TIMBRE STUDY ON MISHEARING PHENOMENA OF SPEECH PHONES IN ROTATIONAL SYNCHRONOUS DISTORTION

(REPORT I: MISHEARING OF VOWELIC TIMBRE)

Yoshiyuki OCHIAI and Noriaki IZUMITACHI

Department of Electrical Engineering

(Received May 13, 1955)

Introduction

To make a full study of formant theory in speech communication, investigations must be done both from a subjective viewpoint as well as a physical viewpoint. Considered from our previous experiences in timbre analysis of vowel, the determination of position and intensity of formant in a physical way was insufficient to interpret the so-called vowel quality. For that purpose we need to follow out a further study on position and intensity of formant in a subjective sense. Articulation-quality experiment in low-cut and high-cut distortions may afford *one* way in that direction. But in that case we should go on into a subjective study on phoneme-timbre — a mishearing study in LCD and HCD — not satisfying ourselves with a merely statistical data of articulation measurements. Strictly speaking, this experiment is rather suitable to detect the subjective intensity, i.e., intensiveness (or, expressed in another term of quality theory, *importance*) of formant. For subjective detection of position of formant peak-center, we should devise *another* way of experiment. Thus, by introducing the so-called transitional distortion to the timbre signal with some formant-structure and by observing the resulted effect in a subjective sense, we endeavoured to make a subjective study on formant position. This is our research scheme in timbre-study attempted by utilizing the distortion RSD.

General Consideration

The problem of mishearing is, in our opinion, nothing more than a problem on hearing of timbre change. When there is no physical cause for timbre change, then the mishearing of phonemes, if occured, is merely a casual event, though it may suggest something on similarity between phonemes. But when a phenomenon of mishearing takes place in cases where there are reasonable circumstances for it, that is, there is a sufficient and suitable distortion, we are duly persuaded that the phenomenon of mishearing thereby deserves not only a careful observation but also a scientific treatment. A faithful description of the phenomenon will be of service to the study of timbre change, and a consideration on the device of the distortion applicable will be led to the idea of timbre-matching. We thus take the percentage of mishearing as a measure of judgment on timbre change. We shall describe below some basic principle that governs positively the mishearing phenomena.

Mishearing of single-formant phonemes

Suppose any existence of phonemes that have really only single formant or that are susceptible to be considered as such, and further suppose that we have here two phonemes X and Y of which formants are illustrated in the frequency range in Fig. 1. If X-formant is lower than Y-formant, then there will be a possibility of mishearing of X-phoneme for Y-phoneme only in U (up) distortion of RSD, and a mishearing of Y-phoneme for X-phoneme only in D (down) distortion of RSD. Further we can infer: if the degree of similarity of phonemes depends utterly on the relative interval of each formant-position which is controllable by RSD, then the phenomena of mishearing will be monotonously increasing only in one direction. We call this type of mishearing *monotonously increasing one-way* phenomenon. And quite similar phenomenon as X-phoneme in U-domain will be repeated in D-domain by Y-phoneme. We can say that the phenomena between phonemes X and Y are *reciprocal*. Such is the characters of mishearing between two phonemes of single formant.

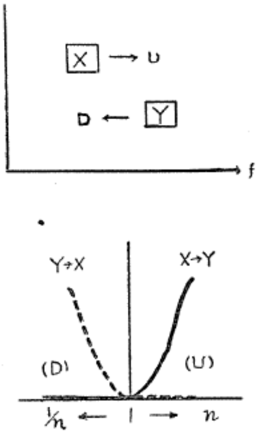


FIG. 1. Illustration of mishearing of single-formant phones, showing one-way type of mishearing which used to happen reciprocally.

Mishearing of phonemes of multi-formants

Mishearing of this type is very complicated, so that to get their general solution we must deal, in convenience, with a simplified case only, i.e., the case where the formants of two phonemes are arranged relatively in a symmetrical way. If we assume that each part of splitted formants is eqal in importance, then we have a *symmetrical* type of *two-way* mishearing — but not of seriously increasing type — as shown in Figs. 2 and 3. The former is a mishearing between single-formant and double-formant phonemes, and the latter between two double-formant phonemes. If the position of formants is not arranged quite symmetrically or the importances of two formants are not equal, then the mishearing comes to be *uusymmetric* as shown with a simplified example

