

## A Model of Coarticulation

Minoru SHIGENAGA, Kiyoshi TANAKA, Hitoshi ARIIZUMI and Kenji SAITO

(Received October 26, 1970)

### Synopsis

In the study of speech, it is one of the principal problems to investigate the mechanism of context in speech and to express it by rule. We have constructed a simple rule for the transition of vocal tract area functions. The rule assumes target configurations  $S_1, S_2, \dots$  for each phoneme, and express the area function  $S$  on the way of transition between  $S_1$  and  $S_2$  as follows:  $S = S_1 + (S_2 - S_1) \cdot T(n)$  ( $n$  = normalized time,  $T(n)$  = normalized locus of second formant frequency  $F_2$  between  $S_1$  and  $S_2$ ). For the transition among  $S_1, S_2$  and  $S_3$ , the locus of  $F_2$  is assumed as the product of the locus of  $F_2$  between  $S_1$  and  $S_2$  ( $T_1(n)$ ) and the one between  $S_2$  and  $S_3$  ( $T_2(n)$ ), and on the articulatory level it is expressed as follows:  $S = S_1 + [S_2 + (S_3 - S_2) \cdot T_2(n - n_a) - S_1] \cdot T_1(n)$  ( $n_a$  = instant of beginning to carry out the command to move on to  $S_3$ ). It has been shown that the results calculated by our rule coincide good enough with the observed effects on  $V_1 V_2 V_1, V_1 C V_2$  and  $CVC$  context ( $V$  = vowel,  $C$  = stop or nasal) and that the results are available for phoneme discrimination.

### 1. Introduction

In the study of speech, it is one of the fundamental problems and plays an important role to make clear the mechanism of coarticulation in speech and express it by rule. This paper is concerning with how the coarticulation effects appear in the context of  $V_1 V_2 V_1, V_1 C V_2$ , etc. ( $V$  = vowel,  $C$  = consonant) and how we should move the vocal tract area function in order to realize the effects. Further this rule may contribute to the discrimination among the cluster of like phonemes.

The problem of coarticulation reported in the past was in the aspects of frequency spectrum<sup>1)</sup>, and when it is discussed on articulatory level<sup>2)</sup> the relation with the frequency spectrum is not clearly solved. Such a problem may not be solved on the acoustical level only but needs to investigate the movement of articulatory organ. However it is not easy to take X-ray photograph, moreover it is not yet enough precise to decide the vocal tract area function from the lateral cineradiograph. Therefore, it is

necessary to make a model for the transition of area function. The writers have made efforts to make a model and the formant patterns calculated by the model have become qualitatively to coincide with the observed ones and to be able to express the coarticulation effect.

### 2. A model for the transition of vocal tract area function—A rule of coarticulation

Our model is made up as follows:—each phoneme has the proper target area function (target configuration) and uttering phonemes in succession is expressed by combining the target configurations of each phoneme by rule. That is, if the target configurations of two phonemes are expressed as  $S_1(x), S_2(x)$  ( $x$  = distance from glottis) respectively then the area function  $S(x)$  on the way of transition from  $S_1(x)$  to  $S_2(x)$  is shown as follows:-

$$S^{1/p}(x) = S_1^{1/p}(x) + [S_2^{1/p}(x) - S_1^{1/p}(x)] \cdot n \quad (1)$$

$n$  = normalized time,  $0 \leq n \leq 1$

Make  $p$  be equal to 2~3 and the second formant

frequencies corresponding to  $S(x)$ ,  $S_1(x)$  and  $S_2(x)$  be  $F$ ,  $F^1$  and  $F^2$  respectively, then the following relation is approximately realized.

$$F = F^1 + (F^2 - F^1) \cdot n \quad (2)$$

This relation is also realized approximately for the transition of first formant frequency and has been also applied to the transitions between any consonant and vowel. Accordingly by making  $p=3$  and using  $T(n)$  instead of  $n$  in Eqs. 1 and 2, which is a normalized function of  $n$  coinciding approximately with the  $F_2$  curve, the following equation is obtained.

