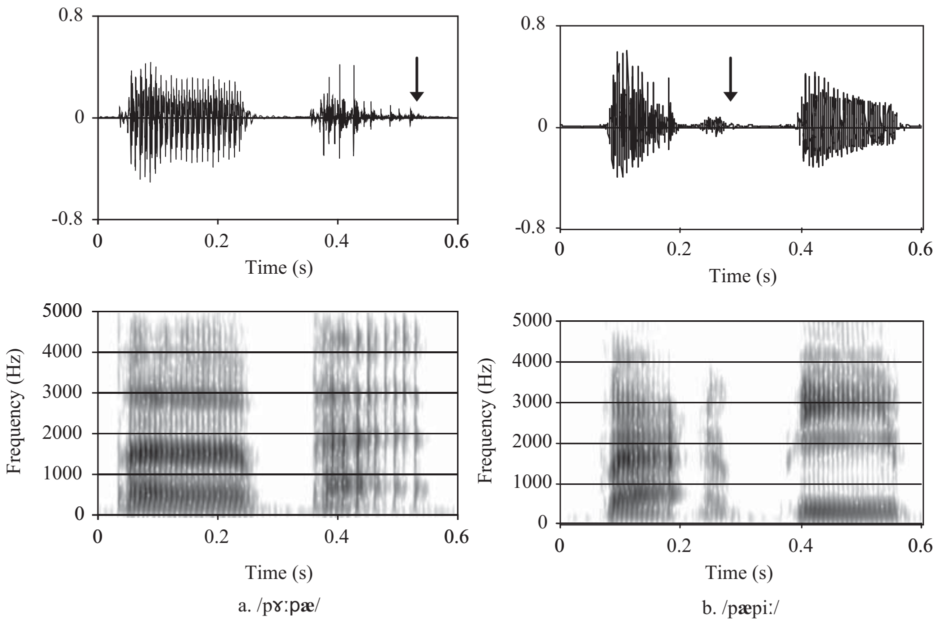

Figure 1. Waveforms and spectrograms of two creaky short vowels /æ/ produced by two male Thai speakers. (a) /pɤːpæ/ produced by male speaker 1; (b) /pæpiː/ produced by male speaker 2

2.4 Acoustic measurements and data analysis

The frequencies of the first two formants (F1 and F2) were measured from 18 pole 25 ms autocorrelation LPC spectra with a Hanning window. All LPC readings were manually checked by reference to the wide band spectrogram and DFT spectra. If there was a difference of more than 50 Hz in the LPC reading and the DFT spectra, formant frequencies were taken from the DFT spectra. About one third of the LPC readings were corrected this way. The temporal locations were identified from the waveform where the beginning and ending of periodic vocalic voicing was taken for the onset and offset of the target vowels. Vowel onset and offset were determined with reference to formant patterns shown in the spectrogram. The midpoint was at the point halfway between the onset and offset. F1 and F2 frequencies were taken from two temporal locations of the target vowels: vowel edge and midpoint. Vowel edge was either the onset or offset of the target vowels, depending on the direction of coarticulation (/pV1pV2paː/): onset of V_2 for assessing carryover coarticulation and offset of V_1 for assessing anticipatory coarticulation. Duration of the target vowels was also measured.

It was, however, not always easy to decide when the vowels end for short vowels in Thai because sometimes they were quite creaky, especially the open ones. Abramson (1962, 1974) mentioned that short vowels commonly end in a glottal stop in Thai. In the present data, glottalization could make up half or more of the phonated portion, see Figure 1a. For a few extreme cases, a weak continuation of the formants or some “detached” pulses appeared after a short period with no periodic voicing at all, see Figure 1b. Previous studies dealt with such cases differently. For example, Gandour (1984) and Sittachit (1972) did not mention them so it is unclear what they did, while Abramson (1962) excluded the “detached” portion as in Figure 1b. Since there were cases with varying degrees of glottalization in the data, the end of the last clear glottal pulse was taken as the end of the vowel for consistency, as indicated by the two arrows in Figure 1. This means that sometimes there was less than 25 ms of voicing for measuring F1 and F2 at vowel offset. Also, the duration of the short vowels reported here will probably be longer than those reported in the literature.

All formant frequency data were normalized using P_{gm} described in Mok (2007b). Each mean F1 (or F2) measurement averaged over all the repetitions for each target vowel in a given context was expressed as a proportion of the grand mean (F1 or F2) of all vowel tokens collected from a particular speaker. There are two steps involved: 1) Calculate the grand mean (F1 or F2) of all target vowels averaged across all conditions for a particular speaker. The grand mean represents the hypothetical center of the speaker's vowel space, for example, 500 Hz for F1. 2) Calculate the context-induced proportional F1 (or F2) difference from that grand mean ($current_{Target\ V}$/grand mean), for example, 700 Hz/500 Hz = 1.4. See the formulae for normalization below. A value bigger or smaller than 1 means that the vowel formant is higher or lower than the hypothetical center. For example, the F1 of /i/ should always be lower than 1, while the F2 of /i/ should always be higher than 1. The deviations of the proportions from the grand mean (i.e., 0.4 in the above example) only represent the normalized distance from the hypothetical center of each speaker's vowel space, but not degree of V-to-V coarticulation. These proportions were used for statistical analysis.

