Ann. Bull. RILP (1983) No. 17, 107-113

THE INFLUENCE OF CONSONANTAL CONTEXT ON VOWEL IDENTIFICATION

Tatsuo Nakagawa, Shuzo Saito and Tomoyoshi Yoshino***

Introduction

In recent vowel identification studies, the perception of vowels in a consonantal context and in isolation has been widely discussed. A series of studies on this subject originated in the experiments performed by Strange and her colleagues.[1] They examined the effects of speaker variation and consonantal context on the identification of naturally spoken vowels. The results indicated that the effects of consonantal context were greater, namely isolated vowels tended to be identified much less accurately than vowels in a consonantal context. They explained this identification advantage for vowels in a consonantal context by saying that the subjects could utilize the additional acoustic information due to the coarticulation effects exhibited by the speech organs. In their study, they suggested that this information is contained in the formant transitions. Subsequent studies[2] have supported their findings and have suggested that duration information also seems to play a critical role in vowel identification.

Howell[3] measured the accuracy of the identification of dipthongs in isolation, in a consonantal context, and in a tonal context. In the tonal context condition, a 40 ms 1 kHz tone was introduced before or after the vowel. In comparison with the identification of isolated vowels, an improved performance was obtained for the tonal context as well as for the consonantal context. He proposed the alternative explanation that tone and consonants provide a reference context.

Several recent studies, however, have failed to support the hypothesis that vowels in consonantal contexts are more accurately identified than vowels in isolation. Macchi[4] tested the effects of vowel intelligibility with and without a consonantal context using an experimental paradigm in which the factors listening condition, dialect, and response alternatives were carefully controlled. The results indicated that subjects could identify vowels quite well whether a consonantal context was present or not.

Diel et al.[5] also could not find a contextual advantage for identifying vowels. In their experimental paradigm, subjects responded to isolated vowels or vowels in a consonantal context either by choosing from written isolated vowel syllables; or by choosing from written vowels in consonantal context syllables; or by vocally mimicking the items. An improvement in the identification of vowels in a consonantal context over isolated vowels was only observed in the second condition. They suggested that contextual effects may be due to "stimulus-response compatibility" and "memory load".

---

* Graduate Program in Special Education, Tsukuba University
** Institute of Special Education, Tsukuba University

Table 1 *A list of the 100 Japanese syllables*

pa	pi	pɯ	pe	po
ta			te	to
ka	ki	kɯ	ke	ko
		tsɯ		
sa		sɯ	se	so
ha	hi	hɯ	he	ho
pja		pjɯ		pjo
kja		kjɯ		kjo
tʃa	tʃi	tʃɯ		tʃo
ʃa	ʃi	ʃɯ		ʃo
hja		hjɯ		hjo
ba	bi	bɯ	be	bo
da			de	do
ga	gi	gɯ	ge	go
dza		dzɯ	dze	dzo
ɾa	ɾi	ɾɯ	ɾe	ɾo
wa				
ja		jɯ		jo
ma	mi	mɯ	me	mo
na	ni	nɯ	ne	no
bja		bjɯ		bjo
gja		gjɯ		gjo
dʒa	dʒi	dʒɯ		dʒo
ɾja		ɾjɯ		ɾjo
mja		mjɯ		mjo
nja		njɯ		njo
a	i	ɯ	e	o

It has been shown by Gottfried and Strange[6] that the advantage in the identification of vowels in a consonantal context over isolated vowels is not always present in every consonantal context. Vowels in /p/-vowel-/p/, /b/-vowel-/b/, /k/-vowel-/k/, /k/-vowel, and vowel-/k/ syllables were identified much more accurately than isolated vowels. But vowels in /g/ consonantal contexts were not more intelligible compared with isolated vowels and vowels in the other consonantal contexts.

The purpose of the present study was to examine the advantage for the identification of vowels in a consonantal context using the results of articulation tests. We investigated whether the contextual advantage occurred for the identification of Japanese vowels under two conditions: a single-speaker condition and a multiple-speaker condition. Using the 100 Japanese syllables, we assessed which consonantal context has an advantage or disadvantage for vowel identification.