$$\begin{aligned} S^{1/3}(x) &= S_1^{1/3}(x) + [S_2^{1/3}(x) - S_1^{1/3}(x)] \cdot T(n), \\ 0 &\leq T(n) \leq 1 \end{aligned} \quad (3)$$

Moreover by adding the term of complementary functions, the following equation is formed for the transition between two phonemes.

$$\left. \begin{aligned} S^{1/3}(x) &= S_1^{1/3}(x) + \{ [K(n, x) \cdot S_2(x)]^{1/3} \\ &\quad - S_1^{1/3}(x) \} \cdot T(n, x) \\ T(n, x) &= T(n) \cdot DLH(n, x) \end{aligned} \right\} \quad (4)$$

Where  $DLH(n, x)$  and  $K(n, x)$  express the inherent features for the coarticulation between the phonemes.<sup>8,9)</sup>

Next let's consider the case uttering three phonemes successively. The target configurations of each phoneme are represented as  $S_1$ ,  $S_2$  and  $S_3$ , then the  $F_2$  curve may be expressed as the product of the normalized time function  $T_1(n)$ , which is the normalized  $F_2$  curve drawn on the way of transition from  $S_1$  to  $S_2$ , and the function  $T_2(n-n_a)$ , which is the normalized  $F_2$  curve for the transition from  $S_2$  to  $S_3$  and  $n_a$  is the instant of beginning to carry out the command to move on to  $S_3$  on the way of transition from  $S_1$  to  $S_2$ . That is

$$\left. \begin{aligned} T_1(n) \cdot \{1 - cT_2(n-n_a)\} \\ T(n) &= 0.0 \text{ for } n \leq 0.0, \\ T(n) &= 1.0 \text{ for } n \geq 1.0 \end{aligned} \right\} \quad (5)$$

where  $c$  is the coefficient of transformation between the normalized  $F_2$  curve from  $S_1$  to  $S_2$  ( $T_1(n)$ ) and the normalized  $F_2$  curve from  $S_2$  to  $S_3$  ( $T_2(n)$ ). This rule has been concluded for the  $F_2$  curve drawn by

successive utterance of three vowels  $V_1 V_2 V_1$ , but we assume that to every contexts this rule is applicable.

Rewriting this relation by formant frequencies and designating the  $F_2$ 's of  $S_1$ ,  $S_2$ ,  $S_3$  and  $S$  as  $F^1$ ,  $F^2$ ,  $F^3$  and  $F$  respectively, we get the following equation.

$$\left. \begin{aligned} \frac{F^1 - F}{F^1 - F^2} &= T_1(n) \cdot \left[ 1 - \frac{F^3 - F^2}{F^1 - F^2} \cdot T_2(n-n_a) \right] \\ \therefore F &= F^1 + [F^2 + (F^3 - F^2) \cdot T_2(n-n_a) \\ &\quad - F^1] \cdot T_1(n) \end{aligned} \right\} \quad (6)$$

This equation is interpreted as follows:—instead of the target frequency  $F^2$  for the transition from  $S_1$  to  $S_2$ , the  $F_2$  on the way from  $S_2$  to  $S_3$ , that is  $F^2 + (F^3 - F^2) T_2(n-n_a)$ , is regarded as the target frequency for the transition from  $S_1$  to  $S_2$ .

While, if we use Eq. 3 for the transition of the area function, the relation between  $F_2$  curve calculated from the area function and the time function for the transition of the area function becomes almost linear. So the area function  $S$  among  $S_1$ ,  $S_2$  and  $S_3$  is shown as follows:—

$$\begin{aligned} S^{1/3} &= S_1^{1/3} + [S_2^{1/3} + (S_3^{1/3} - S_2^{1/3}) \cdot T_2(n-n_a) \\ &\quad - S_1^{1/3}] \cdot T_1(n) \end{aligned} \quad (7)$$