Formulae for normalization (P_{gm}, proportions of the grand mean).

$$\bar{x} = \frac{\sum_{j=1}^{n} x_j}{n}$$

$$P_{gm_i} = \frac{x_i}{\bar{x}}$$

where
x_j = F1 (or F2) of the target vowel
n = total number of vowel tokens for each speaker
$\bar{x}$ = the grand mean of a given formant averaged across all target vowels and conditions
x_i = context-dependent F1 (or F2) of current vowel i
P_{gmi} = P_{gm} for the current vowel i

In interpreting the statistical results in the following sections, if the sphericity assumption of any main effect or interaction in the repeated measures ANOVAs is violated, the degree of freedom is adjusted with the Hyunh-Feldt epsilon in generating the F ratios and p values. When appropriate, the interpretation of the data is supplemented by comparing the partial η^2 of different factors, which shows how much of the variance in the dependent variable can be attributed to the factor(s) in question. Also, the Dunn-Sidak adjustment is used for all post hoc multiple comparisons to control for family-wise Type I error. Two-tailed t-tests are used to compare simple main effects.

3 Results

3.1 Vowel duration

In order to simplify the analysis, target vowel duration (ms) was collapsed across context vowels and was submitted to a 3-way repeated measures ANOVA (Length × Direction × Target). Figure 2 shows the durations of all the long and short target vowels averaged across speakers and contexts. As expected, the main effect of Length is significant, $F(1, 5) = 22.389$, $p = 0.005$. Paired t-tests show that all long vowels are significantly longer than short ones, /i/, $t(5) = 4.118$, $p = 0.009$; /æ/, $t(5) = 3.845$, $p = 0.012$; /a/, $t(5) = 3.794$, $p = 0.013$; /ɤ/, $t(5) = 5.225$, $p = 0.003$; /ɔ/, $t(5) = 6.003$, $p = 0.002$; /u/, $t(5) = 5.580$, $p = 0.003$. But on average, long vowels are only around 1.24 times longer than short vowels. The difference is the biggest for target /u/ (1.32) and smallest for target /æ/ (1.17), with the Length × Target interaction being significant, $F(5, 25) = 3.756$, $p = 0.011$.

The main effect of Direction is also significant, $F(1, 5) = 23.596$, $p = 0.005$, with vowels in the anticipatory direction (V_1, 142 ms) longer than those in the carryover direction (V_2, 128 ms).

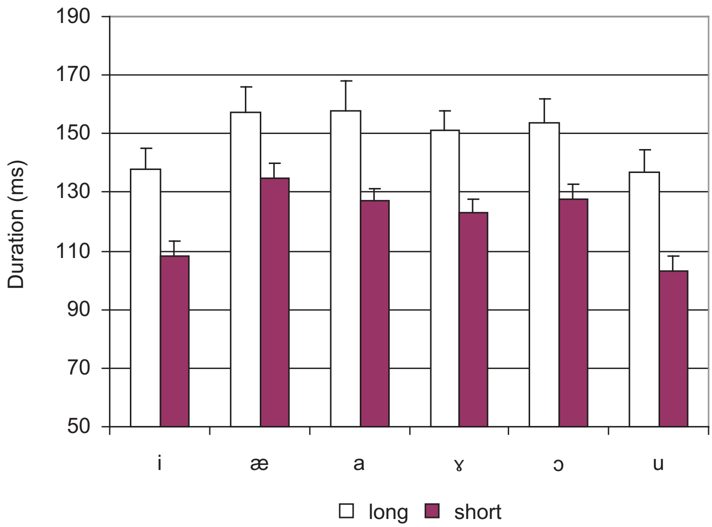

Figure 2. Duration (ms) of long and short target vowels in Thai. Ratios for long/short target vowels: /i/: 1.28; /æ/: 1.17; /a/: 1.24; /ɤ/: 1.23; /ɔ/: 1.20; /u/: 1.32

3.2 Formant frequencies

The normalized formant frequency data were submitted to four (2 formants × 2 temporal locations) 4-way repeated measures ANOVAs: Length (long, short) × Direction (anticipatory, carryover) × Target vowels (/i æ a ɤ ɔ u/) × Context vowels (/i æ a ɤ ɔ u/). Because there are many levels of target and context vowels, separate 3-way repeated measures ANOVAs (Length × Direction × Context) were also conducted for each target vowel in order to facilitate the interpretation of the 4-way ANOVA results. Figure 3 shows the twelve target vowels averaged across all speakers and contexts at both vowel edge and midpoint. Their distributions in the normalized F1/F2 plot are in the expected relationship. The vowels at midpoint are a bit more peripheral than at vowel edge, as the formant frequencies at midpoint are more canonical.

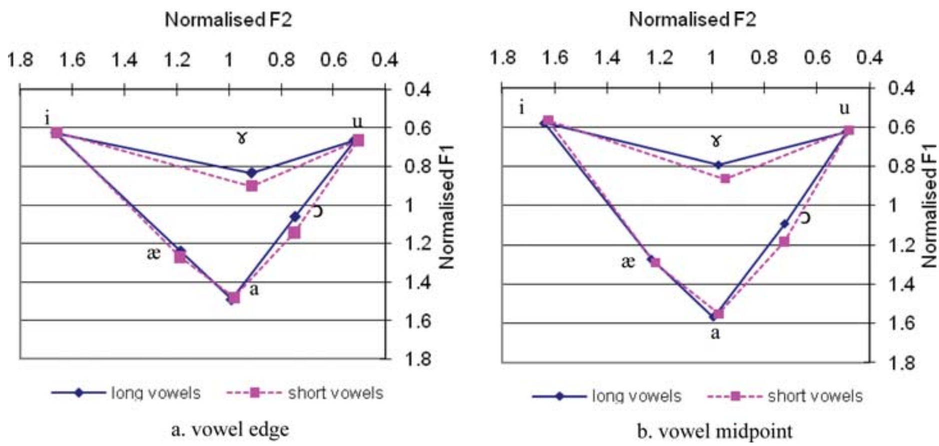

Figure 3. Normalized F1 and F2 of the 12 target vowels at (a) vowel edge and (b) vowel midpoint averaged over two coarticulatory directions

