ISCA Archive
http://www.isca-speech.org/archive

International Symposium on Chinese Spoken Language Processing (ISCSLP 2004)
Hong Kong
December 15-18, 2004

EFFECTS OF PHONEMIC VS ALLOPHONIC DENSITY AND STRESS ON VOWEL-TO-VOWEL COARTICULATION IN CANTONESE AND BEIJING MANDARIN

Pik Ki Peggy Mok and Sarah Hawkins

Department of Linguistics, University of Cambridge

ABSTRACT

Effects of phonemic vs. allophonic vowel distribution, stress and direction of coarticulation on V-to-V coarticulation were examined in Cantonese and Beijing Mandarin (BM). Cantonese has more vowel phonemes but BM has more allophones. Cantonese should show less V-to-V coarticulation than BM if phonemic contrast determines degree of V-to-V coarticulation. The vowels used were /i a u/ in /pVpVpV/ structures. Phonemic vowel space density did not influence V-to-V coarticulation differentially in Cantonese and BM. Effects of stress and direction were not consistent. Generally, there was more carryover coarticulation, and more coarticulation on unstressed vowels, but exceptions were common. No one factor appears to determine patterns of V-to-V coarticulation in different languages. Other potential phonological influences are discussed.

1. INTRODUCTION

Vowel-to-Vowel (V-to-V) coarticulation refers to the coarticulatory effects extending from one vowel to another across intervening consonant(s). Many factors affect V-to-V coarticulation in different languages. Manuel suggested that linguistic contrasts of vowel phonemes could determine the degree of coarticulation exhibited [1] [2]. For example, more coarticulation might occur in a sparse vowel space because big shifts in vowel quality should not cause perceptual problems. But in a crowded vowel space, massive coarticulation would blur or even obliterate contrast between different vowels. However, her proposal is primarily based on vowel phonemes. Allophonic versus phonemic variation appears not to have been considered.

Other factors affecting V-to-V coarticulation include stress and relative degree of anticipatory vs. carryover coarticulation (henceforth, direction of coarticulation). Stressed vowels are less likely to coarticulate with flanking vowels than unstressed ones in English, e.g. [3]. Anticipatory coarticulation is believed to involve articulatory preplanning while carryover coarticulation may be due to inertia, e.g. [4]. Both anticipatory and carryover coarticulation are present in the same language, but languages tend to show greater coarticulation in one direction e.g. [1][2][5].

This study assessed the relative influence of phonemic vs. allophonic contrast in Cantonese and BM V-to-V coarticulation. Effects of stress and direction of coarticulation were also investigated.

2. CANTONESE AND BEIJING MANDARIN VOWEL SYSTEMS

The number of phonemic vowels in Cantonese and especially for BM is controversial. The reason for the controversy is that syllable structure and syntagmatic relationship between different segments are important in Chinese. The main controversy in Cantonese phonemic vowels lies in the distinction of vowel length. Phonetically, there are 7 long [iː yː ɛː œː aː ɔː uː], 7 half-long [i y ɛ œ a ɔ u] and 4 short [ɪ ɐ ɵ ʊ] distinctive vowels in Cantonese, although the formant frequencies for [ɪ] and [ɛː], [ʊ] and [ɔː] and [ɐ] and [aː] are quite similar [6]. Most analyses, however, treat length as a redundant allophonic feature in Cantonese, because for except [a] and [ɐ], all short and half-long vowels are in complementary distribution with their long counterparts. Thus, Cantonese is usually said to have 8 phonemic vowels: /i y u ɛ œ ɔ a ɐ/.

