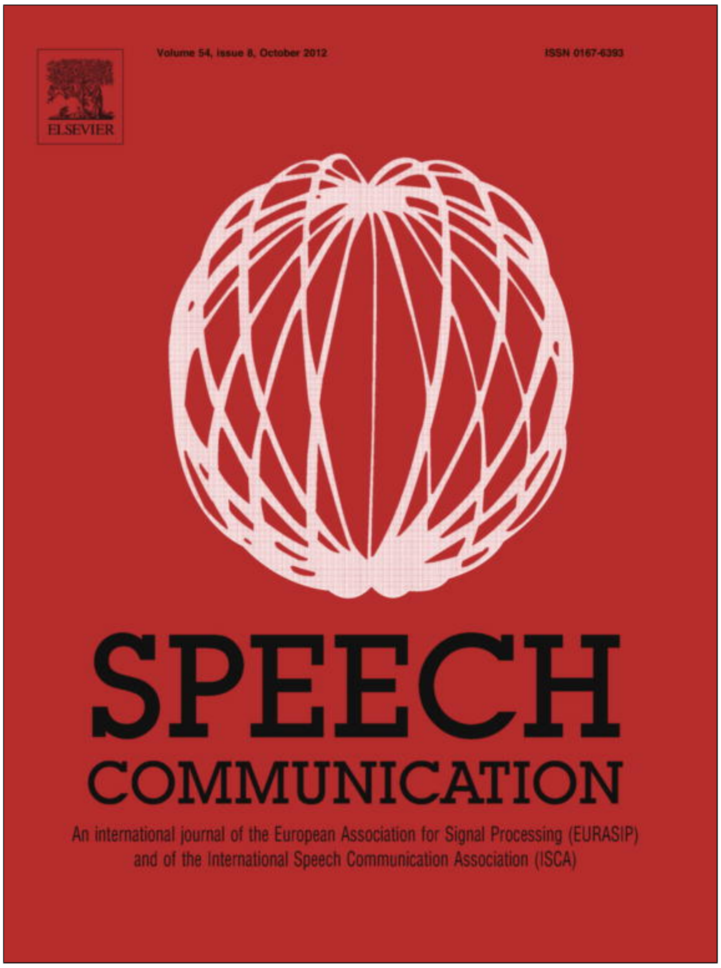

Volume 54, Issue 8, October 2012
ISSN 0167-6393
ELSEVIER
SPEECH
COMMUNICATION
An international journal of the European Association for Signal Processing (EURASIP)
and of the International Speech Communication Association (ISCA)

Available online at www.sciencedirect.com

SciVerse ScienceDirect

ELSEVIER

Speech Communication 54 (2012) 946–956

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

Effects of consonant cluster syllabification on vowel-to-vowel coarticulation in English

Peggy P.K. Mok *

Department of Linguistics and Modern Languages, Leung Kau Kui Building, The Chinese University of Hong Kong, Shatin, Hong Kong

Received 29 June 2011; received in revised form 4 April 2012; accepted 4 April 2012
Available online 10 April 2012

Abstract

This paper investigates how different syllable affiliations of intervocalic /st/ cluster affect vowel-to-vowel coarticulation in English. Very few studies have examined the effect of syllable structure on vowel-to-vowel coarticulation. Previous studies show that onset and coda consonants differ acoustically, articulatorily, perceptually and typologically. Onsets are stronger, more stable, more common and more distinguishable than codas. Since codas are less constrained, it was hypothesized that coda /st./ would allow more vowel-to-vowel coarticulation than onset /.st/. Three vowels (/i ɑ u/) were used to form the target sequences with the /st/ cluster in English: onset /CV.stVC/, heterosyllabic /CVs.tVC/, coda /CVst.VC/. F1 and F2 frequencies at vowel edges and the durations of the first vowel and the intervocalic consonants were measured from six speakers of Standard Southern British English. Factors included in the experiment are: Direction, Syllable Form, Target, Context. Results show that coda /st./ allows more vowel-to-vowel coarticulation than onset /.st/, and heterosyllabic /s.t/ is the most resistant among the Syllable Forms. Vowels in heterosyllabic /s.t/ are more extreme than in the other two Syllable Forms in the carryover direction. These findings suggest that vowel-to-vowel coarticulation is sensitive to different syllable structure with the same segmental composition. Possible factors contributing to the observed patterns are discussed.

Keywords: Vowel-to-vowel coarticulation; Syllable structure; English; Consonant clusters

1. Introduction

Vowel-to-vowel (V-to-V) coarticulation refers to the coarticulatory influence of one vowel on another across intervening consonant(s). Since Öhman (1966) seminal study on V-to-V coarticulation, the effects of many factors on V-to-V coarticulation have been investigated, for example, stress (Fowler, 1981), prosodic hierarchy (Cho, 2004), vowel harmony (Beddor and Yavuz, 1995) and vowel inventory density (Manuel, 1990; Mok, accepted for publication). V-to-V coarticulation also differ in dimensions like coarticulatory direction (Recasens, 2002), coarticulatory resistance (Recasens et al., 1997) and extent of coarticulation (Recasens, 1989). However, the effects of syllable structure on V-to-V coarticulation remain poorly understood. Most studies on V-to-V coarticulation deal with only one syllable type, i.e., open syllables. Only very few studies examined the effects of different syllable types across singleton consonants, i.e., open vs. closed syllables (Modarresi et al., 2004a; Mok, 2010); and even fewer studies involved consonant clusters (e.g. Recasens and Pallarès, 2001). This study investigates how different syllabifications of intervocalic consonant clusters in English with the same segmental composition (/st/) as onset /.st/, a heterosyllabic sequence /s.t/ and coda /st./ affect V-to-V coarticulation (where '.' denotes a syllable/word boundary).

Despite not having strong evidence for phonetic correlates of the syllable boundary, there are many studies showing that syllable onset and coda are different acoustically, articulatorily, perceptually and typologically. Acoustically, onset consonants are relatively longer than coda consonants

* Tel.: +852 3943 1768; fax: +852 2603 7755.
E-mail address: peggymok@cuhk.edu.hk

0167-6393/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.specom.2012.04.001

in non-prepausal position (Anderson and Port, 1994; de Jong et al., 2004; Lehiste, 1970; Quené, 1992). The closure duration of intervocalic stop is a good indicator of syllabification (Boucher, 1988; Tuller and Kelso, 1991). Syllabification of two-consonant clusters is also characterized by different acoustic durational patterns of the two consonants, e.g. long–short (onset), long–long (heterosyllabic) or short–long (coda) (Christie, 1974; Haggard, 1973a,b). In addition to duration, spectral properties are also related to syllable structure. CV transitions are generally more informative about stop consonant identity than VC transitions, which may be due to CV unit motor programming (see review in Pickett et al., 1995). Locus equations are different for onset versus coda stops, with stronger consonant–vowel coarticulation for CV than VC units (Sussman et al., 1997). Also using locus equations, Modarresi et al. (2004b) demonstrated that stop + vowel coarticulation across a syllable or word boundary (C.V) was weaker than within a CV syllable, but were still greater than in VC in both English and Persian (CV > C.V > VC).