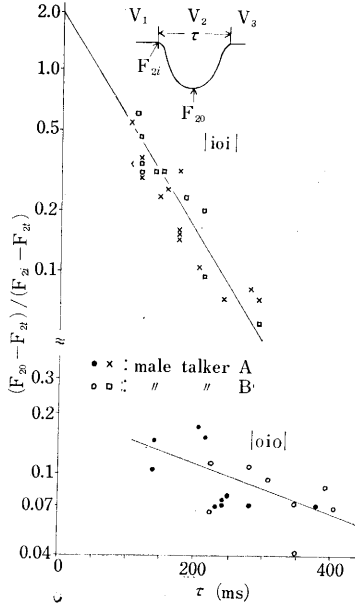
In the case of more than three phonemes, the above relation may be easily applicable on expanded forms.

Moreover in the case of context of three vowels  $V_1 V_2 V_1$ , if the  $F_2$  is traced and each symbol is defined as shown in Fig. 1 and  $F_{2t}$  is the target frequency of the second formant frequency of  $V_2$ , the following relation has been confirmed by tracing the  $F_2$  curves for various  $V_1 V_2 V_1$ , for example as shown in Fig. 1.

$$\begin{aligned} F_{20} &= \kappa(F_{2t} - F_{2t}) \cdot \exp(-\beta\tau) + F_{2t}, \\ \kappa, \beta &= \text{constants} \end{aligned} \quad (8)$$

This relation shows the effect of coarticulation by observing the variation of the maximum or minimum value of  $F_2$  curve, and have to coincide with the result obtained by Eq. 5, consequently Eq. 7.

Now, if we approximate both  $T_1(n)$  and  $T_2(n)$  by  $[1 - \exp(-\alpha n)]$ , the maximum or minimum point of  $F_2$  curve occurs at the cross point of  $T_1(n)$  and  $T_2(n-n_a)$  at which  $n = \tau/2$ , and the value at the



**Fig. 1** Relation between  $(F_{20} - F_{2t}) / (F_{2t} - F_{2i})$  and transitional duration  $\tau$  for /ioi/ and /oio/ uttered by A and B (male).

extreme point, which is the normalized  $F_{20}$ , is shown as follows:—

$$[1 - \exp(-\alpha\tau/2)]^2 = 1 - 2 \exp(-\alpha\tau/2) + \exp(-\alpha\tau)$$

If  $\alpha\tau$  is not so small then

$$[1 - \exp(-\alpha\tau/2)]^2 \approx 1 - 2 \exp(-\alpha\tau/2) \quad (9)$$

while by Eq. 8 the normalized  $F_{20}$  is expressed as follows:—

$$\frac{F_{20} - F_{2t}}{F_{2t} - F_{2i}} = 1 - \kappa \cdot \exp(-\beta\tau) \quad (10)$$

Namely the normalized  $F_{20}$ 's in Eqs. 9 and 10, which are derived from Eqs. 5 and 8 respectively, take approximately the same form. Therefore, by moving the target configuration as shown in Eq. 7, the relation of Eq. 8 may be almost consistent.

The above equations have been made sure for vowels, but by taking the inherent features of each phoneme into considerations Eq. 7 may be applicable to the transition among any phonemes. After all, the problem of coarticulation is expressed by Eq. 4 for the transition between two phonemes

and by Eqs. 5, 6 or 7 for the transition among three phonemes quantitatively, and for the transition among more phonemes the rule may be expanded.

In the above relations there are some points which are not proved strictly but they are simple expressions and consistent enough at first approximation, and we will explain by a few examples that these relations are consistent and also useful in speech recognition.

### 3. Examples

#### 3.1 Vowels

In the case that three vowels,  $V_1 V_2 V_1$ , are uttered successively at different speeds, Eq. 8 is established fairly well if  $F_{2t}$ ,  $F_{20}$  and  $\tau$  are defined as in Fig. 1 and  $F_{2t}$  is equal to the target frequency of  $F_{20}$ .

