Gestural overlap across word boundaries: Evidence from English and Mandarin speakers*

SHAN LUO

University of Victoria

shanluo@uvic.ca

Abstract

This article examines how three factors determine the surface forms of English stop-stop co-articulation across word boundaries in both native and nonnative speech: place of articulation, frequency, and speech rate. The release percentage and closure duration ratio produced by English (L1) and Mandarin (L2) speakers were measured. The results showed that a place order effect was only partially supported in L1 speech but not shown at all in L2 speech. The results also confirmed a gradient lexical effect, finding a significant correlation between self-rated frequency and overlap. In addition, the results showed that increased speech rate did not induce increased overlap, given that speakers from both groups had either more or less overlap at the fast speech rate than at the slow rate.

Keywords: cluster coarticulation, place order effect, frequency, speech rate, Mandarin speakers

Résumé

Cet article examine trois facteurs qui déterminent les formes de surface des coarticulations obstruante-obstruante à travers des frontières des mots en anglais natif et en anglais non natif : le lieu d'articulation, la fréquence, et le rythme de la parole. Nous avons mesuré la fréquence de relâchement des obstruantes et la proportion de la durée de leurs fermetures chez des locuteurs natifs de l'anglais (L1) et chez des locuteurs natifs du mandarin (L2). Les résultats ont montré qu'un effet de l'ordre des lieux d'articulation a été partiellement attesté chez les locuteurs L1, mais que cet effet n'a aucunement été attesté chez les locuteurs L2. Les résultats ont également confirmé un effet lexical progressif; une corrélation significative entre la fréquence autoévaluée et le chevauchement a été observée. De plus, les résultats ont montré qu'un rythme plus élevé n'a pas entraîné un degré élevé de chevauchement, étant

donné que les locuteurs des deux groupes ont produit soit plus, soit moins, de chevauchement au rythme plus élevé.

Mots-clés: coarticulation des suites consonantiques, effet de l'ordre des lieux d'articulation, fréquence, rythme, mandarin

1. INTRODUCTION

The physical realizations of sounds are often conditioned by the surrounding segmental context in connected speech. For instance, it is typical in English for coronals to be perceived as assimilated to the following consonant or deleted: /t/ in *must be* is not perceived and the phrase is thus heard as [mʌsbi] (Brown 1977). Extensive studies within the framework of Articulatory Phonology have discovered that the occurrence of such coarticulation is due to substantial overlap of adjacent gestures in speech, and that the gesture that is not perceptually available is in fact not deleted but hidden by adjacent gestures (e.g., Browman and Goldstein 1990, Barry 1991, Nolan 1992, Recasens et al. 1993, Zsiga 1994, Byrd 1996a, Gafos 2002, Davidson and Stone 2003). Many non-native speech errors are thus assumed to arise from failing to apply the appropriate degree of overlap in the target language (Solé 1997; Davidson 2003, 2006, 2010; Zsiga 2003, 2011; Colantoni and Steele 2008). As part of the ongoing process of understanding speech production, the current study examines how three factors (i.e., place of articulation, frequency, and speech rate) affect the degree of overlap in both native (L1)[1] and non-native (L2) speech. In particular, I investigate surface forms of stop-stop coarticulation spanning across words (C1#C2) produced by English and Mandarin speakers.

The reasons for choosing the two languages are as follows. Mandarin has both unaspirated (/p, t, k/) and aspirated stops (/p^h, t^h, k^h/) (Cheng 1966). One major difference between Mandarin and English is that all Mandarin stops are voiceless, while English stops may be voiceless or voiced, although the six stops are sometimes transcribed as the same [p, b, t, d, g, k]. Another major difference is that Mandarin has strong phonotactic constraints on syllable shape, categorically prohibiting word-final stops and stop-stop clusters (Cheng 1966, Lin 2001), while English is well known for allowing heavy syllables, with up to three consonants in onset position and four in coda position (Harris 1994, Shockey 2008). In addition, stressed segments in English usually have less coarticulatory effects than unstressed segments, such as more lenited characteristics (e.g., reduced duration, deletion) (Öhman 1967, Crystal 1976, Avery and Ehrlich 1992, Cummins and Port 1998). The duration of each vowel or syllable in Mandarin, however, is stably fixed so that each syllable receives relatively equal timing across a sentence (Grabe and Low 2002, Lin and Wang 2007, Mok and Dellwo 2008). There is no distinctive vowel reduction or cross-word

---

[1]ANOVA: analysis of variance; C: consonant; ESL: English as a Second Language; IELTS: International English Language Testing System; L1: first language; L2: second language; NE: native English; NM: native Mandarin; POA: place of articulation; POE: place order effect; SD: standard deviation; TOEFL: Test of English as a Foreign Language; V: vowel.

coarticulation in Mandarin, with few exceptions (e.g., r-retroflexion) (Cheng 1966). Due to these differences between the two languages, Mandarin speakers are found to have substantial difficulty in acquiring English clusters (e.g., Anderson 1987, Weinberger 1987, Broselow and Finer 1991, Hansen 2001, Chen and Chung 2008).

Most studies investigate coda production and only a few have dealt with English cross-word coarticulation produced by Mandarin learners (e.g., Tajima et al. 1997, Chen and Chung 2008). Chen and Chung discovered that the lower the English proficiency of the Mandarin learners, the longer the duration of the consonant clusters that they produced. However, it remains poorly understood how cross-word coordination is implemented in English by Mandarin speakers. Moreover, there is extensive evidence that three factors systematically determine the overlap degree in native speech: place of articulation, frequency, and speech rate. This study aims to combine these factors concurrently to develop a more comprehensive account of what affects consonantal coarticulation in both L1 and L2 speech.

2.   ENGLISH CLUSTER COARTICULATION

This section discusses English cluster coarticulation, first in native speech and then in non-native speech.

2.1   English cluster coarticulation in native speech

Recent research findings have suggested that English consonant clusters have approximately 20% to 60% overlap (Catford 1977, Barry 1991, Zsiga 1994, Byrd 1996a). Three primary factors that have been shown to have systematic effects on gestural overlap of English clusters are discussed in detail below.

The most consistent result involves the place of articulation of C1 and C2 in stop-stop sequences. Earlier instrumental studies have discovered that in most sequences containing two adjacent English stops, either within words or across words (VC1C2V or VC1#C2 V), the closure for C1 is not released until the closure for C2 is formed (Catford 1977, Hardcastle and Roach 1979, Ladefoged 2001). Hardcastle and Roach (1979) report that in only 32 cases out of 272 was C1 released before the onset of the second closure. In English homorganic clusters, the lack of a release results in the whole sequence being interpreted as a long consonant rather than as a sequence of two segments, involving "a prolongation of the articulatory posture" (Catford 1977: 210). In heterorganic clusters, the lack of a release is more evident when C2 is more anterior than C1 than vice versa: fewer acoustic releases are produced in back-front clusters (e.g., coronal-labial, dorsal-coronal, or dorsal-labial) than in front-back sequences (e.g., labial-coronal, coronal-dorsal, or labial-dorsal) (Henderson and Repp 1982, Byrd 1996a, Zsiga 2000, Davidson 2011; see different results in Ghosh and Narayanan 2009). For example, Henderson and Repp found that 58% of C1s in [-C1C2-] (e.g., *napkin, abdomen*) were released, and release percentages were strongly conditioned by the place of articulation of C2 in the sequence. An average of 16.5% of C1s followed by labial C2s were released, compared to 70% of those followed by alveolar C2s and 87.5% of those followed

by velar C2s. That is, C1 was released more often when followed by a "back" consonant. Similar results were reported in Zsiga (2000: 78), who concluded that in English, differences in place of articulation "seem to account for release patterns in a straightforward way: clusters are more likely to have an audible release if C1 is further forward than C2".

In addition to affecting release patterns, this place order effect (POE) determines the degree of overlap. Hardcastle and Roach (1979) found that the interval between the onset of the C1 closure and the onset of the C2 closure was significantly shorter for /Vt#kV/ compared to /Vk#tV/ sequences, suggesting more overlap in /t#k/ (also see Henderson and Repp 1982, Barry 1991, Zsiga 1994). Byrd (1996a) discovered that the sequence overlap and C1 overlap were both much greater for [d#g] than for [g#d]. C2 also started much later relative to C1 for [g#d] than for [d#g], suggesting greater latency in [g#d]. The POE (more releases and more overlap in front-back clusters than in sequences of reversed order) is also evident in other languages (for Taiwanese, see Peng 1996; for Georgian, see Chitoran et al. 2002; for Russian, see Kochetov et al. 2007; for Greek, see Yip 2013).

Different syllable positions might result in different sequence timing and thus confound the POE. Existing analyses have demonstrated that English stop sequences have more timing variants and allow more overlap in codas than in onsets (Byrd 1996b, Browman and Goldstein 2000, Nam and Saltzman 2003, Byrd and Choi 2010). Speakers also display a greater amount of overlap across word boundaries than within (Sproat and Fujimura 1993, Byrd 1996b, Fougeron and Keating 1997, Zsiga 2000). However, more overlap across words does not mean the POE would be stronger in such positions. Davidson and Roon (2008) hypothesized that both consonants in C1#C2 are longer than their counterparts in #C1C2 because of the intervening word boundary. Their analysis confirmed that both monolingual Russian and bilingual Russian-English speakers produced #C1əC2 with the longest duration, followed by C1#C2 sequences, then by #C1C2 sequences. The follow-up study of Davidson (2011) on English found that neither English word-final clusters (CC#) nor clusters across words (C#C) allow more overlap in back-front clusters than in front-back clusters; only the word-internal clusters ([-CC-]) did. She posited that a coordination relation across words is not stable enough for a significant place order effect to emerge.