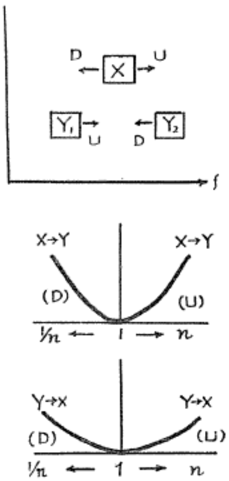
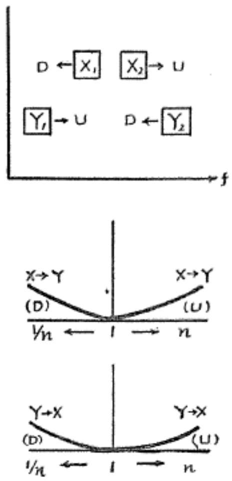
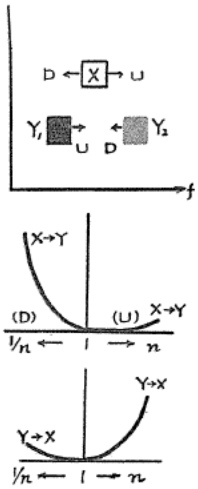



FIG. 2 (*left*). Illustration of symmetrical two-way type of mishearing between single-formant and double-formant phones.

FIG. 3 (*right*). Illustration of symmetrical two-way mishearing between double-formant speech phones.



in Fig. 4. It is to be noted that, in this case, the *reciprocity* of phenomena between phonemes are more or less still maintained, though the *symmetricity* of phenomena is almost lost. Considered from our ordinary experiences, the mishearing of this type deserves to be noticed. Because, there is few cases that the importances of each formant of double-formant phonemes are equal. In Fig. 4, for example, we illustrate an instance where Y_1-formant has higher importance than Y_2-formant.

FIG. 4. Illustration of unsymmetrical two-way mishearing which used to happen reciprocally.

Symbolic Method for Description of Mishearing

For the simplicity of description of so complicated phenomena such as mishearing, we can advantageously adopt a symbolic method. When the phone X is confused with the phone Y in the distortion D, we denote this fact by the expression.

$$\boldsymbol{H} \cdot \mathrm{D}(X) = \boldsymbol{H}(Y)$$

where we cannot neglect principally the sign of subjective operation $\boldsymbol{H}$ because of its subjective nature of hearing. If the judgement in this case is mainly based upon timbre judgment on phoneme, we can describe more precisely by

$$\boldsymbol{H}_{t_1} \cdot \mathrm{D}(X) = \boldsymbol{H}_{t_1}(Y),$$

remembering that here is brought into play the first kind of timbre judgement. But our concerns are, for the time being, with a simplified description, not with a precise description. For that purpose we denote the above fact by $\mathrm{D}(X) = (Y)$ or $X \xrightarrow{(D)} Y$ only. If it is needful to differentiate the up distortion (U) from down distortion (D), we must add distortion-suffix in the notation. Therefore, we can denote the symmetrical two-way mishearing by $X \xrightarrow[(D)]{(U)} Y$, and the reciprocal one-way mishearing by $X \xrightarrow{(U)} Y$ and $Y \xrightarrow[(D)]{} X$, or in a combined manner by $X \underset{(D)}{\overset{(U)}{\rightleftarrows}} Y$.

We had better prepare a description symbol for case where the mishearing of X for Y never happen: $X \nrightarrow Y$. Because we must be careful none the less for phenomena that never happen. Not only a phenomenon liable to happen but also a phenomenon never happen or happens only with much difficulty are both very important for us.

Mishearing of Vowel-Sounds

The mishearing phenomena of Japanese vowel-sounds are illustrated in Fig. 5. Observations and measurements were made at seven steps of distortion, i.e., $r.r.=$