Articulatorily, many studies show that gestures for onset consonants are stronger and more distinct than those for coda consonants (Browman and Goldstein, 1988, 1995; Byrd, 1996; Macchi, 1988; Sproat and Fujimura, 1993; Tuller and Kelso, 1991). Onset position is stronger than coda position because it is generally associated with tighter articulatory constrictions which involve greater articulatory effort and with less variability. Krakow (1999) thoroughly reviewed the literature on nasal, lateral and stop articulations and syllable structure in American English. She concluded that onset and coda consonants are associated with different characteristic articulatory patterns. Two more recent studies concur with Krakow's conclusion. Gick et al. (2006) investigated the intergestural timing of liquids in different syllable positions in six languages using ultrasound imaging. Likewise, Kochetov (2006) examined the syllable-position effects of three Russian consonants /p^j p j/ using EMMA data. These two studies found that syllable position could indeed be distinguished and characterised articulatorily in each language, but there was no single pattern that characterised syllable positions across all the languages. For example, the timing and the presence or absence of tongue dorsum gestures in liquids are not uniform across syllable positions in the six languages examined in Gick et al. (2006). Although the intergestural patterns were language-specific, they were consistently different in onset and coda positions within each language. These studies show that there is a clear relationship between articulatory gestures and syllable position.

Onset and coda consonants also differ in their coordination with the vowel. Browman and Goldstein (1988) and Honorof and Browman (1995) suggested that for prevocalic consonants or consonant clusters, it was the center of the consonant sequence (the C-center) that was most tightly coordinated with the following vowel gesture. For postvocalic consonants or consonant clusters, however, it was the left edge, rather than the C-center, of the consonant sequence. Browman and Goldstein (1995) argued that differences in stability of intergestural phasing relation patterns between American English onsets and codas can provide a unified account of some seemingly unrelated phonological phenomena of nasals and laterals. Other articulatory studies also reveal that the variability of consonant cluster coordination with the vowels in VC_nV sequences forms a continuum with respect to syllable affiliation, with onset being the most stable: tautosyllabic onset > tautosyllabic coda > heterosyllabic coda + onset (across word/syllable boundary) (Byrd, 1995; Nam and Saltzman, 2003).

In addition, there are perceptual differences between onset and coda consonants. Redford and Diehl (1999) found that onset consonants had greater acoustic distinctiveness than coda consonants. In their study, onset consonants were consistently identified more accurately than coda consonants in noise. A robust adaptation effect specific to syllable position was demonstrated by Samuel (1989). His results indicated that the perceptual system was sensitive to syllable structure. Selective adaption refers to the repetitive presentation of a stimulus in order to induce changes in the perception of related stimulus. For example, when subjects were presented with many /ba/ stimuli, they would identify fewer /ba/ stimuli when hearing ambiguous stimuli between clear /ba/ and /pa/. Samuel found that /aeb/ and /aed/ adaptation shifted the identification of /aeb/ – /aed/ (with the consonant in the same syllable position), but not /bae/ – /dae/ (with the consonant in different syllable position). Moreover, VC syllables are often perceived as CV syllables when repeated at fast rates (Stetson, 1988; Tuller and Kelso, 1990, 1991). Tuller and Kelso suggested that the perceptual shift was a result of the loss of stability of the VC pattern because speakers tended to move from the less stable VC coordination to the more stable CV coordination. The shift in syllabification perception in fast rates is also caused by the loss of important juncture cues (de Jong, 2001).

The above acoustic, articulatory and perceptual differences between onset and coda are also echoed by typological data. Syllables with onsets are much more common than syllables with codas in the world's languages. CV is the only and the most frequent syllable type that occurs in all languages. The most frequent syllable type with a coda (CVC) also has an onset. Structures with no onset (VC) are rare. In addition, phonologically onset-less stressed syllables are often produced with a glottal stop as an empty onset in many languages. Some languages do not even allow structure with no onset, e.g. Arabic. However, structures with no coda (CV) are very frequent and some languages prohibit codas altogether, e.g. Hawaiian. In addition, many languages allow more consonants in the onset than in the coda position and impose stricter constraints on coda (Bell and Hooper, 1978; Greenberg, 1978; Maddieson, 1984).

The preference for CV syllables in the world's languages can also be found in infants' babbling patterns. CV syllables predominate in early babbling, even for children whose languages have coda consonants. Normally developing children up to three years old often leave out coda consonants (Vihman, 1996). These developmental results demonstrate the preference for CV over VC or other syllable types, which may be motivated by biomechanical constraints (MacNeilage and Davis, 2000).

The many studies discussed above strongly point to the conclusion that onsets and codas are different in various aspects, and that the syllable is an important unit in both production and perception. Onset consonants are longer, stronger, more stable, and have tighter coordination with the vowel than coda consonants. Onsets are more common than codas, and are also perceived more accurately. If syllable structure is indeed an influential factor in speech production and perception, then it is reasonable to expect that syllable structure would affect V-to-V coarticulation differently. However, the effects of syllable structure on V-to-V coarticulation are still unclear as very few researchers have investigated this issue.

Two studies had examined V-to-V coarticulation across singleton consonants with different syllable affiliations. Modarresi et al. (2004a) compared nonsense sequences /CV.CV/ with /tVC.Vt/ in American English using six intervocalic stops /b p d t g k/ and four Target vowels /i e u ɔ/. They found that in general, VCV sequences with a closed syllable (VC.V) had slightly more overall V-to-V coarticulation than VCV sequences with an open syllable (V.CV), i.e., the coda allowed more coarticulation than the onset. They explained such results by the different temporal intervals between open and closed syllables in their data: they measured F2 frequencies at two temporal locations for carryover coarticulation. They asked their subjects to fully release all the final stops in /tVC.Vt/ sequences. They measured F2 frequency at the release of the intervocalic stop burst for closed syllables /tVC.Vt/ but at the onset of periodicity of the second vowel for open syllables /CV.CV/, so the measuring points for closed syllables /tVC.Vt/ were nearer to the Context vowels (i.e., the first vowel). Therefore, it is unclear whether syllable structure or proximity to contextual variations contributes to the larger degree of V-to-V coarticulation for coda. Mok (2010) examined the effects of syllable structure on V-to-V coarticulation using Thai and English data. She also compared /CV.CV/ with /CVC.Vt/ using two vowels (/i a/ in Thai, /i ɑ/ in English) and two intervocalic consonants (/p t/). She avoided the temporal confound for carryover coarticulation in Modarresi et al. (2004a) by measuring formant frequency for both syllable types at the beginning of periodicity of the second vowel. Mok found that coda (closed syllables) allowed more V-to-V coarticulation than onset (open syllables) for the vowel /i/ in Thai, but there was no consistent pattern of syllable structure effect on V-to-V coarticulation in English. The results of these two studies show that syllable structure may indeed affect V-to-V coarticulation, but their inconclusive results warrant more investigation on this topic.

A literature search suggests that there is no study investigating V-to-V coarticulation across consonant clusters with different syllable affiliations. Recasens and Pallarès (2001) studied V-to-V coarticulation across consonant clusters in Catalan, but only in /aC.Ca/ sequences. The vowel combination and syllabification of the consonant clusters were not varied. Therefore, the effects of V-to-V coarticulation across consonant clusters with different syllable affiliations are still unknown. This study aims to fill this gap and extends the research on the effects of syllable structure on V-to-V coarticulation by investigating V-to-V coarticulation across /st/ clusters in English. The /st/ cluster was chosen because it is homorganic in English, thus reducing conflicting influences of intervocalic consonants on formant transitions. Also, it is the only consonant cluster in English that can be used to study the effects of syllable structure with real words comprehensively. The /st/ cluster can be syllabified in three ways in English: onset /.st/, heterosyllabic /s.t/ and coda /st./. The same cluster /st/ was used for all sequences so the data is well controlled in terms of formant transition. Any Context vowel effect on the Target vowels is therefore attributable to V-to-V coarticulation instead of formant transition only.