The second variable that could affect sequence timing is word frequency. A word frequency effect on English word reduction is attested in previous research, and these studies agree in finding that more frequent words have a variety of phonetically reduced characteristics (e.g., shorter durations, reduced vowels, deleted codas, more tapping, palatalization, etc.; see Fosler-Lussier and Morgan 1999, Gregory et al. 1999, Bybee 2000, Jurafsky et al. 2001, Bell et al. 2009). Bybee found that the schwa in frequent words like *memory* is more likely to delete than the schwa in the nearly identical, but infrequent, *mammary*. Bybee's data also showed a significant relationship of word frequency with rate of /t, d/ deletion. The most frequent word *told* had /d/ deleted in 68% of tokens, while the least frequent *meant* never had the /t/ deleted. Comparably, the results in Jurafsky et al. (2001) indicated that /t, d/ are deleted more often in high-frequency words than in low-frequency

words and that, if they are not deleted, their durations are significantly shorter in high-frequency than in low-frequency words. As much of the research deals with word-final deletion, there is a lack of research on how word frequency affects English cluster coordination (but see Yanagawa 2006, Davidson 2011). A notable exception is the study of Bush (2001), showing that palatalization of /tj/ and /dj/ sequences in English conversation occurs only between pairs of words that occur together most frequently (e.g., *did you, don't you, would you*).

One question that needs addressing is this: does frequency affect degree of overlap in consonant sequences, and if so, how? More specifically, to what extent are sound patterns like the POE evident in high-frequency versus low-frequency words? One way to answer this question is to test speakers' production of non-words. This idea has its root in research that claims that speakers possess implicit knowledge, and that this internalized working system would be able to select the right form in new cases (Berko 1958, Ernestus and Baayen 2003, Pater and Tessier 2006, Wilson 2006). The application of native phonological knowledge, however, may be incomplete (see Fleischhacker 2005, Zuraw 2007, Zhang and Lai 2010, Zhang et al. 2011).

Another variable that has been shown to affect consonantal coarticulation is speech rate (see Munhall and Löfqvist 1992, Byrd and Tan 1996, Tjaden and Weismer 1998, Fosler-Lussier and Morgan 1999, Jurafsky et al. 2001). These studies indicate that greater fluency is associated with segment shortening, such that the faster the speech rate, the more overlap occurs. For example, Munhall and Löfqvist found that glottal openings were blended into one single movement at fast rate when producing adjacent alveolar fricatives and alveolar stops. Similarly, Byrd and Tan reported that faster speech was associated with reduced magnitude in C1#C2. Individual consonants were shortened in duration as speaking rate increased, regardless of the place or manner of articulation of the consonants.

Other studies have reported inconsistent results. In the study of Zsiga (1994), comparison of cluster overlap in slow and fast tokens of the same utterance was not indicative of a direct relationship. The interaction of consonant and rate was "rarely significant" (1994:54) (also see Dixit and Flege 1991, Kochetov et al. 2007). The assumption that speaking rate change induces varying amounts of overlap has been further discouraged by a number of studies outlining substantial variability among speakers (Kuehn and Moll 1976, Ostry and Munhall 1985, Allen et al. 2003, Tsao et al. 2006). The study by Shaiman et al. (1995) showed that the timing between the jaw closing and the upper lip lowering gestures of a Vowel + Bilabial Stop sequence became significantly shorter as speakers moved from normal to fast speech. However, this effect was not consistent across speakers. They concluded that how speech rate changes were implemented varied across individual speakers (see Lubker and Gay 1982, Barry 1992, Nolan 1992, Tjaden and Weismer 1998).

Building upon previous research showing the influence of three factors (i.e., place of articulation, frequency, and speech rate) on consonantal coarticulation in English, a concurrent examination is needed because it remains to be seen first, whether the POE will show up in cross-word positions in English; second, how frequency (ranging from high- to low- to extremely low-frequency) affects cluster

coordination; and third, whether speech rate changes (reading phrases vs. embedded in conversation) affect overlap degree. More important, no previous studies have examined lexical frequency combining the POE associated with overlap degree in L1 and L2 speech. The current study remedies the lack of comparable data by testing speakers' productive extension with non-words in both L1 and L2 speech. In the next section, I discuss studies reporting on L2 production of English clusters.

2.2 English cluster coarticulation in non-native speech

Due to the extensive overlap in English clusters, most L2 learners encounter difficulty in coordinating adjacent gestures in an appropriate way, such that an insufficient overlap pattern arises (e.g., schwa insertion occurs, see Major 1987; Broselow and Finer 1991; Davidson 2003, 2006; Zsiga 2003, 2011; Colantoni and Steele 2008; Sperbeck 2010), or inappropriate assimilation occurs (see Weinberger 1994, Solé 1997, Cebrian 2000). Solé found that Catalan speakers produced one glottal gesture for a whole cluster, when the cluster members differ in voicing (e.g., [sn-] in *snail* becomes [zn-]), while this anticipatory effect is not shown in English speakers. Solé argued that the observed anticipatory effect in the Catalan subjects was derived from difficulty in changing the fossilized articulatory habits that govern anticipatory voicing in their L1 (for Hungarian regressive voicing assimilation transfer in English, see Altenberg and Vago 1983; for Polish regressive voicing assimilation in English, see Rubach 1984).

Interestingly, Cebrian (2000) found Catalan subjects failed to apply the L1 rule of regressive voicing assimilation at word boundaries in English; for example, *Swiss girl* did not undergo such rule. Cebrian then proposed the Word Integrity Principle, positing a tendency to preserve segments within one word, which implied a failure to coarticulate segments across word boundaries (also see Weinberger's (1994) Recoverability Principle). Such a principle predicts no specific coordination relationship ensuring overlap across words, and so that gestures across words produced by L2 speakers would drift apart. This hypothesis is confirmed by Zsiga (2003), who found that both English and Russian speakers preferred little overlap at word boundaries when producing L2 clusters. Meanwhile, the carryover of L1 is shown to be strong in Zsiga's study. English speakers produced Russian clusters with fewer C1 releases and greater overlap than Russian speakers, carrying over from the English patterns. Russian speakers produced English clusters with more C1 releases and smaller overlap, comparable to their articulatory habits in Russian. This finding is in line with other studies discussing L1 transfer that gives rises to L2 mistiming alignment (e.g., Sato 1987, Tajima et al. 1997, Chen and Chung 2008).

Aside from word-integrity by the Recoverability Principles, a number of studies have relied on Markedness Theory to account for L2 speech errors in producing clusters, particularly in three aspects: voicing (see Broselow and Finer 1991, Major and Faudree 1996, Broselow et al. 1998, Hansen 2001), cluster size (Anderson 1987, Weinberger 1994, Hansen 2001), and sonority sequencing (Broselow and Finer 1991, Eckman and Iverson 1993, Carlisle 2006, Cardoso 2008). For example, Spanish learners of English in Carlisle (2006) were found to modify /#sn/ more

than /#sl/ clusters, which is due to /#sn/ being more marked than /#sl/ according to the sonority distance. However, the extent to which sonority sequencing can account for L2 consonant cluster processing has been called into question. Davidson et al. (2004) reported that English speakers have varying accuracy when producing Polish /zm/ and /vn/ clusters (63% vs. 11%), although both have the same sonority distance. In a follow-up study, Wilson et al. (2014) discovered that other phonetic details (e.g., voicing) can be equally important for predicting L2 production patterns.

In sum, the reviewed studies on L2 gestural mistiming of English clusters suggest that a purely articulatorily motivated account is not sufficient. Rather, theories of word integrity, L1 influence, and markedness are all active in explaining L2 cluster production (see Davidson 2011 for a review). Given the conflicting results found in L2 studies, as well as the limited number of languages that have been studied, it is important to extend our cross-linguistic comparisons.

2.3   Research questions

Overall, three primary research questions concerning the degree of overlap in English clusters are raised.

First, does place of articulation (i.e., homorganic, front-back, and back-front clusters) affect English and Mandarin speakers similarly in their English stop-stop co-articulation? The general assumption is that the native Mandarin (NM) group will deviate from the native English (NE) group. I raise two hypotheses, listed in (1).

(1)   a.   The NE group will have fewer releases and more overlap than the NM group in producing homorganic clusters.

       b.   The NE group will show the POE, but the NM group will not.

Second, does frequency affect English and Mandarin speakers similarly in their English stop-stop coarticulation? This gives rise to the hypotheses in (2).

(2)   a.   Both groups will have fewer releases and more overlap in meaningful high-frequency words than in meaningful low-frequency words.

       b.   The NE group will show the POE in both real words and non-words, while the NM group will not show the POE in non-words.

Finally, do changes in speech rate affect English and Mandarin speakers similarly in their English stop-stop coarticulation?

3.   EXPERIMENTS

To answer the three research questions raised above, two studies were designed in which English and Mandarin speakers produced a set of clusters in slow and fast speech rates.

3.1   Participants

The experiment included two groups: one group of 25 native Mandarin speakers who use English as a second language (ESL), and a group of 15 native English speakers.

