Title	発話時における調音運動に基づいた調音結合の分析
Author(s)	鈴木，丈晴
Citation	
Issue Date	2005-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1939
Rights	
Description	Supervisor:党　建武，情報科学研究科，修士

Analysis of coarticulation based on articulatory movements at the time of utterance

Takeharu Suzuki (310054)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 10, 2005

Keywords: coarticulation,articulatory movements,speech production,physiological articulatory model.

1 Introduction

Coarticulation during speech production takes place in both the physiological level concerned with articulator's properties and the planning stage for generating motor commands. This study focuses on investigation and modeling of coarticulation at planning stage of speech production to implement human mechanism in controlling a physiological model. A "carrier model" was proposed to describe the mechanisms of coarticulation, in which the vocalic movement is considered to be the primary component as a "carrier wave" and the consonantal movement as a "modulation wave". We attempt to quantify the parameters of the model using articulatory observations obtained from the electromagnetic articulographic system.The articulatory point of the central vowel was dragged towards the articulatory point of the bilateral vowel. Parameters for the proposed mode were evaluated using the articulatory data.

2 Separation of articulatory movements about consonants and vowels

Carrier model was verified by separating the movement ingredient about vowels and consonants and reconstructing each articulatory movements. Moreover, the interaction of articulatory movements about vowels and consonants is examined. In this research, much phoneme environments was taken into consideration by a creating average phoneme environment by the technique of a frequency addition average, and reconstructing the waveform of articulatory movement.

The waveform of articulatory movements in from frequency domain to time domain was reconstructed by giving a double complex Fourier series from the spectrum which it is led by frequency addition average.

$$\mathrm{T}_k y(t) = a_0 + \sum_{i=1}^{N/2} a_i cos(2\pi f_0 it) + b_i sin(2\pi f_0 it) \tag{1}$$

$$a_i = 2real(C_i)/N \quad b_i = -2imag(C_i)/N \tag{2}$$

$$a_0 = C_0/N \quad k = 1, 2, 3 \quad i = 1, , , , , , N/2 - 1 \tag{3}$$

C_nThe value calculated by the addition-average of frequency,NThe number of points of Fourier transform,f_0Frequency resolution.Separation of articulatory movements about vowels and a consonants was aimed at using the waveform of typical articulatory movements. The waveform of articulatory movements relatively seen from the reconstructed speed of an articulatory waveform and the central point between observing points was shown. They can be considered that articulatory movements of prompt consonant is what is depended on superimposing after that a quiet vowel's articulatiory movements.

3 carrier model

This model adopts spatial targets in the control strategy. Modeling of coarticulation is focused on the coarticulation taking place in the target planning stage while the coarticulation in the physiological level is expected to be realized by the physiological properties in a physiological articulatory model. To plan articulation based on the concept of a "carrier" movement superimposes "modulation" movement, it requires us to separate a given utterance into vowel and consonant sequences. For a given utterance, therefore, the vowels and consonants are separated into two phoneme sequences as the following. To construct the "carrier wave", articulatory movement

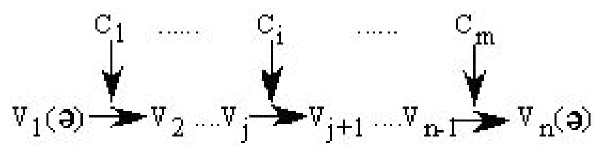

Figure 1: carrier model

is considered as a continuous movement from one vowel to another. The resultant target of consonant C_i is dependent on a "tug-of-war" of the neighboring vocalic targets, and thus a virtual vocalic target G_i should be generated in the position of C_i. This procedure is described by this Figure

$$G_i = \alpha V_j + \beta V_{j+1} \quad \alpha \leq \beta \tag{4}$$

Based on the look-ahead mechanism, the following target has stronger effects on the virtual target than the preceding one in the targetplanning stage. Therefore, the coefficient β should be larger than α. The second process accounts for the effect of the consonants on the vowels, where only the immediately adjacent phonemes are taken into account. Here, the crucial feature in the consonantal target is first constructed according to the typical feature C_i and virtual feature G_i, where no change happens in the indecisive features.This processing is carried out by the following formula.

$$C_i' = \frac{(d_{ci} C_i + G_i)}{d_{ci} + 1} \tag{5}$$

where d_{ci} is a weight coefficient based on the degree of the articulatory constraint of crucial C_i. Second processing in this step is to account for the effects of the following consonant on the preceding vowel via the look-ahead mechanism.

$$G'_i = \frac{\gamma d_{ci} \mathrm{C'}_i + \tau d_{vj} \mathrm{V}_j}{(\gamma d_{ci} + \tau d_{vj})} \quad \tau \approx \gamma \tag{6}$$

In this study, the values for both γ and τ are set as 0.5.We have to lead the degree of degree of the articulatory constraint for applying a model to a physiological model.