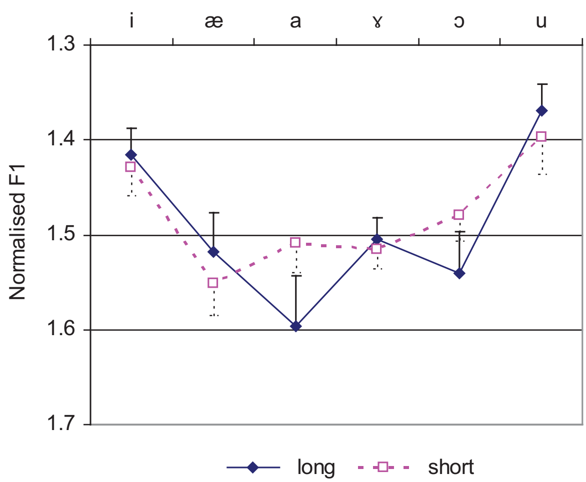

Figure 4. Normalized F1 of target /a/ at vowel edge with different context vowels

In F1, the Length × Target interaction is significant at both vowel edge, $F(5, 25) = 2.864$, $p = 0.035$, and midpoint, $F(5, 25) = 2.794$, $p = 0.039$. The 3-way ANOVA results for each target vowel indicate that the Length difference is only significant for target /ɤ/, with short vowels having a higher F1 than long vowels at both vowel edge, $F(1, 5) = 40.437$, $p = 0.001$, and midpoint, $F(1, 5) = 37.812$, $p = 0.002$. There is also a non-significant trend for target /ɔ/ at the vowel edge, $F(1, 5) = 5.545$, $p = 0.065$, for the same pattern. These patterns can be seen in Figure 3. The results suggest that short /ɤ/ is more open than long /ɤ/.

The main research question is whether vowel length affects V-to-V coarticulation. The Length × Target × Context interaction at vowel edge is significant in F1, $F(25, 125) = 2.389$, $p = 0.001$. The 3-way ANOVA results for each target vowel reveal that the interaction is only significant for target /a/, $F(4.439, 22.197) = 3.832$, $p = 0.014$ (see Figure 4), but two-tailed *t*-tests indicate that vowel length does not induce any significant difference on target /a/ for all of the context vowels.

In F1, a high order interaction involving both Length and Direction (Length × Direction × Target × Context) is significant at both vowel edge, $F(25, 125) = 2.010$, $p = 0.007$, and midpoint, $F(25, 125) = 2.613$, $p < 0.001$, indicating that both factors jointly affect V-to-V coarticulation. The 3-way ANOVA results indicate that this interaction is only significant for target /æ/ (vowel edge, $F(5, 25) = 3.177$, $p = 0.024$; midpoint, $F(5, 25) = 4.624$, $p = 0.004$). Figure 5 shows the Length × Direction × Context interactions for target /æ/ at vowel edge and midpoint. The effects of the two factors in this interaction are discussed separately below.

Vowel length is only significant for context /u/ in the carryover direction at both vowel edge and midpoint (compare the 2nd and 4th bars at ‘u’ in Figure 5a and Figure 5b; vowel edge: $t(5) = -3.781$, $p = 0.013$; vowel midpoint: $t(5) = -3.209$, $p = 0.024$). Context /u/ lowers F1 of long /æ/ more than short /æ/ in both cases, which means that, contrary to our expectation, there is more V-to-V coarticulation for long /æ/ than short /æ/ with a /u/ context. However, there is no consistent pattern of long vowels allowing more V-to-V coarticulation than short vowels in the data.

As for the effect of Direction, there is more carryover than anticipatory coarticulation for target /æ/. F1 of long /æ/ is lowered more by the two high carryover context vowels /i/ and /u/ at vowel edge and midpoint (compare the 1st and 2nd bars at ‘i’ and ‘u’ in Figure 5a; context /i/: $t(5) = 4.063$, $p = 0.010$; context /u/: $t(5) = 7.892$, $p = 0.001$; and compare the 1st and 2nd bars at ‘u’ in Figure 5b,

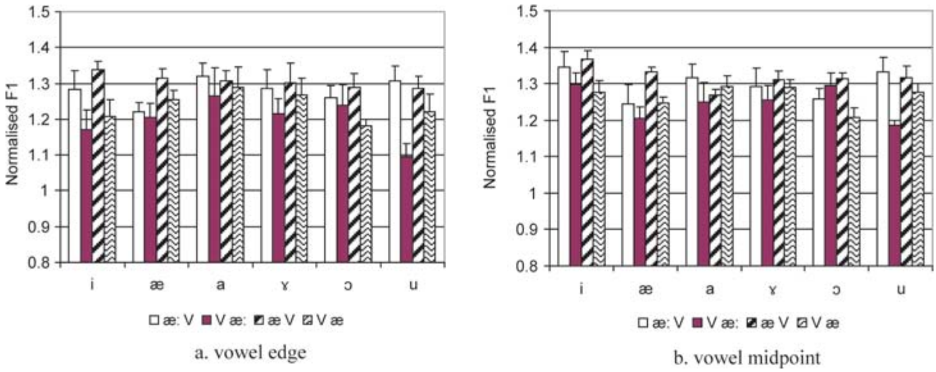

Figure 5. Normalized F1 of target /æ/ showing contextual effects interacting with Length and Direction at (a) vowel edge and (b) midpoint. 'æ: V' = long /æ/ with anticipatory vowel context; 'V æ:' = long /æ/ with carryover vowel context; 'æ V' = short /æ/ with anticipatory vowel context; 'V æ' = short /æ/ with carryover vowel context