Method

The following two experimental conditions were designed. In the single-speaker condition, four female subjects served as speakers and listeners; when one talked, the other three listened. They all spoke Tokyo dialect. None had any speech or hearing defects. They were well trained for the articulation tests. Each speaker uttered the 100 Japanese syllables shown in Table 1, and the listeners were forced to respond for every syllable. With three listeners, there were 300 responses per speaker; pooling the four speakers gives 1,200 observations for which confusions were calculated.

In the multiple-speaker condition, the lists of the 100 Japanese syllables spoken by each of the four speakers were recorded in advance. 25 syllables spoken by the same person were presented to all four listeners in random order. With four listeners there were 400 responses per speaker; and the total amount of observations for which confusions were calculated was 1,600.

The speech materials were presented through the Japanese Master Reference Telephone Transmission System. The stimuli used here were low-pass filtered at 1,000, 1,400, 1,700 and 2,000 Hz, and high-pass filtered at 2,500, 2,000, 1,700, 1,400 and 1,000 Hz.

Results and Discussion

The identification data, low-pass filtered at 2,000 Hz, were used for the analysis. The average articulation scores for the consonants and vowels were calculated separately. Table 2 shows these values for the single-speaker and the multiple-speaker conditions. It is apparent that there was no significant difference in the identification performance between the two speaker conditions. These results indicate that speaker variation does not significantly affect the articulation scores for vowels and consonants.

Since there was no significant difference in articulation scores between the two speaker conditions, a subsequent analysis was performed for the data combining the single-speaker condition with the multiple-speaker condition. Fig. 1 shows the mean articulation scores for the consonants and vowels. The vowel articulation

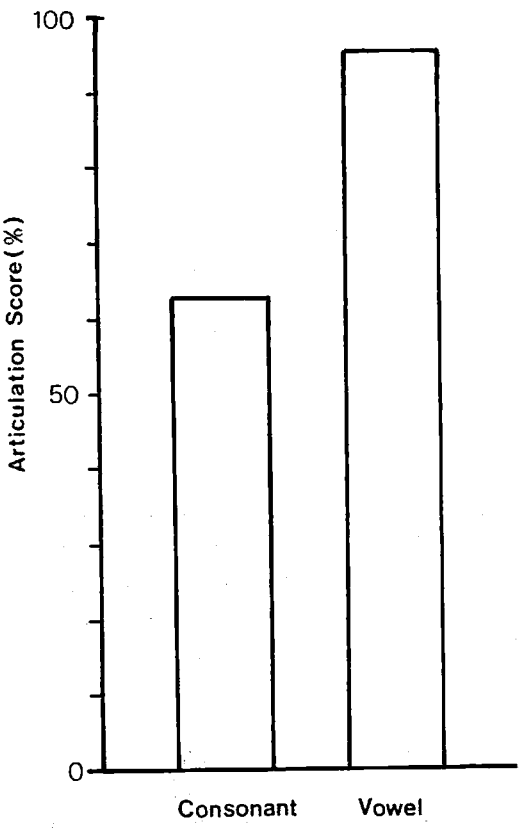

Fig. 1 *Mean articulation scores of consonants and vowels for the data combining the single-speaker condition and the multiple-speaker condition.*

Table 2 *Average articulation scores for consonants and vowels in the single-speaker and multiple-speaker conditions*

(%)

Talker	Consonant	Vowel
single－talker	63.3	95.2
multi－talker	61.8	96.3

values were much higher than the consonantal articulation values. Therefore, it can be suggested that the consonantal contexts did not always aid in the specification of the vowel identity.

To examine whether any particular consonantal context reduced the vowel articulation score an error analysis was performed. The mean error rate for the vowels was 4.2%. Table 3 presents the confusion metrices for the five vowels. As shown in the table, confusions among /i/ and /u/ occurred more often than for any other vowels. Errors on these vowels seem to be explained by the first two formants' relationship represented in a two dimensional space defined by their frequencies. The vowels /i/ and /u/ have similar first formant frequencies, but they have different second formant frequencies. As one of the present authors has reported,[7] when the frequency components above the first formant frequency are eliminated, confusions are seen between vowels having similar first formant frequencies such as /i/ and /u/, and /e/ and /o/.