For each group, a questionnaire was administered to collect background information (e.g., age, gender, knowledge of additional languages, etc.) prior to the experiment. To control within-group homogeneity, Mandarin volunteers were selected only if they scored on/above 70% on TOEFL or IELTS. There were 14 females and 11 males with an average age of 27 in the NM group (23–40 years old). To control regional dialect influence, only speakers that had been born and raised in British Columbia, Canada, served as the NE group. The NE group consisted of 15 native English speakers, including nine females and six males with an average age of 26.3 years old (19–40 years old). Participation in this study was completely voluntary. All participants reported no hearing or speech problems. Participants were told that they were taking part in a speech study comparing native versus non-native production, but were given no other details until after the recording session was complete.

3.2 Materials

The central criteria for building the data were to consider first, place of articulation (i.e., homorganic, front-back, and back-front clusters); second, frequency (i.e., ranging from high- to low- to extremely low-frequency); and third, speech rate (reading phrases and spontaneous speech). Three lists were designed to incorporate these criteria.

List 1 contained 24 items, yielding a C1#V, V#C2 context (see Zsiga, 2003). Each stop appeared as onset and coda twice, to calculate the baseline closure duration where no consonant-consonant coarticulation occurs. List 2 included 144 disyllabic items with all stop-stop combinations, yielding a VC1#C2V environment. Seventy-two real words were chosen, including 36 compounds (noun-noun combinations, e. g., *soup pot*) and 36 non-compounds (e.g., *keep pace*). Another 72 phrases were designed as corresponding non-words. By corresponding, I mean that the vowel contexts between real words and non-words were controlled as similarly as possible (e.g., *peep pate* vs. *keep pace*), and their initial consonants and codas agreed in voicing and manner (e.g., /d/ in *dak kit* agrees with /b/ in *back kick* in voicing, and they are both obstruents). That is, two aspects of lexical frequency were considered: high-frequency versus low-frequency items in real words (e.g., *take care* vs. *dock gate*), and real words versus non-words (e.g., *keep pace* vs. *peep pate*). By doing this, we can make inferences about the ability of speakers to generalize from known forms to less familiar forms, then to novel forms. To control the stress effect, all testing syllables were monosyllabic words (CVC). Using these 144 disyllabic phrases, List 3 included 144*4 dialogues combining four kinds of focus contrasts where C1 and C2 were embedded in a target sentence (more in 3.3).

3.3 Procedure

A pre-test was conducted to assess Mandarin participants' English proficiency in addition to the self-reported TOEFL/IELTS scores. The NM group was asked to speak about a topic for two to three minutes; the question was "What do you feel the most passionate about?" (see Lin and Wang 2007). The accentedness of the

monologue was judged by two native English speakers with a Likert scale from one to seven (Likert, 1932), where a score of one indicates native-like and a score of seven indicates heavy accent. No participants received scores higher than six (i.e., extremely heavy accent). The mean rating was 3.38/7 for Judge 1 and 3.06/7 for Judge 2, suggesting that the NM group was perceived to be moderately accented, where 3.5 is the midpoint of the scale. The Pearson correlation test shows that the two judges' scores are significantly correlated with one another [r = .58, p < .002]. I am confident that the NM group in this study represents intermediate or better proficiency.

The experiment itself was composed of four tasks, administered using E-prime 2.0 Professional (version 2.0.1.97). I first needed to confirm that the participants did make a distinction between high- and low-frequency words. Therefore, Task 1 asked participants to rate word familiarity on all 144 items on a seven-point Likert scale, where one indicates the least familiarity and seven indicates extremely familiar. Prior to Task 1, participants were reminded to make judgements based on orthography rather than sound because some non-words might sound like real words in English (e.g., *tik doun*).

E-prime collected a total of 5760 familiarity responses from Task 1 (40 participants * 144 stimuli items). Table 1 summarizes the lexical judgment scores for the two groups with standard deviations (SD) in parentheses. The results show that for the NE group, the average score for the 72 real words was 5.66, and for the non-words 1.06. For the NM group, the average score for the meaningful items was 4.74, while for the non-words 1.16. Table 1 also includes the results of two-tailed t-tests, showing that both groups significantly distinguished real words from non-words. The categorical frequency effect was thus validated.

Task 2 asked participants to read the 24 phrases contained in List 1 aloud, clearly. Task 3 asked participants to read the 144 items contained in List 2 aloud, in natural speech. Prior to Task 3, the participants were instructed to produce the words as naturally as possible; the researcher also explicitly told participants to utter all the items in an English way when they ran into words they did not know. In Task 3, stimuli were marked by underlining the appropriate lexical stress in both real words and non-words (e.g., _soup pot, foop pok_). Each phrase appeared once in Task 2 and once in Task 3, requiring participants to repeat each phrase three times in each task.

Group	Real Words	Non-words	t-test
NE (n = 15)	M = 5.66 (.62)	M = 1.06 (.14)	t (14) = 30.118, p < .001
NM (n = 25)	M = 4.74 (.95)	M = 1.16 (.38)	t (24) = 19.911, p < .001

Table 1: Summary of lexical frequency ratings for both groups[2]

---

[2]In the ANOVA and t-test analyses reported in this paper, a level of .05 was used in all cases to determine significance.

To further study the effect of speech rates, a fourth task was included. Task 4 asked participants to answer questions based on a given dialogue. Specifically, using the 144 disyllabic phrases in Task 3, Task 4 was designed as having 144*4 dialogues. Each dialogue included a question and a given answer. Participants were asked to read the question silently, then produce the given answer aloud. The answers were designed to elicit four sentential stress patterns: focused versus unfocused, unfocused versus focused, focused versus focused, and unfocused versus unfocused.[3] Only the embedded clusters were subject to analysis in Task 4.

The experiment was self-paced. Participants were presented with a new item only after they had indicated that they were ready by pressing the keyboard. All tested materials in each task were in a randomized order generated by E-prime 2. The data were collected using a head-mounted microphone with a fixed distance from the mouth. All participants were recorded in the phonetic laboratory of the Department of Linguistics at the University of Victoria. All .wav files were recorded at 44.1 kHz/16 bits.

Speech rate was calculated by using a speech rate script, measuring how many syllables were produced per second (De Jong and Wempe 2009). The results are summarized in Table 2 below (SD in parentheses). The table indicates that the NE group produced 1.48 syllables /per second at the word level and 2.22 syllables at the sentence level, and the NM group produced 1.28 and 1.76 syllable/per second, respectively. Also included in the table are the results of the two-tailed t-tests, which indicate that both groups had a significantly faster speech rate at the sentence level than at the word level. The *Speech Rate* effect was thus validated for both groups.

3.4 Data analysis

A total of 29,760 sound files were acoustically analyzed (40 participants * (24 in Task 2 + 144 in Task 3 + 576 in Task 4)) using Praat software (Boersma and Weenink 2009). Duration features were extracted from the manually labelled segments using a duration script developed for Praat. To exclude disfluent tokens, any phrase in which there was a period of silence of 750 ms or more in C1#C2 was discarded. A total of 58 tokens were excluded based on this criterion to avoid skewing the data (0.2% of the total collected).

Group	Word Level	Sentence Level	t-test
NE (n = 15)	M = 1.48 (.51)	M = 2.22 (1.02)	t (14) = −4.155, p < .001
NM (n = 25)	M = 1.28 (.4)	M = 1.76 (.53)	t (24) = −6.329, p < .001

Table 2: Speech rate summary for both groups

---

[3]An example of Task 4 is given below.
Focused – Unfocused
Prompt: Did you say "**good** day" just now?
Target: No, I said "**bad** day" just now.

In the present study, C1#C2 articulatory relationships were measured through two dependent variables. The first dependent variable was release percentage. A cluster of C1#C2 was counted as released if C1's burst of energy was visible on a spectrogram. In Task 3, the middle token was subject to analysis when the release identification agreed with the first or the third token; otherwise, I chose the third token to analyze. All the data were coded by the author. Another trained researcher coded 507 sound files that were randomly chosen from Task 3 (8.8% of the total collected in Task 3). Twenty tokens did not agree in release identifications (mainly quasi-release vs. no release; for quasi-release, see Davidson 2011), making the disagreement 3.9% (20/507) between judges. The inter-coder reliability was 96.1%, suggesting high consistency. Release percentage was then calculated as the number of tokens that had released C1 divided by the total tokens. The second variable was closure duration ratio, computed for each phrase for each speaker. The ratio is defined as "*the mean duration of the C1#C2 cluster divided by the sum of the mean closure durations of $C1_b$ and $C2_b$ occurring intervocalically*" (Zsiga 2003:441), so that a smaller value indicates a larger degree of overlap.

4. RESULTS

Based on the experimental design, both release percentage and duration ratio were analyzed using four-way repeated measures ANOVAs with a 3*2*2*2 design. The three within-subjects factors were PLACE OF ARTICULATION (POA, 3 levels: homorganic, front-back, and back-front clusters), FREQUENCY (2 levels: real words vs. non-words) and SPEECH RATE (2 levels: word level vs. sentence level). GROUP was the between-subjects factor (2 levels: the NE group vs. the NM group). Table 3 below provides the overall statistics with significant effects in bold. There are two main effects, and two two-way interactions are significant on the release percentage (POA; FREQUENCY; POA by GROUP; SPEECH RATE by GROUP); one main effect and one 2-way interaction are significant on the duration ratio (FREQUENCY; FREQUENCY by GROUP). In the following subsections, I will divide and collapse data to focus on and investigate each specific research question.