$t(5) = 3.443$, $p = 0.018$). In addition, F1 of short /æ/ is also lowered more by contexts /i/, /æ/ and /ɔ/ in the carryover direction (compare the 3rd and 4th bars at 'i', 'æ' and 'ɔ' in Figure 5b; 'i': $t(5) = 3.708$, $p = 0.014$; 'æ': $t(5) = 6.078$, $p = 0.002$; 'ɔ': $t(5) = 3.208$, $p = 0.024$). These results show that carryover coarticulation is stronger than anticipatory coarticulation for both long and short target /æ/.

In F2, Length did not significantly interact with Context, meaning that vowel length did not affect V-to-V coarticulation. The main effect of Length, $F(1, 5) = 21.926$, $p = 0.005$, and the Length × Target interaction, $F(5, 25) = 2.646$, $p = 0.047$, are significant at vowel midpoint. The 3-way ANOVAs indicate that the Length difference in F2 is significant for target /a/ and target /ɤ/ only, both with a higher F2 for long vowels (/a/, $F(1, 5) = 46.502$, $p = 0.001$; /ɤ/, $F(1, 5) = 14.889$, $p = 0.012$). However, the difference is rather small (see Figure 3b). There is no other significant result involving Length in the F2 dimension.

The Thai data show that, contrary to the hypothesis that a shorter duration will induce more V-to-V coarticulation, short vowels do not show more coarticulation than long vowels. Long /æ/ instead allows more carryover coarticulation than short /æ/ in F1 with a /u/ context, but no consistent pattern of long vowels allowing more V-to-V coarticulation than short vowels was found. There is no Length effect on V-to-V coarticulation in F2.

3.3 Vowel qualities

Not every vowel is equally susceptible to coarticulation. Using the Thai data, we can roughly determine their maximum "coarticulability" by comparing the largest range of significant contextual variations of each target vowel and using partial η^2 of the Context main effects for each target vowel as a supplementary measure. The Context main effects of each target vowel show evidence of V-to-V coarticulation averaged across context vowels, Length and Direction. Partial η^2 is an estimate of the effect size of a factor or an interaction of factors. It shows how much of the overall variability of the dependent variable can be attributed to the factor in question. Its value ranges from zero to one; the greater the value, the greater the effect size of the factor. However, it is only used as a supplementary measure to show the *relative* strength of the averaged "coarticulability" of various target vowels here. No claim to the suitability of using partial η^2 as a measure of the total effect size in a multifactorial ANOVA model is being made here. Since V-to-V coarticulation is the

Table 2. Effect size of the contextual variations on different target vowels at vowel edge for F1 (upper panel) and F2 (lower panel) in Thai

Target vowel	Most extreme contexts with significant pairwise comparison	Normalized range of contextual variation	Partial η^2 for Context main effect
F1			
/a/	/a/ vs. /u/ (p = 0.013)	0.170	0.764
/æ/	/a/ vs. /u/ (p = 0.020)	0.068	0.359
/ɔ/	/ɤ/ vs. /u/ (p = 0.026)	0.056	0.492
/i/	/ɔ/ vs. /u/ (p < 0.001)	0.010	0.322
/u/	Non-significant	–	0.343
/ɤ/	Non-significant	–	0.268
F2			
/a/	/i/ vs. /u/ (p = 0.016)	0.0712	0.798
/æ/	/i/ vs. /u/ (p = 0.003)	0.0501	0.657
/ɔ/	/i/ vs. /ɔ/ (p = 0.004)	0.0492	0.777
/ɤ/	/a/ vs. /u/ (p = 0.027)	0.0384	0.667
/u/	/i/ vs. /u/ (p = 0.051)	0.0230	0.523
/i/	Non-significant	–	0.387

The table is sorted according to the normalized range of contextual variation. See the main text for explanation.

most prominent at vowel edge because of the proximity to the source of contextual influence, comparing differences at vowel edge would suffice.

There are six context vowels for each target vowel. The partial η^2 shows the averaged effects of the context vowels. The maximum "coarticulability" of each target vowel can be compared by using the most extreme contexts which theoretically should result in the strongest V-to-V coarticulation. The most extreme contexts in Table 2 are determined by two principles. First, the two contexts must be significantly different in the post hoc pairwise comparisons with Sidak adjustment. Second, if there is more than one pair of significantly different contexts, then the pair that is farthest apart in either F1 or F2 is chosen. For example, contexts /i/ and /u/, and /i/ and /ɔ/ are both significantly different in F2 for target /a/. The pair /i/ and /u/ were used as the extreme contexts. The first context vowels listed in Table 2 represent the context vowels which induce the highest F1 (or F2) for a particular target vowel; the second context vowels induce the lowest F1 (or F2). As expected, the most extreme contexts are usually /a/ and /u/ for F1, and /i/ and /u/ for F2, but there are exceptions. The range of contextual variation represents the difference in normalized formant frequencies caused by the two most extreme context vowels, for example, context /i/ minus context /u/ in the above example. A larger difference indicates more V-to-V coarticulation. The difference between contexts /i/ and /u/ of target /u/ in F2 just misses significance (p = 0.051). It is nonetheless included for a more comprehensive picture of the data.