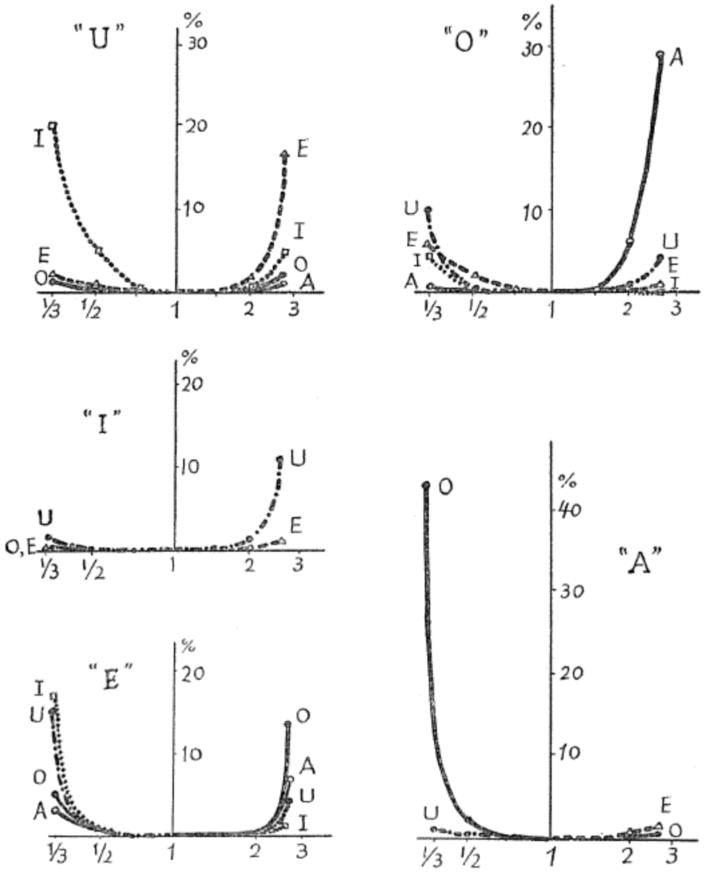


FIG. 5. Mishearing characteristics of five Japanese vowels; ordinate shows the percentage of mishearing which the vowel given in head of each figure is heard amiss to the vowels inscribed on each curve in figure.

1/3, 1/2, 3/4, 1, 3/2, 2, 8/3. As numerical values of mishearing are taken percentages of mishearing, because the numbers of observation per vowel are not equal among five vowels. Generally the mishearing does not happen in small range, that is, at the steps of $r.r. = 3/4$, 1, 3/2, but beginning to rise at steps of condition extending beyond the above range, come to set up their character only at extreme steps of $r.r. = 1/3$ and $r.r. = 8/3$. As for some vowels, their characters are not clear enough in this range of distortion. For instance, for vowel "A" it looks necessary to add further condition of $r.r. = 4$, and for vowel "I" further condition of $r.r. = 1/4$.

By the result of this figure, we see firstly as the most clear feature: $\mathrm{O} \underset{(D)}{\overset{(U)}{\rightleftarrows}} \mathrm{A}$, i.e., there is a cleancut one-way, reciprocal mishearing between "O" and "A," as shown in Fig. 6. And secondly $\mathrm{U} \underset{(D)}{\overset{(U)}{\rightleftarrows}} \mathrm{E}$, i.e., there is also a one-way reciprocal character, but in slighter degree, between vowels "U" and "E." Needless to say: these results indicate, of course, that vowel "A" is higher in its position than

"O" which is reasonable from the result of vowel analysis, and further vowel "E" is higher than "U," which must be taken for in a subjective and collective sense. The above-mentioned two phenomena will be possibly observable in the concerned vowels of foreign language. But the following is interesting to show the special character of vowel "U" (ɯ) in Japanese pronunciation, $I \underset{(D)}{\overset{(U)}{\rightleftarrows}} U$: it means U is higher than I, naturally in a collective sense. We shall come back again to this phenomenon. It must be also noted that there is almost no mishearing between "A" and "I," and between "O" and "I," that is, A ↮ I, and O ↮ I.

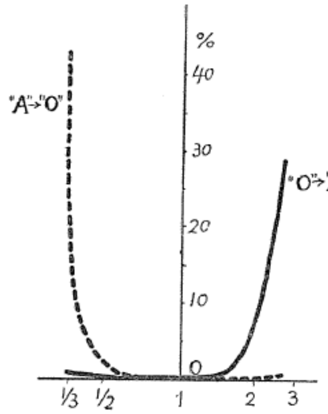


FIG. 6. Typical characteristic of one-way mishearing between vowels "A" and "O," showing clearly higher position of "A" than "O."

Effect of voices on phoneme-mishearing

In facing the subjective timbre treatment such as a problem of mishearing, we must be precautious to detect the effect of talkers on mishearing phenomena. It is rather to the relationship between vocal and phoneme qualities that we must pay attention even in the timbre treatment in a phonemic aspect. To verify the general adequacy of a given tendency in mishearing, we must obtain individual characteristics per voice (i.e., per talker) and check up the fact that the tendency also holds good no matter what the talker is. We show such individual characteristics of mishearing of vowel phones "A" and "O" in Fig. 7. We can see there almost always the one-way characteristic which happened reciprocally. Only in the phonated case of male, we see a too slowly building-up characteristic of O→A in U-domain which must be attributed to the speciality of this subject.