Effect	Release Percentage	Duration Ratio
POA	**$F_{(2,76)} = 73.606$, p < .001**	$F_{(2,76)} = .551$, p > .572
FREQUENCY	**$F_{(1,38)} = 46.719$, p < .001**	**$F_{(1,38)} = 54.592$, p < .001**
SPEECH RATE	$F_{(1,38)} = .598$, p > .444	$F_{(1,38)} = .008$, p > .928
POA * GROUP	**$F_{(2,76)} = 4.675$, p < .016**	$F_{(2,76)} = .912$, p > .403
FREQUENCY * GROUP	$F_{(1,38)} = .122$, p > .729	**$F_{(1,38)} = 4.825$, p < .034**
SPEECH RATE * GROUP	**$F_{(1,38)} = 6.085$, p < .018**	$F_{(1,38)} = .046$, p > .832

Table 3: Summary of statistical results (significant effects in bold)

4.1   Place of articulation

The first research question addresses whether place of articulation (i.e., homorganic, front-back, and back-front clusters) affects English and Mandarin speakers similarly in their English stop-stop coarticulation.

The statistical results indicate that POA has significant effect on release percentage but not on duration ratio for either group. The overall POA effect averaged over both speech rates on release and duration ratio patterns for each group is plotted in Figure 1 below, which shows a similar release trend for both groups. This trend is visible in that both groups had the highest release percentage in front-back clusters, followed by back-front clusters, then homorganic clusters (i.e., front-back > back-front > homorganic, where the symbol ">" indicates "more releases than"). The NE group had the most overlap in front-back clusters, followed by back-front clusters, then homorganic clusters (i.e., front-back >> back-front >> homorganic, where the symbol ">>" indicates "more overlap than"); the NM group did not exhibit much difference in producing the three kinds of clusters.

Multiple comparisons, with Bonferroni corrections, were conducted to further examine group differences. In terms of hypothesis (1a), despite the visible trend that the NE group had fewer releases than the NM group, the overall group differences were not statistically significant in releasing homorganic clusters at either speech rate [both p > .072]. Similarly, the NE group did not have significantly more overlap than the NM group in producing homorganic clusters at either speech rate [both p > .163]. In terms of hypothesis (1b), the results showed, on the one hand, that the NE group released significantly more often in front-back than in back-front clusters at both speech rates [both p < .005], while the NM group did not distinguish the two types of clusters significantly at either speech rate [both p > .37]. On the other hand, the analyses indicated that neither of the groups had significantly more overlap in front-back than in back-front clusters at either speech rate [all p > .849].
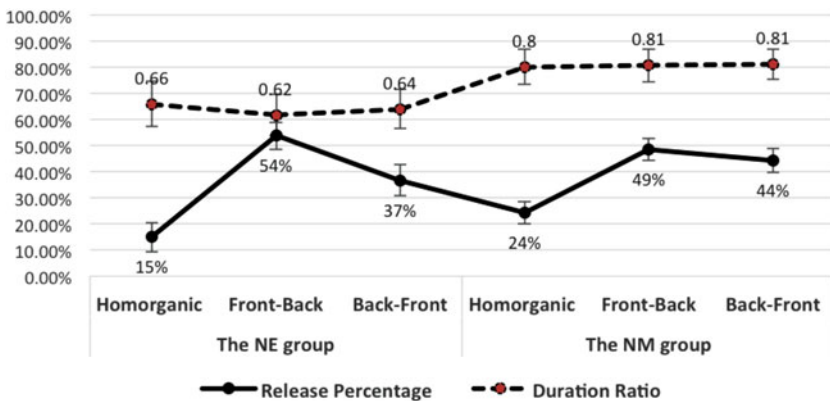

Figure 1: POA effect on release percentage and duration ratio for both groups
Note: The larger the duration ratio, the smaller the gestural overlap.

These findings do not support hypothesis (1a), although the trend was in line with the prediction. Hypothesis (1b) was partially supported, since the predicted release pattern of the POE was statistically evident in L1 speech but not in L2 speech. However, hypothesis (1b) was not supported in that the predicted overlap order of the POE was not statistically shown in either group.

4.2   Lexical Frequency

The second research question asked whether and how lexical frequency determines the cluster realizations: Will English and Mandarin speakers exhibit the same frequency effect on overlap degree?

The results showed that for the 72 real words, the NE group released 31% and produced an average duration ratio of 0.61 (i.e., 39% overlap); the NM group released 36% and produced an average duration ratio of 0.73 (i.e., 27% overlap). For the 72 real words, Pearson two-tailed correlation tests were conducted to examine whether there was a correlation between lexical frequency and release percentage, and between lexical frequency and overlap degree. The results showed that lexical frequency and release percentage were not correlated in either of the groups [both $p > .657$], while the frequency rating and the closure duration ratio were significantly negatively correlated for both groups [both $p < .003$]. This result suggested that the more frequent the item was, the shorter the closure duration was (i.e., the greater the overlap). Hypothesis (2a) was therefore partially supported: it was supported in terms of a significant correlation between frequency and duration overlap, but not supported in terms of lacking a correlation between frequency and release percentage.

As for hypothesis (2b), Table 4 summarizes the release percentages for both groups averaged over speech rates, subcategorized by the cluster type in both lexical contexts (SD in parentheses). The table shows that the NE group released 31% in real words and 40% in non-words, while the NM group released 35% and 43%, respectively. The main effect of FREQUENCY was significant on release percentage, indicating that both groups released significantly less often in real words than in non-words.

Group	The NE Group			The NM Group		
Cluster Type	Homorganic	Front-Back	Back-Front	Homorganic	Front-Back	Back-Front
Real Words	11% (.23) M = 31%	49% (.32)	33% (.31)	20% (.25) M = 35%	45% (.19)	41% (.23)
Non-Words	19% (.3) M = 40%	58% (.32)	41% (.35)	29% (.25) M = 43%	52% (.2)	48% (.25)

Table 4: FREQUENCY effect on release percentage for both groups

Despite the fact that both groups showed the same release trend of front-back > back-front > homorganic clusters regardless of lexical status, the statistical analyses found that the NE group had significantly more releases in front-back than in back-front clusters for real words as well as for non-words at both speech rates [all p < .002], while the NM group did not statistically distinguish the two types in real words or non-words at either speech rate [all p > .614]. This result directly says that the NE group had a consistent and strong POE not subject to the FREQUENCY effect.

The main effect of FREQUENCY was also significant on the duration ratio, with both groups having significantly smaller overlap in non-words than in real words. To further explore the significant interaction of FREQUENCY * GROUP, Table 5 summarizes the descriptive statistics of duration ratios in each cluster type averaged over speech rates produced by both groups.

The table shows that the NE group had the same overlap order across lexical conditions. This pattern is visible: front-back clusters had the most overlap, followed by back-front clusters, then homorganic clusters in both lexical environments (i.e., front-back >> back-front >> homorganic). The NM group had inconsistent organizations. In real words, the NM group had the most overlap in homorganic clusters, followed by front-back and then back-front clusters (i.e., homorganic >> front-back >> back-front). In non-words, back-front clusters allowed more overlap than front-back clusters (i.e., homorganic >> back-front >> front-back). The analyses indicated that neither of the groups had significantly more overlap in front-back than in back-front clusters in either lexical context at either speech rate [all p > .8]. This result means the hypothesis 2b) was not supported.

Post hoc analyses showed that group duration ratio differences were significant in non-words [F(1.38) = 4.121, p < .049] but not in real words [F(1.38) = 2.728, p > .107]. These significant differences came from different organizational patterns in nonsense front-back and back-front clusters [both p < .037]. This finding suggests that both groups coordinated the real words in a similar manner; however, the NM group failed to coordinate the non-words, mainly nonsense front-back and back-front clusters, in the same fashion that the NE group did. These findings are of considerable

Group	The NE Group			The NM Group		
Cluster Type	Homorganic	Front-Back	Back-Front	Homorganic	Front-Back	Back-Front
Real Words	62% (.39) M = 61%	59% (.34)	61% (.32)	75% (.28) M = 75%	74% (.26)	76% (.36)
Non-Words	70% (.41) M = 67%	65% (.35)	67% (.32)	86% (.27) M = 87%	87% (.33)	87% (.32)

Table 5: FREQUENCY effect on duration ratio for both groups

significance as they show that unlike the NE group, the NM group was unable to extract phonetic details and apply their phonetic knowledge to novel words.

4.3   Speech rate

The third research question asked whether gestural overlap increases when speech rate is increased. Table 6 provides the descriptive statistics, including the mean release percentage and mean duration ratio for each group at each speech rate (SD in parentheses). The table shows that the NE group released more often at the fast rate than at the slow rate (37% vs. 33%) while the NM group released less often at the fast rate (35% vs. 43%).

As shown in Table 3, the main effect of SPEECH RATE was not significant for duration ratio. In order to further examine how gestural overlap is implemented to differing degrees by individual speakers, the performance of each participant was examined. A breakdown of each participant's data indicates that there are two types of speakers: those that had more overlap in fast speech (19 speakers, nine NE and 10 NM) and those that had less overlap in fast speech (21 speakers, six NE and 15 NM). For the first type, participants' release percentage was significantly correlated with the duration ratio slow speech [r = .56, n = 19, p < .014], while no significant correlation was found between release percentage and duration ratio in fast speech [r = .26, n = 21, p > .277]. For the second type, speakers' release percentage was significantly correlated with the duration ratio regardless of speech rates [both p < .036]. These results showed that predetermined gestural patterns were not simply accelerated by faster speech rates.