The studies reviewed above show that, in general, coda consonants are more sensitive to change and coarticulatory effects than onset consonants. Since onsets are shown to be more stable and to have a tighter coordination with vowels, and codas to be more variable, it was hypothesized that coda /st./ would allow more V-to-V coarticulation than onset /.st/. The case for heterosyllabic /s.t/ is less clear. Since heterosyllabic /s.t/ is a combination of coda + onset, degree of V-to-V coarticulation allowed may lie between canonical onset /.st/ and coda /st./ (/st./ > /s.t/ > /.st/). However, it can also be argued that since heterosyllabic /s.t/ contains a syllable/word boundary within the cluster, it may exhibit the least V-to-V coarticulation compared to the other two structures without a boundary inside the cluster (/st./ > /.st/ > /s.t/), since heterosyllabic /s.t/ is higher up in the prosodic hierarchy than tautosyllabic onset and coda (Beckman and Pierrehumbert, 1986).

2. Method

2.1. Subjects

Six native speakers of Standard Southern British English (two male, four female) were recorded. All speakers were graduate students from the University of Cambridge and had no history of speech or hearing impairment. They were all in their twenties, and were paid to participate in the experiment.

2.2. Materials

Three vowels, /i ɑ u/, were used to form the target sequences with the /st/ cluster in three forms: onset /CV.stVC/, heterosyllabic /CVs.tVC/ and coda /CVst.VC/.

Table 1
Experimental materials with the /st/ clusters in English (and the corresponding English words).

Vowel sequence	Onset CV. stVC		Heterosyllabic CVs. tVC		Coda CVst. VC	
i i	/bi stid/	Bee Steed	/pis tit/	Peace Teat	/bist it/	Beast Eat
i ɑ	/bi stɑd/	Bee Starred	/pis tɑt/	Peace Tart	/bist ɑt/	Beast Art
i u	/bi stup/	Bee Stoop	/pis tut/	Peace Toot	/bist uz/	Beast Ouse[a]
ɑ i	/b ɑ stid/	Bar Steed	/pɑs tit/	Pass Teat	/pɑst it/	Past Eat
ɑ ɑ	/b ɑ stɑd/	Bar Starred	/pɑs tɑt/	Pass Tart	/pɑst ɑ t/	Past Art
ɑ u	/b ɑ stup/	Bar Stoop	/pɑs tut/	Pass Toot	/pɑst uz/	Past Ouse
u i	/bu stid/	Boo Steed	/mus tit/	Moose Teat	/bust it/	Boost Eat
u ɑ	/bu stɑd/	Boo Starred	/mus tɑt/	Moose Tart	/bust ɑt/	Boost Art
u u	/bu stup/	Boo Stoop	/mus tut/	Moose Toot	/bust uz/	Boost Ouse

[a] The River Ouse is a river in Northern England.

All possible combinations of the vowels were used (see Table 1). Each individual monosyllabic word is a real word in English but the resultant two-word combinations are mostly nonsense sequences. The sequences were presented to the speakers as two separate words starting with a capital letter like this: ‘Past Art’, so the intended syllabifications of the target sequences should be unambiguous to the speakers. The sequences were embedded in a carrier phrase: ‘Not a ______, it’s a ______ again’, with the target sequences embedded in the second half of the carrier phrase in order to elicit contrastive stress. Carrier phrases were constructed so that the target syllables were always defocused bearing no contrastive stress, but still produced with full vowels. Both the first and the second words in the target sequences were possible target syllables, depending on the direction of coarticulationary influence: the first vowel for investigating anticipatory coarticulation and the second vowel for investigating carryover coarticulation. For instance, in ‘Not a Tape Art, it’s a Past Art again’, the sequence ‘Past Art’ is the one used. Contrastive stress falls on the contextual syllable ‘Past’ which is not analysed, while ‘Art’ is the defocused target syllable for investigating carryover coarticulation. Anticipatory coarticulation on the target syllable ‘Past’ is elicited by using ‘Not a Past Form, it’s a Past Art again’, with contrastive stress falling on the contextual syllable ‘Art’.

Five repetitions of the materials were intended to be collected (27 target sequences $\times$ 2 coarticulatory directions $\times$ 5 repetitions = 270 tokens), but seven repetitions were recorded in case there was any problem with the recordings. Since most of the tokens were usable with no problem, each data point for the repeated measures ANOVAs was an average of at least five to no more than seven repetitions. Many previous studies on V-to-V coarticulation, both acoustic and articulatory, also used the same method by averaging formant frequencies across multiple tokens for analysis (e.g. Beddor et al., 2002; Cho, 2004; Magen, 1997; Manuel, 1990; Recasens, 2002; Recasens et al., 1997).

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 some of the materials as many times as they liked. All speakers were instructed to read the materials at a normal speaking rate. The speech was recorded using a DAT tape (48 kHz sampling rate) via a Sennheiser MKH 40 P48 microphone and a Symetrix SX 202 amplifier into a Sony DTC-60ES recorder and later down-sampled to 16 kHz using *Xwaves* for acoustic analysis.

2.4. Acoustic measurements

Frequencies of the first two formants (F1 and F2) of the Target vowels (either V1 or V2 depending on coarticulatory direction), and the acoustic durations of the first vowel and the intervocalic consonants were measured. Formant frequencies 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. The beginning and ending of periodicity was taken for the onset and offset of the Target vowels. The onset and offset of aperiodic noise was taken as the onset and offset of /s/. The silence between the /s/ offset and the /t/ burst was taken as the /t/ closure. The portion from the release of the /t/ burst to the onset of periodicity of the second vowel was taken as the /t/ aspiration, which includes the occasional silence in some coda /st./ produced with a clearly released /t/ followed by a short glottal stop. Since such tokens only appeared occasionally as the speakers produced the materials fluently and often released the final /t/ onto the following vowels, they were not treated separately. Occasionally, speakers produced a very fricated or lenited /t/ in onset /.st/ or coda /st./ so that there was no closure for /t/ at all. Such tokens were excluded from analysis. F1 and F2 frequencies were measured at the offset of periodicity of the first vowel without contrastive stress (for anticipatory coarticulation) and at the onset of periodicity of the second vowel without contrastive stress (for carryover coarticulation). The formant frequency data were taken at similar locations comparable to other studies in V-to-V coarticulation (e.g. Beddor et al., 2002; Manuel, 1990; Öhman, 1966; Recasens, 1987).

All formant frequency data were normalised for individual differences. A straight-forward Bark-transformation (which is vowel-intrinsic) was not used because the focus of this study was not on the perception of V-to-V coarticulation. The normalisation procedure in this study is also used in Mok (2011), which is vowel-extrinsic and formant-intrinsic, and is essentially the same as the *S*-centroid method (Fabricius et al., 2009; Watt and Fabricius, 2002), but the two methods were developed independently. Adank et al. (2004) found that vowel-extrinsic normalisation procedures preserve phonemic variation and reduce anatomical/physiological variation more effectively than vowel-intrinsic procedures. Information comparing the *S*-centroid method with other normalisation procedures can be found in Watt et al. (2010).