The analysis of the vowel system of BM is yet more controversial. Most analyses agree that /i y a u/ are all phonemic, albeit with many distinct allophones. But there is little agreement about the mid-vowels, where many qualities are found: [e ɛ ə ɤ œ ɔ o]. The quality used depends on the overall structure of the syllable. For example, [ɤ] only appears in an open syllable. All other qualities occur either before or after glides, e.g. [t^jɛn] and [t^wɔ]. Although phonetically very distinct, the distribution of the mid-vowels is clearly complementary. The most convincing proposal is thus that there is an unspecified mid-vowel phoneme /E/ in BM [7]. This analysis results in only 5 vowel phonemes in BM: /i y E a u/.

Stress in Chinese is primarily cued by longer duration and secondarily by more extended pitch range and a complete pitch contour. Amplitude is the least important cue. Unstressed vowel qualities are not systematically reduced, unlike English. Most studies showing larger coarticulatory effects on unstressed than stressed vowels were based on English and other non-tonal languages in which pitch is an important indicator of stress. Stress may not have the same effect in Chinese.

3. METHOD

Eight native speakers (four male, four female) were used for each language. The Cantonese speakers (all in their 20s) were born and grew up in Hong Kong. The BM speakers (all in their 20s or 30s) were from Beijing or areas around Beijing and spoke BM with a northern accent.

The three vowels used were [i], [a] and [u] because they are the only common vowels in both dialects that can appear after [p] in open syllables ([pu] is a marginal syllable in Cantonese). Bilabial [p] was chosen because it phonetically allows most V-to-V coarticulation. The three syllables, [pi], [pa] [pu], were combined into nonsense trisyllables (high level tone), with stress in the middle syllable, e.g. [paˈpipi]. Both [pa] and [pi] were designated the target syllable in different analyses: [pa] for unstressed [a] with anticipatory coarticulation from context [i]; [ˈpi] for stressed [i] with carryover coarticulation from context [a]. They were embedded in short carrier phrases which were phonetically similar in both languages, presented to the speakers in Chinese characters. Stressed syllables were underlined.

Subjects read ten randomized lists of the phrases at a comfortable speed. The speech was recorded in a sound-treated room directly into a Silicon Graphics Indigo computer using *Xwaves* (sampling frequency 16 kHz). The frequencies of the first two formants (F1 and F2) were measured from 18 pole 25ms autocorrelation LPC spectra with a Hanning window. Wide band spectrogram and DFT spectra were used when necessary. Measurements were made at two places in each target vowel: vowel edge (either onset for carryover coarticulation or offset for anticipatory coarticulation) and midpoint. The vowel-edge spectra were centred 12.5 ms inwards from the beginning or end of periodicity in the waveform.

3.1. Data normalization

In order to minimize individual differences, each measurement was expressed as a proportion from the mean (F1 or F2) of a given vowel across contexts and stress for that particular speaker. For example, mean onset F1 of target /a/ in all /u/ context ([puˈpapa], [paˈpapu], [paˈpupu] and [puˈpupa]) was 732 Hz for Cantonese Speaker 1. His mean F1 frequency in onset target /a/ across all contexts and stress was 752 Hz. The normalized measure was: 732/752-1. Degree of coarticulation is thus expressed as a deviation from zero. A positive value means a formant was higher than the mean due to the context vowels; negative values mean it was lower. The greater the absolute value, the more the coarticulation exhibited.