To sum up, a breakdown of speakers suggests a large amount of speech freedom: speakers could either have more or less gestural overlap at faster rates. The prediction related to the third research question is thus not supported; instead, individual differences are strongly shown in the current study.

5.   SUMMARY AND DISCUSSION

This study investigates how gestural overlap is conditioned by three possible influences: place of articulation, frequency, and speech rates. The data raise some important questions about the nature of coordination, how it is implemented and how it is stored cognitively.

Group	The NE Group		The NM Group	
Variables	Release Percentage	Duration Ratio	Release Percentage	Duration Ratio
Slow Rate	33% (.36)	64% (.3)	43% (.38)	82% (.29)
Fast Rate	37% (.34)	64% (.43)	35% (.28)	81% (.32)

Table 6: SPEECH RATE effect on release percentage and duration ratio for both groups

The overall results showed that release percentage and closure duration ratio were significantly positively correlated [r = .47, n = 40, p < .002], that is, the greater the release percentage, the greater the closure duration ratio (i.e., the smaller the gestural overlap). This effect was in fact due to the NE group [r = .58, n = 15, p < .02]. The NM group did not have a significant correlation [r = .37, n = 25, p > .07]. This result indicates that when the NE group had more releases, they had longer closure durations of C1#C2, resulting in less overlap. However, this systematicity was not shown in non-native speech. The coordination of English clusters by the NM group seems fairly random.

On average, 35% of C1s were released in the NE group and 39% of C1s were released in the NM group. The English speakers' release percentage falls within the range predicted by previous studies (Ghosh and Narayanan 2009, Davidson 2011). The NE group's overlap is also congruent with the overlap range suggested in previous research (Barry 1991, Zsiga 1994, Byrd 1996a). Not surprisingly, the NE group's overlap degree was almost twice that of the NM group (36% vs. 19%).

5.1 Place of articulation

This study included all stop-stop combinations across words to examine the effect of POA. The main effect of POA was found to be significant on release percentage, but not on closure duration ratio.

The fact that both groups had fewer releases in homorganic than in heterogenic clusters (see Figure 1) is not unexpected, in light of existing cross-linguistic findings, which suggest fewer releases in clusters where the adjacent consonants share the same place of articulation than in clusters where the adjacent consonants have different articulation places (Elson 1947, Clements 1985, Pouplier 2003, Zsiga 2003). Even in L2 speech, this prolongation is strongly evident (Zsiga 2003). The articulatory rationale behind this, as Pouplier (2003) and Goldstein et al. (2007) argue, is that the consonants in heterorganic clusters, which use different subsets of articulators, are largely independent of each other; the constriction gestures are therefore compatible and can be produced concurrently. For homorganic clusters that share the same constriction gesture, the consonantal gestures are in competition and can only be manipulated sequentially. The results from both groups in this study suggest a universal grounded articulatory constraint, such that speakers, regardless of their native language, tend not to release C1 in homorganic clusters.

In the current study, the results showed that the NE group had a closure duration ratio of 0.66 (34% overlap) and the NM group had 0.8 (20% overlap) in producing homorganic clusters. This finding provides strong evidence for the blending hypothesis of Browman and Goldstein (1989). Browman and Goldstein discussed gestural blending in terms of place compromise: the location of the constriction should fall somewhere in between C1 and C2 when they share the same articulatory tier, as reflected by the acoustic consequences corresponding to the influence of both consonants. Following this idea, I argue for temporal blending between adjacent gestures when they share the same articulator, which is reflected as the timing influence of both consonants. Specifically, if homorganic clusters were geminate sequences,

they would have a single gesture, and the other gesture would be deleted. However, the current study found that the two groups shortened the closure duration of each stop and produced an overlap relation of 34% and 19%, respectively. If one gesture were deleted, a 0.5 closure duration ratio would be expected. Given that both groups yielded overlap figures less than 50%, I conclude that they did not delete a gesture entirely; rather, they shortened each gesture, resulting in less than 50% overlap. The hypothesis of temporal blending is supported by other studies such as Byrd (1995), and Munhall and Löfqvist (1992). Byrd discovered that two lingual gestures are canonically present in geminated sequences (C1#C2), and that the coproduced movement for C1#C2 is longer than the non-coproduced movement for a single gesture.

In this study, the NE and NM speakers exhibited a trend of the POE: more releases and more overlap in front-back than in back-front clusters. This result is consistent with crosslinguistic studies of the POE (Hardcastle and Roach 1979, Henderson and Repp 1982, Zsiga 1994, Byrd 1996a, Peng 1996, Chitoran et al. 2002, Kochetov et al. 2007, Davidson 2011). However, findings here were somewhat mixed: only the release pattern was statistically shown in native speech, and the overlap order was not statistically shown in either group.

We have to consider the fact that the consonant clusters used in this study were in boundary positions. Gestures in such cross-word positions stand in a substantially looser relationship than those formed within words in English, where speakers are likely controlling a more stable pattern (Byrd and Saltzman 1998, Cebrian 2000, Davidson 2011). A loose coordination of consonant clusters would result in greater splittability (e.g., excrescent vowels) (Zuraw 2007), longer duration in C1#C2 (Davidson and Roon 2008), and gradient overlap degree (Zsiga, 1995). Under this assumption, gestures are more likely to drift apart when in a boundary position. The findings in the current study appear to show both the POE and the juncture effects. Both groups exhibited the trend of the POE; but the POE was only partially shown in L1 speech. The significant release pattern of the POE in the NE speech indicates that a constriction transition from front to back articulators is strongly favored and articulatorily grounded even in an English boundary position where gestures are relatively loosely coordinated. Compared to NE speakers, NM speakers were more likely to eliminate contact across words, showing a tendency of preserving each word as a unit. What these results suggest is weaker juncture effects in native speech than in non-native speech. Considering that no significant overlap orders emerged in either group, future research is needed to further explore the relation between the POE and the juncture effects with different measurements of overlap.

5.2 Cognitive processing

This study considered two aspects of frequency in order to assess its role in cluster realization: high-frequency versus low-frequency items in real words (e.g., *take care* vs. *dock gate*), and the real words (i.e., high-frequency) versus the non-words (i.e., extremely low-frequency) (e.g., *keep pace* vs. *peep pate*).

In terms of the gradient effect (i.e., high vs. low frequency items in real words), the results found no significant correlation between frequency and release percentage in either group, but a significant correlation between frequency and overlap for both groups. Both groups rated real words such as *take care* and *look good* highly familiar; these had up to 50% overlap. Words such as *hug pug* and *dock date* were rated as much less familiar, and showed only up to 20% overlap. This result suggested a frequency-matching behavior for both groups: more reductions occur for more frequent words in meaningful items. This finding is in accordance with existing research on the effect of lexical frequency in sound patterns (Gregory et al. 1999; Bybee, 2000, 2007; Bush 2001). Gregory et al. report that tapping of a word-final /t/ or /d/ in an intervocalic context is highly affected by the probabilistic variable "mutual information", which measures the likelihood that two words will occur together. As mutual information increases, reflecting a stronger cohesion between the two words, the likelihood of tapping increases. In this study, as lexical frequency increases, the overlap increases.

Moreover, both groups showed frequency-matching behavior by essentially differentiating real words from non-words. The categorical *FREQUENCY* effect was found to be significant on both release percentage and closure duration ratio. Both groups had significantly fewer releases and more overlap in clusters with real words than with non-words (e.g., *keep pace* >> *peep pate*). The NE group had 39% overlap in producing real words and 33% in non-words averaged over both speech rates; the NM group had 24% and 15% overlap, respectively.

The group differences were observed in the interaction of Frequency * POA. In the current study, NE speakers' treatment of real words versus non-words exhibited a high degree of consistency. In terms of release percentage, the NE group had significantly more releases in front-back than in back-front clusters in both lexical contexts. In terms of duration ratio, the trend of the NE group was to organize the three cluster types in the same fashion (i.e., front-back >> back-front >> homorganic) in both lexical environments. The results indicated that NE speakers maintained a stable timing pattern in new forms even when they slightly modified the degree of overlap in non-words. This observed consistency supports previous studies showing that native speakers have the ability to identify regularities and systematic alternations, and extend their phonological knowledge to unfamiliar environments (see Berko 1958, Wilson 2006, Zhang and Lai 2010). English speakers' ability to articulate using their knowledge in this study further suggests that they have productively internalized the phonological process. As Wilson (2006: 946) comments, speakers have "detailed knowledge of articulatory and perceptual properties". It is obvious that their grammatical systems can make reference to that knowledge.

The rich and detailed representation found in the NE group also essentially corroborates a fundamental hypothesis of Articulatory Phonology; phonological contrasts are represented in terms of articulatory gestures, not discrete features (Browman and Goldstein 1989, 1990, 1992; Goldstein et al. 2006). One implication of the gestural unit hypothesis is that this kind of representation is apparently what speakers make use of for their lexicon. The findings reported here suggest a step further: the stored gestural units are available for native speakers to reconstruct

and redeploy combinatorially in new contexts. If these units could not be reorganized, the English speakers in this study would coordinate non-words differently from real words, perhaps having a different overlap order among the three cluster types. Yet in non-words, they still followed the pattern of front-back >> back-front >> homorganic. It seems that native speech not only encodes gestural organizations manifested in the lexicon, it also allows for the recombination of new possibilities with the same structure (also see Kita and Özyürek 2003).