Table 2 shows that, for both F1 and F2, target /a/ is the most susceptible to V-to-V coarticulation, followed by targets /æ/ and /ɔ/. The two high vowels /i/ and /u/ are the least susceptible. This pattern is consistent for both formants, although the effects seem to be more robust for F2 because of more differences reaching significance.

A general conclusion from these results is that, the lower the vowel, the more susceptible it is to V-to-V coarticulation. The vowel /a/, with the most open jaw position, shows the most V-to-V coarticulation, followed by /ɔ/ and /æ/ which require a closer jaw position than /a/. The reason why high vowels /i/ and /u/ are the least susceptible to coarticulation seems likely to be related to the high jaw position and also a high tongue body position, rather than the "anchor" effect or stability of them being point vowels or "quantal vowels", because the low point vowel /a/ allows the most coarticulation.

4 Discussion

4.1 Duration

The main issue in this paper is to examine whether a shorter duration caused by phonological vowel length distinction induced more V-to-V coarticulation. The results show that it did not. Nevertheless, in order to accept this conclusion, we need to make sure that the crucial factor, vowel duration, is within the normal range of the language first.

There is quite a large difference between the ratios for long/short vowels in this study (ranging from 1.17 to 1.32) and those reported in the literature, for example, 2.4 and 3.2 for two speakers in Abramson (1962), ranging from 2.2 to 3.3 for five speakers in Gandour (1984), and 2.6 for the normal speakers in Gandour et al. (1987). However, smaller ratios were also reported in the literature, for example, 1.7 for a single speaker in Sittachit (1972) and 1.8 for six speakers in Roberson (1982). A recent study by Abramson (2001) showed that the ratios can change under different speaking conditions, for example, from 1.5 for words embedded in carrier phrases spoken at a fast rate to 2.2 in conversational excerpts (similar to citation forms in other studies) read by the same speakers. Similar to the present data, Abramson's data showed a fair amount of overlap of long and short vowel durations in different speaking conditions. This overlap is contrary to earlier studies which showed no or only a very slight overlap for long and short vowel duration (e.g., Abramson, 1962; Gandour, 1984; Gandour et al., 1987).

There are a number of reasons why the ratios for long/short vowels are smaller in the present study than those in the literature. First, weak glottalization at the end of short vowels was included as part of the vowels in the present study. Not all previous studies have clearly indicated how they dealt with such cases. It is possible that this measurement difference has contributed to a longer duration of some short vowels. However, this cannot be a major reason because not all short vowels were equally creaky, and there were only a few extreme cases. A more important reason is that, the long/short vowel ratios reported in the literature mainly compared vowels in citation forms or vowels embedded in short carrier phrases. Abramson (2001) stressed that citation form or dictionary entry cannot always be taken as the basis to determine vowel length in Thai because there are some rule-governed shifts in connected speech which can affect the actual realizations of long and short vowels. Since the materials in this study were in trisyllables embedded in carrier phrases of medium length, the durational differences between long and short vowels will be smaller than those measured in (near) citation forms.

Another possible reason for the smaller ratios in this study is related to vowel positions. The results show that vowel duration is longer in the anticipatory direction (V_1) than in the carryover direction (V_2). This is because the experimental materials are in the form of trisyllables. The second syllable in a trisyllable, that is, in the carryover direction in the present study, is often shorter than the first and the last ones, a phenomenon also found in Chinese. This explains why there is a directional difference in vowel duration. Experimental materials in previous studies, however, were mainly monosyllabic words, and often under focus. Therefore, a difference in the ratios for long/short vowels based on different experimental materials can be expected. All in all, although there is some difference between the long/short vowel ratios in the present data and those reported in earlier studies, the vowel length distinction is still robustly maintained by the speakers.

Although it is intuitively appealing to expect that a shorter duration could lead to more coarticulation, it is not too surprising to find otherwise as some studies reviewed in the Introduction do point to the possibility that duration may only exert a minimal effect on vowel reduction and coarticulation. There are two factors which may have affected the durational effects on V-to-V coarticulation.

The first one is the percentage or magnitude of durational differences. The long vowels in Thai are only about 1.24 times longer than short vowels. Strong vowel undershoot was reported by Moon and Lindblom (1994) with vowels shortened by 40%–60%. However, no vowel undershoot was reported in several studies with smaller durational differences, for example, Gay (1978) with an overall 22% vowel shortening in fast rate; Van Son and Pols (1992) with an overall 13% vowel shortening; and Fourakis (1991) with an overall 29% vowel shortening. Mok (2007b) compared V-to-V coarticulation in normal and fast speaking rates using Cantonese and Mandarin data. She also found that speaking rate did not significantly affect degree of V-to-V coarticulation when the rate difference is about 20%. All these results suggest that vowel undershoot (and an increase in coarticulation) would only occur with a large durational difference, probably when the speakers are close to their maximum speed of articulation.

Flege (1988, p. 909) reported an interesting phenomenon which supports the explanation of the magnitude of durational difference. In his first experiment with only one female subject, the subject showed significant lingual undershoot for both /ɪ/ and /ɑ/. The averaged vowel shortening at fast rate was about 45%. However, the same subject shortened the vowel /ɪ/ much less in the second experiment (23% vs. 51%), and consequently, also showed much less undershoot for /ɪ/. Instead, she shortened /ɑ/ more in the second than the first experiment (49% vs. 40%), and also showed more undershoot in the second experiment.