The normalisation procedures in this study are as follows. 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 tokens for a particular speaker. The grand mean represents the hypothetical centre of the speaker's vowel space, e.g. 500 Hz for F1. (2) Calculate the Context-induced proportional F1 (or F2) difference from that grand mean ($current_{Target\ V}$/grand mean), e.g. $700\ Hz/500\ Hz = 1.4$. A value bigger or smaller than 1 means that the vowel formant is higher or lower than the hypothetical centre. 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 normalised distance from the hypothetical centre of each speaker's vowel space, but they do not show degree of coarticulation. The normalised proportions are analogous to formant frequencies and should be interpreted in the same way.

These normalised proportions were used for statistical analysis. In interpreting the statistical results, if the sphericity assumption of any main effect or interaction in the repeated measures ANOVAs was violated, the degree of freedom was adjusted with the Huynh–Feldt epsilon in generating the p values for more conservative results. In case of a significant main effect or interaction involving a factor with more than three levels, post hoc pairwise comparisons were conducted with the Sidak adjustment to control for family-wise Type I error. Further details of the statistical tests are given in the Results section for easy reference.

3. Results

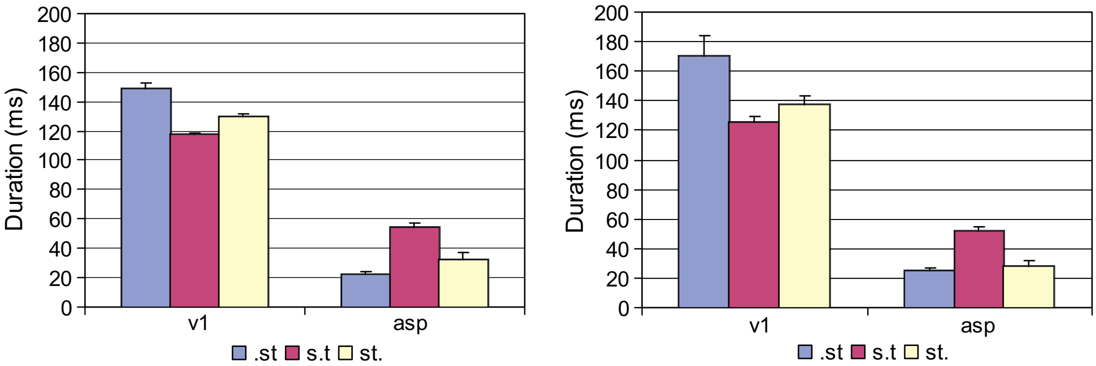
3.1. Duration

Syllabification is the crucial factor in this study. However, since syllabification of intervocalic consonants can be tricky in English, and since the experimental materials are nonsense sequences, it is important to ensure that the speakers did produce the target sequences with the intended syllabifications as shown in Table 1. Two types of durational data can give us evidence for syllabification. Vowels in English are subjected to pre-fortis clipping when they are followed by a fortis (voiceless) consonant within the same syllable (Wells, 1990). Therefore, duration of the first vowel in onset /.st/ (i.e., in an open syllable) should be longer than in heterosyllabic /s.t/ and coda /st./ (the vowel is subjected to clipping in the latter two conditions). The aspiration duration of /t/ is another cue to syllabification. It should be the longest in heterosyllabic /s.t/ because /t/ is strongly aspirated as the onset of the second syllable. It should be shorter in onset /.st/ and coda /st./.

To test whether these duration patterns showed evidence of syllabification, duration of the first vowel (V1) and the /t/ aspiration collapsed across Target and Context vowels were submitted to two one-way ANOVAs comparing Syllable Forms (onset /.st/, heterosyllabic /s.t/, coda /st./) in the anticipatory and the carryover directions respectively because of different stress placements. Fig. 1 shows the durational data for different Syllable Forms. As expected, V1 duration is longer in the carryover than anticipatory direction because V1 is stressed in the carryover direction. Syllable Form is significant for V1 duration in both Directions (anticipatory: [$F(1.10, 5.499) = 79.342$, $p < 0.0001$]; carryover: [$F(1.051, 5.257) = 22.347$, $p = 0.005$]). Post-hoc comparisons with Sidak adjustment confirm that V1 duration in the three syllable structures is significantly different from each other (anticipatory: onset vs. heterosyllabic, $p = 0.001$; onset vs. coda, $p = 0.001$; heterosyllabic vs. coda, $p = 0.002$; carryover: onset vs. heterosyllabic, $p = 0.014$; onset vs. coda, $p = 0.020$; heterosyllabic vs. coda, $p = 0.009$). Syllable Form is again significant for the duration of /t/ aspiration in both Directions (anticipatory: [$F(2, 10) = 40.678$, $p < 0.0001$]; carryover: [$F(2, 10) = 50.448$, $p < 0.0001$]). Post-hoc comparisons with Sidak adjustment show that the aspiration duration in heterosyllabic /s.t/ is significantly longer than both onset /.st/ (anticipatory, $p < 0.0001$; carryover, $p < 0.0001$) and coda /st./ (anticipatory, $p = 0.008$; carryover, $p = 0.002$), but onset and coda are not significantly different (anticipatory, $p = 0.215$; carryover, $p = 1.00$). The two durational patterns conform to the expectations based on syllable structure mentioned above very well. Hence, we can safely conclude that the speakers did produce the target sequences with the intended syllabifications as shown in Table 1.

3.2. Formant frequency

The normalised F1 and F2 frequencies data measured at the offset of the first vowel (anticipatory direction) and the onset of the second vowel (carryover direction) were submitted to two 4-way repeated measures ANOVAs with factors Direction (anticipatory, carryover), Syllable Form (onset /.st/, heterosyllabic /s.t/, coda /st./), Target (/i ɑ u/)



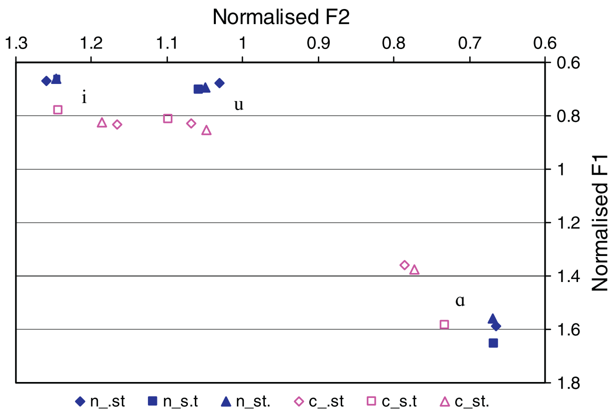
a. anticipatory direction (stress on V2)

b. carryover direction (stress on V1)

Fig. 1. Duration of the first vowel (V1) and the /t/ aspiration (asp) in three Syllable Forms. Error bars show standard error of mean.

and Context (/i ɑ u/). Separate 3-way repeated measures ANOVAs (Direction × Syllable Form × Context) for each Target vowel were also conducted to facilitate the interpretation of the 4-way results. The results of the 3-way ANOVAs for each Target vowel would be discussed if the corresponding 4-way ANOVAs were significant.