In contrast, there is no evidence that Mandarin speakers have acquired nonnative phonological knowledge; rather, how they organize L2 clusters seems to be largely dependent on familiarity with lexical knowledge. The NM group had different overlap orders in different lexical conditions. They produced front-back >> back-front in real words, and back-front >> front-back in non-words. The former pattern of the NM group was consistent with that of the NE group, showing that the two groups had similar organizations with familiar lexical items. However, the NM group's responses for non-words showed that they did not generalize beyond memorized items. This response shows that Mandarin speakers were not sensitive to fine-grained phonetic details in English and thus were unable to reorganize English gestural units that they were familiar with, at least within the confines of the difficult task of reading and orthographically decoding the non-words in L2. As a result, L2 speakers have a limited repertoire for reconstructing new possibilities within L2 phonotactic constraints.

5.3   Stable pattern

This study not only examined how coordination is implemented and how sound patterns are stored in L1 and L2 speech, it also investigated whether changes in speech rate induce greater overlap. The validation analyses established that both groups produced more syllables per second at the fast rate (i.e., Task 4) than at the slow rate (i.e., Task 3); however, the effect of *SPEECH RATE* was not significant on either release percentage or duration ratio. The results showed that the NE group released more often when speech rate increased (33% vs. 37%), and had slightly more overlap at the slow rate (36.3% vs. 35.9%). Meanwhile, the NM group released less often when speech rate increased (43% vs. 35%), and had slightly more overlap at the fast rate (18% vs. 19%). The results suggest that the *SPEECH RATE* effect was opposite for the two groups.

Further analyses indicated that individual speakers differed in having either *more* or *less* overlap at the fast rate than at the slow rate, pointing towards a speaker-dependent effect. Nine NE speakers and ten NM speakers had more overlap in fast speech, while six NE speakers and 15 NM speakers had less overlap in fast speech. Worth noting is that there was a larger proportion in the NE group (9/15) who had more overlap in faster speech rate than in the NM group (10/25). This finding is consistent with the only study, to my knowledge, that has examined the speech rate effect on overlap produced by Mandarin speakers, Chen and Robb (2004). They found that Mandarin speakers produced English with a slower speaking rate and articulation rate compared to American speakers. Their analyses also

indicated that the Mandarin speakers inserted significantly more pauses between syllables. Comparably, the current study found that the NM group tended to have a longer duration (i.e., smaller overlap) than the NE group even when their general speech rate increased.

The overall results indicate a stable pattern along with speaker variations in both slow and fast speech rates, which is consistent with other work showing individual differences in this regard (e.g., Kuehn and Moll 1976, Ostry and Munhall 1985, Zsiga 1994, Tsao et al. 2006). I outline here three possible accounts for the lack of correlation between increased overlap and increased speech rate.

One view is to consider the phonetic nature of stops. As Kessinger and Blumstein (1997) point out, stops have intrinsic limitations in their production. Specifically, stops involve a pressure build-up behind the constriction and a rapid release of the oral closure. As seen from the acoustic consequences, stops have a rapid change and an abrupt amplitude increase at their release when preceding a vowel (Stevens 1980). These physical gestures must occur to implement the obstruent manner of articulation in either slow or fast speech. With this obstruent manner, stops are less vulnerable than other consonants (e.g., affricates) in that changes in speaking rate are less likely to affect them. In fact, acoustic studies have shown that speaking rate changes affect the duration of glides and affricates, but do not affect stops (Miller and Baer 1983, Shinn 1984). Also, changes in speech rates do not induce differing voice onset time of English stops (Kessinger and Blumstein 1997; but see different results in Pind, 1995 and Theodore et al. 2007).

An alternative view is that the rate is simply not fast enough to have exerted much influence. Tjaden and Weismer (1998) suggested that the increase in speech rate that participants self-produced may not have been great enough to shift coordination patterns. The transition between adjacent gestures cannot be affected if the rate change is not large enough. The speech organs would rather remain relatively stable during the course of production. Lindblom (1983) showed that speakers do not execute extreme displacements or velocities in coarticulation and vowel reduction; they neither hyper-articulate, nor constantly whisper or mumble. Instead, speakers prefer behaviours that minimize motoric demands.

An additional view is that people always tacitly find alternative means to reach target values without compromising a steady speech rhythm. For example, Moisik and Esling (2011) discovered that speakers raise their larynx to supplement the raising pitch value supposedly induced by a voiced target without laryngeal constriction. Ellis and Hardcastle (1999) showed that individual speakers may exploit either categorical or gradient assimilation when producing /n#k/. The EPG data showed that two speakers never exhibited tongue-palate contact in the velar region during the /n/, whereas four other speakers consistently produced /ŋ/. Two more speakers alternated between complete assimilation and no assimilation. Only two speakers showed a range of tongue-palate contact from full alveolar closure to partial velar contact to total velar closure. Ellis and Hardcastle (1999: 2428) concluded that speakers' preferred strategies "can be fundamentally different". In the present study, the fact that speakers had either more or less overlap when speech rate increased suggested

that gesture coordination was self-controlled; speakers could easily modify the overlap pattern without compromising a steady speech pattern.

6. CONCLUSION

This study extends our understanding of gestural overlap in both L1 and L2 speech. Although articulatory contact was not directly measured, a detailed examination of two-stop combinations was conducted. I tested the effects of three factors concurrently on the amount of overlap in English stop-stop sequences across words in both L1 and L2 speech: place of articulation, frequency, and speech rate. The first two were found to affect the coordination of gestural units. The POE was only partially shown in L1 speech, but not shown at all in L2 speech. I discussed the implications and proposed to account for the results in terms of juncture effects where gestures become larger and longer, weakening the POE. I also found that the NE group extended the release pattern of the POE into non-words and had consistent overlap orders across lexical conditions; in contrast, the NM group deviated significantly from the NE group especially in coordinating nonsense front-back and back-front clusters, suggesting that they cannot extend their L2 phonetic awareness (if any) into unfamiliar phonological production. The consistency of the NE group showed that the POE is learnable in native processing, indicating that sound patterns are registered in memory along with their phonetic details. Further analyses indicated that a faster speech rate did not induce more overlap in either group, indicating that overlap was not simply accelerated at faster rates.

REFERENCES

Allen, J. Sean, Joanne L. Miller, and David DeSteno. 2003. Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America* 113(1): 544–552.

Altenberg, Evelyn, and Robert M. Vago. 1983. Theoretical implications of an error analysis of second language phonology production. *Language Learning* 33(4): 427–447.

Anderson, Janet. 1987. The markedness differential hypothesis and syllable structure difficulty. In *Interlanguage Phonology: The Acquisition of a Second Language Sound System*, ed. Georgette Ioup and Steven Weinberger, 279–291. Cambridge, MA: Newbury House.

Avery, J. Peter, and Susan Ehrlich. 1992. *Teaching American English pronunciation*. Oxford: Oxford University Press.

Barry, Martin. 1991. Temporal modelling of gestures in articulatory assimilation. In *Proceedings of the 12th ICPHS*, 14–17. Aix-en-Provence.

Barry, Martin. 1992. Palatalisation, assimilation and gestural weakening in connected speech. *Speech Communication* 11(4–5): 393–400.

Bell, Alan, Jason M. Brenier, Michelle Gregory, Cynthia Girand, and Dan Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60(1): 92–111.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14: 150–177.

Boersma, Paul, and David Weenink. 2009. Praat: doing phonetics by computer (Version 5.1. 05)[Computer program]. Retrieved May 1, 2009.

Broselow, Ellen, Su-I Chen, and Chilin Wang. 1998. The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* 20(2): 261–280.

Broselow, Ellen, and Daniel Finer. 1991. Parameter setting in second language phonology and syntax. *Second Language Research* 7(1): 35–59.

Browman, Catherine, and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6(2): 201–251.

Browman, Catherine, and Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, ed. George N. Clements, John Kingston, and Mary Beckman, 341–376. Cambridge: Cambridge University Press.

Browman, Catherine, and Louis Goldstein. 1992. Articulatory phonology: An overview. *Haskins Laboratories Status Report on Speech Research* 112: 23–42. New Haven: Haskins Laboratories.

Browman, Catherine, and Louis Goldstein. 2000. Competing constraints on intergestural co-ordination and self-organization of phonological structures. *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5: 25–34.

Brown, Gillian. 1977. *Listening to spoken English*. London: Longman.

Bush, Nathan. 2001. Palatalization in English. In *Frequency and the emergence of linguistic structure*, ed. Joan L. Bybee and Paul Hopper, 255–280. Amsterdam: John Benjamins.

Bybee, Joan. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-Based Models of Language*, ed. Michael Barlow and Suzanne Kemmer, 65–85. Stanford: CLSI Publications.

Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Byrd, Dani. 1995. Articulatory characteristics of single and blended lingual gestures. In *Proceedings of the 13th ICPHS*, ed. Kjell Elenius and Peter Branderud, 438–441. Stockholm.

Byrd, Dani. 1996a. Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24(2): 209–244.

Byrd, Dani. 1996b. A phase window framework for articulatory timing. *Phonology* 13(2): 139–169.