Inversely, the vowel  $V_2$  in  $V_1 V_2 V_1$  may be distinguished by means of calculation of  $F_{2t}$ , substituting measured  $F_{2t}$ ,  $F_{20}$  and  $\tau$  into Eq. 8<sup>4)</sup>. But in this case, instead of expressing  $\kappa$  and  $\beta$  as function of  $(F_{2t} - F_{20})$ , give the time function  $T(n)$ , then  $F_{2t}$  may be also calculated by Eq. 11.

$$(F_{20} - F_{2t}) / (F_{2t} - F_{2i}) = T^2(n) \quad (11)$$

The same relation can be discussed as to  $F_1$ . And with the same calculation of  $F_{1t}$  versus  $F_{10}$ ,  $V_2$  can be distinguished even when  $V_1 V_2 V_1$  is uttered rapidly. But, there can be some cases in which inaccuracy is left in measuring  $\tau$ .

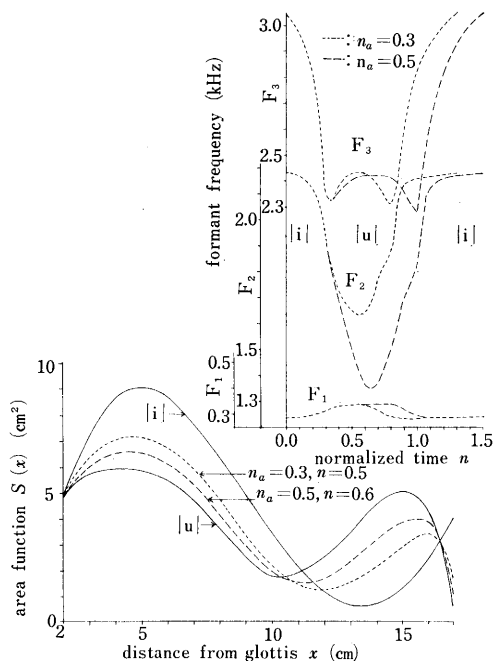
Fig. 2 shows the calculated vocal tract area functions and the formant frequencies in the cases where we make  $n_a$  in Eq. 7 be equal to 0.3 and 0.5 for the context of /i u i/. It shows that the vocal tract returns to /i/ without reaching /u/ sufficiently. And we can see that the faster the speed of the utterance (the smaller  $n_a$ ) is this phenomenon becomes clearer.

#### 3.2 Nasal consonants

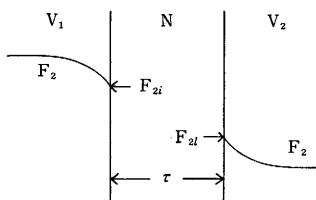
Let us consider the case that  $V_1 N V_2$  ( $V_1 = /i, o/, N = /m, n, \eta/$  and  $V_2 = /i, e, a, o, u/$ ) are uttered at different speeds. If we define  $F_{2t}$  and  $\tau$  as shown in Fig. 3 and  $F_{2t}$  designates the target frequency of  $F_{2i}$ , we will find that the next relation is consistent almost sufficiently as shown in Fig. 4.<sup>6)</sup>

$$F_{2t} = \kappa(F_{2i} - F_{2t}) \cdot \exp(-\beta\tau) + F_{2i}, \quad (12)$$

$\kappa, \beta = \text{constants}$



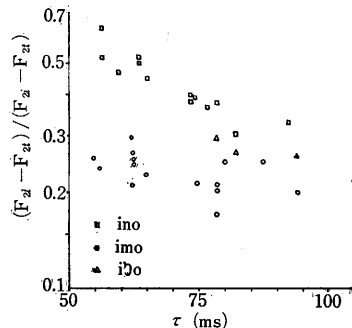
**Fig. 2** Transitions of area functions and formant frequencies of /iui/ for  $n_a = 0.3$  and  $0.5$  obtained by the model. The smaller  $n_a$  corresponds to the faster utterance.