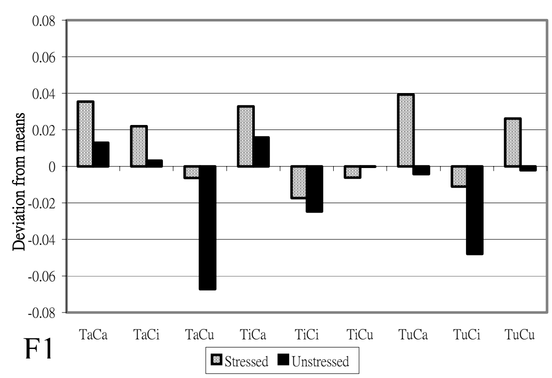
4. RESULTS

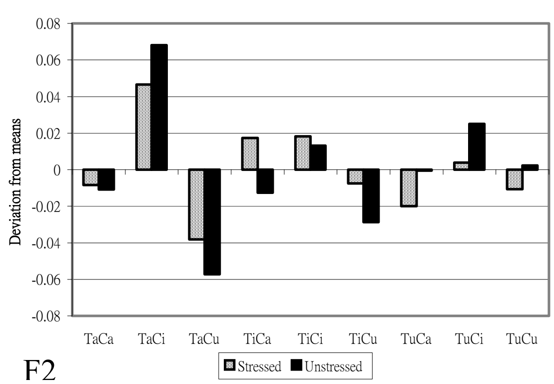
Four (two formants and two temporal locations) 5-way repeated measures ANOVAs with factors Language (Cantonese, BM), Direction (anticipatory, carryover), Stress (stressed, unstressed), Target (/a i u/) and Context (/a i u/) were conducted on the mean normalized data.

4.1. Language

No Language main effect was significant. The only significant interaction involving Language was Language x Direction at the vowel edges for F2 [$F(1,14) = 4.801$, $p<0.05$] but post hoc tests showed no Language difference. It is concluded that, despite their differences in phonemic and phonetic vowel inventory, Cantonese and BM behave similarly in degree of V-to-V coarticulation.

Fig. 1 Coarticulatory effects of Stress at vowel edges on F1 (upper panel) and F2 (lower panel), for different Target and Context combinations pooled across Language and Direction. "T": target vowel; "C": context vowel.

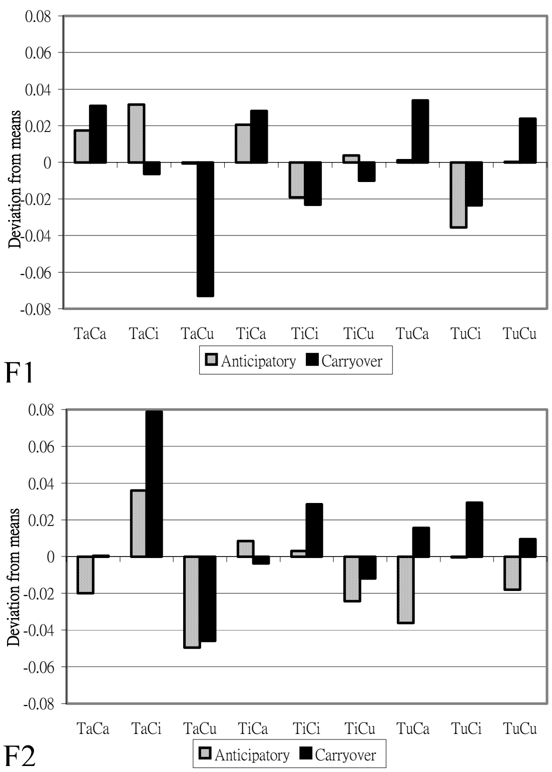




4.2. Stress

The effect of Stress is best understood in its interaction with other factors. Fig. 1 shows the Stress x Target x Context interactions at vowel edges: F1 [F(4, 56) = 3.953, p<0.01]; F2 [F(4, 56) = 2.926, p<0.05]. Conservative post-hoc tests (paired-sample 2-tailed T-tests with modified Bonferroni adjustments) showed unstressed vowels coarticulated more than stressed ones for F1 of TaCu [t(15) = 4.438, p = 0.000] and TuCi [t(15) = 3.865, p = 0.002] and F2 of TiCu [t(15) = 4.639, p = 0.000]. There was a trend for F2 of TaCu and TuCi to show the same pattern. In contrast, stressed vowels coarticulated more than unstressed ones for F1 of TaCa [t(15) = 3.073, p = 0.008], TuCa [t(15) = 3.375, p = 0.005] and TuCu [t(15) = 3.296, p = 0.004], with a non-significant trend in the same direction for and F2 of TiCa. Thus, Stress does not have a uniform effect on the degree of V-to-V coarticulation in Cantonese and BM.