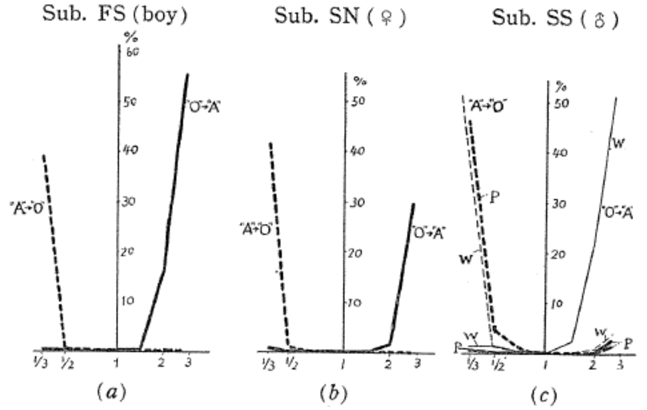


FIG. 7. Individual mishearing-characteristics per voice. (*a*) Sub. FS (boy) in phonated utterance. (*b*) Sub. SN, female voice in phonated utterance. (*c*) Sub. SS, male, where is added whispered utterance(*W*) to phonated utterance (*P*).

Mishearing of Special Phones

Mishearing of whispered vowels

In Fig. 8, we show the characteristics of mishearing of whispered vowels of male subject, compared with those of phonated (or voiced) vowels of the same subject. Generally speaking, the mishearing of phonated vowels of male utterance with a relatively low pitch is considerably different in characteristics from those of two other subjects who uttered with a relatively high pitch. But the mishearing of whispered vowels (i.e., pitchless vowels) of the male subject is rather similar to the others. This means : When the vowels of this male subject become pitchless, the characteristics of vowel-mishearing become similar to those of high-pitched vowels of other subjects.

(*a*) phonated

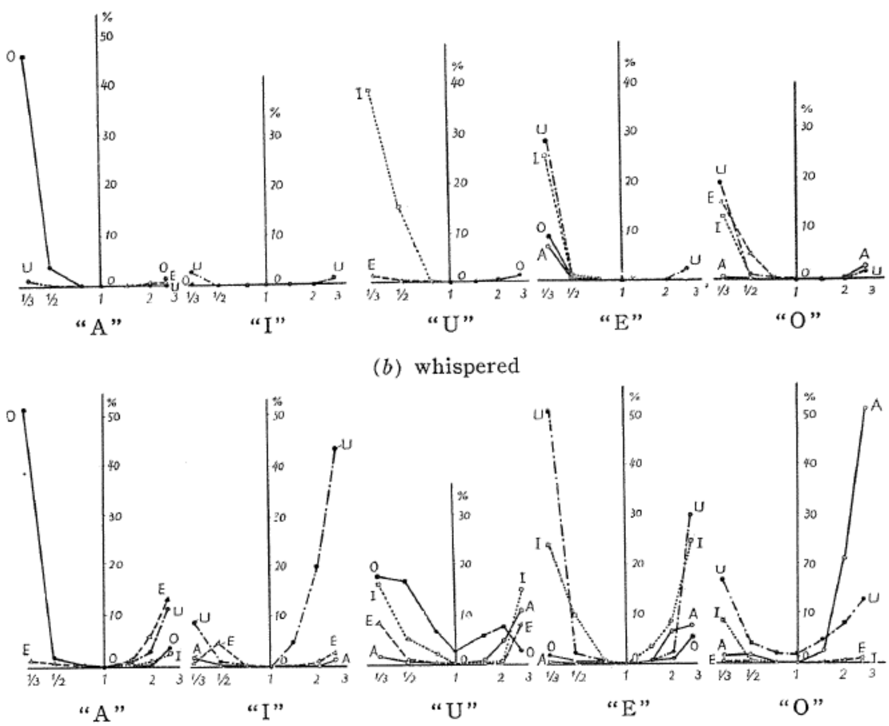


FIG. 8. Mishearing characteristics of phonated and whispered vowels of male subject.

It is to be pointed out: so far as the vowels of male subject is concerned, the mishearing of whispered vowels is conspicuously characterized by the symmetrical two-way type such as $\underset{\circ}{E} \xrightarrow[(D)]{(U)} \underset{\circ}{U}$, $\underset{\circ}{E} \xrightarrow[(D)]{(U)} \underset{\circ}{I}$, $\underset{\circ}{U} \xrightarrow[(D)]{(U)} \underset{\circ}{I}$: The first two are non-reciprocal that means $\underset{\circ}{U} \nrightarrow \underset{\circ}{E}$, $\underset{\circ}{I} \nrightarrow \underset{\circ}{E}$, and the last is reciprocal, that is, $\underset{\circ}{I} \longrightarrow \underset{\circ}{U}$. For a fuller explanation for this, we need probably further studies.