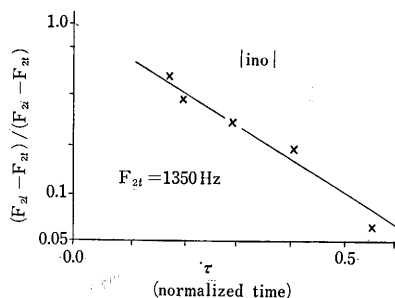
**Fig. 3** Illustration of symbols.

Moving the vocal tract area function by Eq. 7 and calculating  $F_{2l}$ ,  $F_{2t}$  and  $\tau$  from the area function, we have plotted  $F_{2t}$  according to Eq. 12 by making  $F_{2t}$  be identical with the  $F_{2l}$  of the monosyllable ( $NV_2$ ) and it is shown in Fig. 5. The result satisfies enough the actualization of Fig. 4 qualitatively.<sup>7)</sup>

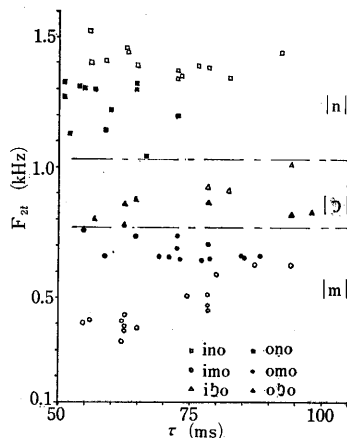
In the case of monosyllables the discrimination among the nasal consonants /m/, /n/ and /ŋ/ as for every same following vowels may be possible by the accurate extraction of  $F_{2t}$ . And if the discrimination only between /m/ and /n/ is required, even in words, it is possible on real time.<sup>5)</sup> But it is impossible to discriminate /m/, /n/ and /ŋ/ in words by the ex-



**Fig. 4** Normalized  $F_{2t}$  plotted against nasal segment duration  $\tau$  in the contexts of /ino/, /ijo/ and /imo/ (male voice).

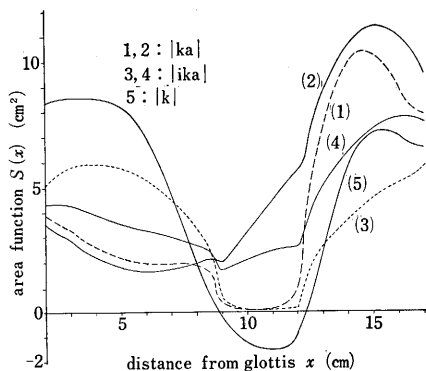


**Fig. 5** Normalized  $F_{2t}$  plotted against nasal segment duration  $\tau$  (in normalized time) calculated for /ino/ by the model.



**Fig. 6**  $F_{2t}$ 's calculated by Eq. 12 assigning 1.0 and 0.017 for  $\kappa$  and  $\beta$  respectively.

traction of  $F_{2t}$  only. Therefore, we tried the transformation by Eq. 12 of  $F_{2t}$ 's into the domain of  $F_{2l}$ . By that,  $F_{2t}$ 's of /m/, /n/ and /ŋ/ related to the same following vowel were separated into each



**Fig. 7** Area functions at the instant of /k/-explosion [(1) and (3)] and the onset of the following vowel /a/ [(2) and (4)] for /ka/ and /ika/ ( $n_a = 0.225$ ) respectively calculated by the model. Curve (5) is the target configuration of /k/. Curves (3) and (4) show the influence of the preceding vowel /i/ upon /ka/.

nasal domain and the discrimination among them became possible as shown in Fig. 6.<sup>6)</sup>

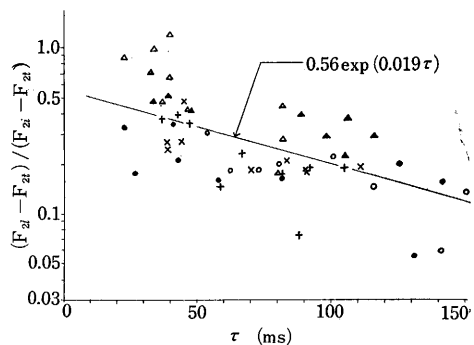
### 3.3 Stop Consonants

In the case of stop consonants, we also move the vocal tract area function according to Eq. 7, which is now under particular considerations taking account of the features of stops i.e. the existence of the period of vocal tract closure and the rapid movement of the place of constriction just after the instant of explosion. Fig. 7 shows the area functions at the instant of explosion and the onset of the following vowel /a/ for /ka/ and /ika/ calculated by the model. They show the effect of coarticulation by /i/ upon /ka/.