Fig. 2 Carryover vs anticipatory coarticulation for different Target and Context combinations at vowel edges for F1 (upper panel) and F2 (lower panel) pooled across Language and Stress. "T": target vowel; "C": context vowel.



4.3. Anticipatory vs Carryover coarticulation

The main effect of Direction was only significant for the vowel edges of F2 [F(1,14) = 24.936, p<0.0001], due mainly to slightly higher F2 frequencies of vowel onset (carryover coarticulation) compared with vowel offset (anticipatory coarticulation). The effect of Direction is best understood in its interactions with other factors. Fig. 2 shows the Direction x Target x Context interactions at vowel edges for F1 [F(4, 56) = 3.658, p<0.05] and F2 [F(4, 56) = 3.752, p<0.01]. Carryover coarticulation was stronger than anticipatory coarticulation for F1 of TaCu [t(15) = 3.916, p = 0.001] and F2 of TiCi [t(15) = 3.758, p = 0.002]. There was also a trend for F1 of TuCa and TuCu to show the same pattern. However, F2 of TaCa [t(15) = 4.770, p = 0.000], TuCa [t(15) = 5.043, p = 0.000] and TuCu [t(15) = 4.369, p = 0.001] instead showed more anticipatory coarticulation. Overall at vowel edges, there was generally more carryover coarticulation than anticipatory coarticulation in F1 and more anticipatory coarticulation in F2. This conclusion is tentative, however.

4.3. Direction of coarticulation x Stress x Context

Both Direction of coarticulation and Stress interacted with Context for F2 at both vowel edges and midpoint (p<0.005 or better). Post hoc tests showed that carryover coarticulation was greater than anticipatory coarticulation for unstressed context /i/ at vowel edge [t(15) = -3.686, p = 0.002]. Unstressed context /i/ at vowel midpoint also approached significance for more carryover coarticulation. In contrast, there was more anticipatory coarticulation at vowel edges for unstressed context /a/ [t(15) = -3.323, p = 0.005] and stressed context /u/ [t(15) = -2.990, p = 0.009].

For Stress, unstressed vowels showed more coarticulation than stressed vowels for carryover context /i/ at midpoint [t(15) = -3.904, p = 0.001]. Carryover context /i/ and carryover context /u/ at vowel onset also approached significance for the same pattern. In brief, for the Stress x Direction x Context interactions, unstressed vowels had slightly more coarticulation than stressed vowels. Unstressed context /i/ allowed more carryover coarticulation while context /a/ and context /u/ have more anticipatory coarticulation.

5. DISCUSSION

The main question of the study was to see whether Cantonese and BM differ in degree of V-to-V coarticulation with respect to their different vowel inventory size, as proposed by Manuel. Results show that Cantonese and BM did not differ from each other either overall or in interaction with other factors, despite their difference in the number of phonemic vowels. In fact, Shona, an African language with just 5 vowels, also did not coarticulate more than English [5]. Manuel's proposal may work well for languages in which phonemic analysis can account for the entire vowel system, as long as the number of phonemes bears a simple relationship with the number of allophones. But this means that in essence, the crucial factor is the number of phonetic allophones in natural speech, not the abstract phonemes.

Studies e.g. [8] also show that vowels in languages with small vowel inventories (e.g. Modern Greek and Spanish, both with a 5-vowel system) do not vary more than vowels in languages with much larger inventories like English and German. These data fail to support the strong assumption of the influence of phonemic contrast based on inventory size, namely that vowels in smaller inventories *can* and *do* vary more freely than larger inventories. Manuel [2] herself pointed out that languages probably have some tendency to constrain coarticulation in order to maintain contrast. We would expect to find counter-examples for the predictions based on number of phonemic contrasts. It thus seems that the crowdedness of the F1-F2 space caused by distribution of allophones may influence patterns of V-to-V coarticulation and that other factors may also be at play.