Mishearing of glided vowels

Interests of mishearing phenomena in the distortion RSD will be focussed at the problem of special phones. For example, Umlaut-vowels in German vowel-system, and neutral vowels in Anglo-Saxon and French languages may afford a very interesting problem. As special phone in Japanese language, we picked up the so-called glided-vowel system (Yoon-Gun in Japanese) and tried to observe its mishearing phenomena. We showed some examples of the results in the tables where the mishearing phenomena was measured at the following four steps of distortion $r.r.=1/3,\ 1/2,\ 2,\ 8/3$ about male-subject's voicing and whispering, and boy-subject's voicing, studying the phenomena of glided vowels in comparison with those of direct vowels. By the way, it is suitable to add some notice here that the general description of vowel-mishearing treated above is all the average result of direct and glided vowels. That is to say, the general vowels (V) are nothing but summated figures of direct vowels (V_d) and glided vowels (V_g). We shall show below in Fig. 9 two examples worthy of study. These figures show the following interesting facts: the apparent two-way phenomenon $U \xrightarrow[(D)]{(U)} I$ is explained by the superposition of two mishearing phenomena both of one-way type, i.e., $U_g \xrightarrow{(U)} I$ and $U_d \xrightarrow[(D)]{} I$, and the phenomenon of two-way type $E \xrightarrow[(D)]{(U)} U$ is also given by the superposition of a two-way type $E \xrightarrow[(D)]{(U)} U_d$ and a one-way type $E \xrightarrow[(D)]{} U_g$.

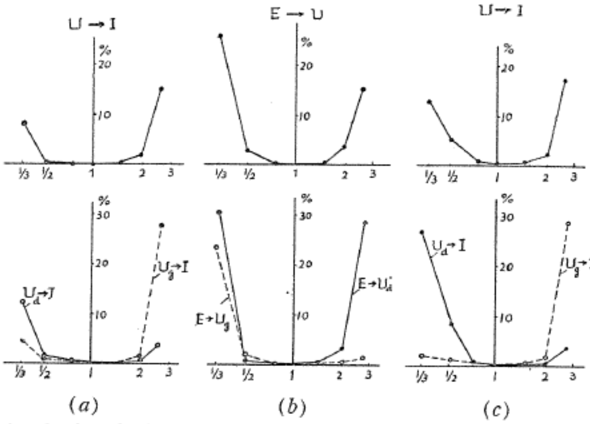


FIG. 9. Analysis of phenomena of apparent two-way mishearing, showing the special character of glided vowels in Japanese phone-system. Uppers: Mean characteristics of vowels in general. Lowers: Analytical characteristics of glided and direct vowels. (*a*): Talker SN (phonated). (*b*): Talker SS (whispered). (*c*): Talker SS (whispered).

Mishearing of semi-vowels i.e., mishearing between "m" and "n"

After the minute study we can say that the mishearing of semi-vowels [m], [n], [j], [w] is characterized by the type of *within-group* mishearing. Among them, we shall report here about two nasal phones, i.e., [m] and [n] of which the mishearing-character proved to be considerably clear and interesting as shown in Fig. 10. This is mean characteristic obtained from the observation of three listeners on three talkers' voices, based upon a total of 5256 observations per condition. It means $m \xrightarrow{(U)} n$, $m \xleftarrow[(D)]{} n$, that is, reciprocal one-way mishearing—but of an imperfect type—implying higher position of [n] phone than [m] phone in a subjective and collective sense Fig. 11 shows individual characteristics per

TABLE. Examples of Mishearing Data of Direct and Glided Vowels of Two Subjects; Numbers in Column are Percentages of Mishearing and Numbers in Bracket are Real Numbers of Occurrence of Mishearing.