Fig. 2 shows the averaged normalised F1 and F2 frequency of the three Target vowels collapsed across Context vowels. It should be noted that the vowel /u/ in Standard Southern British English is fronted (Ferragne and Pellegrino, 2010; Hawkins and Midgley, 2005), which explains its proximity to /i/ in the F2 dimension in Fig. 2. Two obvious differences between vowels in the carryover direction (V2, empty symbols) and vowels in the anticipatory direction (V1, solid symbols) can be observed. First, vowels in the carryover direction are more centralized in the vowel space. Second, there is a larger difference among the three Syllable Forms for vowels in the carryover direction. Moreover, vowels in heterosyllabic /s.t/ are more extreme than vowels in both onset /.st/ and coda /st./ in the carryover direction (compare the empty squares with the empty diamonds and empty triangles in Fig. 2), i.e., F2 for /i, u/ and F1 for /ɑ/ are higher for the /s.t/ vs. /.st/ and /st./ conditions.



“n” = anticipatory direction; “c” = carryover direction

Fig. 2. Averaged normalised F1 and F2 frequency of the three Target vowels collapsed across Context vowels.

These differences are confirmed by the significant Syllable Form × Direction × Target interactions in both F1 [$F(4,20) = 3.554$, $p = 0.024$] and F2 [$F(4,20) = 17.642$, $p < 0.0001$]. Post hoc pairwise comparisons with Sidak adjustment among the three Syllable Forms show that as for F1, heterosyllabic /s.t/ is significantly different from onset /.st/ for Target /ɑ/ only ($p = 0.006$). For F2, heterosyllabic /s.t/ is significantly different from onset /.st/ for all three Target vowels (/i/ $p = 0.042$/; /ɑ/ $p = 0.030$; /u/ $p = 0.034$/), and it is significantly different from coda /st./ for the high vowels /i/ ($p = 0.014$) and /u/ ($p = 0.004$).

The main question in this study is whether Syllable Form affects V-to-V coarticulation differently. Since the dependent variable is normalised formant frequency, evidence of V-to-V coarticulation is shown by a significant Context effect. Any interaction with Context indicates that V-to-V coarticulation is affected by other factors.

Regarding F1, there is no significant Context main effect or any interaction involving Context, which means that V-to-V coarticulation is only minimal. As for F2, the effect of Syllable Form on V-to-V coarticulation is reflected in the Syllable Form × Context interaction [$F(4,20) = 3.160$, $p = 0.036$] (see Fig. 3). The data points in Fig. 3 all cluster around 1 (i.e., the hypothetical centre of the vowel space) because they show the averaged Context effects collapsed across all Target vowels. Again, the normalised F2 values of Context /u/ fall between Contexts /i/ and /ɑ/ because the vowel /u/ is fronted in Standard Southern British English. Since there are three levels for both factors involved in this significant interaction, one-way ANOVAs comparing the main effects of Context and Syllable Forms separately, and post hoc pairwise comparisons with Sidak adjustment were conducted to find out what contributed to the significant interaction. The Context main effects are significant for all three Syllable Forms (onset: [$F(2,10) = 33.122$, $p < 0.0001$], heterosyllabic: [$F(2,10) = 6.683$, $p = 0.014$], coda: [$F(2,10) = 39.507$, $p < 0.0001$]). Post hoc comparisons show that the three Context vowels are significantly different from each other with coda /st./ (/ɑ/ vs. /i/, $p = 0.001$; /ɑ/ vs. /u/, $p = 0.015$; /i/ vs. /u/, $p = 0.034$). The largest difference in normalised F2 is 0.033 (between

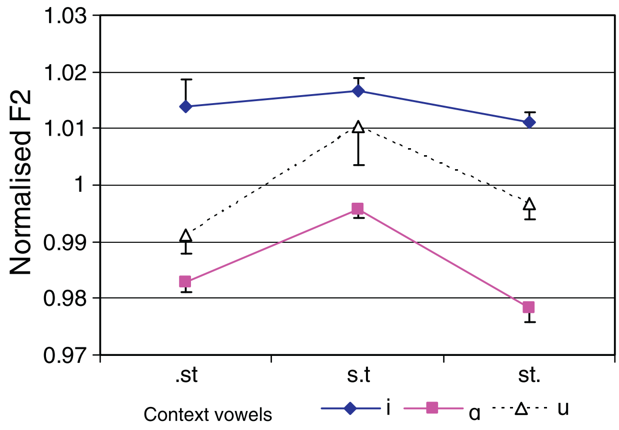


Fig. 3. Normalised F2 for the three Context vowels collapsed across Target vowels in each Syllable Form. Error bars show standard error of mean.

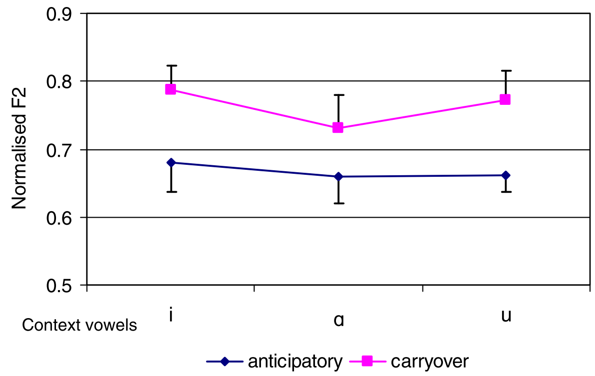


Fig. 4. Averaged normalised F2 frequency of Target /ɑ/ in three vowel Contexts and two coarticulatory directions. Error bars show standard error of mean.

Contexts /ɑ/ and /i/). As for onset /.st/, the largest significant difference in normalised F2 (0.031) is between Contexts /ɑ/ and /i/ ($p = 0.002$). The difference between Contexts /i/ and /u/ is also significant ($p = 0.012$), while the difference between Contexts /ɑ/ and /u/ is not ($p = 0.091$). For heterosyllabic /s.t/, only Context /ɑ/ is significantly different from Context /i/ ($p = 0.005$), and the normalised F2 difference is 0.021. There is no significant difference between Contexts /i/ and /u/ ($p = 0.810$), and Contexts /ɑ/ and /u/ ($p = 0.174$). These results show a continuum of coarticulatory effect for the three Syllable Forms: coda /st./ > onset /.st/ > heterosyllabic /s.t/, supporting the expectation that coda /st./ allows the most V-to-V coarticulation.

Nevertheless, the difference of the Context effects between coda /st./ and onset /.st/ is not too extensive. This is further supported by the result comparing Syllable Form for each Context vowel separately. Syllable Form is significant for Context /ɑ/ [$F(2,10) = 16.183$, $p = 0.001$] and Context /u/ [$F(2,10) = 5.194$, $p = 0.028$], but not for Context /i/ [$F(2,10) = 0.846$, $p = 0.458$]. However, post hoc comparisons only show significant differences for Context /ɑ/, probably because of the corrected alpha by Sidak adjustment. For Context /ɑ/, heterosyllabic /s.t/ is significantly different from onset /.st/ ($p = 0.005$) and coda /st./ ($p = 0.021$), but the difference between onset /.st/ and coda /st./ is not significant ($p = 0.543$). Fig. 3 shows that the F2 lowering effect of Context /ɑ/ is weaker in heterosyllabic /s.t/. In line with the above results, heterosyllabic /s.t/ allows less coarticulation than the other two Syllable Forms, while there is only a small difference between onset /.st/ and coda /st./.