Byrd, Dani, and Susie Choi. 2010. At the juncture of prosody, phonology, and phonetics – The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In *Laboratory Phonology* 10, ed. Cécile Fougeron, Barbara Kuehnert, Mariapaola Imperio, and Nathalie Vallée, 31–60. Berlin: De Gruyter.

Byrd, Dani, and Cheng Cheng Tan. 1996. Saying consonant clusters quickly. *Journal of Phonetics* 24(2): 263–282.

Byrd, Dani, and Elliot Saltzman. 1998. Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics* 26(2): 173–199.

Cardoso, Walcir. 2008. The development of sC onset clusters in interlanguage: Markedness vs. frequency effects. In *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference*, ed. Roumyana Slabakova, Jason Rothman, Paula Kempchinsky, and Elena Gavruseva, 15–29. Somerville, MA: Cascadilla Proceedings Project.

Carlisle, Robert. 2006. The sonority cycle and the acquisition of complex onsets. In *English with a latin beat: Studies in Portuguese/Spanish-English interphonology*, ed. Barbara Baptista and Michael Watkins, 105–137. Amsterdam: John Benjamins.

Catford, John. C. 1977. *Fundamental problems in phonetics*. Bloomington: Indiana University Press.

Cebrian, Juli. 2000. Transferability and productivity of L1 rules in Catalan–English interlanguage. *Studies in Second Language Acquisition* 22(1): 1–26.

Chen, Hsueh-Chu, and Raung-Fu Chung. 2008. Interlanguage analysis of phonetic timing patterns by Taiwanese learners. *Concentric: Studies in Linguistics* 34: 79–108.

Chen, Yang, and Michael P. Robb. 2004. A study of speaking rate in Mandarin speakers of American English. *Asia Pacific Journal of Speech, Language and Hearing* 9(3): 173–188.

Cheng, Robert. 1966. Mandarin phonological structure. *Journal of Linguistics* 2(2): 135–158.

Chitoran, Ioana, Louis Goldstein, and Dani Byrd. 2002. Gestural overlap and recoverability: Articulatory evidence from Georgian. In *Laboratory Phonology* 7, ed. Carlos Gussenhoven and Natasha Warner, 419–448. Berlin: Mouton de Gruyter.

Clements, George N. 1985. The geometry of phonological features. *Phonology* 2: 225–252.

Colantoni, Laura, and Jeffrey Steele. 2008. Integrating articulatory constraints into models of second language phonological acquisition. *Applied Psycholinguistics* 29(3): 489–534.

Crystal, David. 1976. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.

Cummins, Fred, and Robert Port. 1998. Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26(2): 145–171.

Davidson, Lisa. 2003. Articulatory and perceptual influences on the production of non-native consonant clusters. In *Proceedings of the 15th ICPHS*, ed. Maria-Josep Solé, Daniel Recasens, and Joaquín Romero, 3–9. Barcelona.

Davidson, Lisa. 2006. Phonotactics and articulatory coordination interact in phonology: Evidence from nonnative production. *Cognitive Science* 30(5): 837–862.

Davidson, Lisa. 2010. Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics* 38(2): 272–288.

Davidson, Lisa. 2011. Characteristics of stop releases in American English spontaneous speech. *Speech Communication* 53(8): 1042–1058.

Davidson, Lisa, Peter Jusczyk, and Paul Smolensky. 2004. The initial and final states: theoretical implications and experimental explorations of Richness of the Base. In *Fixing Priorities: Constraints in Phonological Acquisition*, ed. René Kager, Joe Pater, and Wim Zonneveld, 321–368. Cambridge: Cambridge University Press.

Davidson, Lisa, and Kevin Roon. 2008. Durational correlates for differentiating consonant sequences in Russian. *Journal of the International Phonetic Association* 38(2): 137–165.

Davidson, Lisa, and Maureen Stone. 2003. Epenthesis versus gestural mistiming in consonant cluster production: an ultrasound study. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, ed. Gina Garding and Mimu Tsujimura, 165–178. Somerville, MA: Cascadilla Press.

Dixit, R. Prakash, and James. E. Flege. 1991. Vowel context, rate and loudness effects of linguopalatal contact patterns in Hindi retroflex /ʈ/. *Journal of Phonetics* 19: 213–229.

Eckman, Fred, and Gregory Iverson. 1993. Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research* 9(3): 234–252.

Ellis, Lucy, and William J. Hardcastle. 1999. An instrumental study of alveolar to velar assimilation in fast and careful speech. In *Proceedings of the 14th ICPHS*, ed. John Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Graville, and Ashlee Bailey, 2425–2428. San Francisco.

Elson, Ben. 1947. Sierra Popoluca syllable structure. *International Journal of American Linguistics* 13(1): 13–17.

Ernestus, Mirjam, and R. Harald Baayen. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79(1): 5–38.

Fleischhacker, Heidi Anne. 2005. Similarity in phonology: Evidence from reduplication and loan adaptation. Doctoral dissertation, University of California, Los Angeles.

Fosler-Lussier, Eric, and Nelson Morgan. 1999. Effects of speaking rate and word frequency on pronunciations in conversational speech. *Speech Communication* 29(2–4): 137–158.

Fougeron, Cécile, and Patricia Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101(6): 3728–3740.

Gafos, Adamantios. 2002. A Grammar of Gestural Coordination. *Natural Language and Linguistic Theory* 20(2): 269–337.

Ghosh, Prasanta Kumar, and Shrikanth S. Narayanan. 2009. Closure duration analysis of incomplete stop consonants due to stop-stop interaction. *Journal of the Acoustical Society of America* 126(1): EL1–EL7.

Goldstein, Louis, Dani Byrd, and Elliot Saltzman. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. In *Action to language via the mirror neuron system*, ed. Michael Arbib, 215–249. Cambridge: Cambridge University Press.

Goldstein, Louis, Marianne Pouplier, Larissa Chen, Elliot Saltzman, and Dani Byrd. 2007. Dynamic action units slip in speech production errors. *Cognition* 103(3): 386–412.

Grabe, Esther, and Es Ling Low. 2002. Durational variability in speech and the rhythm class hypothesis. In *Laboratory phonology* 7, ed. Carlos Gussenhoven and Natasha Warner, 515–546. Berlin: Walter de Gruyter.

Gregory, Michelle, William Raymond, Allen Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production, In *Proceedings of the 35th Regional Meeting of the Chicago Linguistic Society*, ed. Sabrina J. Billings, John P. Boyle, and Aaron M. Griffith, 151–166. Chicago: Chicago Linguistic Society.

Hansen, Jette G. 2001. Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics* 22(3): 338–365.

Hardcastle, William, and Peter Roach. 1979. An instrumental investigation of coarticulation in stop consonant sequences. In *Current issues in the phonetic sciences 1*, ed. Harry Hollien and Patricia Hollien, 531–540. Amsterdam: John Benjamins.

Harris, John. 1994. *English sound structure*. Oxford: Blackwell.

Henderson, Janette, and Bruno Repp. 1982. Is a stop consonant released when followed by another stop consonant? *Phonetica* 39: 71–82.

De Jong, Nivja, and Ton Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41(2): 385–390.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency and the emergence of linguistic structure*, ed. Joan Bybee and Paul Hopper, 229–254. Amsterdam: John Benjamins.

Kessinger, Rachel, and Sheila E. Blumstein. 1997. Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics* 25(2): 143–168.

Kita, Sataro, and Asli Özyürek. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48(1): 16–32.

Kochetov, Alexei, Marianne Pouplier, and Minjun Son. 2007. Cross-language differences in overlap and assimilation patterns in Korean and Russian. In *Proceedings of the 16th ICPhS*, ed. Jürgen Trouvain and William Barry, 1361–1364. Saarbrücken.

Kuehn, David, and Kenneth L. Moll. 1976. A cineradiographic study of VC and CV articula-
    tory velocities. *Journal of Phonetics* 4: 303–320.

Ladefoged, Peter. 2001. *A course in phonetics*. 4th ed. San Diego: Harcourt Brace.

Likert, Rensis. 1932. A technique for the measurement of attitudes. *Archives of Psychology*
    140: 1–55.

Lin, Hua. 2001. *A grammar of Mandarin Chinese*. München: Lincom Europa.

Lin, Hua, and Qian Wang. 2007. Mandarin rhythm: An acoustic study. *Journal of Chinese
    Linguistics and Computing* 17: 127–140.

Lindblom, Björn. 1983. Economy of speech gestures. In *The production of speech*, ed. Peter F.
    MacNeilage, 217–245. New York: Springer.

Lubker, James, and Thomas Gay. 1982. Anticipatory labial coarticulation: Experimental,
    biological, and linguistic variables. *Journal of the Acoustical Society of America* 71(2):
    437–448.

Major, Roy C. 1987. Phonological similarity, markedness, and rate of L2 acquisition. *Studies
    in Second Language Acquisition* 9(1): 63–82.

Major, Roy C., and Michael C. Faudree. 1996. Markedness universals and the acquisition of
    voicing contrasts by Korean speakers of English. *Studies in Second Language
    Acquisition* 18(1): 69–90.

Miller, Joanne, and Thomas Baer. 1983. Some effects of speaking rate on the production of /b/
    and /w/. *Journal of the Acoustical Society of America* 73(5):1751–1755.

Moisik, Scott, and John Esling. 2011. The 'whole larynx' approach to laryngeal features. In
    Proceedings of the 17th ICPhS, ed. Wai-Sum Lee and Eric Zee, 1406–1409. Hong Kong.