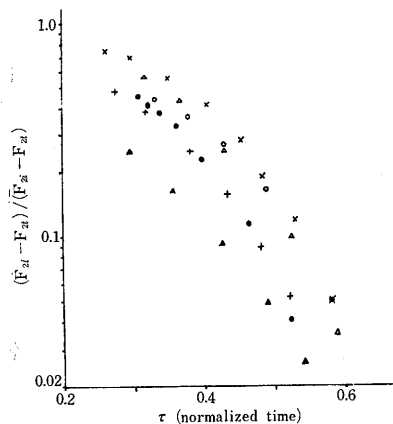
Now we compare the actual values measured as to the influences of coarticulation on  $V_1 C V_2$  and  $C V C$  ( $C = \text{stop}$ ) with the results of calculation by our model.

(i) In the case of  $V_1 C V_2$

For  $V_1 C V_2$  ( $V_1 = /i, o/, C = /k, t, p/, V_2 = /a/$ ) uttered at different speeds we have measured the second formant frequencies at the onset of  $V_2$  ( $F_{2l}$ ). The  $F_{2l}$ 's of /ka/, /ta/ and /pa/, which are separated in the case of monosyllables, overlap each other because of the coarticulation. But when we investigate the relation between  $\log [(F_{2l} - F_{2t}) / (F_{2i} - F_{2t})]$  and  $\tau$ , defining  $F_{2l}$ ,  $F_{2t}$  and  $\tau$  as same as in Fig. 3 (putting



**Fig. 8** Normalized  $F_{2l}$  plotted against consonant segment duration  $\tau$  for various contexts of  $V_1 C V_2$  ( $V_1 = /i, o/, V_2 = /a/, C = /k, t, p/$ ) (male voice). In Fig. 8 and 9  $\circ$ : /ika/,  $\cdot$ : /oka/,  $\triangle$ : /ita/,  $\blacktriangle$ : /ota/,  $+$ : /ipa/ and  $\times$ : /opa/.



**Fig. 9** Normalized  $F_{2l}$  plotted against consonant segment duration  $\tau$  (in normalized time) calculated by the model. Notations are the same as in Fig. 8.

$C$  in place of  $N$ ), we find almost linear relation as shown in Fig. 8. So the next relation can be concluded as same as in the case of nasals.

$$F_{2l} = \kappa(F_{2i} - F_{2t}) \cdot \exp(-\beta\tau) + F_{2t} \quad (13)$$

$\kappa, \beta = \text{constants}, F_{2t} = \text{target of } F_{2l}$

We have calculated  $F_{2l}$ 's inversely from the measured  $F_{2i}$ ,  $F_{2t}$  and  $\tau$ , fixing  $\kappa$  and  $\beta$  to the mean values for each context as shown in Fig. 8. Then  $F_{2l}$ 's are divided into each domain of /ka/, /ta/ and /pa/, which make possible to distinguish each other.<sup>8)</sup>

Our model also shows similar results with the

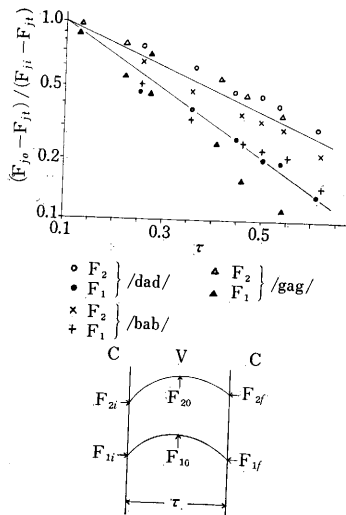


Fig. 10 Normalized  $F_{j0}$  ( $j=1,2$ ) plotted against vowel segment duration  $\tau$  (in normalized time) calculated for CVC contexts by the model.

above observation as shown in Fig. 9<sup>9</sup>). So we may conclude that our model is able to represent the influence of the coarticulation well enough.