Table 3 *Confusion matrices for the five vowels*

Intended Vowel	Respose i	e	a	o	ɯ	Total Error(%)
i	283	1			52	44.9
e		339	2	18	5	21.2
a		4	720	2	2	6.8
o			4	696		3.4
ɯ	18			10	644	23.7

With the identification errors for /i/ and /u/, and /e/ and /o/ excluded from the analysis, the mean error rate for vowels is reduced to 1.1%. The error rate obtained in the present study is substantially lower than that reported in previous studies.

Most of the perceptual confusions reported in the paper by Strange et al. involved vowel pairs which have similar formant frequencies. For example, / /-/ /, / /-/ /, and / /-/ /-/ / were often substituted for each other. Japanese has fewer vowels than English. The vowel space defined by the first two formant frequencies

Table 4 *Error analysis for the consonantal contexts in which the vowels were misidentified. The symbol "." means no consonantal context accompanied.*

Consonantal Context	Error Number	Percent	Consonantal Context	Error Number	Percent
p	10	8.5	dz	2	1.7
t	4	3.4	r	6	5.1
k	7	5.9	w	0	0
ts	0	0	j	0	0
s	1	0.8	m	14	11.9
h	2	1.7	n	6	5.1
pj	5	4.2	bj	3	2.5
kj	2	1.7	gj	7	5.9
tʃ	8	6.8	dʒ	7	5.9
ʃ	5	4.2	rj	4	3.4
hj	0	0	mj	5	4.2
b	3	2.5	nj	1	0.8
d	3	2.5	·	5	4.2
g	8	6.8	total	118	100

seems to be wider in Japanese than in English. Therefore, it may not be necessary to decide on subtle differences in vowel tonality for vowel identification. The consonantal contexts in which the vowels were misidentified were also examined. Table 4 shows the results of the analysis. From this table, it can be seen that there were relatively many identification errors in the /p/ and /m/ consonantal contexts. But, in general, there seem to have been no consonants that affected the vowel articulation scores. And it can not be concluded that the identification performance for the isolated vowels was poorer than that for the vowels in the consonantal contexts. In the future we will analyze the data with other filtering conditions.

Summary

The present study tested the advantage for the identification of vowels in consonantal contexts over vowels in isolation. Listening tests were administered under two conditions: a single-speaker condition and a multiple-speaker condition. The 100 Japanese syllables, low-pass filtered at 2,000 Hz, were identified by four subjects. The results of the tests indicated that there was no significant difference in the identification performance between the two speaker conditions. Moreover, the tests could not define the effects of the consonantal contexts on vowel identification.

References

1. Strange, W., R.R. Verbrugge, D.P. Shankweiler and T.R. Edman (1976); Consonant environment specifies vowel identity, J. Acoust. Soc. Am., 60, 213-224.
2. Strange, W., T.R. Edman and J.J. Jenkins (1979); Acoustic and phonological factors in vowel identification, J. Exp. Psychol., Human Percept. Perf. 5, 643-656.
3. Howell, P. (1981); Identification of vowels in and out of context, J. Acoust. Soc. Am., 70, 1256-1260.
4. Macchi, M. (1980); Identification of vowels spoken in isolation versus vowels spoken in consonantal context, J. Acoust. Soc. Am., 68, 1636-1642.
5. Diel, R.L., S.B. McCusker and L.S. Chapman (1981); Perceiving vowels in isolation and in consonantal context, J. Acoust. Soc. Am., 69, 239-248.
6. Gottfried, T.L., and W. Strange (1980); Identification of coarticulated vowels, J. Acoust. Soc. Am., 68, 1626-1635.
7. Saito, S. (1959); Study on the evaluation of the articulation score of the Japanese speech sounds, Rep. ECL, NTT, 7, 207-212.