Mok, Peggy P. K., and Volker Dellwo. 2008. Comparing native and non-native speech
    rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English.
    In *Proceedings of the Fourth International Conference on Speech Prosody*, ed. Plínio
    A. Barboso, Sandra Madureira, and Cesar Reis, 423–426. Campinas: Editora RG/CNPq.

Munhall, Kevin, and Anders Löfqvist. 1992. Gestural aggregation in speech-laryngeal ges-
    tures. *Journal of Phonetics* 20: 111–126.

Nam, Hosung, and Elliot Saltzman. 2003. A competitive, coupled oscillator model of syllable
    structure. In *Proceedings of the 15th ICPHS*, ed. Maria-Josep Solé, Daniel Recasens, and
    Joaquín Romero, 2253–2256. Barcelona.

Nolan, Francis. 1992. The descriptive role of segments: evidence from assimilation. In *Papers
    in laboratory phonology II: Gesture, segment, prosody*, ed. Gerard Docherty and Robert
    Ladd, 261–280. Cambridge: Cambridge University Press.

Öhman, Sven. 1967. Numerical model of coarticulation. *Journal of the Acoustical Society of
    America* 41(2): 310–320.

Ostry, David, and Kevin G. Munhall. 1985. Control of rate and duration of speech movements.
    Journal of the Acoustical Society of America 77(2): 640–648.

Pater, Joe, and Anne-Michelle Tessier. 2006. L1 phonotactic knowledge and the L2 acquisition
    of alternations. In *Inquiries in linguistic development: In honor of Lydia White*, ed.
    Roumyana Slabakova, Silvina Montrul, and Philippe Prévost, 115–131. Amsterdam:
    John Benjamins.

Peng, Shu-Hui. 1996. Phonetic implementation and perception of place coarticulation and tone
    sandhi. Doctoral dissertation, Ohio State University.

Pind, Jörgen. 1995. Constancy and normalization in the perception of voice offset time as a cue
    for preaspiration. *Acta Psychologica*, 89(1): 53–81.

Pouplier, Marianne. 2003. Units of phonological encoding: Empirical evidence. Doctoral
    Dissertation, Yale University.

Recasens, Daniel, Jordi Fontdevila, Maria Dolors Pallarès, and Antoni Solanas. 1993. An electropalatographic study of stop consonant clusters. *Speech Communication* 12(4): 335–355.

Rubach, Jerzy. 1984. Rule typology and phonological interference. In *Theoretical issues in contrastive phonology*, ed. Stig Eliasson, 37–50. Heidelberg: J. Groos.

Sato, Charlene. 1987. Phonological processes in second language acquisition: another look at interlanguage syllable structure. In *Interlanguage phonology: The acquisition of a second language sound system*, ed, Georgette Ioup and Steven H. Weinberger, 248–260. Cambridge, MA: Newbury House.

Shaiman, Susan, Scott G. Adams, and Mikael Kimelman. 1995. Timing relationships of the upper lip and jawacross changes in speaking rate. *Journal of Phonetics* 23: 119–128.

Shinn, Phil. 1984. A cross-language investigation of the stop, affricate, and fricative manners of articulation. Doctoral Dissertation, Brown University.

Shockey, Linda. 2008. *Sound patterns of spoken English*. Malden, MA: Blackwell Publications.

Solé, Maria-Josep. 1997. Timing patterns in the production of a foreign language. In *Views on the Acquisition and Use of a Second Language: Proceedings of the European Second Language Association 7th annual Conference*, ed. Lourdes Díaz and Carmen Pérez, 539–551. Barcelona: Universitat Pompeu Fabra.

Sperbeck, Meiko. 2010. The acquisition of consonant clusters by Japanese learners of English: Interactions of speech perception and production. Doctoral dissertation. City University of New York.

Sproat, Richard, and Osamu Fujimura. 1993. Allophonic variation in English /1/and its implications for phonetic implementation. *Journal of Phonetics* 21: 291–311.

Stevens, Kenneth N. 1980. Acoustic correlates of some phonetic categories. *Journal of the Acoustical Society of America* 68(3): 836–842.

Tajima, Keiichi, Robert Port, and Jonathan Dalby. 1997. Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics* 25(1): 1–24.

Theodore, Rachel M., Joanne Miller, and David DeSteno. 2007. The effect of speaking rate on voice-onset-time is talker-specific. In *Proceedings of the 16[th]ICPHs*, ed. Jürgen Trouvain and William Barry, 473–476. Saarbrücken.

Tjaden, Kris, and Gary Weismer. 1998. Speaking-rate-induced variability in F2 trajectories. *Journal of Speech, Language, and Hearing Research* 41(5): 976–989.

Tsao, Ying-Chiao, Gary Weismer, and Kamran Iqbal. 2006. Interspeaker variation in habitual speaking rate: Additional evidence. *Journal of Speech, Language, and Hearing Research* 49(5): 1156–1164.

Weinberger, Steven. 1987. The influence of linguistic context on syllable simplification. In *Interlanguage phonology: The acquisition of a second language sound system*, ed. Georgette Ioup and Steven Weinberger, 401–417. Cambridge, MA: Newbury House.

Weinberger, Steven. 1994. Functional and phonetic constraints on second language phonology. In *First and second language phonology*, ed. Mehmet Yavaş, 283–302. San Diego, CA: Singular.

Wilson, Colin. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30(5): 945–982.

Wilson, Colin, Lisa Davidson, and Sean Martin. 2014. Effects of acoustic–phonetic detail on cross-language speech production. *Journal of Memory and Language* 77: 1–24.

Yanagawa, Mariko. 2006. Articulatory timing in first and second language: a cross-linguistic study. Doctoral dissertation, Yale University.

Yip, Jonathan. 2013. Phonetic effects on the timing of gestural coordination in Modern Greek consonant clusters. Doctoral dissertation, University of Michigan.

Zhang, Jie, and Yuwen Lai. 2010. Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology* 27(1): 153–201.

Zhang, Jie, Yuwen Lai, and Craig Sailor. 2011. Modeling Taiwanese speakers' knowledge of tone sandhi in reduplication. *Lingua* 121(2): 181–206.

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

Zsiga, Elizabeth C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*, ed. Bruce Connell and Amalia Arvaniti, 282–302. Cambridge: Cambridge University Press.

Zsiga, Elizabeth C. 2000. Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28: 69–102.

Zsiga, Elizabeth C. 2003. Articulatory timing in a second language. *Studies in Second Language Acquisition* 25(3): 399–432.

Zsiga, Elizabeth C. 2011. External sandhi in a second language: The phonetics and phonology of obstruent nasalization in Korean-accented English. *Language* 87: 289–345.

Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language* 83: 277–316.

APPENDIX: TESTING MATERIAL

Single Coda $C1_b$#V	Single Onset V#$C2_b$
keep art, pep art, job art, rib art, hot art, shut art, look art, rock art, food art, red art, leg art, big art	saw pain, saw pen, saw balm, saw boy, saw tie, saw take, saw care, saw car, saw day, saw down, saw good, saw game

List 1: 24 Items of /C1#V/ and /V#C2/

		Nonsense C1#C2	C1#C2	C1#C2	Nonsense C1#C2
homorganic clusters	p#p	foop pok	soup pot	keep pace	peep pate
	p#b	mip bong	lip balm	deep bite	beep bipe
	t#t	shot tud	hot tub	get tip	det tik
	t#d	zet dep	vet debt	beat down	deat doun
	k#k	dak kit	back kick	take care	pake kare
	k#g	bek gaep	deck gate	look good	mook goob
	b#p	lib peim	rib pain	grab pen	brab pem
	b#b	zob banp	job bank	rub back	mub bap
	d#t	fade tenk	shade tent	need time	meed tine
	d#d	nud dei	mud day	bad day	gad dei
	g#k	grag kar	drag car	smug king	swug kim
	g#g	brug gam	drug gang	big game	dig gane
front-to-back clusters	p#t	kep tot	pep talk	skip test	stip tesp
	p#d	rap dant	lap dance	deep down	beep doun
	p#k	mip kare	lip care	keep calm	teep kalm
	p#g	pape gum	tape gun	top gear	pop geal
	t#g	biet gaib	diet guide	great guy	breat gai
	t#k	doot kamk	boot camp	fat cat	shat kap
	b#t	rab tesp	lab test	glib talk	blib tot
	b#d	prab dis	crab dish	rub down	lub doun
	b#k	rab coak	lab coat	grab cash	brab kaf
	b#g	plub girm	club girl	fab gift	sab gisp
	d#g	fead gane	head game	good game	bood gane
	d#k	sood kork	food court	bad kid	gad kig
back-to-front clusters	t#p	fuit panf	suit pants	bright pink	grait pinp
	t#b	srat boi	frat boy	get back	det bap
	k#p	feek paim	cheek pain	look pale	rook peil
	k#t	mek tai	neck tie	black tea	glack tee
	k#b	lock banb	rock band	quick boost	kwik booft
	k#d	bok dape	dock date	kick down	tik doun
	d#p	fead paim	head pain	red pen	led pem
	d#b	ded bud	bed bug	need beer	meed beel
	g#t	gog tael	dog tail	big time	dig tine
	g#d	dag dear	bag deal	big deal	kig deer
	g#p	glog peiz	blog page	hug pug	sug puz
	g#b	reg bome	leg bone	big boy	dig boi

List 2: 144 items of C1#C2