Finally, Direction also affects V-to-V coarticulation in F2, as evidenced by the Direction × Context [$F(2,10) = 4.407$, $p = 0.042$] and Direction × Target × Context [$F(4,20) = 3.194$, $p = 0.035$] interactions. The 3-way ANOVAs for each Target vowel show that Direction only affects V-to-V coarticulation in Target /ɑ/ [$F(2,10) = 9.258$, $p = 0.005$] (see Fig. 4). We can use the data points of Context /ɑ/ in Fig. 4 as the baseline for comparison (i.e., the 'uncoarticulated' reference for Target /ɑ/ in the /ɑ/ Context). Post hoc pairwise comparisons with Sidak adjustment show that in the anticipatory position, Context /i/ is significantly different from Context /ɑ/ ($p = 0.011$, difference in normalised F2: 0.02). In the carryover position, Context /ɑ/ is significantly different from both Context /i/ ($p = 0.004$) and Context /u/ ($p = 0.004$). The largest difference in normalised F2 is 0.056 (between Contexts /i/ and /ɑ/). The results show stronger carryover than anticipatory coarticulation for Target /ɑ/. The same conclusion is reached with the Direction × Context interaction collapsed across Target vowels.

4. Discussion

The results show that syllable structure does affect V-to-V coarticulation with this continuum: coda /st./ > onset /.st/ > heterosyllabic /s.t/. Coda /st./ allows more coarticulation than onset /.st/ in F2 while heterosyllabic /s.t/ allows the least V-to-V coarticulation in F2 among all Syllable Forms. The Target vowels are also more peripheral in heterosyllabic /s.t/ than in the other two Syllable Forms in the carryover direction. The results confirm our hypothesis that the coda, being more variable, is more transparent to V-to-V coarticulation than the onset. This is the first study clearly demonstrating that V-to-V coarticulation is sensitive to syllable affiliation of the intervocalic clusters with the same segmental composition. The results confirm the conclusion from previous literature that the syllable is an important unit in speech production, and have furthered our understanding of syllable structure as a factor affecting V-to-V coarticulation.

In the present study, Target vowels in the carryover direction (V2) are more centralised than in the anticipatory direction (V1), although both positions were defocused and were produced with full vowels. This may relate to the strong trochaic tendency of disyllabic words in English (Culter and Carter, 1987). Although the target sequences were nonsense sequences consisting of two separate monosyllabic words and great care was taken regarding stress placement, the target sequences resembled compound words in the carrier phrase. Subjects may have produced the vowels in the carryover direction (V2) with slightly more centralization.

There is more carryover than anticipatory V-to-V coarticulation in the present study, which concurs with previous findings in English (e.g. Magen, 1997). In addition to language-specific patterns of coarticulatory direction, the prominence of carryover V-to-V coarticulation over anticipatory V-to-V coarticulation may also be related to the segmental composition of the /st/ clusters. Although both /s/ and /t/ are homorganic and do not directly involve the tongue dorsum in production, the precise articulatory requirements of forming a medial groove for fricatives render /s/ to be more highly constrained than /t/ (Recasens, 2002; Recasens and Espinosa, 2009; Recasens et al., 1997). As a result, the degree of articulatory constraints decreases across the course of the /st/ clusters. The strong trochaic tendency mentioned above renders V1 to be stronger than V2. More articulatory specification and effort are thus required for the first half of the /VstV/ sequence than towards the end of it. It is convincible that because of these reasons the effects of V1 on V2 across the /st/ cluster will be more prominent than vice versa.

The results show that heterosyllabic /s.t/ is the most resistant to V-to-V coarticulation and has more peripheral vowel positions than onset /.st/ and coda /st./. A reason for such patterns would be related to word boundary. Although all three structures involve a word boundary, heterosyllabic /s.t/ involves a word boundary within the /st/ cluster which is higher up in the prosodic hierarchy than tautosyllabic onset and coda (Beckman and Pierrehumbert, 1986). Heterosyllabic /s.t/ is a prosodically stronger structure than onset /.st/ and coda /st./. It is thus not surprising to find heterosyllabic /s.t/ being more resistant to coarticulation than the other two Syllable Forms, and having more extreme vowel Targets. The present results on lingual consonant clusters are particularly interesting since the tongue needs to respond to various articulatory demands from the vowels and consonants and the differences in syllable structure in a short period of time. The results demonstrate that the seemingly mechanical V-to-V coarticulation can be influenced by abstract linguistic structure. This corresponds quite well with recent research on the relationship between prosody and articulation (e.g. Cho, 2001, 2004; Keating et al., 2003; Turk and Shattuck-Hufnagel, 2000).

In addition, the effects of heterosyllabic /s.t/ may also be related to the differences in articulatory trajectory. The /t/ in the onset position in heterosyllabic /s.t/ is strongly aspirated compared with /t/ in onset /.st/ and coda /st./ (see Fig. 1). If the jaw starts to lower for the following vowel at around the same time after the /t/ closure for the three Syllable Forms, then it is possible that the strong aspiration in heterosyllabic /s./t/ will affect the jaw position at the onset of periodicity of the following vowel. Since there is a longer time lag between the /t/ burst and the onset of periodicity of the following vowel in heterosyllabic /s.t/, the jaw will have reached a more open position at the onset of periodicity than in onset /.st/ and coda /st./. Conceivably, the tongue can also have reached a more extreme position. Since the onset of periodicity of the following vowel in heterosyllabic /s.t/ is further into the vowel gesture compared with onset /.st/ and coda /st./, it may explain the weaker coarticulation and more extreme vowel Targets in heterosyllabic /s.t/. In addition, as suggested by one reviewer, /t/ is articulated with more tongue contact for heterosyllabic /s.t/ than for onset /.st/ and coda /st./ (evidenced by the longer /t/ closure and burst duration data in Fig. 1). This can also explain the resistance of heterosyllabic /s.t/ observed in the data. Further studies with articulatory data are needed to confirm the above proposal.

The difference in intervocalic duration may also contribute to the observed coarticulatory difference among the three Syllable Forms, since the two vowels are farther apart in time in heterosyllabic /s.t/ than onset /.st/ and coda /st./ (see Section 3.1). While this is possible, it is worth pointing out that the degree of V-to-V coarticulation does not necessarily correlate with the intervocalic duration, as demonstrated in Mok (2010). The different allophonic realizations of /t/ are an integral part of syllable structure in English. In order to examine the effects of prosodic boundary, articulatory trajectory and intervocalic duration separately, we need to compare V-to-V coarticulation with different syllable structures in a language with more consistent realizations of the /st/ cluster at different syllable positions.

Although the results confirm our hypothesis about the effects of syllable structure on V-to-V coarticulation that coda /st./ would allow more V-to-V coarticulation than onset /.st/, there is only a small coarticulation difference between the two Syllable Forms. The /st/ cluster was chosen because it is homogenous in English, and also because only the /st/ cluster can form a complete set of materials with real words as shown in Table 1. However, as mentioned above, the tongue is maximally constrained in a /st/ cluster (Recasens et al., 1997). The magnitude of V-to-V coarticulation allowed is likely to be reduced. Therefore, it is possible that the effects of syllable structure on V-to-V coarticulation may be more prominent in other clusters in which the tongue is less constrained than in a /st/ cluster, e.g. /sp/ or /sk/. Using heterorganic clusters would probably require the use of nonsense words, but it can give us valuable insights into the effects of syllable structure on V-to-V coarticulation. In fact, Byrd (1996) findings support this possibility. She found that onset, coda and heterosyllabic clusters differ in their intergestural timing and reduction patterns, but the exact nature depends on the consonants involved. All this suggests that it is important to investigate the effects of syllable structure on V-to-V coarticulation with more clusters to verify the results found in this study.