(ii) In the case of CVC

Dr. B. Lindbrom<sup>1)</sup>, according to his detail experiments, confirmed the next relation to be almost consistent, defining the symbols as shown in Fig. 10.

$$F_{j0} = \kappa(F_{jt} - F_{jt}) \cdot \exp(-\beta\tau) + F_{jt}, \quad (14)$$

where,  $j=1,2$ ,  $\kappa, \beta = \text{constants}$ ,

$F_{jt} = \text{target of } F_{j0}$

Calculating by our model,  $F_{jt}$ ,  $F_{j0}$  and  $\tau$  for the various contexts of  $C=[g,d,b]$  and  $V=[a]$ , the relations between  $\log [(F_{j0} - F_{jt}) / (F_{jt} - F_{jt})]$  and  $\tau$  at various speeds of the utterance become as shown in Fig. 10<sup>9</sup>). These results satisfy well the ones of the experiments shown in the reference 1)-(ii) except that the correction of  $T(n)$  in /bab/ is a little small and that we have dealt with  $F_{10}$  as same as  $F_{20}$ .

#### 4. Discussion and Conclusion

We have shown some examples that the effects of coarticulation upon the second formant frequencies of  $V_1 V_2 V_1$ ,  $V_1 C V_2$  ( $C=\text{stops or nasals}$ ) and  $C V C$  ( $C=\text{stops}$ ) are represented by Eqs. 8, 12, 13 and 14 and that, inversely, if the target frequencies are calculated by these relations, even in the case that

the frequency spectrum is considerably deformed by the influence of the coarticulation, the phonemes can be distinguished. And we have shown that these relations may be realized almost sufficiently by the way that the transition between two phonemes is represented by Eq. 4 and among three phonemes simply by Eqs. 5 or 6 in the frequency domain and by Eq. 7 in the articulatory level.

In the articulation discussed here it is assumed that the command from the brain always works in order to lead the area function to the new target and that the performance of the command aiming at the new target will be done with invariant effort. But it may be hard to assume that the same effort is made even in short sentences, because of stress or intonation activity. And the phenomenon like 'laziness' may take place. So we should take such influences into consideration in our model.

Although the concept of  $F_2$  locus is advocated as for the stop consonant and our model can give the frequency which is similar to the  $F_2$  locus given up to now, it is difficult to consider that the way obtaining the  $F_2$  locus discussed up to the present gives the frequency corresponding to the real articulation of the stop consonant as initially intended. According to our model, however, the formant frequency corresponding to the target configuration of the stop consonant (we call it target frequency) can be decided as  $F^1$  calculated by substituting  $F$  and  $F^2$  into Eq. 2.<sup>10)</sup> (There is a premise that Eqs. 1 and 2 still hold formally during the closure of vocal tract.)

With the use of this target frequency, we can represent the effect of coarticulation upon the case containing not only vowels but also both vowels and consonants simply by the forms like Eqs. 5 or 6. But the study of coarticulation should be discussed at the articulatory level and by doing so we can solve simultaneously the effect on each formant and antiformant frequencies of the vocal tract.

But in order to proceed the study, it is necessary to decide the accurate target configuration of each phoneme and to get accurate data of the vocal tract area function during the utterance, which may make us improve our model further.

Since lessening  $n_a$  than 0.3 corresponds to a very rapid utterance, in the utterance at natural speed it is enough to consider that the two preceding and the two following phonemes have the influence of coarticulation on the phoneme. And in the case of rapid utterance, it will be enough to take three phonemes which precede or follow into consideration.