Talker S.S. (Voiced)

r.r. = 0.33

Emitted→		A		I	U		E	O	
Received amiss ↓		A_d (270)	A_g (198)	(216)	U_d (234)	U_g (198)	(234)	O_d (252)	O_g (198)
A	A_d						6.4(15)		
	A_g						2.6(6)	0.4(1)	0.5(1)
I					47.5(111)	28.8(57)	26.9(63)	21.8(55)	1.5(3)
U	U_d	0.4(1)		1.4(3)			20.5(48)	20.2(51)	7.1(14)
	U_g		2.5(5)	1.4(3)			6.9(16)	1.6(4)	9.1(18)
E					3.0(7)			25.4(64)	4.0(8)
O	O_d	28.1(76)	4.5(9)	0.5(1)	1.7(4)	1.0(2)	3.8(9)		
	O_g	4.1(11)	60.6(120)				3.8(9)		

r.r. = 2.66

Emitted→		A		I	U		E	O	
Received amiss ↓		A_d (270)	A_g (198)	(216)	U_d (234)	U_g (198)	(234)	O_d (252)	O_g (195)
A	A_d							.2.4(6)	
	A_g								3.0(6)
I						0.5(1)		0.4(1)	
U	U_d	0.4(1)		1.9(4)			0.4(1)	0.2(3)	0.5(1)
	U_g		0.5(1)				1.7(4)		1.0(2)
E		0.4(1)	3.0(6)			0.5(1)			
O	O_l	0.7(2)							
	O_g		2.5(5)			0.5(1)			

r.r. = 0.5

Emitted→		A		I	U		E	O	
Received amiss ↓		A_d (585)	A_g (429)	(468)	U_d (507)	U_g (429)	(507)	O_d (546)	O_g (429)
A	A_d				0.4(2)				
	A_g								1.2(1)
I					23.5(109)	10.5(45)	1.0(5)		
U	U_d							1.6(9)	1.2(5)
	U_g								0.5(2)
E					0.4(2)	0.2(1)		8.1(44)	1.6(7)
O	O_d	2.2(13)			0.4(2)				
	O_g		4.0(17)						

r.r. = 2.0

Emitted→		A		I	U		E	O	
Received amiss ↓		A_d (500)	A_g (396)	(432)	U_d (468)	U_g (396)	(468)	O_d (504)	O_g (396)
A	A_d								
	A_g								
I					0.6(3)				0.3(1)
U	U_d							0.6(3)	
	U_g								
E			0.8(4)	0.5(2)					
O	O_d								
	O_g								

Talker S.S. (Whispered)

r.r. = 0.33

Emitted→ / Received amiss↓		A: A_d (360)	A: A_g (264)	I (288)	U: U_d (312)	U: U_g (264)	E (312)	O: O_d (336)	O: O_g (264)
A	A_d			0.4(1)			0.3(1)	1.9(6)	
	A_g			1.4(4)		3.0(8)	0.6(2)		1.5(4)
I					26.3(82)	1.9(5)	25.3(79)	16.1(54)	0.4(1)
U	U_d	4.2(15)	0.8(2)	6.6(19)			29.8(93)	20.2(78)	1.9(5)
	U_g	0.6(2)	2.7(7)	2.1(6)			22.8(71)	3.6(12)	3.0(8)
E		0.6(2)		0.7(2)	15.4(48)	0.8(2)		0.9(3)	
O	O_d	35.8(129)	5.3(14)		19.9(62)	1.5(4)	1.0(3)		
	O_g	8.6(31)	56.8(150)		2.6(8)	10.6(28)	1.0(3)		

r.r. = 2.66

Emitted→ / Received amiss↓		A: A_d (360)	A: A_g (264)	I (288)	U: U_d (312)	U: U_g (264)	E (312)	O: O_d (336)	O: O_g (264)
A	A_d			1.0(3)	13.8(43)	3.4(9)	7.4(23)	43.5(146)	21.2(56)
	A_g			0.7(2)	1.6(5)	0.8(2)	1.0(3)	4.2(14)	34.8(92)
I		2.2(8)	4.2(11)		3.2(10)	28.0(74)	25.6(80)	0.3(1)	1.1(3)
U	U_d	4.7(17)	6.1(16)	43.1(124)			28.8(90)	15.8(53)	3.0(8)
	U_g	7.8(28)	6.8(18)	1.7(5)			1.3(4)	1.2(4)	4.8(13)
E		21.7(78)	2.7(7)	3.1(9)	6.1(19)	7.5(25)		0.9(3)	1.1(3)
O	O_d	1.6(2)	0.4(1)		2.6(8)	1.1(3)	5.5(17)		
	O_g		8.0(21)			0.4(1)	0.3(1)		