3.1 Distance using the distribution of an articulatory position which paid its attention to the vowel

To estimate the parameters for the carrier model, phoneme sequence of $\mathrm{V}_b\mathrm{C}\mathrm{V}_c\mathrm{C}\mathrm{V}_b$, where V_b is a vowel in both sides, V_c is the central vowel, and C is a consonant, are segmented out from the read sentences. The bilateral vowels move the central vowel from its average position. To evaluate the interaction between the adjacent vowels, the displacement of the central vowel apart from its average location is calculated and shown in Table. The value of each column indicates how large of the effect the central one

Table 1: the bilateral vowels move the central vowel from its average position

$\mathrm{V}_b \backslash \mathrm{V}_c$	/a/	/i/	/u/	/e/	/o/
/a/	0.152	0.058	0.106	0.124	0.107
/i/	0.288	0.059	0.236	0.281	0.389
/u/	0.238	0.122	0.196	0.180	0.184
/e/	0.085	0.130	0.117	0.029	0.101
/o/	0.226	0.016	0.143	0.092	0.201

is affected. The value of each row shows the effect of the bilateral vowel on the central vowel. The values in the rows correspond to the degree of articulatory constraint (DAC). The lager of the value, the stronger of the DAC. If using the average displacement of each row as a measure for the DAC, after normalizing,DAC is this Table.

Table 2: vowel's DAC

	/a/	/i/	/u/	/e/	/o/
DAC	1.2	2.7	2.0	1.0	1.5

3.2 Distance using the distribution of an articulatory position which paid its attention to the alveolar

To estimate the parameters for the carrier model, phoneme sequence of $V_bCV_cCV_b$, where C is a vowel in both sides, V_c is the central vowel, and C is a consonant, are segmented out from the read sentences.The bilateral consonants move the central vowel from its average position. To evaluate the interaction between the adjacent vowels, the displacement of the central vowel apart from its average location is calculated.But, since there were few samples – even [vowel] – one – receiving – the gum – the distance of a consonant was found.In order to treat equally the number of the distribution which determined the consonant of both ends, all the things of the group which determined the consonant of both ends were added, and the average value was equalized by dividing by the total number of samples.

- d_i:distance × samples of the selected phoneme
- N:samples of all alveolar's consonants
- n:number of groups of alveolar's consonants

$$D = \frac{1}{N}\sum_{i=1}^{n} d_i \tag{7}$$

When close to the value as compared with the distance in a vowel, it

Table 3: Distance between the average value calculated using the parameter near tongue tip

C \ V	/a/	/i/	/u/	/e/	/o/
C:alveolar's consonants	0.289	0.020	0.407	0.007	0.246

considered as the degree of the same restraint.Moreover, when value is a small value, the degree of articulatory constraint was set to 1.

3.3 Evaluation of a parameter

Articulatory targets for which it led by the carrier model was led to the distribution using the articulation position for which it led by the actual

Table 4: tongue tip the degree of articulatory constraint about alveolar's consonants

DAC \ V	/a/	/i/	/u/	/e/	/o/
DAC	2.7	1	2.7	1	2.7

physiological utterance model.In order to confirm the adjustment of this distribution for which it led, it investigated how much it would pile up with a distribution and would overlap.Moreover, analytic data differs the data when building a physiological utterance model, and this time. Therefore, I think that compensation of a grade is required.Then, the central point of a distribution was arranged this time.

$$\lambda = \overline{\mathrm{T}_1} - \mu \tag{8}$$

- μ:Average as a result of simulation
- $\overline{\mathrm{T}_1}$:Average of observation data

$$\hat{\mathrm{T}}_1 = \acute{\mathrm{T}}_1 + \lambda \tag{9}$$

- $\acute{\mathrm{T}}_1$:Value acquired through imitation
- $\hat{\mathrm{T}}_1$:Value of the imitation to which compensation was applied

The result of simulation showed carrying out behavior which is appropriate to /a/, /i/, and /e/.However, it is thought that it was not able to confirm since the physiological utterance model did not correspond to what mainly speaks only in one utterance organ which is changed in case lips also articulate like /u/and /o/.

4 Summary

In this research, the coarticulation based on articulation movement was analyzed.Consequently, the same result as precedence research was brought.Moreover, it was shown clearly that vowel's articulatory movement is not quiet, and its influence is storng relatively but consonant's articulatory movement is quiet and its influence is not storng relatively. Furthermore, it was shown clearly that we can reappears positions of articulation which considers coarticulation through this simulation although there is exceptional phonemes.