Another reason why phonemic contrast based on inventory size cannot satisfactorily account for the language-specific patterns of V-to-V coarticulation is that the number of vowel phonemes does not necessarily indicate the number of vowel qualities captured as F1-F2 space in a language. Many languages have different series of vowels, e.g. long and short, oral and nasalized, with similar vowel qualities captured as F1-F2 space in each series. There can be many phonemic vowels but only few vowel qualities involved. Other languages distinguish vowel length, but include quality distinctions for at least some of the vowel pairs. Such things complicate predictions based on phonemic contrast by adding other dimensions which cannot be accounted for by the 2-dimensional vowel space. Table 1 schematizes the relationship between the number of phonemes and the crowdedness of F1-F2 space.

Table 1 Relationship between number of phonemes and crowdedness of F1-F2 space, with examples of languages

F1-F2 space	Relative number of phonemes	
	Few	Many
Crowded	BM	English
Sparse	Spanish	Czech

There are many other possible influences on the degree of coarticulation allowed. One is stress. Unstressed syllables tend to be more susceptible to coarticulation than stressed ones in studies of English and other non-tonal languages e.g. [3]. By analogy, stress-based languages, which use lexical and metrical stress to convey differences in meaning, might be expected to show more V-to-V coarticulation, if only in their unstressed syllables, than languages that accord syllables relatively similar stress. However, even this point is challenging to investigate, because the principal acoustic correlates of stress do not always involve vowel quality, e.g. in Chinese as discussed in Section 2. Likewise, stress in Shona is cued by duration and amplitude and generally not by other acoustic attributes of prominence, and consistent with this reasoning, unstressed vowels in Shona do not show more coarticulation than stressed ones [5].

It seems necessary to take a more system-sensitive approach for cross-linguistic study of influences on V-to-V coarticulation than an approach simply based on phonemic, or even allophonic contrasts. Factors that could be taken into consideration are: stress (as long as its acoustic correlates include changes in vowel quality), vowel harmony and syllable structure. Languages that place strong constraints on the tongue at syllable edges, due to complex consonant clusters, might allow more variation in vowel quality than languages, like Chinese, that have relatively simple syllable structures. Further research is needed to investigate these ideas.

ACKNOWLEDGEMENTS

This research is supported by the Sir Edward Youde Memorial Fellowships for Overseas Studies in HK, the Overseas Research Studentship in the UK and the Cambridge Overseas Trust.

6. REFERENCES

[1] S. Manuel, "The role of contrast in limiting vowel-to-vowel coarticulation in different languages," *Journal of the Acoustical Society of America* 88, pp.1286-1298. 1990

[2] S. Manuel, "Cross-language studies: relating language-particular coarticulation patterns to other language-particular facts," In W. J. Hardcastle and N. Hewlett (Eds), *Coarticulation: Theory, Data and Techniques*. Cambridge: Cambridge University Press, pp. 179-198. 1999

[3] C. A. Fowler, "Production and perception of coarticulation among stressed and unstressed vowels," *Journal of Speech and Hearing Research* 46: 127-139. 1981

[4] D. Recasens, "Vowel-to-vowel coarticulation in Catalan VCV sequences," *Journal of the Acoustical Society of America* 76(6): 1624-1635. 1984

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

[6] E. Zee, "Frequency analysis of the vowels in Cantonese from 50 male and 50 female speakers," *Proceedings of The 15th ICPhS,* Barcelona, pp. 1117-1120. 2003

[7] R. Wiese, "Underspecification and the description of Chinese vowels," In J. L. Wang (Eds), *Studies in Chinese Phonology*, Berlin: Mouton de Gruyter, pp.219-249. 1997

[8] A. R. Bradlow, "A comparative acoustic study of English and Spanish vowels," *Journal of the Acoustical Society of America* 97, pp.1916-1924. 1995