The present results have interesting implications for Articulatory Phonology on syllable structure and gestural coordination. As mentioned in the Introduction, Articulatory Phonology assumes that gestures are timed and coordinated with respect to each other. Onset consonants are phased with the vowel as a unit (the C-center effect) while only the beginning of a coda cluster is phased with the vowel (Browman and Goldstein, 1988, 1995; Byrd, 1995). There is no clear prediction of V-to-V phasing across word/syllable boundaries since previous studies mainly focused on vowel phasing within a syllable. However, an effect of syllable structure on V-to-V coarticulation can be expected because according to Articulatory Phonology, coarticulation is a result of gestural overlap. Vowels have different phasing relations with onset and coda clusters, i.e., they overlap differently. Byrd (1995) found that the underlying syllable affiliations of intervocalic consonant sequences influenced the articulatory organization of the consonant sequences and their adjacent vowels. All this implies that clusters with different syllable affiliations may affect V-to-V coarticulation differently. Our acoustic data seem to support this idea quite well. Of course, in order to model V-to-V coarticulation explicitly, higher level extra-syllable V-to-V or syllable-to-syllable phasing relations must be included, in addition to the local phasing relations between consonants and vowels (e.g. Nam and Saltzman, 2003). Further investigations using articulatory data are needed to verify this proposal. Also, since consonant and vowel coordination patterns can be language-specific (e.g. Smith, 1995), it is important to compare V-to-V coarticulation across consonant clusters cross-linguistically.

In fact, Czech was also considered for the present study because of its complex syllable structure and consistent realizations of consonants in different syllable positions. All stops are phonologically unaspirated in both onset and coda positions, and in /s/ + stop clusters. Nevertheless, there are other complications in Czech, as it is a heavily inflected language with strong polysyllabic tendency. It is impossible to have a set of experimental materials with monosyllabic real words like those in Table 1. In addition, stress is fixed on the first syllable of a word in Czech. Thus, no unstressed syllables can be used to examine carryover coarticulation. Issues like these highlight the difficulties of investigating the effects of syllable structure on V-to-V coarticulation in different languages. Nevertheless, the positive results on English in this study, together with the results in Modarresi et al. (2004a) and Mok (2010), strongly suggest that the relationship between syllable structure and V-to-V coarticulation is worth further exploration.

The present study contributes to our understanding of the relationship between syllable structure and coarticulation. However, it still has some limitations which can be improved in future studies. The data is based on six native speakers of Southern British English. It would be desirable to collect data from more speakers, and if possible, from different accents of English. It would be interesting to compare the effects of syllable structure in different varieties of English as the same phonological phenomenon can have different realizations in different accents (e.g. see Low and Grabe, 1999; Low et al., 2000). The present study used only acoustic data. Articulatory data like EMMA data across the /st/ cluster can provide more dynamic and comprehensive picture of the effects of syllable structure. Articulatory data on tongue movement will be particularly insightful. Finally, only one speech rate (normal) was used in this study. Speech rate can affect the realizations of consonant clusters and juncture in English (Byrd, 1996; Byrd and Tan, 1996; Schwab et al., 2008). The investigation of speech rate variation on syllable structure effects and V-to-V coarticulation would be very interesting as only very few studies have examined the effects of speech rate and duration on V-to-V coarticulation (Hertrich and Ackermann, 1995; Mok, 2011).

The present study is the first attempt to examine the effects of syllabification of consonant cluster (/st/) on V-to-V coarticulation in English. Previous studies of consonant clusters are mainly concerned with how the intervocalic consonants themselves are coproduced under different syllabifications, their inter-gestural coordination and timing. They have not addressed how syllabification of the intervocalic clusters affects vowel production and coarticulation. For example, only one vowel was used in Recasens and Pallarès (2001) and Byrd (1996). Zsiga (1994) did vary the vowels in her study on consonant clusters in order to allow for possible effects of V-to-V coarticulation across /C.C/ sequences, but she only mentioned the results very briefly. Also, syllabification of the consonant cluster was not varied. The results of the present study suggest that syllable structure can indeed affect coarticulatory patterns. Further investigations with more consonant clusters in both English and other languages that have more consistent realizations of consonant clusters at different syllable positions are needed in order to explore the effects of syllable structure on coarticulation more thoroughly.

Acknowledgments

The author would like to thank Sarah Hawkins for help and guidance throughout the project. Thanks also go to Rachel Smith and Francis Nolan for helpful discussion. She thanks the Editor and two anonymous reviewers for their constructive comments. This research was supported by the Sir Edward Youde Memorial Fellowship for Overseas Studies from Hong Kong and an Overseas Research Studentship from the United Kingdom.

References

Adank, P., Smits, R., van Hout, R., 2004. A comparison of vowel normalization procedures for language variation research. J. Acoust. Soc. Am. 116, 3099–3107.

Anderson, S., Port, R., 1994. Evidence for syllable structure, stress and juncture from segmental durations. J. Phonetics 22, 283–315.

Beckman, M.E., Pierrehumbert, J., 1986. Intonational structure in Japanese and English. Phonology Yearbook 3, 255–309.

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

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

Bell, A., Hooper, J.B., 1978. Issues and evidence in syllabic phonology. In: Bell, A., Hooper, J.B. (Eds.), Syllables and Segments. North-Holland Publishing Co., Amsterdam, Oxford, pp. 3–22.

Boucher, V.J., 1988. A parameter of syllabification for VstopV and relative-timing invariance. J. Phonetics 16, 299–326.

Browman, C.P., Goldstein, L., 1988. Some notes on syllable structure in Articulatory Phonology. Phonetica 45, 140–155.

Browman, C.P., Goldstein, L., 1995. Gestural syllable position effects in American English. In: Bell-Berti, F., Raphael, L.J. (Eds.), Producing Speech: Contemporary Issues. American Institute of Physics, New York, pp. 19–33.

Byrd, D., 1995. C-centers revisited. Phonetica 52, 285–306.

Byrd, D., 1996. Influences on articulatory timing in consonant sequences. J. Phonetics 24, 209–244.

Byrd, D., Tan, C.C., 1996. Saying consonant clusters quickly. J. Phonetics 24, 263–282.

Cho, T., 2001. Effects of morpheme boundaries on intergestural timing: evidence from Korean. Phonetica 58, 129–162.

Cho, T., 2004. Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. J. Phonetics 32, 141–176.

Christie, W.M., 1974. Some cues for syllable juncture perception in English. J. Acoust. Soc. Am. 55, 819–821.

Culter, A., Carter, D.M., 1987. The predominance of strong initial syllables in the English vocabulary. Comput. Speech Lang. 2, 133–142.

de Jong, K., 2001. Rate-induced resyllabification revisited. Lang. Speech 44, 197–216.

de Jong, K., Lim, B., Nagao, K., 2004. The perception of syllable affiliation of singleton stops in repetitive speech. Lang. Speech 47, 241–266.

Fabricius, A., Watt, D., Johnson, D.E., 2009. A comparison of three speaker-intrinsic vowel formant frequency normalisation algorithms for sociophonetics. Lang. Variat. Change 21, 413–435.

Ferragne, E., Pellegrino, F., 2010. Formant frequencies of vowels in 13 accents of the British Isles. J. Internat. Phonetic Assoc. 40, 1–34.

Fowler, C.A., 1981. Production and perception of coarticulation among stressed and unstressed vowels. J. Speech Hear. Res. 46, 127–139.