The difference in magnitude of durational differences can explain the contradicting results on vowel reduction and duration reported in the literature quite well because the critical factor, that is, the percentage or magnitude of durational change, was not precisely controlled in all of the studies, including the present one. There is no pre-defined percentage or magnitude of durational change for various speaking rates in previous studies. This is understandable because as Gay (1978) reported, speakers often have difficulties in controlling different levels of speaking rates precisely by themselves. Also, the magnitude of durational difference between long and short vowels differs among languages (Lehiste, 1970), so it is hard to ensure a pre-defined percentage of phonological durational change.

What seems to be crucial but lacking in the literature, though, is the minimal threshold of durational difference or the articulatory limit of maximum speed that can "guarantee" target undershoot and more coarticulation. No such information is found in Lindblom (1963) because he plotted the formant frequencies against all vowel durations. There was a significant negative correlation between the two in his data, but there was no critical threshold. From the several studies quoted above, it seems that vowel duration needs to be shortened by at least 40% in order to induce target undershoot, which suggests that this percentage of durational change may be close to the maximum articulatory speed limit. In order to identify such a threshold, researchers may need to manipulate several levels of rate changes in the experiments or control rate changes explicitly, for example, by using a metronome. De Jong (2001a, 2001b) showed that speakers were able to follow the rhythm of the metronome closely. Using a metronome allows researchers to explicitly manipulate the percentage difference between various speaking rates.

Another possible factor modulating the effect of duration on V-to-V coarticulation is related to clarity. Gay (1978) found that when speakers were told to maintain phonetic identity of the vowels, unstressed vowels were not substantially reduced. Matthies et al. (2001) tested this hypothesis to see whether coarticulation is influenced by listeners' need for clarity. They found that clarity did influence formant frequencies, but the effect was small.

In this study, subjects were not told explicitly to speak clearly. They were told to speak naturally. However, the speakers were recorded in experimental settings. It is possible that, even with ample practice, speakers were still consciously speaking clearly. Flege (1988) also suggested that different factors affecting the effects of speaking rate on the tongue movements are likely to be influenced by how clearly the subjects would like to speak. Speakers can speak quickly and yet precisely (Zsiga, 1994). The trade-off between clarity and vowel reduction or coarticulation echoes Gay's (1981) claim that speakers do not control their rate of speech either by a single mechanism or along a single dimension. Indeed, speakers' desire for clarity can override other factors affecting vowel production. This important point is also taken into account by Lindblom (1983, 1990) in his H&H theory which suggested that there is a trading relation between clarity and the economy of effort.

This study tested the hypothesis that a shorter vowel duration would induce more V-to-V coarticulation. The results disconfirmed this hypothesis. Short vowels did not allow more V-to-V coarticulation than long vowels in Thai. The magnitude of durational difference and the speakers' effort to maintain clarity may explain the lack of durational effect. The present data agree with many earlier studies that there is no simple one-to-one relationship between duration and coarticulation, at least when the durational difference is less than 40%.

4.2 Vowel qualities

The Thai data show a continuum of susceptibility to V-to-V coarticulation: /a/ > /æ/, /ɔ/ > /ɤ/ > /i/, /u/. This suggests that the lower the vowel, the more susceptible it is to V-to-V coarticulation. Besides /i/ and /u/ in the present study, another high vowel /e/ was also quite resistant to V-to-V coarticulation. In Manuel (1987), the vowel /e/ did not allow much V-to-V coarticulation in three African languages examined (Ndebele, Shona, and Sotho). Her results are comparable with the present study because the consonant /p/ was used in both experiments. Her results can be taken as extra evidence supporting the proposal of a high jaw position limiting the degree of V-to-V coarticulation a vowel can accommodate. Because the low vowel /a/ allows more V-to-V coarticulation than the high vowels /i/ and /u/ in many languages, it is probably related to physical constraints of the articulators.

The reason why a high jaw position affects the susceptibility to V-to-V coarticulation is possibly because when the jaw is high, there is simply less space and freedom for the tongue to move around, so the tongue is less affected by the context. The oral cavity is more open for low vowels, so probably the tongue is also freer to coarticulate. Also, because low vowels involve the movement of a sluggish articulator (the jaw), possibly the tongue can respond to the neighboring contexts before the jaw has fully completed its trajectory. In addition, compared to high vowels, low vowels have a longer distance to travel to and from the adjacent consonants which have higher jaw and possibly tongue positions, so the effects of coarticulation would be more prominent in low vowels than high vowels.

Nevertheless, it should be noted that the continuum of vowel susceptibility to V-to-V coarticulation can also be explained in terms of tongue dorsum involvement as proposed by the DAC model (Recasens et al., 1997). The two accounts of jaw height and tongue dorsum involvement are compatible. It is quite likely that both factors jointly contribute to the susceptibility of different vowels to V-to-V coarticulation, as when the jaw gets lower, the tongue dorsum will be less involved in forming a constriction. More studies on vowels involving different degrees of jaw opening in various languages are needed to further verify the results found here. It will be useful to compare vowels that have a similar degree of jaw opening but differ in the front/back dimension or tongue dorsum constriction. Further, it will be interesting to separate the effects of jaw height and tongue dorsum constriction by using vowels having different tongue dorsum involvement with bite blocks. Such results will give insightful answers to the "coarticulability" of different vowels.