r.r. = 0.5

Emitted→ / Received amiss↓		A: A_d (360)	A: A_g (264)	I (288)	U: U_d (312)	U: U_g (264)	E (312)	O: O_d (336)	O: O_g (264)
A	A_d				1.3(4)			2.1(7)	0.4(1)
	A_g						0.3(1)		2.3(6)
I					8.3(26)	1.5(4)	12.5(39)	1.8(6)	
U	U_d	3.6(13)		0.7(2)			0.6(2)	6.9(23)	
	U_g		0.8(2)	0.4(1)			1.9(6)		1.1(3)
E				5.2(15)	1.6(5)			0.9(3)	
O	O_d	3.6(13)			25.0(78)		0.3(1)		
	O_g	0.3(1)	0.4(1)		1.3(1)	8.9(23)			

r.r. = 2.0

Emitted→ / Received amiss↓		A: A_d (360)	A: A_g (264)	I (288)	U: U_d (312)	U: U_g (264)	E (312)	O: O_d (336)	O: O_g (264)
A	A_d				6.7(21)	0.8(2)	4.5(14)	17.0(57)	11.4(30)
	A_g				1.9(6)	1.5(4)		1.5(5)	14.4(38)
I		0.3(1)	1.9(5)			1.5(4)	9.0(28)	0.3(1)	
U	U_d	1.7(6)	0.4(1)	24.6(71)			3.5(11)	14.0(47)	0.4(1)
	U_g	0.6(2)	3.4(9)				0.3(1)		0.8(2)
E		3.6(13)	8.7(23)			0.4(1)			0.4(1)
O	O_d				2.9(9)	1.1(3)	1.3(4)		
	O_g	0.3(1)	1.5(4)		1.3(4)	10.6(28)			

Talker F.S. (Voiced)

r.r. = 0.33

Emitted→ / Received amiss↓		A		I	U		E	O	
		A_d (360)	A_g (264)	(288)	U_d (312)	U_g (264)	(312)	O_d (336)	O_g (264)
A	A_d						0.3(1)	0.6(2)	
	A_g						0.6(2)		
I					15.0(50)	7.2(19)	23.4(73)	0.6(2)	
U	U_d		0.4(1)				2.2(7)	3.9(13)	4.5(12)
	U_g		1.1(3)				1.6(5)	0.3(1)	3.8(10)
E					6.7(21)			4.8(16)	
O	O_d	29.2(105)	3.8(10)		1.9(6)	0.4(1)			
	O_g	3.9(14)	45.2(119)			0.4(1)			

r.r. = 2.66

Emitted→ / Received amiss↓		A		I	U		E	O	
		A_d (360)	A_g (264)	(288)	U_d (312)	U_g (264)	(312)	O_d (336)	O_g (264)
A	A_d				1.6(5)		5.4(17)	59.0(198)	12.9(34)
	A_g				1.6(5)	1.1(3)	5.8(18)	5.7(19)	31.0(82)
I					1.0(3)	1.5(4)	1.6(5)	0.6(2)	0.4(1)
U	U_d		0.4(1)	7.0(20)			2.9(9)	4.2(14)	0.4(1)
	U_g		0.4(1)	1.4(4)			2.6(8)	0.9(3)	3.0(8)
E		0.8(3)		0.7(2)	12.5(39)	4.5(12)		0.3(1)	0.4(1)
O	O_d				7.7(24)		35.0(109)		
	O_g		1.1(3)		2.9(9)	1.1(3)	0.6(2)		

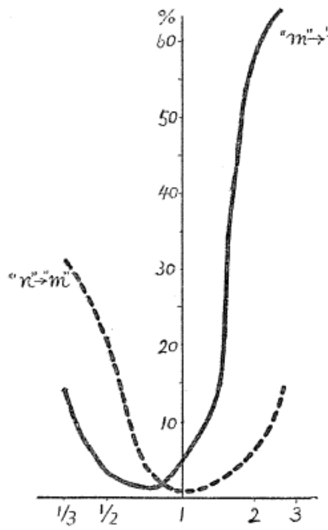
r.r. = 0.5

Emitted→ / Received amiss↓		A		I	U		E	O	
		A_d (360)	A_g (264)	(288)	U_d (312)	U_g (264)	(312)	O_d (336)	O_g (264)
A	A_d					0.4(1)		0.6(2)	
	A_g					0.4(1)			
I					0.6(2)	0.4(1)	0.3(1)		
U	U_d							0.3(1)	
	U_g								0.4(1)
E									
O	O_d				1.6(5)	0.8(2)			
	O_g				0.3(1)	1.1(3)			

r.r. = 2.0

Emitted→ / Received amiss↓		A		I	U		E	O	
		A_d (360)	A_g (264)	(288)	U_d (312)	U_g (264)	(312)	O_d (336)	O_g (264)
A	A_d							19.1(64)	1.9(5)
	A_g			0.4(1)	0.3(1)	0.4(1)	1.0(3)	0.9(3)	8.7(23)
I					0.6(2)		1.0(3)		
U	U_d			1.4(4)			0.3(1)	1.2(4)	
	U_g			0.4(1)			0.6(2)		
E		0.8(3)	0.8(2)		1.3(4)				
O	O_d				0.6(2)		1.0(3)		
	O_g		0.4(1)			0.8(2)			

voice in three phonated cases of three talkers and one whispered case of male subject.