Gick, B., Campbell, F., Oh, S., Tamburri-Watt, L., 2006. Toward universals in the gestural organisation of syllables: a cross-linguistic study of liquids. J. Phonetics 34, 49–72.

Greenberg, J., 1978. Some generalisations concerning initial and final consonant clusters. In: Greenberg, J. (Ed.), Universals of Human Languages, . In: Phonology, vol. 2. Stanford University Press, Stanford, pp. 243–280.

Haggard, M., 1973a. Abbreviation of consonants in English pre- and post-vocalic clusters. J. Phonetics 1, 9–24.

Haggard, M., 1973b. Correlations between successive segment durations: values in clusters. J. Phonetics 1, 111–116.

Hawkins, S., Midgley, J., 2005. Formant frequencies of RP monophthongs in four age groups of speakers. J. Internat. Phonetic Assoc. 35, 183–199.

Hertrich, I., Ackermann, H., 1995. Coarticulation in slow speech: durational and spectral analysis. Lang. Speech 38, 159–187.

Honorof, D.N., Browman, C.P., 1995. The center or edge: how are consonant clusters organized with respect to the vowel? In: The 13th International Congress of Phonetic Sciences (ICPhS), Stockholm, pp. 552–555.

Keating, P., Cho, T., Fougeron, C., Hsu, C., 2003. Domain-initial articulatory strengthening in four languages. In: Local, J., Ogden, R., Temple, R. (Eds.), Papers in Laboratory Phonology 6. Cambridge University Press, Cambridge, pp. 143–161.

Kochetov, A., 2006. Syllable position effects and gestural organization: Articulatory evidence from Russian. In: Goldstein, L., Whalen, D.H., Best, C. (Eds.), Papers in Laboratory Phonology 8. Mouton de Gruyter, Berlin, New York, pp. 565–588.

Krakow, R.A., 1999. Physiological organization of syllables: a review. J. Phonetics 27, 23–54.

Lehiste, I., 1970. Suprasegmentals. MIT Press, Cambridge, MA.

Low, E.L., Grabe, E., 1999. A contrastive study of prosody and lexical stress placement in Singapore English and British English. Lang. Speech 42, 39–56.

Low, E.L., Grabe, E., Nolan, F., 2000. Quantitative characterisations of speech rhythm: syllable-timing in Singapore English. Lang. Speech 43, 377–401.

Macchi, M., 1988. Labial articulation patterns associated with segmental features and syllable structure in English. Phonetica 45, 109–121.

MacNeilage, P.F., Davis, B.L., 2000. On the origin of internal structure of word forms. Science 288, 527–531.

Maddieson, I., 1984. Patterns of Sounds. Cambridge University Press, Cambridge.

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

Manuel, S., 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. J. Acoust. Soc. Am. 88, 1286–1298.

Modarresi, G., Sussman, H., Lindblom, B., Burlingame, E., 2004a. An acoustic analysis of the bidirectionality of coarticulation in VCV utterances. J. Phonetics 32, 291–312.

Modarresi, G., Sussman, H., Lindblom, B., Burlingame, E., 2004b. Stop place coding: an acoustic study of CV, VC#, and C#V sequences. Phonetica 61, 2–21.

Mok, P., 2010. Language-specific realizations of syllable structure and vowel-to-vowel coarticulation. J. Acoust. Soc. Am. 128, 1346–1356.

Mok, P., 2011. Effects of vowel duration and vowel quality on vowel-to-vowel coarticulation. Lang. Speech 54, 527–545.

Mok, P., accepted for publication. Does vowel inventory density affect vowel-to-vowel coarticulation? Language and Speech.

Nam, H., Saltzman, E., 2003. A competitive, coupled oscillator model of syllable structure. In: The 15th International Congress of Phonetic Sciences (ICPhS), Barcelona, pp. 2253–2256.

Öhman, S.E.G., 1966. Coarticulation in VCV utterances: spectrographic measurements. J. Acoust. Soc. Am. 39, 151–168.

Pickett, J.M., Bunnell, H.T., Revoile, S.G., 1995. Phonetics of intervocalic consonant perception: retrospect and prospect. Phonetica 52, 1–40.

Quené, H., 1992. Durational cues for word segmentation in Dutch. J. Phonetics 20, 331–350.

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

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

Recasens, D., 2002. An EMA study of VCV coarticulatory direction. J. Acoust. Soc. Am. 111, 2828–2841.

Recasens, D., Espinosa, A., 2009. An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. J. Acoust. Soc. Am. 125, 2288–2298.

Recasens, D., Pallarès, M.D., 2001. Coarticulation, assimilation and blending in Catalan consonant clusters. J. Phonetics 29, 273–301.

Recasens, D., Pallarès, M.D., Fontdevila, J., 1997. A model of lingual coarticulation based on articulatory constraints. J. Acoust. Soc. Am. 102, 544–561.

Redford, M., Diehl, R., 1999. The relative perceptibility of initial and final consonants in CVC syllables. J. Acoust. Soc. Am. 106, 1555–1565.

Samuel, A.G., 1989. Insights from a failure of selective adaptation: syllable-initial and syllable-final consonants are different. Percept. Psychophys. 45, 485–493.

Schwab, S., Miller, J., Grosjean, F., Mondini, M., 2008. Effect of speaking rate on the identification of word boundaries. Phonetica 65, 173–186.

Smith, C.L., 1995. Prosodic patterns in the coordination of vowel and consonant gestures. In: Connell, B., Arvaniti, A. (Eds.), Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence. Cambridge University Press, Cambridge, pp. 205–222.

Sproat, R., Fujimura, O., 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. J. Phonetics 21, 291–311.

Stetson, R.H., 1988. R. H. Stetson's Motor Phonetics. In: Kelso, J.A.S., Munhall, K.G. (Eds.), A Retrospective edition. College-Hill, Boston, p. 235.

Sussman, H., Bessell, N., Dalston, E., Majors, T., 1997. An investigation of stop place of articulation as a function of syllable position: a locus equation perspective. J. Acoust. Soc. Am. 101, 2826–2838.

Tuller, B., Kelso, J.A.S., 1990. Phase transitions in speech production and their perceptual consequences. In: Jeannerod, M. (Ed.), Attention and Performance XIII. Lawrence Erlbaum, NJ, pp. 429–452.

Tuller, B., Kelso, J.A.S., 1991. The production and perception of syllable structure. J. Speech Hear. Res. 34, 501–508.

Turk, A., Shattuck-Hufnagel, S., 2000. Word-boundary-related duration patterns in English. J. Phonetics 28, 397–440.

Vihman, M., 1996. Phonological Development. Blackwell Publishers, Oxford.

Watt, D., Fabricius, A., 2002. Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1–F2 plane. Leeds Working Papers in Linguistics and Phonetics 9, 159–173.

Watt, D., Fabricius, A., Kendall, T., 2010. More on vowels: plotting and normalization. In: Di Paolo, M., Yaeger-Dror, M. (Eds.), Sociophonetics: A Student's Guide. Routledge, London, pp. 107–118.

Wells, J.C., 1990. Syllabification and allophony. In: Ramsaran, S. (Ed.), Studies in the pronunciation of English: a commemorative volume in honour of A.C. Gimson. Routledge, London, pp. 76–86.

Zsiga, E.C., 1994. Acoustic evidence for gestural overlap in consonant sequences. J. Phonetics 22, 121–140.