5 Conclusion

The effects of two lesser-known factors on V-to-V coarticulation are tested in this study: vowel duration and vowel quality. Experimental results show that a shorter duration did not induce more V-to-V coarticulation. The percentage or magnitude of durational change is important in modulating the durational effects on V-to-V coarticulation. Speakers' effort to maintain clarity can also override any durational effects. There is no simple one-to-one relationship between duration and V-to-V coarticulation. Low vowels are more susceptible than high vowels to V-to-V coarticulation cross-linguistically. Results show that there is a continuum of vowel quality on coarticulation: the lower the vowel is, the more susceptible to coarticulation it is. The jaw position is proposed to account for this pattern, in addition to tongue dorsum involvement. Since this pattern is caused by physiological constraints and is found in many languages, the effects of vowel quality may be a universal influence on V-to-V coarticulation. More studies using data from different languages are needed to verify this conclusion.

Acknowledgements

The author is grateful to Sarah Hawkins for help and guidance throughout the project. She also thanks Kirsty McDougall and Francis Nolan for helpful discussions during the process, and Jiranthara Srioutai for helping with the Thai materials and finding Thai speakers. Thanks also go to Arthur Abramson for his comments on earlier versions of the texts, and to the Editor Timothy Bunnell, Harriet Magen, and an anonymous reviewer for their helpful suggestions on the manuscript. This research was supported by the Sir Edward Youde Memorial Fellowships for Overseas Studies from Hong Kong and an Overseas Research Studentship from the United Kingdom.

References

Abramson, A. S. (1962). The vowels and tones of standard Thai: Acoustical measurements and experiments (Bloomington: Indiana University Research Centre in Anthropology, Folklore and Linguistics. Publication No. 20). *International Journal of American Linguistics*, *28*, Part III.

Abramson, A. S. (1974). Experimental phonetics in phonology: Vowel duration in Thai. *Pasaa: Notes and News about Language Teaching and Linguistics in Thailand*, *4*, 71–90.

Abramson, A. S. (2001). The stability of distinctive vowel length in Thai. In K. Tingsabadh & A. S. Abramson (Eds.), *Essays in Tai linguistics* (pp. 13–26). Bangkok: Chulalongkorn University Press.

Abramson, A. S., & Ren, N. (1990). Distinctive vowel length: Duration vs. spectrum in Thai. *Journal of Phonetics*, *18*, 79–92.

Beddor, P. S., Harnsberger, J. D., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates. *Journal of Phonetics*, *30*, 591–627.

Beddor, P. S., & Yavuz, H. K. (1995). The relationship between vowel-to-vowel coarticulation and vowel harmony in Turkish. *Proceedings of the 13th International Congress of Phonetic Sciences (ICPhS)*, 44–51. Stockholm, Sweden.

Boyce, S. E. (1990). Coarticulatory organization for lip rounding in Turkish and English. *Journal of the Acoustical Society of America*, *88*, 2584–2595.

Butcher, A., & Weiher, E. (1976). An electropalatographic investigation of coarticulation in VCV sequences. *Journal of Phonetics*, *4*, 59–74.

Byrd, D., & Tan, C. C. (1996). Saying consonant clusters quickly. *Journal of Phonetics*, *24*, 263–282.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *32*, 141–176.

Cho, T. (2008). Prosodic strengthening in transboundary V-to-V lingual movement in American English. *Phonetica*, *65*, 45–61.

De Jong, K. (2001a). Effects of syllable affiliation and consonant voicing on temporal adjustment in a repetitive speech-production task. *Journal of Speech, Language and Hearing Research*, *44*, 826–840.

De Jong, K. (2001b). Rate-induced resyllabification revisited. *Language and Speech*, *44*, 197–216.

Diehl, R. (1989). Remarks on Stevens' quantal theory of speech. *Journal of Phonetics*, *17*, 71–78.

Engstrand, O. (1988). Articulatory correlates of stress and speaking rate in Swedish VCV utterances. *Journal of the Acoustical Society of America*, *83*, 1863–1875.

Flege, J. E. (1988). Effects of speaking rate on tongue position and velocity of movements in vowel production. *Journal of the Acoustical Society of America*, *84*, 901–916.

Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *Journal of the Acoustical Society of America*, *90*, 1816–1827.

Fowler, C. A. (1977). *Timing control in speech production*. Bloomington: Indiana University Linguistics Club.

Fowler, C. A. (1981). Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research*, *46*, 127–139.

Fowler, C. A., & Brancazio, L. (2000). Coarticulation resistance of American English consonants and its effects on transconsonantal vowel-to-vowel coarticulation. *Language and Speech*, *43*, 1–41.

Gandour, J. (1984). Vowel duration in Thai. *Crossroads*, *2*, 59–64.

Gandour, J., Weinberg, B., Petty, S. H., & Dardarananda, R. (1987). Vowel length in Thai alaryngeal speech. *Phonetica*, *39*, 117–121.

Gay, T. (1968). Effects of speaking rate on diphthong formant movements. *Journal of the Acoustical Society of America*, *44*, 1570–1573.

Gay, T. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America*, *63*, 223–230.

Gay, T. (1981). Mechanisms in the control of speech rate. *Phonetica*, *38*, 148–158.

Gay, T., Ushijima, T., Hirose, H., & Cooper, F. S. (1974). Effect of speaking rate on labial consonant–vowel articulation. *Journal of Phonetics*, *2*, 47–63.

Hardcastle, W. J. (1985). Some phonetic and syntactic constraints on lingual co-articulation during /kl/ sequences. *Speech Communication*, *4*, 247–263.

Hawkins, S., & Slater, A. (1994). Spread of CV and V-to-V coarticulation in British English: Implications for the intelligibility of synthetic speech. *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*, 57–60.