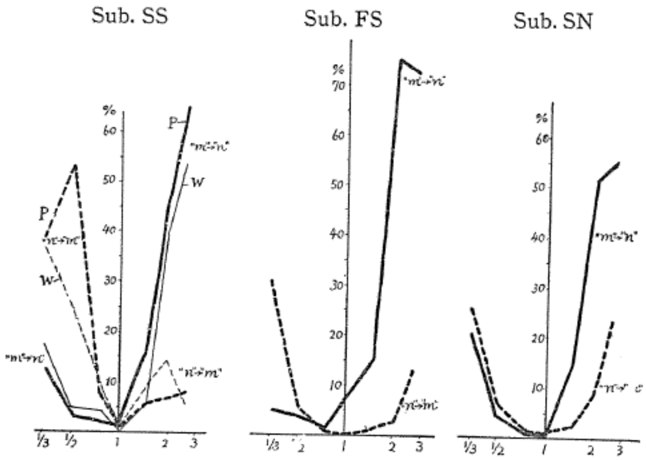


FIG. 10 (*left*). Characteristics of mishearing between phones "m" and "n," each of which needed 5256 observations for determination of one point of each characteristic.

FIG. 11 (*right*). Individual characteristics of mishearing ($m \rightleftarrows n$) per voice.

Discussion

The research of mishearing of vowels by means of the distortion RSD is of considerable significanse as a sort of the position study of vowel-formant in a subjective sense. A similar study by aid of the distortion CSD (carrier synchronous distortion) was carried out and reported previously. These two distortions are indispensable for the position study of phoneme-formant, and every one has its proper character which we must utilize in our quality study. By either distortion, for example, we could most clearly ascertain a relatively perfect reciprocal, one-way mishearing relation between vowels " O " and " A," and further, a relatively imperfect reciprocal one-way mishearing between [m] and [n]: so that we can denote $O \underset{(D)}{\overset{(U)}{\rightleftarrows}} A$, $O \underset{(-)}{\overset{(+)}{\rightleftarrows}} A$, $m \underset{(D)}{\overset{(U)}{\rightleftarrows}} n$, $m \underset{(-)}{\overset{(+)}{\rightleftarrows}} n$. (In either case the relations $m \xrightarrow{(U)} n$, $m \xrightarrow{(+)} n$ are clearly one-way, but both the relations $n \xrightarrow[(D)]{} m$, $n \xrightarrow[(-)]{} m$ are imperfectly one-way, considerably bearing thereby two-way inclination. Judged from these tendencies, the up-distortion in RSD and the (+) shift distortion in CSD are considered *forward-shifted* transitional distortions, and both the down-distortion in RSD and the (−) shift distortion in CSD are admitted as *backward-shifted* transitional distortions. But the distortion RSD is characterized by shifting of speech phones without losing harmonicity and pitchedness, and the distortion CSD by shifting of speech phones at the sacrifice of degradation in pitchedness and harmonicity. Study of whispered phones by these two distortions comes to be interesting under these circumstances.

Our original and fundamental interpretation of "mishearing-phenomena" as a kind of subjective judgement on timbre change was ascertained through these experiments, although not only the measuring technique of mishearing is rough and coarse in the present stage, but also the experimental boundary condition in introducing RSD and CSD in our experiment is not perfect (far from the ideal state of pure distortion). If we are successful in realizing an exact and precise measurement of mishearing phenomena, then we shall be in a position to determine the "subjective interval" of various formants in several phones on one hand, and further to obtain a reasonable placement of speech-phones in psychological space on the other hand.

After finishing our experiment we feel somewhat critically that there remained unsatisfactory points: Firstly, the transmission characteristics of tape-recorder was not too good to permit a pure position-study of timbre; Secondly, the shift range of transitional distortion was not wide enough to give informations on any characteristics other than monotonously increasing type to which our present observation was mainly restricted. Thirdly, in this study we were destitute of loudness-matching process before going on into the study of mishearing. It resulted probably in some imperfectness for "timbre" treatment where the problem of mishearing must be considered.