From the viewpoint of physiological level our model does not explain sufficiently the option of complementary terms used in it or some parameters proper to context, but they represent well the physiological constraints of articulatory manner and the formant pattern calculated by the model coincides sufficiently well with the observed one.

Furthermore, the transition of the area function may be decomposed into the movements of each articulatory organ such as tongue, lips and mandible. And we are now constructing a new model<sup>1D</sup> in which these organs move by some command such as time optimal control satisfying the acoustical demand according to a sequence of discrete phonemic instructions.

Although many discussions are made in the study of the recognition of speech, we consider it reasonable to transform the observed patterns to the target frequency domain by the method such as mentioned in this article and, if possible, into the recognition space of target configurations.

### References

- 1) (i) K.N. Stevens, A.S. House and A.P. Paul: Acoustical Description of Syllabic Nuclei: An Interpretation in Terms of a Dynamic Model of Articulation, *J. Acoust. Soc. America*, **40**, 1, p. 123-132 (1966).  
 (ii) B. Lindblom: Spectrographic Study of Vowel Reduction, *J. Acoust. Soc. America*, **35**, 11, p. 1773-1781 (1963).  
 (iii) S.E.G. Öhman: Coarticulation in VCV Utterances: Spectrographic Measurements, *J. Acoust. Soc. America*, **39**, 1, p. 151-168 (1966) etc.
- 2) (i) S. E. G. Öhman: Numerical Model of Coarticulation, *J. Acoust. Soc. America*, **41**, 2, p. 310-332 (1967).  
 (ii) R. A. Houde: A Study of Tongue Body Motion during selected Speech Sounds, SCRL Monograph No. 2, Speech Communications Research Laboratory, INC. (1968-8).  
 (iii) J. S. Perkell: Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study, The MIT Press, Cambridge (1969) etc.
- 3) M. Shigenaga, H. Ariizumi and T. Tanaka:  
 (i) A Model for Transitions of Vocal Tract Area Functions of Vowels, Report of Committee on Automaton, Inst. Electronics and Communication Eng. Japan, (1968-2).  
 (ii) 6th International Congress on Acoust.: A Model for Transitions of Vocal Tract Area Functions, B-5-9 (1968-8).
- 4) H. Ariizumi and M. Shigenaga: On the Transition of the Second Formant Frequency in Three Vowels  $V_1$   $V_2$   $V_1$ , Report of Joint Meeting of Elect. Eng. No. 2951 (1967-4).
- 5) M. Shigenaga and H. Ariizumi: Automatic Recognition of Nasal Consonants, *J. Acoust. Soc. Japan*, **21**, 5, p. 263-271 (1965).
- 6) H. Ariizumi and M. Shigenaga: On the Discrimination of Nasal Consonants, Report of Faculty of Eng., Yamanashi Univ. No. 17, p. 126-138 (1966-12).
- 7) H. Ariizumi, T. Tanaka and M. Shigenaga: Transitions of Vocal Tract Area Functions for Nasals, Report of Meeting of Acoust. Soc. Japan, p. 47 (1968-4).
- 8) F. Ogawa and T. Kobayashi: Effects of the Preceding Vowels upon Voiceless Stops, Graduate Thesis of Dept. of Elect. Eng., Yamanashi Univ. (1967-3).
- 9) T. Tanaka and M. Shigenaga: Transition of Vocal Tract Area Functions for Stop Consonants, Report of Committee on Speech, Acoust. Soc. Japan (1968-12) and *J. Acoust. Soc. Japan*, **25**, 3, p. 144 (1969-5).
- 10) M. Shigenaga, T. Tanaka and Y. Sugiura: On the Second Formant Target Frequency of the Stop Consonant, Report of Faculty of Eng., Yamanashi Univ. No. 20, p. 135-139 (1969-12).
- 11) H. Ariizumi and M. Shigenaga: On the Movement of Articulatory Organs, Report of Committee on Speech, Acoust. Soc. Japan (1969-11).