Hertrich, I., & Ackermann, H. (1995). Coarticulation in slow speech: Durational and spectral analysis. *Language and Speech*, *38*, 159–187.

Hirata, Y. (2004). Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics*, *32*, 565–589.

Johnson, K., & Martin, J. (2001). Acoustic vowel reduction in Creek: Effects of distinctive length and position in the word. *Phonetica*, *58*, 81–102.

Joos, M. (1948). Acoustic phonetics. *Language*, *24*, 1–136.

Keating, P., & Huffman, M. K. (1984). Vowel variation in Japanese. *Phonetica*, *41*, 191–207.

Kondo, Y., & Arai, Y. (1998). Prosodic constraint on V-to-V coarticulation in Japanese. *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP)*, 1755–1758. Sydney, Australia.

Kuehn, D. P., & Moll, K. (1976). A cinefluorographic investigation of CV and CV articulatory velocities. *Journal of Phonetics*, *3*, 303–320.

Lee, T. (1983). The vowel system in two varieties of Cantonese. *UCLA Working Papers in Phonetics*, *57*, 97–114.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *35*, 1773–1781.

Lindblom, B. (1983). Economy of speech gestures. In P. F. Macneilage (Ed.), *The production of speech* (pp. 217–245). New York: Springer.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht: Kluwer.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Magen, H. S. (1984). Vowel-to-vowel coarticulation in English and Japanese. *Journal of the Acoustical Society of America*, *75*, *Suppl*. S1, T11.

Magen, H. S. (1997). The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *25*, 187–205.

Magen, H. S., & Blumstein, S. E. (1993). Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics*, *21*, 387–409.

Manuel, S. (1987). *Acoustic and perceptual consequences of vowel-to-vowel coarticulation in three Bantu languages* (Unpublished PhD thesis). Yale University.

Manuel, S. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America*, *88*, 1286–1298.

Manuel, S. (1999). Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In W. J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 179–198). Cambridge: Cambridge University Press.

Matthies, M. L., Perrier, P., Perkell, J. S., & Zandipour, M. (2001). Variation in anticipatory coarticulation with changes in clarity and rate. *Journal of Speech, Language and Hearing Research*, *44*, 340–353.

Modarresi, G., Sussman, H., Lindblom, B., & Burlingame, E. (2004). An acoustic analysis of the bidirectionality of coarticulation in VCV utterances. *Journal of Phonetics*, *32*, 291–312.

Mok, P. P. K. (2007a). Effects of syllable structures on v-to-v coarticulation (Thai vs English). *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, 421–424. Saarbrücken, Germany.

Mok, P. P. K. (2007b). *Influences on vowel-to-vowel coarticulation* (Unpublished PhD thesis). Department of Linguistics, University of Cambridge.

Mok, P. P. K. (2010). Language-specific realizations of syllable structure and vowel-to-vowel coarticulation. *Journal of the Acoustical Society of America*, *128*, 1346–1356.

Moon, S. J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America*, *96*, 40–55.

Nord, L. (1987). Vowel reduction in Swedish. In O. Engstrand (Ed.), *Papers from the Swedish Phonetics Conference* (pp. 16–21). Uppsala.

Öhman, S. E. G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, *39*, 151–168.

Przezdziecki, M. (2005). *Vowel harmony and coarticulation in three dialects of Yorùbá: Phonetics determining phonology* (Unpublished PhD thesis). Department of Linguistics, Cornell University.

Purcell, E. T. (1979). Formant frequency patterns in Russian VCV utterances. *Journal of the Acoustical Society of America*, *66*, 1691–1702.

Recasens, D. (1984). Vowel-to-vowel coarticulation in Catalan VCV sequences. *Journal of the Acoustical Society of America*, *76*, 1624–1635.

Recasens, D. (1985). Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences. *Language and Speech*, *28*, 97–114.

Recasens, D. (1987). An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences. *Journal of Phonetics*, *15*, 299–312.

Recasens, D. (1989). Long range coarticulation effects for tongue dorsum contact in VCVCV sequences. *Speech Communication*, *8*, 293–307.

Recasens, D. (2002). An EMA study of VCV coarticulatory direction. *Journal of the Acoustical Society of America*, *111*, 2828–2841.

Recasens, D., & Pallarès, M. D. (2001). Coarticulation, assimilation and blending in Catalan consonant clusters. *Journal of Phonetics*, *29*, 273–301.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, *102*, 544–561.

Roberson, J. (1982). *Bangkok Thai tones* (Unpublished PhD thesis). University of Southern California.

Scripture, E. W. (1902). *The elements of experimental phonetics*. London: Edward Arnold.

Sittachit, K. (1972). Observations on short and long vowels in Thai (in Thai). In J. G. Harris & R. B. Noss (Eds.), *Tai phonetics and phonology* (pp. 28–32). Bangkok: Central Institute of English Language.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*, 3–45.

Svastikula, M. L. K. (1986). *A perceptual and acoustic study of the effects of speech rate on distinctive vowel length in Thai* (Unpublished PhD thesis). University of Connecticut, ProQuest Information and Learning, Ann Arbor.

Tsao, Y. C., Weismer, G., & Iqbal, K. (2006). The effect of intertalker speech rate variation on acoustic vowel space. *Journal of the Acoustical Society of America*, *119*, 1074–1082.

Van Son, R. J., & Pols, L. C. (1992). Formant movements of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustical Society of America*, *92*, 121–127.

Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics*, *22*, 121–140.